EPFL

Thèse n° 10 631

Reprogramming G Protein-Coupled Receptor Structure, Function And Signaling By Computational Design

Présentée le 22 avril 2021

Faculté des sciences de la vie
Unité du Prof. Barth
Programme doctoral en biotechnologie et génie biologique

pour l'obtention du grade de Docteur ès Sciences

par

Dániel KÉRI

Acceptée sur proposition du jury

Prof. M. Dal Peraro, président du jury
Prof. P. D. Barth, directeur de thèse
Prof. I. André, rapporteur
Prof. C. Tate, rapporteur
Prof. B. Correia, rapporteur

École polytechnique fédérale de Lausanne

2021

Acknowledgements

I don't think many graduate students in the biosciences can say they traveled the world during their PhD. I have been lucky enough to do just that! Initially I left the deserts and beaches of California for the tropical swamps, bayous and hurricanes of Houston to start my degree at Baylor College of Medicine. After spending a few years in the tropics, the lab moved to the Alps in Switzerland, where we have a great view of Lake Geneva. It has definitely been an interesting journey!

I am grateful for the help and support from my old friends Cynthia and Sam, both of you helped me get where I am and chances are I would not be where I am now without you.

I am also grateful to have met some great friends who made it a pleasure to be in Houston. Alex, Diana, and Kehan, I will always cherish our time together; all our lunches, hikes, walks, workout sessions, drinks, talks about code and everything else! I am glad I met you all and hopefully we can see each other in person soon!

I want to thank Melvin, Mike, Xiang, and Jason for being perfect labmates. It was an honor working with you all and I am glad to have had your company in lab for your friendship and mentorship!

Thamani, you were a great welcoming party and a guide to Switzerland. You made the transition so much easier! Thanks!

And everyone else in the Barth lab, thanks for your inspiration!

I also want to thank Patrick, as none of this would have been possible without you. Thanks for the opportunities to publish great papers, and to travel to places where I may have not gone otherwise.

Anyu, köszönöm a segítségedet es a türelmet. Tudom hogy sokszor nem vedtem fel a telefont amikor hivtál. Ezen dolgoztam.

Abstract

G protein-coupled receptors (GPCRs) are 7-transmembrane alpha-helical integral membrane proteins on which cells heavily rely to receive information regarding their external environment. These receptors are able to transfer information to intracellular downstream effectors upon stimulation from their cognate ligands, which can be considerably diverse. They can include light, small molecules, peptides, protons; Due to its evolutionary success, over millions of years of evolution, this protein family has expanded to include over 800 members in humans. Despite the large size of the family and the diverse inputs they accept, the structure and the mechanism for relaying signals to their downstream partners, G proteins and β-arrestins, remains largely the same. As these receptors are involved in a great deal of biological processes such as sight (rhodopsin), cognition (dopamine, serotonin, opioid receptors), immunity (chemokine receptors), heart function (adrenergic, angiotensin receptors), ad nauseam; a great deal of effort has gone into understanding their structure and function.

Our efforts have gone into understanding their signaling properties and rationally modifying them. Most proteins are fairly flexible entities and their structures tend to oscillate and move around, sampling a number of conformations. Nature has taken advantage of these natural motions to regulate protein function from distant binding sites, in a phenomenon called allostery. GPCRs are unique in the sense that their primary function relies on allostery. By tweaking the allosteric signaling of GPCRs, we hope to have a better understand allostery, use the principles we learn to create designer GPCR biosensors with the desired sensitivity, and to eventually create *de novo* allosteric proteins.

In this work, I present our progress in this effort, which has been significant. We have created a highly accurate *in silico* method to allosterically alter relative stabilities of the active and inactive states of the GPCR dopamine D2 receptor, but also to independently control its allosteric signaling strength. With our methodology we have been able to create receptors which have high basal activities by selectively stabilizing the active state of the receptor. This has already facilitated the structure determination of the active, G protein-bound dopamine D2 receptor; a first for this GPCR. Additionally, we have created a number of dopamine D2 variants with roughly 100 times higher sensitivity to dopamine than the WT receptor. These efforts continue to characterize the functional effects of these allosteric mutations.

Furthermore, we are also working on a collaboration to rewire the signaling pathway of the

adenosine A2A receptor. Adenosine is a common immunosuppresant in solid tumor microenvironments (TMEs). By rewiring the downstream signaling of the A2A receptor, we hope to reverse the response of T-cells in these TMEs to activate and attack the solid tumor.

Keywords: G protein-coupled receptor, GPCR, G protein, allostery, protein design, biosensor, CAR T-cell, immunotherapy, Rosetta, biased signaling

Résumé

Les récepteurs couplés aux protéines G (GPCR) sont des protéines membranaires intégrales alpha-hélicoïdales transmembranaires 7 sur lesquelles les cellules dépendent fortement pour recevoir des informations concernant leur environnement externe. Ces récepteurs sont capables de transférer des informations à des effecteurs intracellulaires en aval lors de la stimulation de leurs ligands apparentés, qui peuvent être considérablement diversifiés. Ils peuvent inclure de la lumière, de petites molécules, des peptides, des protons ; En raison de son succès évolutif, au cours de millions d'années d'évolution, cette famille de protéines s'est élargie pour inclure plus de 800 membres. Malgré la grande taille de la famille et les diverses entrées qu'ils acceptent, la structure et le mécanisme de relais des signaux vers leurs partenaires en aval, les protéines G et les β -arrestines, restent largement les mêmes. Comme ces récepteurs sont impliqués dans de nombreux processus biologiques tels que la vue (rhodopsine), la cognition (dopamine, sérotonine, récepteurs opioïdes), l'immunité (récepteurs des chimiokines), la fonction cardiaque (récepteurs adrénergiques, angiotensine), ad nauseam ; beaucoup d'efforts ont été consacrés à la compréhension de leur structure et de leur fonction.

Nos efforts ont porté sur la compréhension de leurs propriétés de signalisation et leur modification rationnelle. La plupart des protéines sont des entités assez flexibles et leurs structures ont tendance à osciller et à se déplacer, échantillonnant un certain nombre de conformations. La nature a profité de ces mouvements naturels pour réguler la fonction protéique à partir de sites de liaison distants, dans un phénomène appelé allostérie. Les GPCR sont uniques en ce sens que leur fonction principale repose sur l'allostérie. En peaufinant la signalisation allostérique des GPCR, nous espérons avoir une meilleure compréhension de l'allostérie, utiliser les principes que nous apprenons pour créer des biocapteurs GPCR avec la sensibilité souhaitée, et éventuellement créer des protéines allostériques *de novo*.

Dans ce travail, je présente nos progrès dans cet effort, qui a été significatif. Nous avons créé une méthode *in silico* très précise pour modifier allostériquement les stabilités relatives des états actifs et inactifs du récepteur GPCR dopamine D2, mais aussi pour contrôler indépendamment sa force de signalisation allostérique. Avec notre méthodologie, nous avons pu créer des récepteurs qui ont des activités basales élevées en stabilisant sélectivement l'état actif du récepteur. Cela a déjà facilité la détermination de la structure du récepteur dopamine D2 actif lié à la protéine G ; une première pour ce GPCR. De plus, nous avons créé un certain nombre de variantes de la dopamine D2 avec une sensibilité à la dopamine environ 100 fois plus élevée que le récepteur WT. Ces efforts se poursuivent pour caractériser les effets fonctionnels de ces

mutations allostériques.

De plus, nous travaillons également sur une collaboration pour recâbler la voie de signalisation du récepteur de l'adénosine A2A. L'adénosine est un immunosuppresseur courant dans les micro-environnements de tumeurs solides (TME). En recâblant la signalisation en aval du récepteur A2A, nous espérons inverser la réponse des lymphocytes T dans ces TME pour activer et attaquer la tumeur solide.

Mots clés : Récepteur couplé aux protéines G, GPCR, protéine G, allostérie, conception de protéines, biocapteur, cellule CAR T, immunothérapie, Rosetta, signalisation biaisée

Contents

List of Figures

List of Tables

Glossary

α allosteric efficacy.
5HT1AR 1A serotonin receptor.
5HT2C 2C serotonin receptor.

A2AR Adenosine A2A receptor.
AC adenylyl cyclase.
ATP Adenosine triphosphate.
ATSM Allosteric two-state model.

B1AR β_1 adrenergic receptor.
B2AR β_2 adrenergic receptor.
BSA Bovine serum albumin.
BW Ballesteros-Weinstein.

cAMP cyclic adenosine monophosphate.
CART Chimeric antigen-receptor T cell.
CNO clozapine-n-oxide.

D2DR Dopamine D2 receptor.
DAG diacyl glycerol.
DDM n-Dodecyl β-D-maltoside .
DMEM Dulbecco's Modified Eagle Medium.
DOR δ opioid receptor.
DREADD designer receptor excusively activated by designer drugs.

ECL extracellular loop.
EDTA Ethylenediaminetetraacetic acid.

GDP guanosine diphosphate.
GFP Green fluorescent protein.
Gi G alpha inhibitory subunit.
GPCR G protein-coupled receptor.
Gq G alpha q subunit.
Gs G alpha stimulatory subunit.
GTP guanosine triphosphate.

H5 The G protein carboxy-terminal helix 5.
HA Hemagglutinin.

ICL intracellular loop.

KNF Koshland-Némethy-Filmer.
KOR κ opioid receptor.

MD molecular dynamics.
MSD Multi-state design.
MWC Monod-Wyman-Changeux.

NMA normal mode analysis.

PBS Phosphate-buffered saline.
PFA para-formaldehyde.
PKA protein kinase A.
PLC phospholipase C.
POPC 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine.

rGFP *Renilla* green fluorescent protein.
RLuc8 *Renilla* luciferase 8.
RU Rosetta units.

TM trans-membrane.
TME Tumor micro-environment.
TMH trans-membrane helix.
TRP Transient receptor potential.

WT wildtype.

Å Ångstrom.

1 Introduction

1.1 G Protein-Coupled Receptors

1.1.1 Significance/Relevance

There are over 800 GPCRs within the human genome, meaning the genes encoding these receptors comprise 3% of protein-coding genes[1,2]. G protein-coupled receptors serve a myriad of critical functions within multicellular organisms, including nutrient sensing and reproduction in single-celled organisms[3]. In multicellular organisms such as humans GPCRs serve to sense our environment via rhodopsin and roughly 350 different olfactory receptors; to control our moods and thought processes via dopamine, opioid, and serotonin receptors; to regulate cardiac function via adrenergic, neurotensin, and angiotensin receptors; to protect us from foreign organisms via CCR5, CXCR, and histamine receptors; just to name a few[4,5]. The involvement in these critical functions makes these receptors attractive drug targets. Consequently, around 30% of all FDA approved drugs target GPCRs[6]. Thus, a complete biochemical and structural understanding of these receptors can lead to more effective, safer, and potentially more personalized therapeutics. Furthermore, our improved understanding can also be used to reprogram GPCRs and control the processes which rely upon them, and even allow us to coerce them into novel functions.

Understanding the *in vivo* function of the protein of interest is crucial for proper protein design of pre-existing proteins. Unlike creating a *de novo* protein which has limited function, natural proteins have evolved under numerous constraints. Evolution has subjected them to a multivariable optimization problem with changing optimal conditions, co-evolving with a network of other interacting proteins all to optimize the cells' or organism's survival and response to the external environment. This network has evolved to respond to perturbed conditions by appending additional binding sites, post-translational modification sites, signal sequences, etc.; in addition to the protein's primary function. By being cognizant of these functional regions, we can modify functions we're interested in with minimal side effects.

1.1.2 Function

Although "receptive substances", molecules which bind small molecules or hormones within cells, have been known since the late 19th century, how these receptors function have only begun to be understood since around the 1950's with the discovery of the downstream second messenger cyclic AMP[7]. Filling in the gaps from this second messenger and building backwards, Gilman and Rodbell show the involvement of a guanosine triphosphate-binding protein (G protein) through a biochemically reductionist experimental approach in the decades following. The missing receptor, the purported signal transducer, was the last functional element to be purified and later cloned by the Lefkowitz group in the 1970's and 1980's. Specifically, the first receptors to be purified and characterized were the α_1, α_2, β_1, and β_2 adrenergic receptors[8–10]. In the decades following, a great deal has been learned about these G protein-coupled receptors' (GPCRs) structure and function.

Receptor Signaling

GPCRs are integral membrane proteins, serving at the boundary between the cell interior and the cell exterior. Here, their primary function is sensing the extracellular environment and listening to cues, to changes the cell must respond to. The repertoire of receptors has exploded during evolution to cope with the myriad of potential molecules to which the cell must respond, with each receptor specifically binding to and being activated by very specific or a narrow range of stimuli. These stimuli can take the form of photons, protons, small molecules, peptides, or other proteins; arguably making GPCRs the most diverse receptors in regards to the type of signaling molecule detected. The sources of external stimuli can either be exogenous (e.g. light or smell) or endogenous (e.g. serotonin or dopamine). Figure 1.1 illustrates the diversity of GPCR ligands[1,11].

In the absence of ligand, GPCRs are in a dynamic equilibrium between at least 2 conformationally distinct states, the most relevant being the active and inactive states. This equilibrium results in some level of basal activity for most receptors (i.e. signaling without ligand stimulation), a notable exception being rhodopsin. This dynamic equilibrium can be shifted differentially by different classes of ligands. Agonists stabilize the active state of the receptor, inverse agonists stabilize the inactive state, while neutral antagonists bind both active and inactive states and do not alter the relative stabilities of the 2 states. Additionally, ligands can either be full or partial agonists, each class seemingly selecting a distinct active conformation, as demonstrated by multiple FRET/BRET, NMR and DEER based experiments[12–14]. Figure 1.3 demonstrates the diversity of ligand functional effects.

The canonical role of G protein-coupled receptors, as the name suggests, is the coupling to, and the activation of, GTP-binding proteins or G proteins. G proteins are active or inactive depending on whether they are bound to GTP or GDP, respectively. The receptor acts as a GTP exchange factor, allowing an inactive G protein to release its bound GDP and bind GTP. G proteins are heterotrimeric complexes, made up of an α, β, and a γ subunit. The α subunit

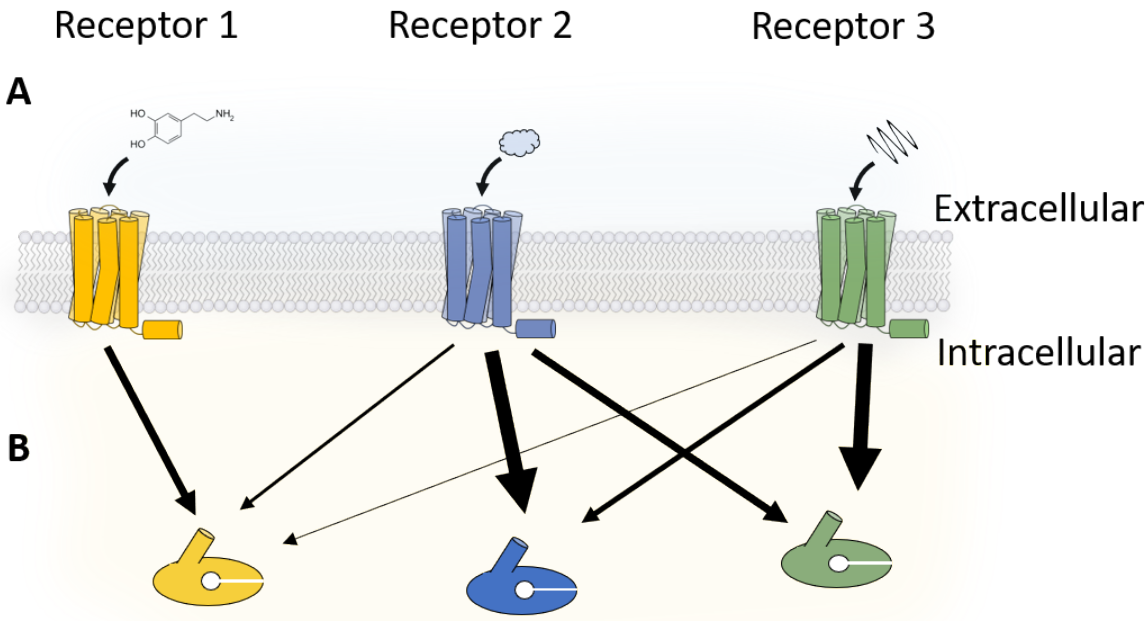


Figure 1.1 – Diversity of GPCR ligands and activation. **A.** GPCRs can bind diverse classes of ligands. Depicted ligands include small molecules (Receptor 1), peptides and proteins (Receptor 2), and light (Receptor 3, e.g. rhodopsin). **B.** GPCRs can bind different G proteins with different propensities, leading to various degrees of signaling cross-over. This activation promiscuity can vary significantly between receptors [15,16].

contains the GTP binding site. Upon activation, the α subunit dissociates from $\beta\gamma$. There are 16 different Gα, 5 β and 12 γ subunits, creating a complex web of possible combinations and functional outputs [17,18]. The most dramatic effect on signaling is determined by the Gα subunit. The most well-studied G proteins are Gα_s, Gα_i, and Gα_q; of which Gα_s's main function is the activation of various adenylyl cyclases and in contrast Gα_i inhibits adenylyl cyclases. Gα_q activates phospholipase C (PLC) on the endoplasmic reticulum (ER) membrane, ultimately resulting in the release of Ca^{2+} from the ER into the cytosol. Figure 1.2 shows an overview of the major pathways in which G proteins are involved. While traditionally GPCRs have been thought to be selective to specific G proteins, today there exists a large body of evidence of promiscuous activation of different G protein classes by many receptors; although usually there still tends to be a preference for a certain class of G protein depending on the receptor [19–23]. This is illustrated in Figure 1.3B. This promiscuity of receptor-G protein coupling has resulted in complicating efforts to delineate the physiological effects of each pathway. Regardless, the effect of receptor activation can result in either the increase or decrease of cellular cyclic adenosine monophosphate (cAMP) or Ca^{2+}, both of which are

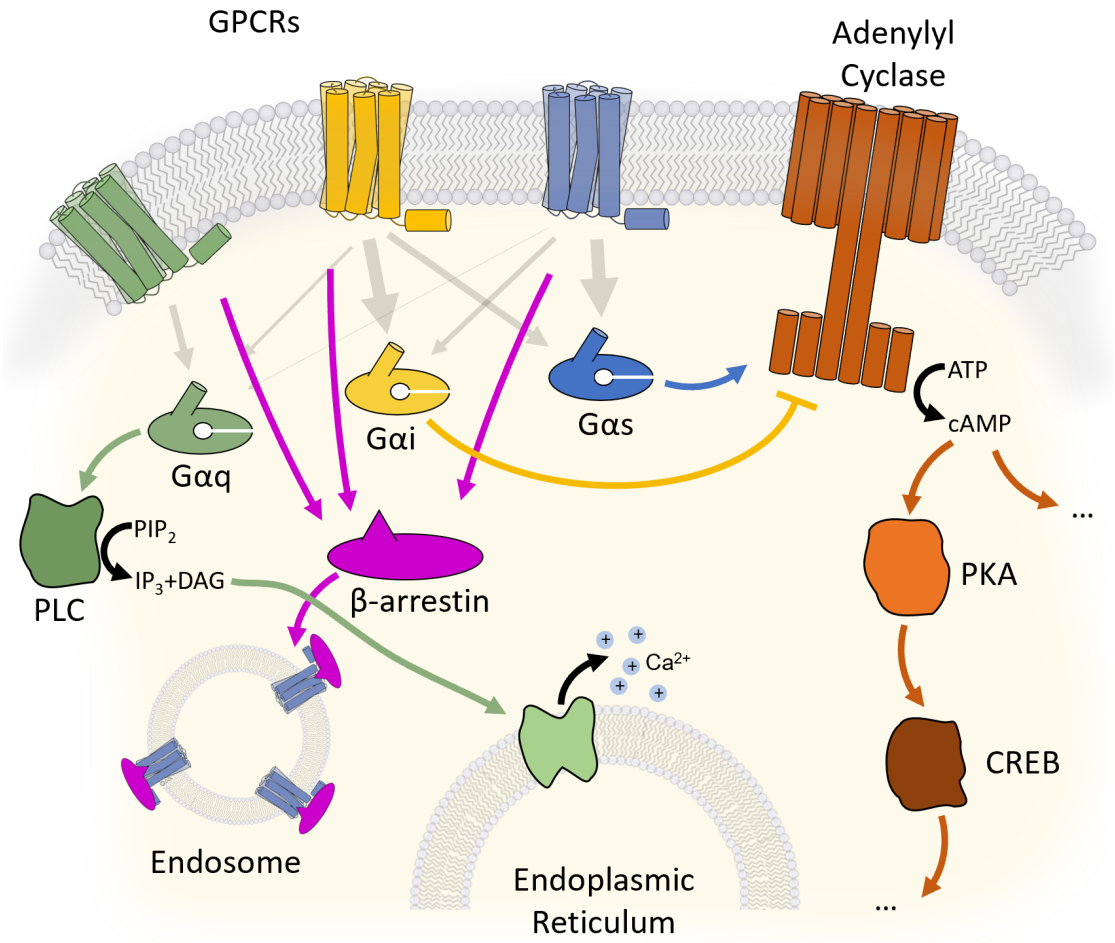


Figure 1.2 – GPCR effector activation pathways. Most common biochemical pathways regulated by GPCRs, color-coded by Gα function. Gq pathway is green. Adenylyl cyclase and its downstream effectors are colored a brown hue. Gs (blue) and Gi (gold) exert their opposing, non-competitive effect on adenylyl cyclase. β-arrestin action (receptor desensitization, recycling, and association with the cytoskeleton) is shown in purple. For simplicity and clarity, the G$\beta\gamma$ subunit is not depicted.

potent second messengers. cAMP primarily mediates the activity of protein kinase A (PKA). PKA and Ca^{2+} can have wide-ranging effects in cells, such as activating or deactivating ion channels, other proteins, and affect gene regulation.

In addition to the effects of Gα, the G$\beta\gamma$ heterodimer also elicits its own signaling. While initially it was thought to be a regulator of the Gα subunit, it has over time been shown to be involved in regulating a large number of proteins, including phospholipace C, adenylyl cyclase[24,25], voltage-gated calcium channels[26], gene transcription[27], and nuclear import and export[28]. It has become clear that both Gα and G$\beta\gamma$ have clear, distinct, and sometimes overlapping cellular functions.

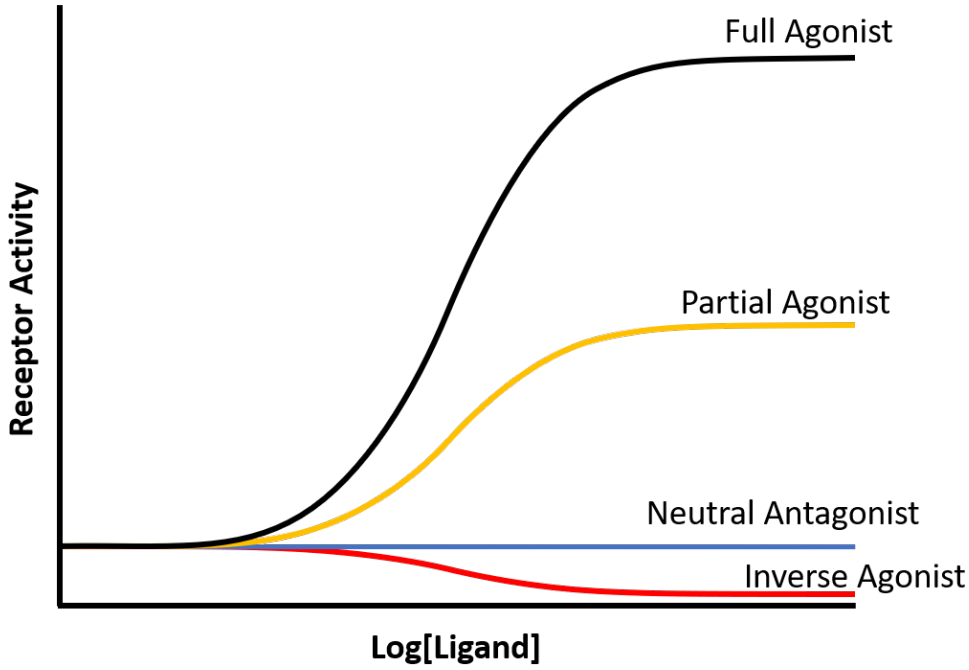


Figure 1.3 – GPCR drug classes. Full agonists result in the highest signal seen with a particular experimental assay (though not necessarily 100% of receptors being activated). This is a relative measurement, for which an endogenous high-efficacy agonist is used as the reference. Anything that results in a positive but lower signal is designated as a partial agonist. Ligands which bind the receptor but do not alter activity are antagonists. Ligands which decrease basal activity of the receptor are inverse agonists, which can also be further subdivided into partial or full inverse agonists.

Another important effector of GPCRs is β-arrestin. The main roles attributed to this effector are receptor desensitization and receptor recycling with the aid of clathrin. But as with G$\beta\gamma$, the many roles played by arrestin were not fully appreciated initially, and there is likely still many unknowns in β-arrestin signaling and function. In contrast to the temporally rapid and transient G protein signaling, β-arrestin signaling tends to be slower and leads to more persistent signaling. Whereas G protein activation occurs within milliseconds following receptor activation while cAMP levels return to basal levels within minutes, arrestin activation occurs minutes following receptor activation and can continue signaling for several hours.

Biased Signaling

Interestingly, GPCR signaling can be discriminatory in regards to what downstream effector is activated, leading to a phenomenon known as biased agonism. A number of factors can contribute to signaling bias: agonists themselves may be biased[29] and stabilize a certain receptor conformation that prefers a certain effector; receptors themselves may be biased[13],

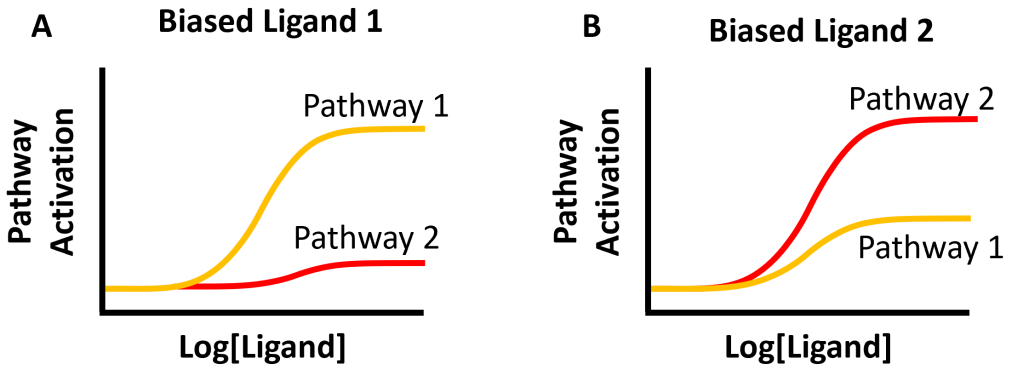


Figure 1.4 – An example of GPCR biased signaling. Both ligand 1(**A**) and ligand 2(**B**) activate the same receptor but, possibly due to subtle changes in receptor conformations and dynamics, the two ligands differentially activate downstream effector pathways.

either through mutation[30], evolution or oligomerization[31], and couple poorly to a particular effector relative to related receptors; or the cellular environment of a receptor may promote coupling to an effector solely due to its high expression. Currently the most studied form of bias is the search for biased agonists as they appear promising as new therapeutics with fewer side effects than nonbiased agonists. For example, μ-opioid receptor agonist morphine (and others such as heroin and fentanyl) activate both the G protein and β-arrestin pathways. By using either knockout mouse lines[32] or using mice with disrupted G protein or β-arrestin pathways, it has been demonstrated that G protein activation results in the analgesic properties of morphine while its undesirable effects such as respiratory depression, tolerance, and dependence are mediated by β-arrestin. In contrast, it has been shown that a biased peptide agonist of the angiotensin II type 1 receptor activating the G protein pathway in mice leads to an increased blood pressure[33], while only activating β-arrestin leads to increased cardiac contractility and performance[34], the desired therapeutic outcome. Though the desirability of activating a certain downstream effector will depend on the cell type and receptor in question, understanding the determinants of biased signaling is key to this goal.

Receptor Isoforms

To increase the already large signaling repertoire of GPCRs, hundreds of receptor genes also possess alternatively spliced variants[35]. According to recent analyses at least 50% of GPCR genes contain at least 2 exons. All class B and C GPCRs have at least 2 exons, while this is only true for 58% of class A receptors. One of the first receptors discovered to have multiple isoforms is the dopamine D2 receptor, which has a long (D_2L) and a short (D_2S) form[36]. This alternative splicing results in the insertion/deletion of 29 residues in the ICL3 of the receptor, likely affecting G protein and β-arrestin coupling. Some of these splicing variants can play a role in normal development[37], though upregulation of certain isoforms may also be involved in cancer[38]. Typical splice variants contain varying lengths of the N-term, ECLs, ICLs, and

C-term; potentially being able to affect all aspects of receptor signaling.

Receptor Oligomerization

To complicate GPCR signaling further, these receptors can not only signal as monomers, but can also do so as dimers, tetramers, or higher order oligomers. These higher order forms can exhibit distinct signaling properties from their monomeric counterparts. Class C receptors can only function as dimers and do not act as monomers. In other cases, dimerization is required for proper localization. For example. CXCR4 receptors can only be exported from the endoplasmic reticulum as a dimer, although they are capable of signaling in monomeric and dimeric states at the cell surface. Interestingly, CXCR4 receptors that are incapable of dimerizing are also unable to recruit β-arrestin (Barth lab, unpublished results). In an intriguing experiment, when Rivero-Müller co-expressed a binding-deficient and a signaling-deficient leutenizing hormone receptors in HEK293 cells, they were able to partially restore WT receptor signaling[39]. Further, they demonstrated the restoration of a WT phenotype in transgenic mice co-expressing the two receptor variants. Although how widepread such an effect is in native systems is not fully known, it clearly has some physiological influence.

1.1.3 Structure

All known GPCRs exhibit a characteristic 7-transmembrane alpha-helical structure. These helices are connected by varying lengths of extracellular (ECL) and intracellular (ICL) loops. Following TM7, the last transmembrane helix, is a shorter intracellular helix that lies parallel to the membrane and snorkels within it. Most receptors have a flexible C-terminus following helix 8 that can be quite long. Numerous post-translational modifications are distributed throughout GPCRs: several N-linked glycosylation sites on the N-term or ECLs; phosphorylation sites on ICL3 and the C-term; and a palmitoylation site on helix 8 to anchor the receptor to the membrane. The number and presence of these modifications vary by receptor. The glycosylations are thought to play a role largely but not exclusively in receptor folding and trafficking[42], while the phosphorylation sites are important in receptor desensitization and internalization post-activation. Several studies have demonstrated the existence of "phosphorylation codes", where different combinations of phosphorylations will result in different downstream events[43,44]. ECL2 and ECL3 typically tend to each contain one disulfide bond. The ECL2 disulfide heavily influences ligand affinity to the receptor and likely plays a role in receptor activation. The ECL3 disulfide most likely provides structural stability. There exists a peculiar sodium binding site in the core of the receptor, where a sodium ion tends to be found in the inactive state of receptors forming an ionic bond primarily with D2.50 (see next paragraph for an explanation of Ballesteros-Weinstein[41] numbering) in high resolution structures. This bound sodium ion acts as a negative allosteric modulator. In addition to sodium, the interior cavity of receptors tend to be highly hydrated.

Due the highly conserved 7TM structures of GPCRs, it is convenient to use a numbering

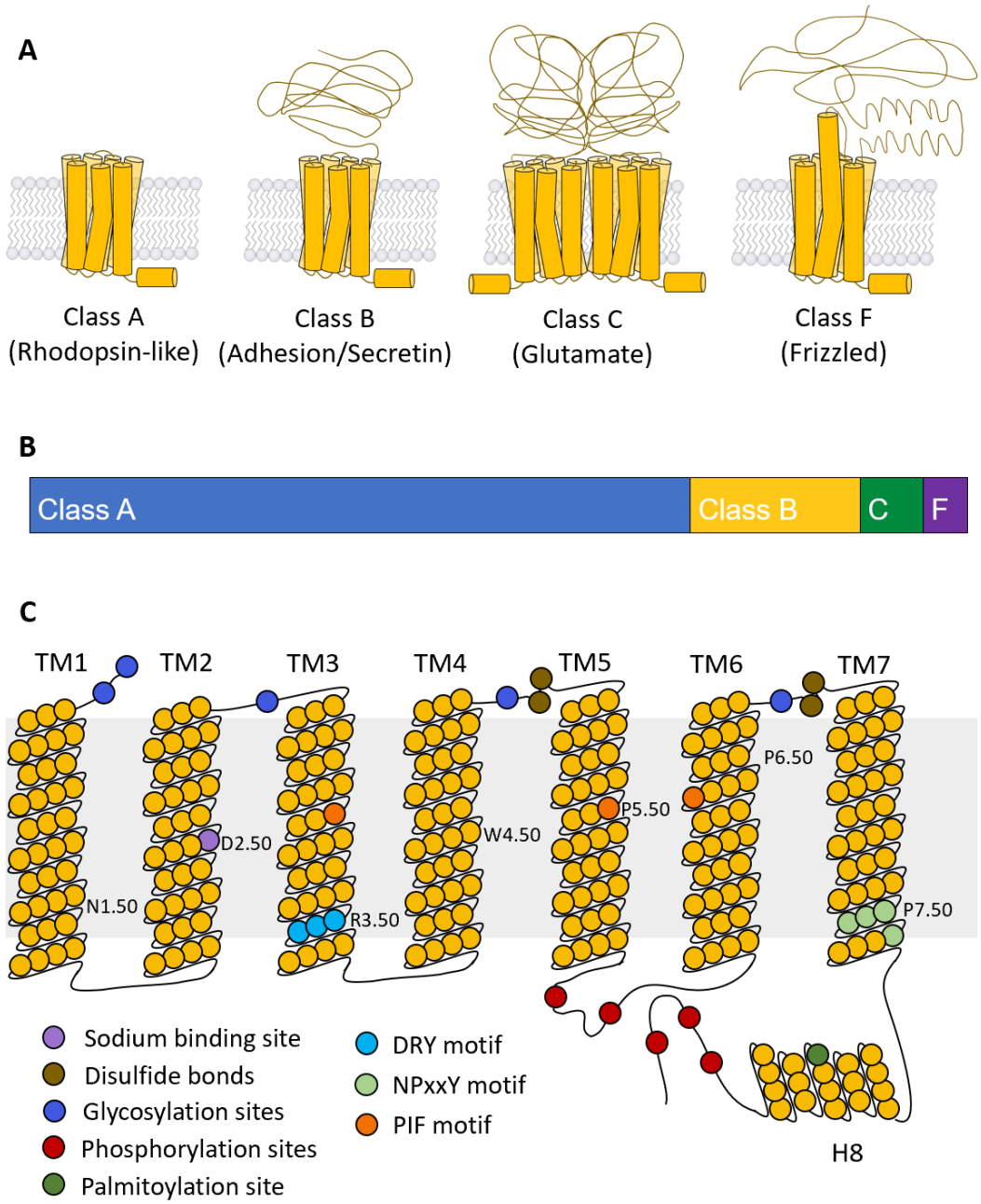


Figure 1.5 – GPCR classification and 2D anatomy. **A.** Representation of structural differences in major GPCR classes. Both A-F[40] and GRAFS[1] classifications are shown. **B.** Relative size of each GPCR class. Values are 388(A), 100(B), 37(C), and 26(F). Olfactory receptors are not included. **C.** Snake plot featuring conserved post-translational modification (PTM) sites, ion-binding sites, and motifs in class A GPCRs. The relative positions of the most conserved sites in each TM and their Ballesteros-Weinstein[41] designations are also shown. The first number refers to the TM while the second refers to the residue position in relation to the most conserved residue in that helix, designated 50.

scheme with which it is easy to compare residue positions across different receptors. The Ballesteros-Weinstein[41] numbering scheme is the most common for this purpose. The numbering follows the format TM.N, where TM refers to the transmembrane helix number and N refers to the residue position in relation to the most conserved residue in that helix, numbered 50. So D2.50 is the most conserved residue in TM2. 2.46 would be the 4th residue going towards the N-term. The relative positions of the most conserved residues in each TM helix is shown in Figure 1.5C.

GPCR Structure Determination

While the basic function of GPCRs has been known for some time, a dearth of information regarding receptor structure has existed up until recently. The first crystal structure solved was that of the inactive state of bovine rhodopsin, a uniquely stable GPCR and a readily-available protein[45]. Another receptor structure would not be solved until 2007, when the inactive state of β_2-adrenergic receptor was crystalised[46]. The conformational flexibility of GPCRs necessitated the use of high-affinity ligands to stabilize receptors in particular states and crystallization in detergent[46,47]. ICL3 is also usually replaced with a cytosolic protein (T4 lysozyme[46,47] or the thermostabilized apocytochrome BRIL[48,49]) to increase intermolecular contacts in the crystal lattice. The use of camelid antibodies (nanobodies) has also greatly facilitated crystallization efforts by further stabilizing specific conformational receptor states. Thermostabilization of a particular receptor state can also aid structure determination by increasing receptor expression and overall stability. This can either be achieved through alanine scanning[50], directed evolution[51,52], or computational methods[53,54].

The first full-length effector-bound active state structure to be solved was the β_2 adrenergic receptor, in complex with the G_s heterotrimer, using a combination of the aforementioned methods[55]. This structure has revealed, for the first time, the structural differences between the fully active G protein-bound and inactive states of the receptor as well as the binding mode of the full length G protein. Although the active state structure of rhodopsin was solved prior, these structures were solved without a bound full-length G protein[56,57]. Structural determination of the relatively unstable GPCR-G protein heterotrimer remained a bottleneck in crystallography. With recent advances in cryo-EM (electron cryo-microscopy) sample preparation, detection and classification algorithms, it has become a viable tool in GPCR structure determination. The two structure determination methods have become complementary to one another. Crystallography has a great track record in resolving inactive receptor structures, which are still beyond the capabilities of cryo-EM due to its small size and lack of asymmetry. Simultaneously, cryo-EM has become indispensable in solving receptor-effector complexes. After the first GPCR-G protein cryo-EM structure, there has been a deluge of new structures with diverse receptors. Even more recently, this phenomenon repeated itself with receptor-arrestin complexes, and this influx of new data shows no sign of slowing as cryo-EM matures as a technology. Figure 1.6 shows the progression of the increasing number of GPCR crystal and cryo-EM structures over the years[58].

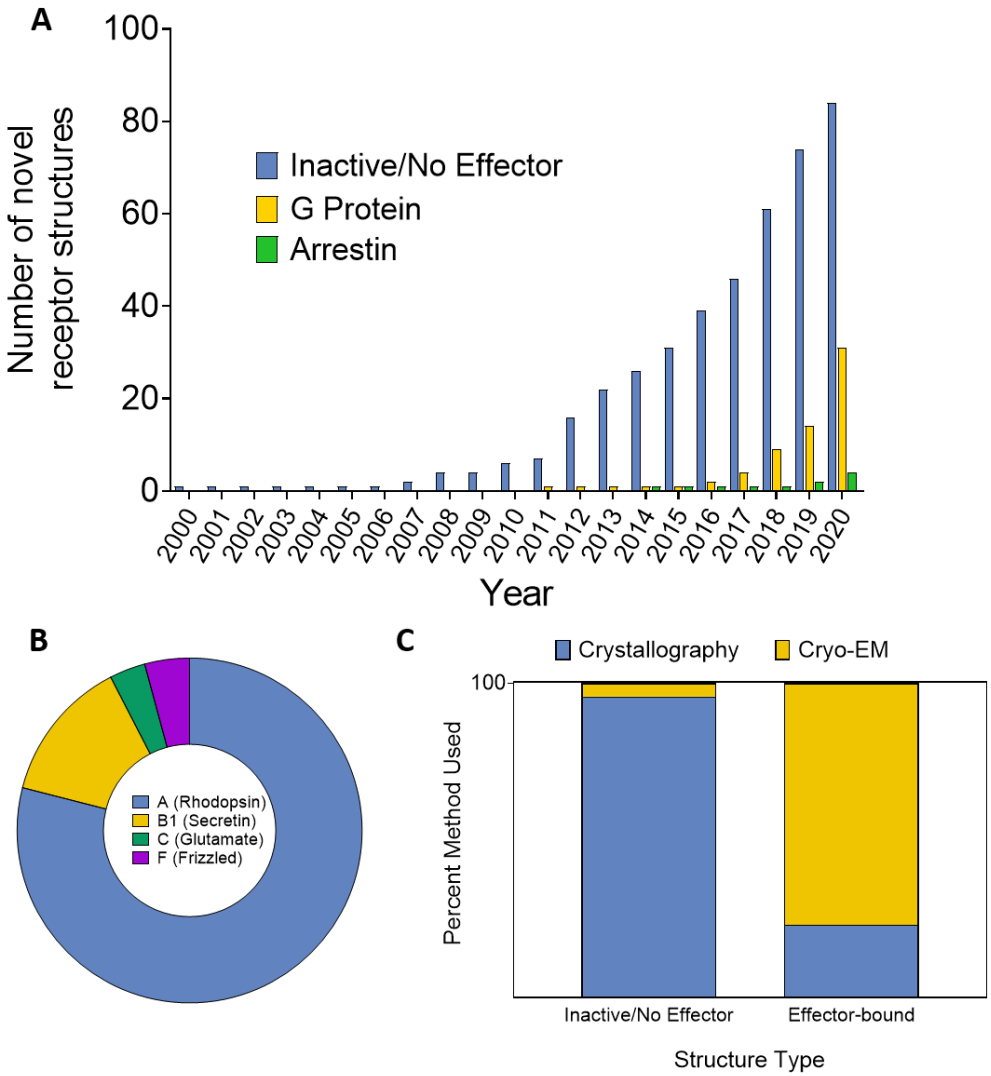


Figure 1.6 – Unique GPCR structures over time. **A.** Number of unique GPCR structures solved since 2000 in either inactive or effector-free (including intermediate and active-like states), active G protein-bound, or active β-arrestin-bound states. **B.** Ratio of receptor class solved from unique structures. **C.** Percent of receptor states solved by specific structure determination method.

GPCR Activation and G Protein Binding

In concert with functional characterization of receptors, we have learned a great deal about what events are involved in receptor signaling and activation and their relation to structure. In aminergic receptors, the most essential ligand binding residues are S5.42, S5.43, S5.46, and D3.32. The serines tend to form hydrogen bonds with ligands, though not all three serines may be involved in binding a particular ligand. D3.32 almost always forms a salt bridge with a positively charged nitrogen on ligands. Alanine mutagenesis at these positions severely impacts the receptors affinity for most ligands. D3.32 is essential, and if mutated, can fully

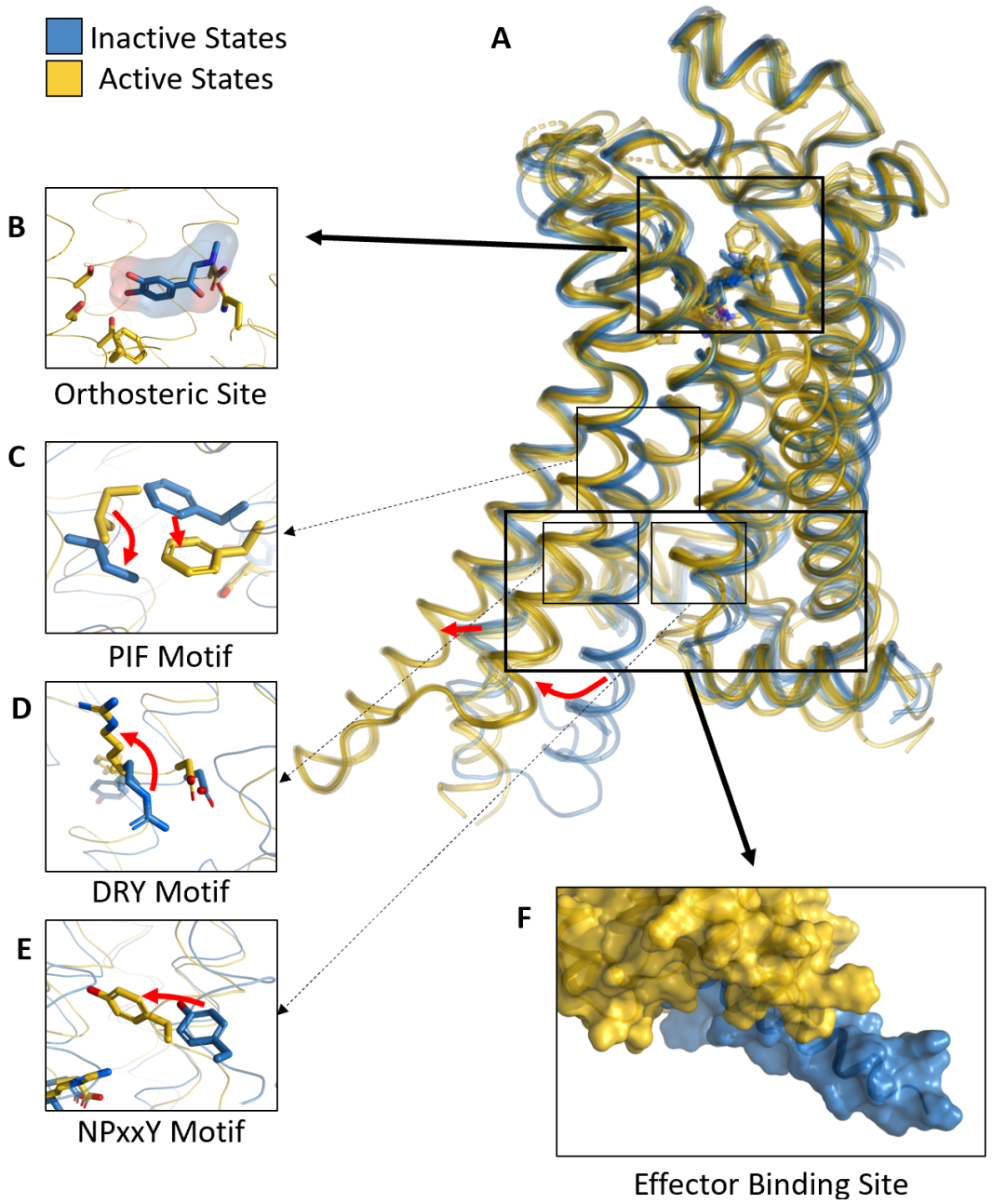


Figure 1.7 – Structurally and functionally conserved regions of GPCRs in active and inactive receptor states. **A.** Multiple superimposed active(gold) and inactive(blue) receptor states demonstrating the high structural conservation of GPCRs. Red arrows indicate typical structural changes during receptor activation, the largest change being the movement of TM6 outward ~15Å in Gs-coupled receptors. **B.** The typical ligand binding site, termed the orthosteric site on the extracellular region of GPCRs. **C-E.** Highly conserved microswitches involved in receptor activation. Red arrows indicate changes in position or rotamer conformation from the inactive to the active state. **F.** The effector binding site of GPCRs on the intracellular region. Both G protein and β-arrestin bind here.

abrogate ligand binding[59,60]. Nearby aromatic residues (F6.55) form π-π interactions with the ring of ligands such as dopamine and adrenaline. These interactions serve to bring the TM helices closer to the orthosteric site. The inward movement of TM5, and specifically P5.50, appears to influence the rotamer conformation of I3.40 and F6.44, stabilizing the active receptor state. Figure 1.8A and B show these interactions in the B2AR with bound epinephrine[61].

What available structural, functional, and computational data reveal is that the ligand-binding and G protein-binding regions of the receptor may each sample the active and inactive states semi-independently of the other. This is clearly shown by crystal and cryo-EM structures with receptor bound to inverse agonist, where the receptor is in a fully inactive state; and receptor bound to agonist, where the ligand-binding region adopts an active conformation while the G protein-binding region remaining in an inactive conformation. Both single molecule FRET experiments and molecular dynamics simulations corroborate these observations[12,62]. Upon agonist binding however, the intracellular region of TM6 becomes more dynamic and will tilt outward, away from TM3 and TM7. Within the core of the receptor, I3.40 and F6.44 of the PIF (proline-isoleucine-phenylalanine) motif "toggle" rotamers. Y7.53 of the NPxxY motif moves closer to TM3 and forms a water-mediated hydrogen bond with Y5.58[63,64]. These events facilitate the binding of an effector, canonically a G protein. The C-term helix (helix 5 or H5) of the G protein extends and rotates into the cavity formed by the receptor. The arginine of the DRY motif which forms a transient salt bridge with E6.30 in the inactive state, will form a cation-π interaction with the G protein C-term H5. The displacement of the C-term helix H5 results in an allosteric conformational change in the nucleotide-binding pocket of the G protein, allowing the separation of the GPCR-bound RAS domain and the helical domain to separate and release the bound GDP cushioned between the two domains, and accelerating the exchange for a GTP molecule. The presence of GTP rather than GDP allosterically dislodges the $\beta\gamma$ dimer. Consequently, Gα and G$\beta\gamma$ can activate downstream effectors.

GPCR-Effector Binding Differences

There are a number of structural differences in Gs-coupled and Gi-coupled receptors. Gi-coupled receptors tend to have smaller TM6 distortions upon G protein binding (~8Åvs 14Å). These findings corroborate the higher G protein selectivity for Gi-coupled receptors that has been experimentally observed. The G protein C-terms also bind their cognate receptors at slightly different angles due to differing steric requirements. The tyrosine at H5.24 on the Gs C-term requires more volume at the receptor-G protein interface compared to the cysteine present in Gi. The tyrosine in Gs additionally forms a cation-π interaction with the conserved arginine in the DRY motif. There is no clear-cut selectivity determinant between the different receptor-G protein pairs, though this is an active area of investigation[19,20].

The effector β-arrestin can bind in the same region on an activated receptor while forming partially overlapping contacts and stabilizing a distinct receptor conformation. In addition to this "tight" binding mode, β-arrestin also binds to the phosphorylated C-term or ICL3,

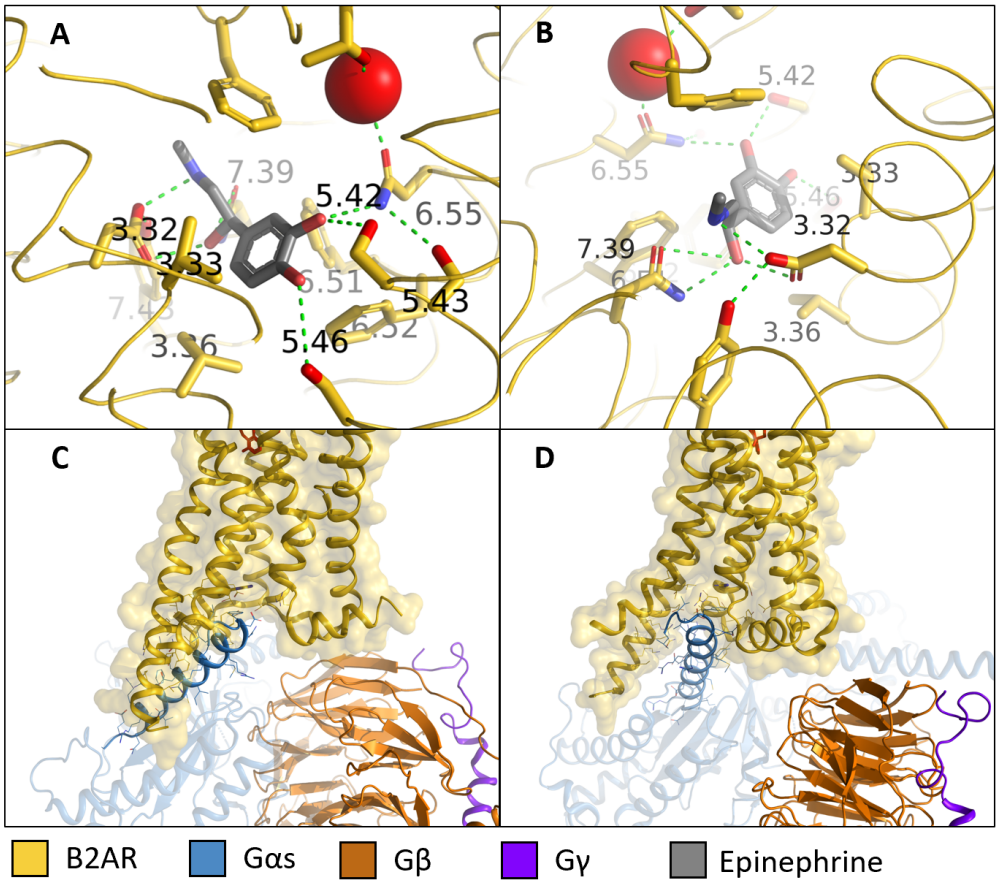


Figure 1.8 – B2AR structures with ligand and G protein binding **A, B.** Active state structure of the B2AR aminergic receptor bound to epinephrine (PDB ID: 4LDO) [61]. Polar contacts between receptor, ligand, and resolved water molecule (red) are shown in green [61]. **C, D.** B2AR bound to the G$\alpha s\beta\gamma$ heterotrimer (PDB ID: 3SN6) [63]. The Gαs subunit makes major contacts with the receptor.

depending on the receptor. Typically this "looser" binding is always present, whereas the additional binding to the receptor effector binding site is not. In contrast to Gs binding to β_1AR where TM6 moves 15A away from the receptor, the binding of β-arrestin stabilizes a conformation in which TM6 is distorted only ~8Å. The DRY motif within the β-arrestin-coupled structure also maintains the intrahelical salt bridge present between D3.49 and R3.50 present in the inactive state while in the Gs-bound structure R3.50 extends towards and forms interactions with Gs. There are additional differences within the orthosteric site which can potentially explain the functional bias seen in certain ligands. β-arrestin itself exhibits interesting allosteric properties. Fluorescence and NMR studies implicate dynamic changes in TM7 playing a role in β-arrestin bias while TM6 plays an outsize role in G protein bias [65,66]. Depending on the phosphorylation code of the receptor, β-arrestin can occupy different conformations and may activate partially distinct downstream effectors even after dissociating from the receptor [67].

1.2 Allostery

The initial interest in allostery, or the phenomenon of altering the conformation of sites distant from the initial site of perturbation, came from studies with hemoglobin, a homo-tetramer oxygen carrier. How binding of each oxygen molecule successively increased affinity of non-bound protomers to oxygen became a point of interest[68]. After its initial discovery, it was realized that allostery is not specific to hemoglobin, but is involved in the function of a wide range of proteins. Mainly, allostery is involved in the regulation of the primary function of proteins. In the notable case of GPCRs, signal transduction across the membrane via allostery *is* the primary function.

1.2.1 Phenomenological Level

The first allosteric model to be proposed with a structural link was by Monod, Wyman, and Changeux[69]. The MWC model posited that allosteric proteins are: 1) symmetric homo-oligomers, 2) each protomer can exist in two states, 3) one state has a higher affinity for ligand than the other, and 4) all binding sites are equivalent. As each protomer binds ligand, it is stabilized in the high-affinity state and consequently stabilizing the whole quaternary structure in this state, further increasing the affinity to subsequent ligands. The simplicity of the MWC makes it a very attractive model, being able to explain the allosteric behavior of a wide range of proteins, with only 4 known variables: basal activity of the protein, affinity of the ligand for each state, and the ligand concentration. There are other models such as the Koshland-Némethy-Filmer (KNF) model[70], which relies on ligand binding to alter the conformation of each protomer (i.e. induced fit). This putatively results in additional stabilizing interactions between adjacent protomers, further stabilizing the complex and consequently reducing the energy required for additional ligand binding. Other more recent models such as the ensemble model[71], of which the MWC and KNF can be considered subsets, also exist. With all the elaborate models of allostery available, the most relevant model to GPCRs is still an adaptation of the initial MWC model[72]. While the initial model was based on oligomers, the model also works well for monomeric receptors.

The allosteric two state model (ATSM) is shown in Figure 1.9 as applied to GPCRs. The model relies on three key parameters: K_A, the association constant of the ligand; L, the basal activity of the receptor; and α, the allosteric efficacy. L is an intrinsic property of the receptor in question, while K_A and α are specific for the receptor-ligand pair. We can determine the ratio of active receptors, E, with[72,73]:

$$E = \frac{R_{A,Lu} + R_{A,Lb}}{R_{A,Lu} + R_{A,Lb} + R_{I,Lu} + R_{I,Lb}} \tag{1.1}$$

where $R_{A,Lu}$ refers to the active unbound receptor, $R_{A,Lb}$ to active ligand-bound receptor, $R_{I,Lu}$ to inactive unbound receptor, and $R_{I,Lb}$ to the inactive ligand-bound receptor. Substituting

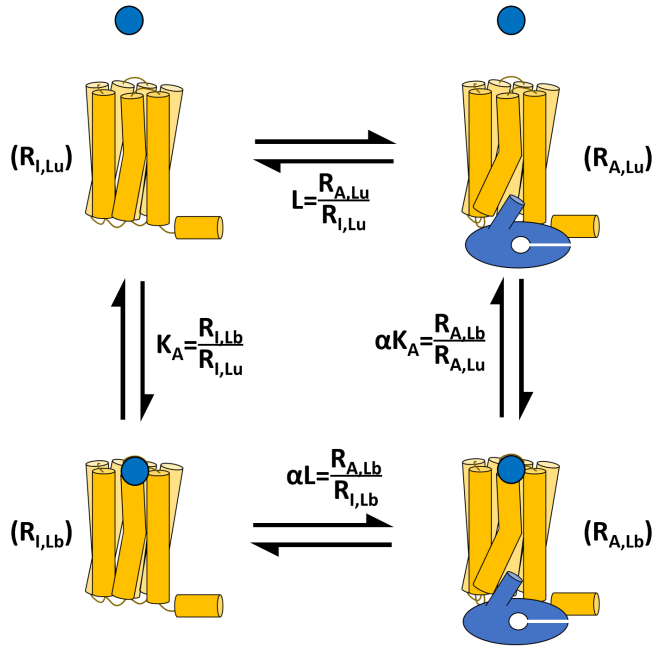


Figure 1.9 – Schematic of the allosteric two-state model. The different states accounted for include the inactive unbound ($R_{I,Lu}$), inactive ligand-bound ($R_{I,Lb}$), active unbound ($R_{A,Lu}$), and active ligand-bound($R_{A,Lb}$) receptor states. The equilibrium constants between these states are also shown.

with the aforementioned reaction constants we can derive the ratio of active receptors as a function of ligand concentration, $[A]$:

$$E([A]) = \frac{L + \alpha L K_A [A]}{1 + K_A [A] + L + \alpha L K_A [A]}. \tag{1.2}$$

Setting $[A]$ to 0 we get the basal activity:

$$E_{basal} = \frac{L}{(1 + L)} \tag{1.3}$$

And the maximum response as $[A]$ approaches infinity:

$$E_{max} = \frac{\alpha L}{1 + \alpha L} \tag{1.4}$$

Solving for the mean of equations 1.3 and 1.4, we obtain the half response ligand concentration EC_{50}:

$$EC_{50} = \frac{1 + L}{K_A (1 + \alpha L)} \tag{1.5}$$

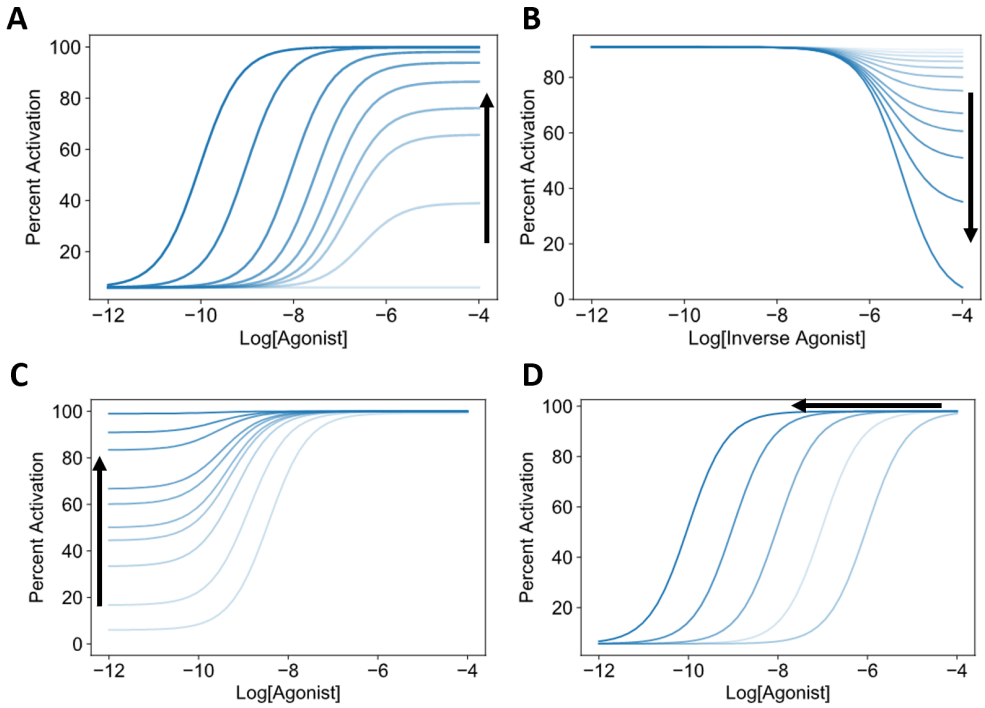


Figure 1.10 – Allosteric two state model variables. **A** Increasing $\alpha > 1$ (8, 16, 32, 64, 160, 800, 8000, 80000). At low α (<800), maximal response increases while EC_{50} decreases. At higher α maximum response is reached but EC_{50} continues to decrease. **B** Decreasing α <=1 (1, 0.9 ,0.8 ,0.7 ,0.6, 0.5, 0.4, 0.3, 0.2, 0.15, 0.1, 0.05, 0.0) results small EC_{50} shifts. The primary effect is a decrease in maximum response. **C** Increasing L (0.064, 0.0001, 0.001, 0.01, 0.2, 0.5, 0.8, 1, 1.5, 2, 5, 10) increases basal activity. **D** Changing ligand affinity K_A (1/450E-9, 1/450E-10, 1/450E-11, 1/450E-8, 1/450E-7) only shifts EC_{50}. Default values are L=0.064, Ka=1/450E-9, α=800 to match experimental results from the dopamine D_2 receptor.

and we can solve for the allosteric efficacy α in terms of K_A, L, and EC_{50} by rearranging equation 1.5:

$$\alpha = \left(\frac{1+L}{K_A EC_{50}} - 1 \right) \cdot \frac{1}{L} \tag{1.6}$$

Figure 1.10 demonstrates how changing each value in equation 1.2 would manifest in experimental dose response curves.

This is the simplest form of the allosteric two state model applied to GPCRs. When taking into account biased signaling (both on part of the ligand and downstream effector), allosteric modulators or multiple ligands; the model would need to be extended by considering additional states. Despite its simplicity, the allosteric two state model is a valuable tool to predict and explain GPCR signaling at a phenomenological, if not structural level.

1.2.2 Structural Level

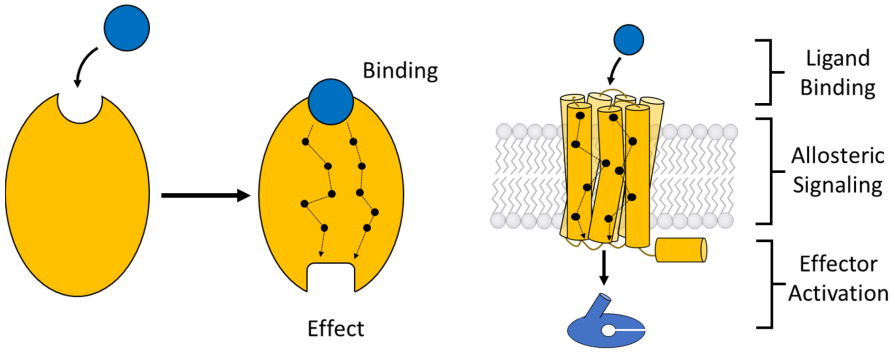


Figure 1.11 – Conceptual illustration of allostery.

The classical thought behind the structural effects of allostery boils down to the bound ligand introducing strain within the binding site of the bound protein, which is then propagated to the allosteric site. In oligomers like hemoglobin, this strain was thought, and later demonstrated, to slightly alter the contacts made between protomers. In the past two decades the flexibility of proteins, irrespective of the number of subunits, and its role in protein function and allostery has gained significant appreciation[74]. Moving away from a rigid view of proteins, more recent views treat proteins as flexible and dynamic; forming ensembles of conformations and continuously shifting between them. The most thermodynamically stable a particular conformation, the more populated it is. Similar to previous models, ligand binding differentially affects conformation stability and results in a population shift within the ensemble. The concept of an ensemble, in a simplified form, can be visualized as an energy landscape as seen in Figure 1.12. Inspired by this idea, Nussinov expanded on the ATSM to include a thermodynamic aspect, creating a more holistic model of allostery[73]. The Gibbs free energies of the ligand-free and ligand-bound states respectively can be calculated as follows:

$$\Delta G_{Lu} = G_{A,Lu} - G_{I,Lu} = -RT \cdot ln\,(L) \tag{1.7}$$

$$\Delta G_{Lb} = G_{A,Lb} - G_{I,Lb} = -RT \cdot ln\,(\alpha L) \tag{1.8}$$

where ΔG_{Lu} refers to the Gibbs free energy difference of the inactive ligand unbound and active ($G_{A,Lu}$) and inactive ($G_{I,Lu}$) states and ΔG_{Lb} refers to their ligand-bound counterparts. The difference in energy of the ligand-bound and ligand-free states can then be related to allosteric efficacy α:

$$\Delta\Delta G = \Delta G_{Lb} - G_{Lu} = -RT \cdot ln\,(\alpha) \tag{1.9}$$

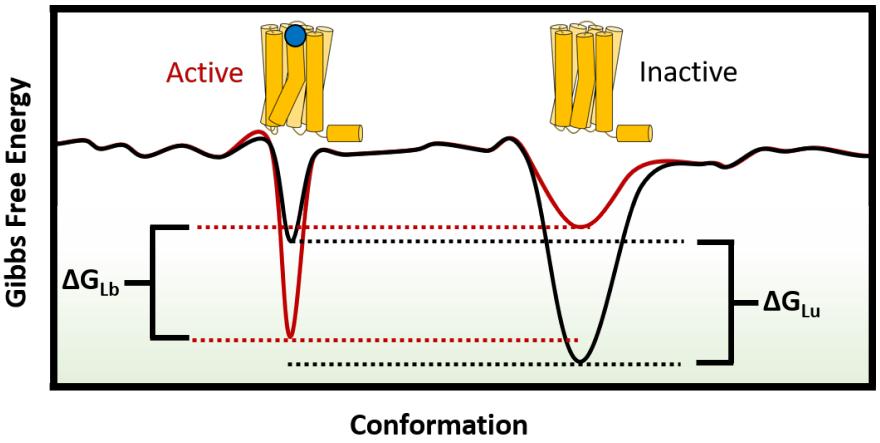


Figure 1.12 – Changes in the energy landscape due to ligand binding. An example of an agonist modifying the energy landscape of a GPCR. The active state (red) is stabilized while the inactive state (black) is destabilized[73].

Modifying the relative energies of functional states is not the only way a ligand can alter the allosteric properties of proteins, however. Dai et al. have shown that for the NMDA receptor it is the curvature of the energy landscape that changes upon agonist binding[75]. While partial agonist and full agonist show a similar energetic differences between the active and inactive conformations, more efficacious agonists result in a more curved energy landscape (i.e. a less dynamic receptor).

The most common experimental technique to directly observe protein ensembles and allosteric effects is nuclear magnetic resonance (NMR). This can be laborious and time-consuming, especially so with our membrane-protein-of-interest. More common methods tend to be computational, such as molecular dynamics (MD) simulations and normal mode analysis (NMA). By analyzing correlated positions or motions of residues, these techniques have been successful at recapturing protein motions and detecting allosteric sites determined experimentally[76,77]. Using MD, the Shaw group predicted an allosteric pathway within the c-src kinase and experimentally validated an allosterically inhibitory mutation 30Å away from the protein's active site. Recently Vaidehi and colleagues have developed a method to analyze correlated changes in backbone dihedral angles, and use this information to predict allosteric "pipelines" between known ligand and effector binding sites by maximizing the measure of mutual information through these pipelines[78,79]. The fact that their proteins-of-interest are GPCRs has allowed us to take advantage of their work to modify the allosteric properties of the dopamine D2 receptor. MD takes into account the actual physical forces that would naturally be present in a system and simulations usually consider all atoms or use parameterized coarse-grained groups that best represent the behavior of a group of atoms. On the other hand, NMA typically considers only Cα atoms connected by springs with specific spring constants representative of the bonds between the residues of these Cαs. This simplification makes NMA magnitudes faster than MD and still offers an accurate picture of protein flexibility and

underlying oscillations present in structures. This speed makes it compatible with existing protein design methods (Rosetta) and offers a complementary perspective that is usually absent in current protein design endeavors.

Other efforts to decode allosteric pathways include analysis of evolutionary conserved residues and co-evolving residues within structures by the Ranganathan lab[80] and others[81]. The Babu group has also been actively working on decoding common allosteric pathways in GPCRs using structure-based analysis of contact differences in active and inactive receptor states from the large number of structures currently available[82–84]. Although these methods can be informative and useful in dealing with naturally evolved proteins, they are less useful in protein design.

1.3 Reprogramming G protein coupled receptor structure and function

D. Keri[1], P. Barth[1,2,3,4]

[1] Swiss Federal Institute of Technology (EPFL), Interfaculty Institute of Bioengineering, 1015 Lausanne, Switzerland
[2] Ludwig Institute for Cancer Research Lausanne Branch, 1066 Lausanne, Switzerland
[3] Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
[4] Department of Pharmacology and Chemical Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Published in Current Opinion in Structural Biology, August 2018
Link to original publication: https://doi.org/10.1016/j.sbi.2018.07.008

Author Contributions: DK performed literature search. DK and PB wrote the manuscript.

1.3.1 Abstract

The prominence of G protein-coupled receptors (GPCRs) in human physiology and disease has resulted in their intense study in various fields of research ranging from neuroscience to structural biology. With over 800 members in the human genome and their involvement in a myriad of diseases, GPCRs are the single largest family of drug targets, and an ever-present interest exists in further drug discovery and structural characterization efforts. However, low GPCR expression and stability outside the natural lipid environments have challenged these efforts. In vivo functional studies of GPCR signaling are complicated not only by the need for specific spatiotemporal activation, but also by downstream effector promiscuity. In this review, we summarize the present and emerging GPCR engineering methods that have been employed to overcome the challenges involved in receptor characterization, and to better understand the functional role of these receptors.

1.3.2 Introduction

G-protein coupled receptors (GPCRs) constitute the largest family of signaling membrane receptors. They are involved in a wide diversity of cellular and physiological processes, including immune responses, vision, neuronal communication and behavior[85]. GPCRs are also associated with severe diseases and represent the target of close to 40% of marketed drugs[6]. GPCRs function as sophisticated allosteric machines. They respond to diverse extracellular stimuli in the form of light, small molecules, peptides, lipids and proteins by transmitting the signal across the membrane and activating a number of intracellular signaling pathways[86]. High conformational flexibility is a hallmark of GPCRs which allow them to sense diverse stimuli and couple to different signaling pathways[64], but represents a challenge for structure characterization which often require conformationally stable proteins. Hence, initial GPCR engineering efforts have focused on developing approaches to identify thermostabilized receptor variants for accelerating X-ray structure determination and rational drug design. In parallel, methods have also been established to create GPCR variants that can be controlled by external cues for better studying cellular signaling. Lastly, computational approaches have recently emerged to rationally design GPCR functions, and pave the road for the design of novel biosensors that should prove useful in cell engineering applications, outlined in Figure 1.13. Below, we first describe empirical, experimentally-driven approaches and then outline recent computational techniques for engineering GPCR structure and function.

1.3.3 Empirical experimentally-driven design of GPCRs

Over the years, a number of experimental approaches have been developed to create GPCR variants for facilitating structural and functional studies. A first line of investigations has focused on modifying and stabilizing receptors to make them more amenable to structural determination and biophysical studies, including drug discovery efforts. A second line of approaches aimed at better understanding the role of GPCRs in neuronal, cellular signaling, and behavior. In each case, the methods similarly involved random or systematic mutagenesis, or a grafting approach to reach the desired molecular properties. The methods used are described below, and highlighted in Figure 1.14.

Structural characterization

Up until 2007, the only GPCR with a solved three-dimensional structure was rhodopsin[45,46]. Due to the low endogenous expression of GPCRs and their inherent instability outside biological membranes, new techniques were necessary to enable their crystallization. Today, over 50 unique GPCRs[58] have been crystallized, thanks, in no small part, to various GPCR engineering methods. Successful receptor stabilization could be achieved by conformationally stabilizing the flexible intracellular loop 3 (ICL3) by antibody fragments recognizing the receptor, or by replacing the ICL3 entirely with different soluble proteins promoting crystal packing such as T4L or BRIL. Conformational thermostabilization was also achieved via scanning mutagenesis,

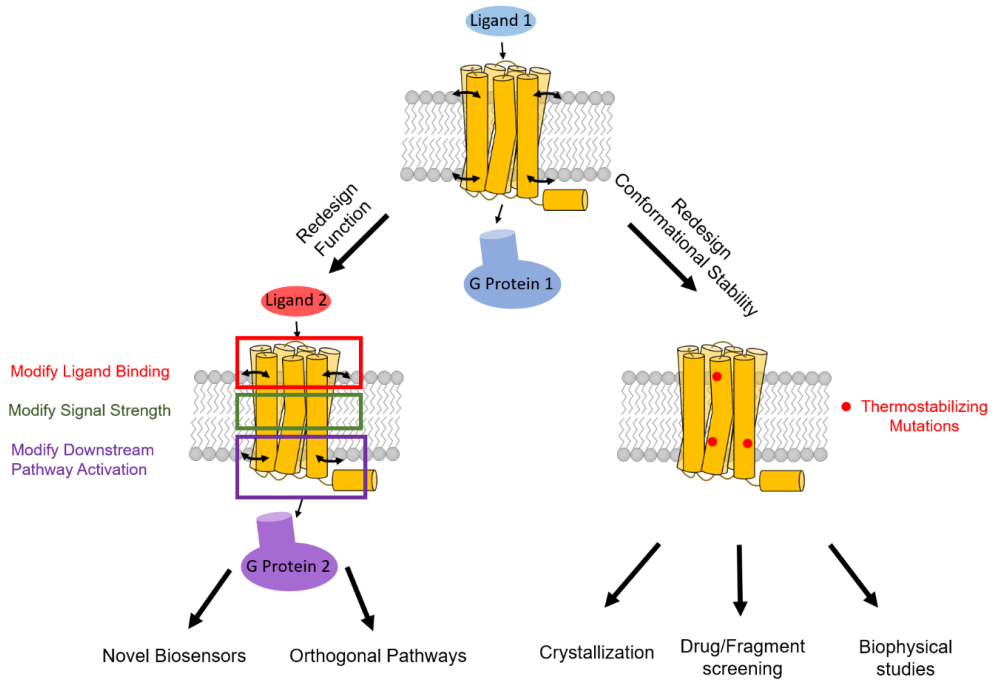


Figure 1.13 – Potential applications of GPCR engineering. A wildtype receptor (top) can be engineered to: Left, create novel receptor functions to respond to different ligands, to transmit ligand-induced signals with different strengths, or to activate a novel effector protein. Right, another route is to modify the wildtype receptor's stability in either the active or inactive state, to generate receptors with higher thermostabilities, which then can be used in other applications. Small arrows on GPCRs represent conformational flexibility. Note its absence on the thermostabilized receptor.

although in many cases a combination of ICL3 insertion and mutagenesis were used[87,88]. Systematic mutagenesis work has been carried out, demonstrating the thermostabilization of receptors by mainly replacing leucines to alanines and alanines to leucines (though other mutations also work), which locked the receptor into a specific conformation, so-called Stabilized Receptors (StaRs)[50]. This technique has proven successful in the generation of antagonist, partial agonist, and agonist-bound structures[89–91].

Various directed evolution techniques such as CHESS[51] and SaBRE[52] have also been applied to GPCRs to screen and select for stabilizing mutations. These techniques rely on the detection of highly expressed mutants using fluorescently labeled ligands and flow cytometry. The increase in expression and ligand binding is thought to be linked to increase in properly folded receptors and increased thermostability[92]. Given the time and effort required to find suitable thermostabilizing mutations and the low success rate of scanning mutagenesis, directed evolution offers a faster route to a more thermostable receptor. Typically, the process to discovering suitable mutations has a hit rate of less than 10 and hundreds of mutations are tested. In contrast, 2 to 3 rounds of CHESS can result in a thermostable receptor with a significantly higher level of expression than the wildtype receptor.

Biophysical studies

Although surface plasmon resonance (SPR) and nuclear magnetic resonance (NMR) are commonly used techniques in drug discovery, their application to GPCRs has been limited due to receptor instability and low expression[91]. There has been some success in applying these methods to the wild-type B2 receptor[93], however the generation of StaRs provides a solution to this obstacle. Unlike wild-type receptors, StaRs exhibit wild-type-like binding affinity only to the class of drug (inverse agonist, antagonist or agonist) which was used during StaR generation, due to conformational selection, with a reduced affinity for other classes. While this may bias drug discovery efforts towards a certain drug class, it may also provide a valuable method in the discrimination between agonist and antagonist hits during screening.

Neuronal Signaling

Engineered GPCRs have been used extensively in optogenetic and chemogenetic applications in neuroscience. The Opto-XR class of engineered GPCRs are chimeric receptors, engineered by grafting the intracellular region of the receptor of interest onto the light-sensitive transmembrane region of rhodopsin. These light-sensitive receptors have been extensively used to study the spatiotemporal effects of receptor activation, and the subsequent behavioral changes elicited by receptor activation. The newest member of this chimeric family, Opto-MOR[94], a rhodopsin-mu opioid receptor chimera, has been shown to recapture the signaling properties of the endogenous mu opioid receptor upon light activation. This new chimera has allowed Suida and colleagues to study receptor activation in mouse models. Another prevalent tool in neuroscience are Designer Receptors Exclusively Activated by Designer Drugs (DREADDs). Whereas Opto-XR design relies on a grafting design method, DREADDs are generated by multiple rounds of directed evolution, to select for mutant receptors that are activated by a normally inert small molecule, while losing the ability to be activated by the endogenous ligand. Initial DREADDs were based on the human muscarinic receptors, which were evolved and selected to bind and be activated by clozapine-N-oxide (CNO). Different DREADDs have been developed to signal through each major downstream effector, including Gq, Gi, Gs and beta-arrestin. There are excellent reviews comparing the two technologies[95,96].

Signaling assays

Other GPCR engineering techniques include the Tango system, which relies on a GPCR-TEV cleavage site-transcription factor fusion protein in conjunction with a beta-arrestin-TEV protease fusion[97] and a recently developed variety of this system using a CRISPR dCas9 in lieu of a transcription factor, capable of multiplex signaling[98]. Kroeze et al. used the Tango system to develop PRESTO-Tango, a high throughput cell-based assay to detect drug activation of the GPCRome[99]. Given the importance of biased signaling in cellular functions and its implication in disease, these techniques constitute an important step toward the discovery of biased ligands.

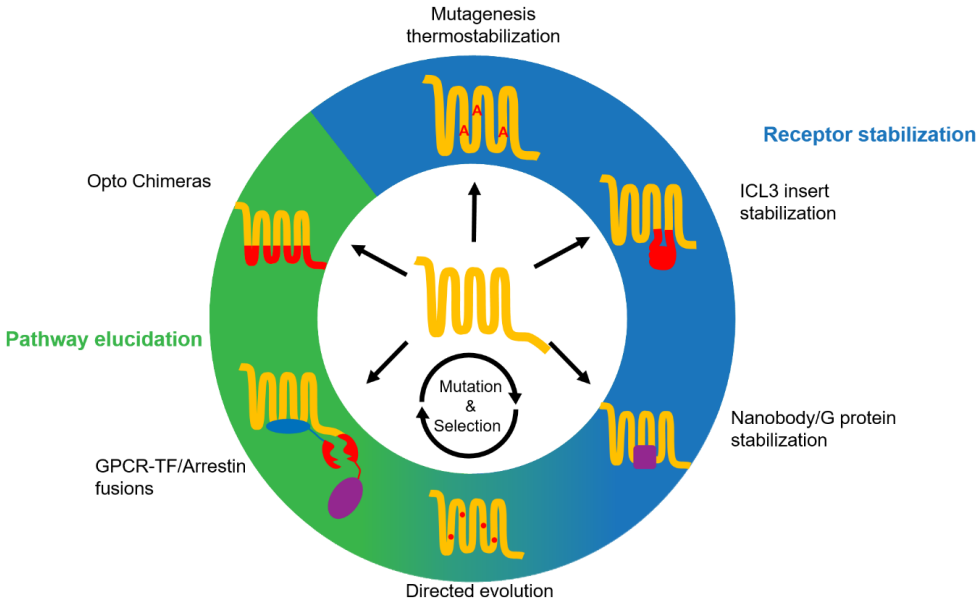


Figure 1.14 – Summary of GPCR engineering experimental methods. From the top going clockwise, the methods are: thermostabilization via alanine scanning mutagenesis, stabilization by replacing intracellular loop 3 (ICL3) with an easily crystallizable soluble protein, using a nanobody or full or partial G protein, directed evolution of receptor through multiple mutagenesis and selection cycles, creation of GPCR-cleavage site-transcription factor/Crispr-dCas9 fusion and arrestin-protease fusion, and the replacement of ICLs of a photosensitive receptor with that of another receptor. Blue background indicates methods used to facilitate crystallization. Green background indicates the methods are used in pathway determination and neuroscience application. Note the blue-green background for directed evolution, which has been used in both applications.

Limitations of experimental methods

While the outlined methods have proven their utility and have provided valuable insights, they suffer from numerous weaknesses. As previously mentioned, scanning mutagenesis for the purposes of thermostabilization usually requires the testing of hundreds of mutants, has a hit rate of roughly 10, and takes months of experimental investigations[50]. Furthermore, these receptors tend to lose their signaling functions. Directed evolution is a possible alternative approach but are currently limited to selection in bacteria or yeast systems, thus preventing the selection for functional signaling properties in situ (i.e. in mammalian cells). Additionally, since known ligands, and their radiolabeled or fluorescently labeled analogs, are mandatory for selection, orphan GPCRs are excluded from such approaches. Opto-XRs, consisting of rhodopsin transmembrane helices and a target receptors ICLs, lose the dimerization or oligomerization propensity and selectivity of the target receptor that depend on specific transmembrane helix associations. Therefore, the chimeric receptor may not recapitulate the interactome profile (e.g. homo and hetero oligomerization[83,100,101]) and associated cellular functions of the target receptor[5]. DREADDs present their own limitation directed evolution methods have only resulted in receptors with CNO as their inert ligand. The latest DREADD

based on the kappa opioid receptor (KOR) were engineered through structure-based rational design approaches, with a new agonist, Salvinorin B[102]. Nonetheless, with only two possible physiologically inert compounds as possible ligands, the usefulness of DREADDs is still limited, especially in applications where a multiplexed approach or sensing native ligands is necessary. As computational design has been used for the creation of KOR, it can offer solutions for biosensors, thermostabilizing receptors, generating orthogonal pathways, and even for fine-tuning receptor sensitivity (see sections below).

1.3.4 Computational design of GPCR structure and function

Computational design approaches have proven quite useful for engineering soluble protein fold, stability, binding and catalysis. Recent successes on membrane proteins include the design of a Zinc transporter[103] and symmetric transmembrane helical (TMH) oligomers[104]. While these designs validated important basic principles underlying membrane protein structure and function, they also ignored key features that constitute the hallmark of naturally-evolved multipass membrane receptors such as GPCRs. Signaling functions require the receptor to rapidly switch between conformations and to enable long-range communication between both sides of the membrane. Such complex functions rely on protein sequences that can intrinsically adopt multiple conformations and shift between these conformations upon external stimuli in the form of ligand, lipid or protein binding. We outline below recent progresses toward the development and application of computational approaches reprograming GPCR structure and function.

Designing GPCRs with enhanced conformational stability

Protein stabilization has been one of the hallmarks of computational protein design. The first computational design method developed for stabilizing multi-pass membrane proteins identified metastable sites in GPCR structures through sequence conservation and measures of suboptimal polar interactions and protein packing defects. Then, the method automatically selected amino-acid substitutions at these sites that enhance protein contacts and conformational stability[105,106]. The approach initially validated on the beta1 adrenergic receptor (B1AR) achieved more than 80% success rate, generated higher thermostabilization than empirical screening approaches and predicted correctly many experimentally selected thermostabilized variants of the adenosine A2A receptor (A2AR) and B1AR (Figure 1.15a). A similar approach combining in silico mutagenesis screening, sequence conservation and machine learning algorithms trained on known GPCR mutations has also been used recently to thermostabilize the 5HT2c receptor[107]. In silico design calculations rely however on high-resolution structural information which is lacking for a large majority of GPCRs. To be useful for structural and biophysical characterization, computational design and protein structure prediction techniques were combined. Homology modeling can generate reliable structural models of a target GPCR, providing GPCR structures with sufficient homology (i.e. 25% sequence identity in the TMH region) are available[106,108]. Using models generated from the B2AR-Gs

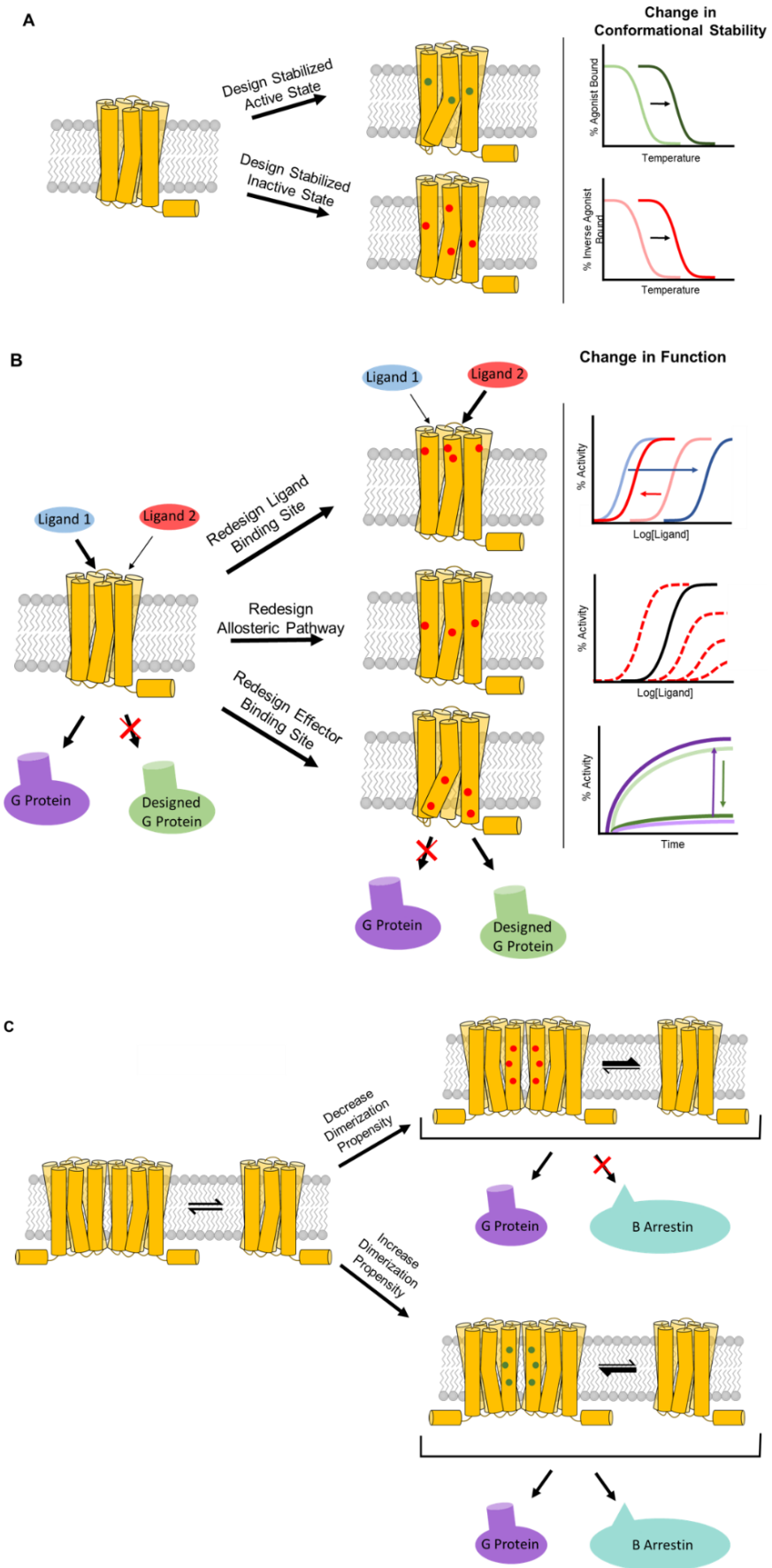

A
Design Stabilized Active State
Design Stabilized Inactive State
Change in Conformational Stability
% Agonist Bound
Temperature
% Inverse Agonist Bound
Temperature
B
Change in Function
Ligand 1
Ligand 2
Redesign Ligand Binding Site
Redesign Allosteric Pathway
Redesign Effector Binding Site
G Protein
Designed G Protein
% Activity
Log[Ligand]
% Activity
Log[Ligand]
% Activity
Time
C
Decrease Dimerization Propensity
Increase Dimerization Propensity
G Protein
B Arrestin
G Protein
B Arrestin

Figure 1.15 – Summary of GPCR properties amenable to computational design. **A.** The stabilities of the active and inactive states of the receptor can be modified independently, to produce receptors that exhibit higher melting temperatures when bound to selective ligands (right). **B.** The wildtype receptor and its upstream and downstream interacting partners (left), and designed receptors with novel functions (right). A modified ligand binding region (top right), transmembrane helical region (center right), and effector binding region can be designed independently, or in combination. The modified functional output for each novel design is depicted in the form of dose response assays (top and center right) or as an activity assay for effector binding (bottom). **C.** The population of receptors present in monomeric or oligomeric states can also be modified (wildtype on left, design on right) by redesigning the interface between the monomers. A weakened interface for the CXCR4 receptor results in decreased recruitment of β-arrestin.

ternary active state structure, dopamine D2 and A2AR variants were designed with enhanced active state stability, agonist binding and constitutive activity[109]. One designed agonist-bound A2AR crystallized in a close to fully active conformation, despite the absence of G-protein Gs (Lai et al., unpublished results). By contrast, A2AR variants previously thermostabilized through mutagenesis screening remained in a partially active conformation and lost their ability to couple to G-proteins and signal[110]. These results suggest that computational design approaches can stabilize GPCR conformations without disrupting ligand binding, G-protein coupling or allosteric signal transduction receptor functions. Hence, these methods should prove useful for accelerating the characterization of receptor structures and their interactions with pharmacologically important, weak or partial agonists which lack the potency required for structure determination.

Toward the computational design of GPCR biosensors

GPCRs have evolved to recognize, sense and respond to a wide diversity of small molecule, peptide, lipid and protein stimuli. These unique properties suggest that the GPCR structure fold constitutes a versatile platform for developing novel biosensors. In recent years, computational design techniques have been developed to manipulate and reprogram ligand binding, allosteric transmission and coupling to intracellular signaling proteins with the goal of designing GPCRs with novel functions.

Designing GPCRs with reprogrammed ligand binding selectivity

In contrast to DREADDs, which are stimulated by inactive drug compounds, GPCRs designed with fine-tuned ability to sense natural ligands should prove useful for studying and rewiring ligand-induced cellular signaling pathways. Due to the receptors high intrinsic conformational flexibility, accurately predicting GPCR-ligand interactions and conformations has remained a challenge especially for receptors without solved structures. By integrating receptor homology modeling, ligand docking, and computational design techniques, Feng and colleagues have developed the software IPHoLD, for modeling and designing GPCR-ligand interactions[108].

Unlike alternative techniques, IPHoLD recapitulated ligand-induced fit effects on receptor conformations and predicted receptor ligand binding selectivity. Functional dopamine D2 variants with novel ligand binding selectivity profiles were successfully engineered using the method (Figure 1.15b). In principle, the method could be expanded to model and design GPCR interactions with peptide, lipid and even proteins.

Designing GPCRs with reprogrammed allosteric signal transduction properties

Extracellular signals triggered by ligand binding are propagated on the other side of the membrane through allosteric communication pathways encoded by networks of highly coupled receptor residues. Residue coupling across long-distance enables efficient transmission of changes in protein structures and dynamics which occur upon ligand binding and can be inferred from correlated movements observed in Molecular Dynamics simulations [78]. Mutations known to affect ligand responses and receptor activation often clustered around the allosteric pathways predicted by conformational dynamics [78]. Keri and colleagues developed a method to rationally engineer GPCRs with altered ligand-induced signaling responses by designing amino-acid microswitches that rewire predicted allosteric pathways [53,111,112]. When applied to the dopamine D2 receptor, designed variants displayed either enhanced or decreased G-protein activation responses to agonists which were largely consistent with the predicted responses (Figure 1.15b). Gain of function mutations switched the D2 receptor from a dopamine to a highly sensitive dopamine and serotonin biosensor. Residue functional coupling can also be inferred albeit more indirectly from sequence covariation analysis [80]. By swapping co-evolving residues between a pair of GPCRs, Sung et al. were able to modulate receptor ligand responses [81]. Since GPCRs can trigger distinct signaling pathways upon different ligand stimuli through a mechanism called biased agonism, current efforts involve the identification and manipulation of functionally selective allosteric pathways [113,114]. Such approaches should enable the design of highly pathway selective biosensors and prove particularly useful in synthetic cell engineering approaches that rely on well-isolated and controllable cellular functions.

Designing GPCRs with reprogrammed signaling output properties

Another avenue to bias or even rewire GPCR-triggered cellular functions consist in designing the binding interactions between the receptor and its downstream transducers, G-proteins and beta-arrestin. The increasing number of GPCR structures bound to Gs, Gi and beta-arrestin uncovered the high conformational diversity of the GPCR-downstream effector binding interfaces and revealed why predicting GPCR-effector recognition from sequence remains a daunting challenge [115]. Capitalizing on existing GPCR-effector structures, Young et al. have developed a method to model by homology and design GPCR-G-protein interactions. Using the approach, they created novel orthogonal dopamine D2-Gi pairs that signal with high selectivity [116]. Unlike the promiscuous D2 WT, the D2 variants did not couple to Gq or beta-arrestin and the Gi variants were solely activated by the designed cognate D2 receptors (Figure

1.15b). Taking advantage of the modular architecture of G-proteins which mainly engage the GPCRs through their C-terminal helix 5, Bourne and colleagues showed that GPCRs can be redirected to activate non-native G-proteins and pathways by swapping a few residues of the non-native G-protein helix 5 to those of the native transducer that optimally couple to the GPCR[117]. Young et al. further extended that approach and designed orthogonal Gs-i chimeric proteins that enabled D2 receptors to trigger activating Gs-mediated instead of the inhibitory Gi-mediated cellular functions upon dopamine stimulus.

Altogether, the above-mentioned computational approaches pave the road for designing novel GPCR-based biosensors and highly selective cellular signaling pathways.

Computational design of GPCR oligomerization

GPCRs can either self-associate or hetero-associate with related GPCRs resulting in multiple combinations of functional units with potentially distinct signaling properties[118,119]. Feng and colleagues have developed a computational approach to model, dock GPCRs and design novel GPCR associations (Feng et al., unpublished results). They applied the technique to modulate the self-association propensity of the chemokine receptor CXCR4 which is known to form constitutive homodimers and higher order oligomers. CXCR4 variants with enhanced or weakened dimerization propensities were designed and displayed changes in dimerization measured by BRET which were consistent with the predictions. Interestingly, all the designs could efficiently activate Gi while those with weakened dimerization propensity were largely impaired in their ability to engage beta-arrestins (Figure 1.15c). The study revealed an unforeseen role of GPCR associations in regulating the receptor signaling selectivity and paves the road for engineering GPCR associations with fine-tuned functions.

Limitations of the computational techniques

The accuracy of GPCR structural models remains the main limitation of current computational techniques which require substantial structural information from not too distant homologs. Current successes have been limited to the design of small ligand-GPCR interactions and additional development will be needed to accurately model peptide, lipid and protein binding to receptors. Lastly, to achieve high computational efficiency, design calculations have often been performed using implicit models of the lipid membrane which neglect the molecular details of solvent and lipid molecules that are known to play important roles in regulating GPCR structure and function. As recently demonstrated by Lai et al. with a hybrid solvation model developed for membrane proteins[120], further developments will have to find appropriate tradeoffs between accurate representation of the lipid environment and calculation efficiency.

1.3.5 Conclusions

The GPCR engineering field has witnessed a rich spectrum of technological developments for reprogramming the structural and functional properties of this large class of receptors. As often, empirical and computational methods have advanced in parallel and complement each other. As major advances in cryo-electron microscopy now enable the routine structure determination of GPCRs in different functional states[121–124], the GPCR engineering field will likely move from a focus on structural stabilization towards accelerating drug discovery and designing new functions. The years to come promise to be an exciting time for GPCR engineering.

2 Description of Methods

Our lab uses a combination of computational and experimental tools. The primary computational tool we employ is Rosetta, a hybrid physics and knowledge-based protein structure prediction software suite. To modify allosteric pathways in GPCRs we supplemented our use of Rosetta with molecular dynamics (MD) simulations and normal mode analysis (NMA). We take advantage of Rosetta's unmatched ability to explore an astronomically large sequence-space at a limited number of conformations. The dynamics of interesting designs from Rosetta are then analyzed with NMA or MD, a function Rosetta mostly lacks. Together these methods have allowed us to generate the allosteric dopamine D2 designs detailed in subsequent chapters. Following design, our structure-based predictions are always followed by experimental validation. This chapter offers a brief description of each of the computational and experimental methods that feature heavily in our work.

2.1 Computational Methods

2.1.1 Rosetta

Rosetta uses a number of weighted physics-based and knowledge-based terms to score protein structures, listed in Table 2.1. The total score of a particular protein structure is the sum of the scores of each individual residue and pairwise interaction in that structure. Rosetta is based on the principle that the native structure of the protein corresponds to the lowest energy conformation a sequence can form (à la Anfinsen's principle). In this vein Rosetta uses a stochastic gradient descent algorithm to minimize the score of a structure given by the energy function (see the subfigure in Table 2.1 for the obligatory energy funnel diagram) [126,127].

A typical energy minimization protocol proceeds by alternating between making small changes to dihedral angles, followed by repacking rotamers (different conformations of sidechains), x number of times. Rotamers are picked based on their probabilities for the current ϕ and ψ backbone dihedral angles from the Dunbrack [128] rotamer library. Rotamers may be either the WT amino acid or anything that is allowed by the protocol. After each step the energy of the

Energy term	Description
fa_atr	attractive energy between two atoms on different residues separated by a distance d
fa_rep	repulsive energy between two atoms on different residues separated by a distance d
fa_intra_rep	repulsive energy between two atoms on the same residue separated by a distance d
fa_sol	Gaussian exclusion implicit solvation energy between protein atoms in different residues
lk_ball_wtd	orientation-dependent solvation of polar atoms assuming ideal water geometry
fa_intra_sol	Gaussian exclusion implicit solvation energy between protein atoms in the same residue
fa_elec	energy of interaction between two nonbonded charged atoms separated by a distance d
hbond_lr_bb	energy of long-range hydrogen bonds
hbond_sr_bb	energy of short-range hydrogen bonds
hbond_bb_sc	energy of backbone−side-chain hydrogen bonds
hbond_sc	energy of side-chain−side-chain hydrogen bonds
dslf_fa13	energy of disulfide bridges
rama_prepro	probability of backbone ϕ, ψ angles given the amino acid type
p_aa_pp	probability of amino acid identity given backbone ϕ, ψ angles
fa_dun	probability that a chosen rotamer is native-like given backbone ϕ, ψ angles
omega	backbone-dependent penalty for cis ω dihedrals that deviate from 0° and trans ω dihedrals that deviate from 180°
pro_close	penalty for an open proline ring and proline ω bonding energy
yhh_planarity	sinusoidal penalty for nonplanar tyrosine χ3 dihedral angle
ref	reference energies for amino acid types

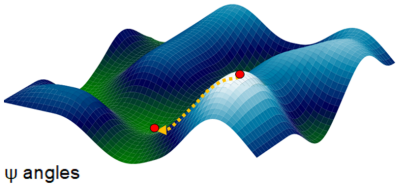


Table 2.1 – Rosetta score function energy terms in ref2015. Descriptions are adapted from Alford et al [125]. The subfigure exemplifies the score function minimization performed by Rosetta.

new structure is calculated, and if lower than the previous conformation's score, the step is accepted. If the conformation's score is higher than its predecessor's the Metropolis criterion is used to accept or reject the move. The probability of accepting a move with the Metropolis criterion is:

$$P(\Delta E) = e^{\Delta E/-Kt} \tag{2.1}$$

ΔE being the change in score and Kt is referred to as the temperature factor and may be adjusted up or down (a.k.a. simulated annealing) to control the strictness of the protocol. Multiple such minimizations are run in parallel due to the stochastic nature of the algorithm to maximize the probability of sampling the global energy minimum.

Through the application of different protocols built on these basic algorithms, Rosetta has seen great success in generating *de novo* protein folds [129], enzymes [130], therapeutic peptides [131], membrane proteins [132], etc. Our work presented in subsequent chapters does not make such drastic changes to protein structure or sequence. Rather, we attempt to make atomistically accurate smaller-scale changes that selectively alter protein function.

2.1.2 RosettaMembrane

RosettaMembrane was developed by our lab specifically to simulate the membrane environment home to integral membrane proteins [133]. In addition to the general energy terms used by Rosetta, RosettaMembrane includes a number of terms to more accurately calculate van der Waals and hydrogen bonding scores within the membrane. The membrane itself is implicitly modeled by three continuum phases: a water phase, lipidic phase, and a phospholipid head-

group phase. The contribution to the solvation term of each phase depends on position within the membrane. With these modifications, RosettaMembrane has more accurately recaptured native rotamer conformations within the membrane during benchmarking.

2.1.3 IPhold

Another Rosetta protocol developed in our lab to better recapture native bound ligand poses during ligand docking is Iphold[108]. Typical ligand docking protocols use a rigid backbone and do not take small conformational selection or induced-fit effects into account. Additionally, a common issue encountered in GPCR homology modeling involves the building of *de novo* loops occluding or occupying the ligand binding site. Iphold solves these issues by removing loops prior to ligand docking and incorporating a flexible backbone ligand docking step. The overall protocol consists of 2 cycles: In cycle 1, loops are removed and a coarse-grained ligand docking is performed, followed by a typical relax protocol and *de novo* loop modeling with ligand present. The 10% lowest energy relaxed decoys are clustered, and the ~5 largest cluster centers are used in cycle 2. This second cycle is a typical ligand docking protocol on each cluster center. We have previously shown this to improve ligand docking accuracy.

2.1.4 Multi-state Design

Typical Rosetta protein design optimizes sequence for one specific backbone conformation (i.e. they're single-state designs). With the multi-state design protocol, sequence can be optimized simultaneously for multiple conformations, each of which may involve different binding partners[134]. Multi-state design then evolves allowed sites in each state to maximize fitness according to a user-specified fitness function for n generations, with p population. The fitness function specifies the fitness of each state. In each generation, each member of the population is mutated randomly according to a genetic algorithm. Then, the fitness of each member is evaluated. The 10% most fit members are kept while the rest are culled. This cycle of mutation and selection go on for n generations. If the fitness function was properly set up the end result should be models matching the desired outcome. For example, we use the multi-state design protocol in Chapter 3 to maximize stability of the active state dopamine D2 receptor while simultaneously destabilizing the inactive state. We have also used it successfully in the past to design an orthogonal receptor-G protein pair[116]. One drawback of multi-state design is its use of rigid backbones.

2.1.5 Normal Mode Analysis

Most proteins are intrinsically dynamic molecules, naturally sampling a range of conformations[135,136]. Multiple of such microstates within this ensemble can be biologically relevant states. Protein structures tend to oscillate between these microstates, proportionally occupying the most stable microstates. These oscillations turn out not to be random thermal

fluctuations, but rather concerted, global movements. The oscillations between these states and their frequencies can be reliably predicted by representing the protein as a network connected by elastic springs and using normal mode analysis to examine the global motion of residues at different modes. Discrete modes are associated with different frequencies of oscillations around a global or local minimum, with lower modes corresponding to lower frequency, large-scale motions; fluctuations on the order of microseconds to milliseconds and above. These lower frequency modes are also more energetically favorable, meaning they are more accessible. The energy of a given normal mode is proportional to the square of its frequency[135]. In the case of rhodopsin, the motions given by lowest 20 modes showed an excellent correlation to the transition seen between 2 distinct, experimentally determined states[137]. Furthermore, it has been shown that alterations in the lowest frequency normal modes via allosteric mutations can have a functional impact on proteins[138]. In the case of pyrimidine attenuator regulatory proteins, allosteric mutations modifying normal modes were able to tune the propensity of the protein towards or away from tetramer formation.

Normal mode analysis is great determining slow microsecond/millisecond global correlated harmonic motions, though not all may be functionally relevant. To remove superfluous residues and only consider the most important allosteric hub residues involved in relaying information between the ligand-binding site and the G protein-binding site in the dopamine D2 receptor, we turned to MD. From MD trajectories we extracted the putative functionally important allosteric hubs and calculated the dynamic cross-correlation matrices (DCCM) from both the active and inactive receptor states, using the lowest 20 modes as mentioned above for rhodopsin. We then searched for mutations *in silico* at the allosteric hubs that alters the DCCM difference between the active and inactive states.

2.1.6 Molecular Dynamics Simulations and Allosteer

In contrast to the elastic network model used in NMA, MD simulations involve full-atom simulations faithfully considering the inter-atomic forces and Newtonian motions involved. The resulting increase in computational cost limits the timescale of the simulations normally to the nanosecond and microsecond ranges, however. Our main interest in MD is its ability to identify residues with concerted motions in an atomically accurate simulation. We use the results from and protocol (Allosteer) developed by Vaidehi[78]. From their microsecond simulations (with explicit solvent and phosphatidylcholine (POPC) membrane) they calculate correlations in dihedral angle changes, using a mutual information-based metric. Allosteric paths are generated from the ligand-binding site to the effector binding site using a weighted shortest-path length (Dijkstra's) algorithm, the weights being the mutual information correlations. These paths are then clustered based on overlap to give the final allosteric "pipelines". We refer to the residues within the allosteric pipelines as allosteric hub residues and modify them in our work.

2.2 Experimental Methods

2.2.1 TRP Channel Assay

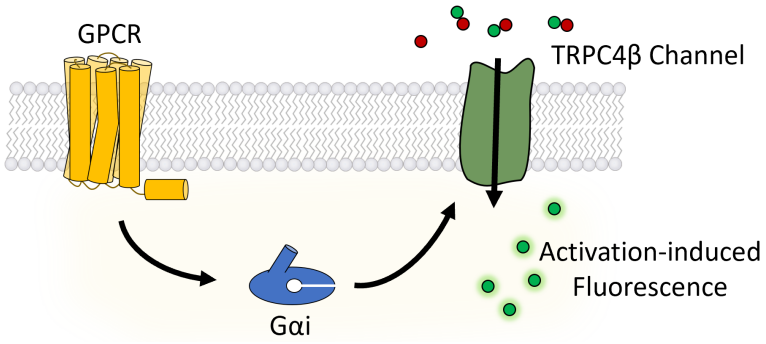


Figure 2.1 – TRP channel assay diagram. Gi-activation of TRP channels results in the influx Ca^{2+} to the cytosol. The modified membrane potential allows the fluorophore (green) to translocate into the cell, away from the quencher (red).

This assay relies on a stably transfected HEK293 cell line which expresses the TRPC4β non-selective cation channel. TRPC4β is activated by (usually) endogenous Gi upon Gi's activation by Gi-coupled GPCRs. The TRP channel itself is a homotetramer, requiring the introduction of only a single gene to function [139]. The assay is performed in a 96-well plate format, with cells transiently transfected with the GPCR of interest 24 hours prior to measurement. A fluorescence dye and quencher solution is applied to wells, incubated, and changes in fluorescence are measured prior to, during, and for a few minutes after drug addition. If properly set up, fluorescence changes in a drug concentration-dependent manner.

The main advantage of this assay is its sensitivity. Very low DNA quantities of receptor is required during transient transfection to generate a strong signal. Many of the receptor variants we generated in lab tend to show poor levels of expression and this assay allows us to measure their activity reliably. The assay is also fairly rapid. A full 96-well plate can be measured in 1 hour.

The high sensitivity of the assay is also its disadvantage, however. If receptor expression is too high (which is very easy to achieve), the TRP channel signal can quickly saturate. In other words, the dynamic range of the assay is very low. This low dynamic range also makes normalization impossible for differing expression receptors, and DNA titrations must be performed instead to equalize surface expression. Assay kinetics are also very sensitive to variations in experimental conditions such as cell number, DNA quantity, well volume, temperature, etc. so extreme care must be taken to ensure replicates are highly homogenous.

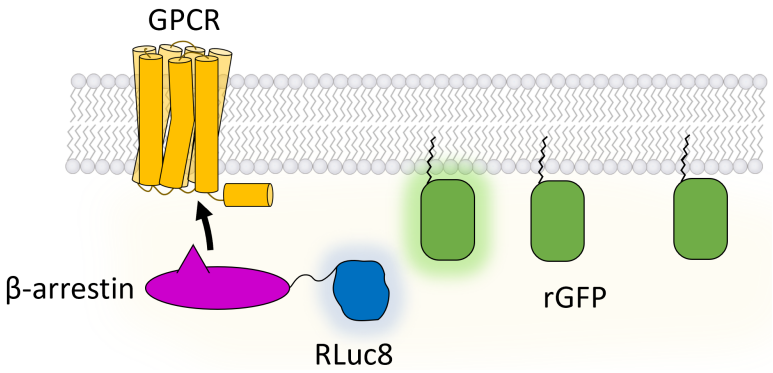


Figure 2.2 – Enhanced bystander BRET β-arrestin recruitment assay.

2.2.2 Enhanced Bystander BRET β-arrestin Recruitment Assay

This assay measures the bioluminescence resonance energy transfer (BRET) between a donor bioluminescent molecule (in this case Renilla Luciferase8 fused to the c-term of Beta arrestin 1 or 2) and a fluorescent acceptor protein, *Renilla* GFP fused to a k-Ras membrane localization sequence (CVIM), separated by a linker. The cysteine in the sequence will be myristoylated, tethering the GFP to the membrane. Activation of receptor on the cell surface will recruit Barr-RLuc8 to the membrane, to the vicinity of the membrane-tethered GFP. This proximity enables the energy transfer from the donor to the acceptor, producing a GFP signal which can then be measured. While the original protocol has been developed by the Bouvier lab [140], we modified it to reduce total assay time from 4 days to 2 days. This assay has recently been also shown to work with G protein or G protein-associated effector recruitment to the plasma membrane [16].

The advantage of this assay is its quick and simple setup. The assay only requires the luciferase substrate coelenterazine and the receptor agonist. The assay is also fairly insensitive to small variations in well conditions, resulting in low error. Rather than measure an intracellular

2.2.3 Cell-based Enzyme-linked Immunosorbent Assay (ELISA)

The ELISA is an immunoassay based on a primary antibody targeting an epitope on the protein of interest and a secondary antibody linked to some form of reporter protein (in this case to horseradish peroxidase) targeting the primary antibody. In our case the primary antibody targets a n-terminal 3xHA tag on our receptors of interest. Transfected cells are grown on a 96-well plate for 24 hours, then fixed within the plate to disallow further trafficking of proteins. Primary and secondary antibodies are applied which can not penetrate the cell membrane. A hydrogen peroxide solution is applied measure the quantity of antibody-bound receptors.

All cell-based assays performed in the lab are accompanied by an ELISA to determine receptor surface expression. To ensure reliable comparison of activity and sensitivity of receptor

variants, DNA titration experiments are performed to determine the DNA quantity required to equalize receptor variant surface expression in subsequent experiments.

3 Computational design of G Protein-Coupled Receptor allosteric signal transductions

Chen, K-Y.[1,‡], **Keri, D.**[2,‡], Barth, P.[1,2,3,4]

1 Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX, USA.
2 Swiss Federal Institute of Technology (EPFL), Institute of Bioengineering, Lausanne, Switzerland.
3 Department of Pharmacology, Baylor College of Medicine, Houston, TX, USA.
4 Structural and Computational Biology and Molecular Biophysics Graduate Program, Baylor College of Medicine, Houston, TX, USA.
‡ **The authors contributed equally**

Adapted from article published in Nature Chemical Biology, December 2019
Link to original publication: https://doi.org/10.1038/s41589-019-0407-2

Author Contributions: P.B. designed the studies. P.B., K.-Y.M.C. and D.K. performed the calculations. K.-Y.M.C. and D.K. performed experiments on D2. P.B., D.K. and K.-Y.M.C. analyzed the data. P.B., D.K., and K.-Y.M.C. wrote the manuscript.

3.1 Abstract

Membrane receptors sense extracellular stimuli and transduce these signals into intracellular signaling responses. Subtle differences in protein sequence often give rise to profound changes in signaling response with no obvious connection to their structure. Here, we report a computational approach for designing signaling function into structurally-uncharacterized G protein-coupled receptors. The method identifies allosteric sites by Molecular Dynamics simulations and engineers amino-acid “microswitches” at these sites using multi-state design and normal mode calculations that modulate receptor stability or long-range coupling, to reprogram specific allosteric signaling properties. Using the approach, we designed 36

dopamine D2 receptor variants, whose constitutive and ligand-induced signaling responses agreed well with our predictions. We repurposed the D2 receptor into a serotonin biosensor, which triggered potent signaling responses upon sensing the non-native serotonin ligand. We also quantitatively predicted the effects of more than 100 known mutations on the signaling activities of several GPCRs. Our results reveal the existence of distinct classes of allosteric microswitches and pathways that define an unforeseen molecular mechanism of regulation and evolution of GPCR signaling functions. Our approach enables the rational design of receptors with enhanced stability and function to facilitate structural, pharmacological characterization, interrogate and manipulate cellular signaling, and the development of biosensors for synthetic biology and cell engineering applications.

3.2 Introduction

Allostery is a fundamental property that enables long-range communication between distant sites on a molecule. It is at the origin of a large diversity of regulatory mechanisms of biological molecular functions but its biophysical underpinnings remain poorly understood. Allosteric communications are thought to be primarily mediated by intraprotein networks of coupled residues (i.e. allosteric residues) physically connecting extracellular and intracellular regions of the receptor [141,142]. Long-range structural coupling can promote the effective communication between distant protein sites through the propagation of even small changes in local structure and dynamics [69,143–145]. Coupled allosteric residues display correlated motions and can in principle be identified by molecular dynamics (MD) simulations [146–148]. Allosteric sites are functionally very sensitive to mutations [149], enabling the modulation of protein function even between closely related protein sequences [150]. These observations suggest that the residue couplings that regulate the intracellular responses to extracellular stimuli largely depend on amino-acid sequence details and fine protein structure and dynamic properties. Hence, predicting how receptor sequences and structures encode specific allosteric responses remains particularly challenging.

G protein-coupled receptors (GPCRs) [151] represent the largest family of membrane-embedded allosteric proteins which receive external stimuli in the form of light, hormones, odorants, neurotransmitters, or peptides; and allosterically transduce these stimuli across the membrane to initiate intracellular signaling pathways [73,152,153]. In the largest family (class A) of GPCRs, signal transductions involve a few highly-conserved allosteric sites (called microswitches) that undergo consensus structural changes triggering receptor activation [154–156]. However, these allosteric sites are not in direct contact and their conserved mechanism of action does not explain the large diversity of allosteric signal transduction responses displayed by class A GPCRs [157,158]. How the conserved microswitches are physically connected throughout the receptor structure and define the allosteric pathways that control the coupling between extracellular and intracellular binding sites and therefore the selectivity of the transduced signals remains poorly understood.

Better understanding and engineering allostery in receptors would greatly benefit not only the study of receptor structure, function, and pharmacology but also the design of novel biosensing receptors for synthetic and cell biology applications. So far, structural biologists have developed empirical approaches for screening stabilized mutants to facilitate structure determination [159]. However, the thermostabilized receptors often no longer exhibited the ligand-induced signal transduction response associated with the wild-type receptor, highlighting the challenges in understanding the determinants regulating receptor stability and function [160–163]. Protein engineering techniques which reprogram conformational stability and long-range allosteric coupling have not yet been reported.

Here, we developed a computational protein design approach linking protein conformational stability, dynamics, ligand binding and allosteric regulation of protein activity to rationally design receptors with reprogrammed signaling activities. We validated this approach by engineering the GPCR, dopamine D2, whose structure in the activated state is not yet known, with reprogrammed stability, constitutive and ligand induced G-protein activation. Remarkably, we were able to design D2 variants with considerably enhanced potency and signaling responses to the native ligand agonist dopamine and the non-native ligand serotonin. The approach is general and should prove useful for designing novel membrane receptor biosensors. Lastly, our study revealed new insights into the role of allosteric pathways in the regulation and evolution of GPCR signaling functions.

3.3 Results

3.3.1 Computational design framework of GPCR signaling properties

To implement our approach, we took advantage of the rich experimental structural information on GPCRs, the prediction of allosteric residues by MD simulations, and theoretical allosteric models linking protein conformational stability, ligand binding and long-range structural coupling to the regulation of receptor activation. Using this knowledge, we developed a general design approach combining protein structure, dynamics and allostery to engineer protein signaling functions described in detail below.

GPCR structures are composed of 3 main regions: the extracellular ligand binding pocket, the intracellular G-protein binding domain and the “transmission” transmembrane (TM) region which connects the two binding regions and allows them to communicate [73,152] (Figure 3.1a). GPCRs typically switch from inactive to active state conformations upon activating agonist ligand and G-protein binding at the extracellular and intracellular domains, respectively. The structural rearrangements upon receptor activation involve large intracellular reorientations of transmembrane helices (TMH) TMH6 and TMH7 and smaller scale movements of individual amino-acids (we define these residues as “microswitches”) across the entire TM domain (Figure 3.1a). As identified by MD simulations, many microswitches display correlated motions, are organized in sparse allosteric networks that couple the extracellular and

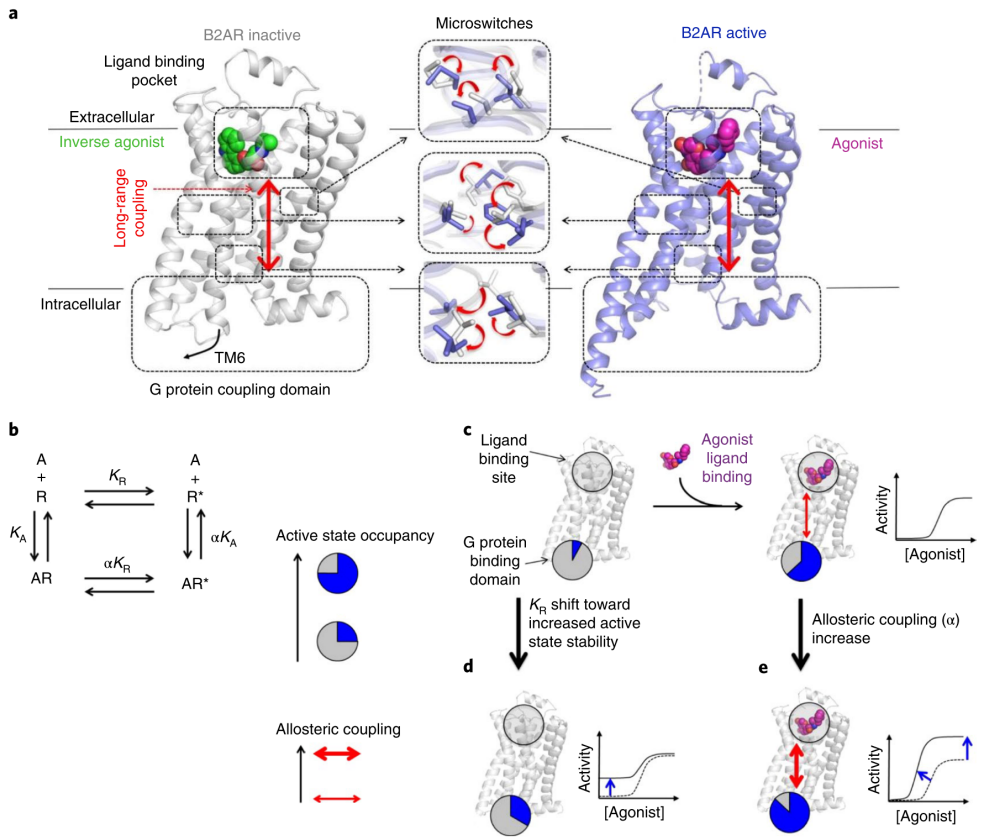


Figure 3.1 – Rational design of GPCR allosteric regulation. **a.** Left, Right. B2AR inactive (grey) and active (blue) state X-ray structures bound to inverse agonist and agonist ligands, respectively. The red arrow depicts schematically the structural coupling connecting the extracellular ligand binding pocket and the intracellular G protein binding domain. Middle. Amino-acid microswitches are highlighted for 3 local TM regions with their movements indicated by curved red arrows. **b.** Allosteric two state model (ATSM) describing the regulation of receptor activity by ligand binding. **c-e.** Effects of ligand binding, protein stability and allosteric coupling on receptor activity according to the ATSM. **c.** Effects of agonist ligand binding on receptor activity. Inactive and active receptor G protein binding site conformation occupancies are indicated in grey and blue pie fractions, respectively. Left. In absence of agonist, the G protein binding site mostly occupies the inactive conformation. Right. The agonist-bound ligand site preferentially couples to the active conformation of the G protein binding site, increasing active state occupancy and enhancing receptor intracellular activity. **d.** Increasing the stability of the active relative to that of the inactive state conformation (i.e. increase K_R) enhances the receptor activity even in absence of ligand (i.e. constitutive). **e.** Increasing the allosteric coupling (α) of the ligand agonist enhances the stabilization by the ligand of the G protein binding site active conformation and the intracellular response (i.e. activity) of the receptor to agonist binding. In this example, K_R and therefore the receptor constitutive activity remain unchanged.

intracellular domains of the receptor across long-distance [146–148], and are defined as allosteric microswitches in our study. More generally, by changing conformations during receptor activation, microswitches form new sets of contacts in distinct receptor functional states. Therefore, microswitches can selectively fine-tune the stability or the long-range structural

coupling within each receptor conformation, which can be employed to reprogram receptor activation and signaling as outlined below.

The regulation of protein signaling can be described using the allosteric two state model (ATSM) which provides a theoretical framework linking protein stability, ligand binding and allosteric coupling to the receptor activity[73] (Figure 3.1b-e, Figure 3.6). In this model, the receptor occupies two main functional states: an inactive (R) and an active (R*) state in which the ligand and G protein binding site conformations are the most coupled. In absence of ligand, most native GPCRs primarily occupy the more stable inactive state (i.e. $K_R < 1$, Figure 3.1b) and therefore display low constitutive activity[163] (Figure 3.1c left). Extracellular ligands often display different affinities for the inactive and active conformations of the ligand binding site and can therefore regulate receptor function by stabilizing distinct receptor states[163]. The ability of a ligand to selectively stabilize a receptor state is at the origin of the allosteric effect and measured by the allosteric coupling alpha (i.e. α, Figure 3.1b), which is related to the structural coupling as described below and in the methods. For example, agonists preferentially interact with the ligand binding active conformation (i.e. $\alpha > 1$) which is most coupled to the active conformation of the G protein binding site. Consequently, agonist binding preferentially stabilizes the receptor active state, leading to increased receptor activity and intracellular signaling (Figure 3.1c right). The reverse scenario occurs with inverse agonists (Figure 3.6).

According to the ATSM, receptor activity can be modulated through three distinct mechanisms (Figure 3.1d-e, Figure 3.6d-k). 1) Increasing the stability of the active compared to that of the inactive ligand-free structures (i.e. increase K_R) will enhance the receptor activity even in the absence of ligand agonist (i.e. constitutive activity) (Figure 3.1d). 2) Modifying the allosteric coupling (α) of a ligand will shift the receptor response to ligand binding. As demonstrated previously[164] and herein (Methods), this can be achieved by altering the long-range structural coupling between the ligand and G protein binding sites in selective receptor conformations. For example, optimizing the structural coupling of the agonist bound ligand site with the active but not with the inactive conformation of the G protein binding site will increase the G protein binding site active conformation stability and occupancy upon agonist binding. Therefore, it will enhance the allosteric coupling (α), the recruitment of G protein by the receptor and the intracellular receptor activity stimulated by the agonist (i.e. ligand response) (Figure 3.1e). Additional regulatory effects can be engineered by manipulating stability (K_R) and allosteric coupling (α) simultaneously (Figure 3.6g-k). 3) Altering the ligand binding affinity (K_A) will affect the ligand concentration triggering receptor activation (i.e. ligand potency, EC50 or IC50, Figure 3.6l). In our study, we reprogrammed constitutive and ligand-induced receptor activities through the first two above-mentioned mechanisms. As described below, these functional properties can be rationally engineered through the design of two specific classes (i.e. stability and allosteric) of microswitches.

To date, roughly 5% of GPCRs have been structurally characterized, thus limiting the application of rational structure-based protein design. Therefore, we developed a general method,

which combines structure prediction and design techniques and is in principle applicable to GPCRs and other membrane receptor classes (Figure 3.1d). First, we use the homology modeling and ligand docking softwares RosettaMembrane and IPHoLD to model ligand-free and ligand-bound inactive and active conformations of the GPCR using experimentally determined homolog structures [165,166] (Figure 3.1d, left). Then, we scan the TM domain in the aforementioned conformations to identify positions where microswitches with novel properties can be designed. Specifically, we mutate in silico multiple positions to all possible 20 amino-acids, and select microswitches that shift the stability or the structural coupling of specific receptor conformations (Figure 3.1d, right). For example, receptors with increased constitutive activity can be engineered through stability microswitches that make more stabilizing contacts in the active versus the inactive ligand-free conformations. On the other hand, receptors with enhanced signaling response to agonist can be engineered through allosteric microswitches increasing the structural coupling of the ligand-bound site to the active G protein site conformation while decreasing that to the inactive G protein site conformation.

The contribution to the receptor stability of a microswitch is calculated by the scoring function RosettaMembrane developed for the prediction and design of membrane protein structures [105,133] and largely depends on the number and strength of short-distance hydrophobic and polar atomic contacts with neighboring residues [133]. Since allosteric protein responses are most sensitive to mutations of coupled residues [142], we approximate the structural coupling by the interactions between the highly coupled allosteric microswitches which are calculated as follows. Depending on their locations, the structural coupling between allosteric microswitches involves local or long-range interactions (Methods). Local interactions are calculated through short-range aforementioned atomic contacts between the microswitches scored by RosettaMembrane [133]. Long-range interactions are calculated by the correlated movements of coupled microswitches which constitute the allosteric networks (Methods). Long-range structural coupling between microswitches is stronger when the correlation in their motions is higher. Perturbations in correlated movements are rapidly calculated by Normal Mode Analysis (NMA) [166]. NMA is an efficient simulation technique that can capture important functional motions in proteins and has been extensively applied to study the sequence, structure and dynamic origins of protein allostery [138,167] (Methods). The combination of techniques implemented in our approach enables the efficient scanning of billions of possible combinations of designed microswitches modulating stability and long-range structural coupling, which would be intractable using classical molecular dynamic simulations.

Below, we first describe the validation of our approach by engineering dopamine D2 receptor (D2) variants with reprogrammed signaling properties.

3.3.2 Design of D2 receptor signaling

D2 is involved in many central neurobiological functions and associated with numerous diseases including Parkinson, Alzheimer and schizophrenia. Engineered D2s with reprogrammed

signaling properties would provide novel molecules to better study downstream signaling and behavioral outcomes for example. We modeled D2 in the inactive and active states from the homolog dopamine D3 receptor (D3) and beta2 adrenergic receptor (B2) structures, respectively (Methods). D2 shares high sequence identity in the TM region (> 40%) with these homologs and suggests that the positions of allosteric residues can be mapped onto D2 using information derived from molecular dynamics simulations of B27. This mapping is then substantially improved using structural details of the D2 models enabling reliable calculations of receptor conformational stability and structural coupling (Methods). We validated the selected designed D2 variants using in vitro G-protein Gi activation, fluorescence measurements, ligand binding and intracellular signaling assays[139] (Methods).

3.3.3 Distinct effect of a stability or allosteric microswitch

Since allosteric protein properties are most sensitive to changes in coupled residues, we first tested the accuracy of our approach by mutating single allosteric residues in TM regions distant from the ligand and G protein binding sites. For example, our calculations indicated that 3 microswitches with distinct effects on stability or structural coupling could be designed at the residue Thr 5.54 (Ballesteros Weinstein notation[168], which corresponds to D2 position 205). Ile 5.54 was designed as a selective stability microswitch, making more optimal contacts in the active than in the inactive conformation as compared to the native Thr without affecting the structural coupling with other allosteric residues. (Figure 3.2a,b, Table 3.1). Consequently, it should selectively stabilize the active state even in the absence of ligand. Consistent with the predictions, we observed an increased constitutive activity with a maximal agonist-induced activity similar to WT (Figure 3.2c). Conversely, Met 5.54 and Val 5.54 were designed as selective allosteric microswitches inducing no effects on the relative stability of the receptor conformations (i.e. similar local contacts in both inactive and active states) but perturbing the structural couplings (Figure 3.2d,e,g,h, Table 3.1). Met 5.54 selectively increased the structural coupling between the agonist and the G-protein binding site in the active state and should therefore enhance the intracellular response to ligand agonists (i.e. increase the allosteric coupling α) (Figure 3.2e). Conversely, Val 5.54 selectively increased the structural coupling between the agonist and the G-protein binding site in the inactive state and should decrease the intracellular response to ligand agonists (i.e. decrease the allosteric coupling α) (Figure 3.2h, Figure 3.6a,f). As predicted, both microswitches had no measurable effects on the receptor constitutive activity. Met 5.54 increased receptor activity in vitro by up to 100% in response to the weak agonist serotonin while Val 5.54 significantly decreased receptor activity in response to both dopamine and serotonin (Figure 3.2f,i). To achieve these opposite effects, Met 5.54 and Val 5.54 perturbed distinct sub-networks of allosteric residues, consistent with alternative allosteric pathways coupling the extracellular and intracellular sites in the inactive and active receptor conformations7 (Figure 3.3).

These results suggested that our approach is sensitive enough to predict the effects of mutations on distinct receptor signaling properties, setting the stage for more extensive protein

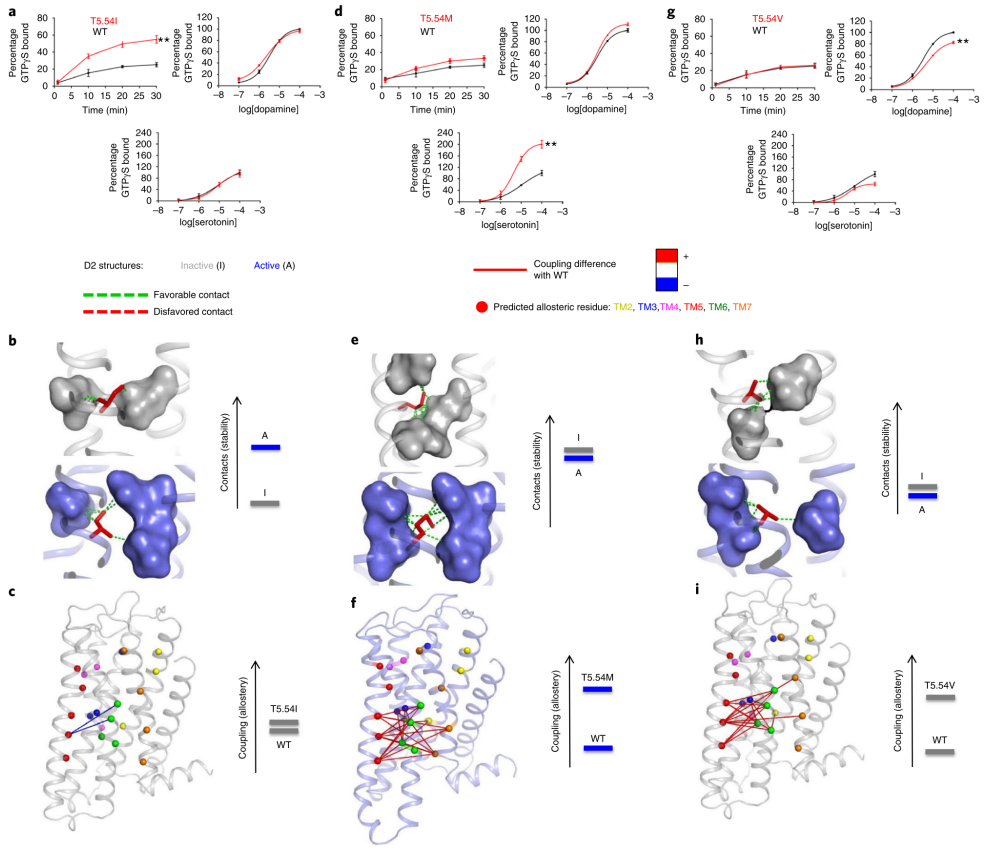


Figure 3.2 – Design of single microswitches modulating stability or long-range structural coupling. 3 microswitches at position 205 (5.54): T5.54I **(a,b,g)**; T5.54M (**c,d,h**); T5.54V (**e,f,i**). **a,c,e.** *In vitro* measurements of D2 activation (WT: black, design: red). Left. Constitutive (time course; N=3 independent experiments, data shown are mean ± s.d.); right, bottom. Agonist induced (dose response) as percent of maximal dopamine induced D2 WT activity (N=2 independent experiments, data shown are mean ± s.d.). ** two-sided unpaired t-test: $p < 0.05$. **b,d,f.** Comparison between local contacts (stability) of active and inactive D2 conformations. **g,h,i.** Changes in dynamic coupling from WT (differences in correlated motions of amplitude > 0.02 measured by Normal Mode Analysis, see method) induced by the designed microswitch are highlighted by a colored line (red and blue for coupling increase and decrease, respectively). The selected state (i.e. inactive, gray or active, blue) is the most perturbed by the mutation. The allosteric residues are indicated as spheres on the structure, colored based on their TM location.

engineering efforts. Unlike directed mutagenesis and deep sequencing approaches which often probe only one substitution at a time[160], structure-based computational design can scan through an astronomical number of combinations of mutations to create large changes in sequence, structure and function. We assessed whether D2 variants with a wide diversity of signaling properties could be designed through novel networks of microswitches.

3.3.4 Designed D2 receptors stabilized in the active state

Engineering D2 stabilized in an activated form would accelerate the structural characterization of intrinsically unstable active states which often requires formation of a ternary complex between the receptor, ligand agonist and downstream effectors[63,169–171]. Highly-constitutively activated receptors would also foster the study of the allosteric properties underlying inverse agonism and the discovery of potent long sought-after receptor inhibitors (inverse agonists).

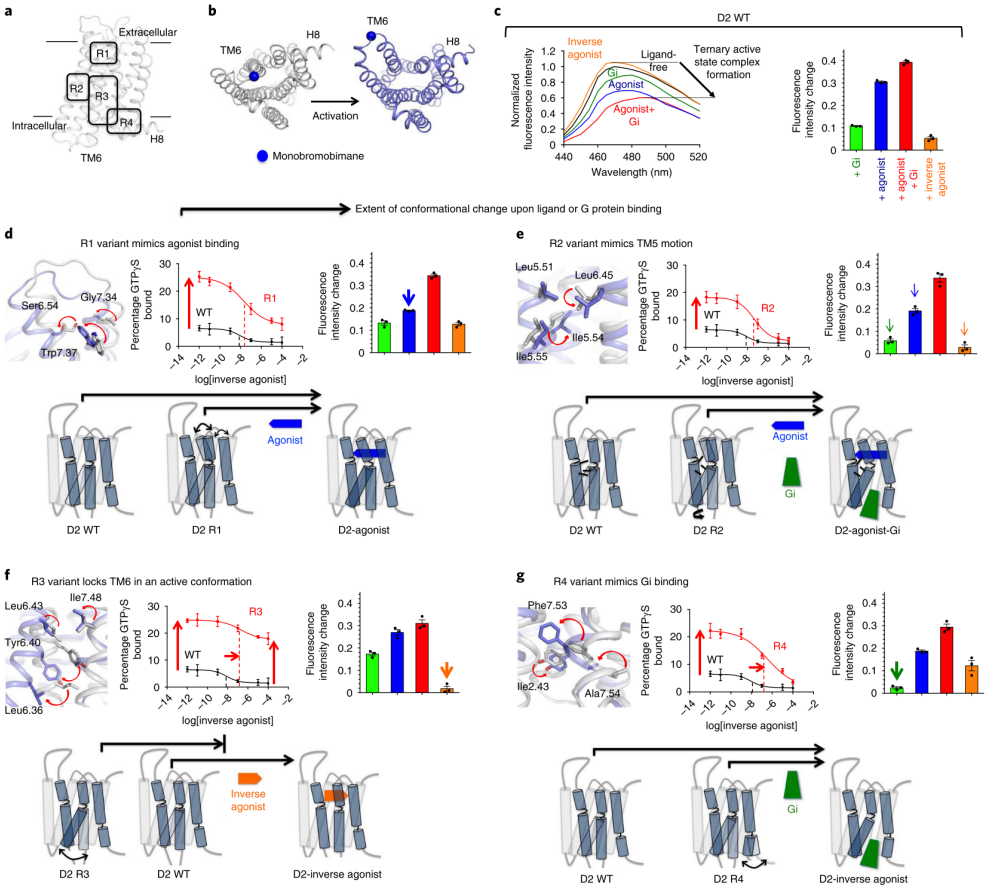


Figure 3.3 – Designed microswitches stabilize distinct local active state structures. **a.** R1-R4 designed regions span the entire TM domain. **b.** The cytoplasmic end of TMH6 is coupled to a fluorophore whose fluorescence is quenched and red-shifted upon receptor activation. **c.** Left. The largest shifts are reached when D2 is bound to both agonist and Gi in a ternary active state complex. One representative experiment is shown. Right. Changes in maximal fluorescence intensity upon binding to ligand or Gi are normalized to that induced by binding of both agonist and Gi. The absolute values of the changes are reported. **d-g.** Changes in receptor activity, fluorescence and proposed local conformational changes are indicated for microswitches designed in regions R1 **(d)**, R2 **(e)**, R3 **(f)** and R4 **(g)**. The designed microswitches are indicated in sticks on the inactive (grey) and active (blue) state D2 structures. In vitro measured downregulation of receptor activity by spiperone are reported for D2 WT (black) and designed variants (red). Receptor activity is reported as percent of maximal dopamine induced D2 WT activity. The largest changes from WT are indicated by a red arrow in the activity plots.

To achieve high conformational stability, we designed stability microswitches in 4 regions

(R1, R2, R3 and R4) spanning the entire TM domain (Figure 3.3a). As assessed by functional assays, fluorescence spectroscopy and structure modeling, each designed region partially activates the receptor by adopting specific local active conformations in absence of ligands or Gi (Figure 3.3). Fluorescence spectra were recorded for a fluorophore monobromobimane (mbb) coupled to the intracellular end of TMH6 which undergoes the largest conformational changes upon receptor activation[172] (Figure 3.3b). As observed for the beta2 adrenergic receptor (B2AR), by shifting the receptor to the active state conformation, binding of agonist, Gi, or both leads to quenching of mbb coupled to WT D239 (Figure 3.3c). The level of spectral change corroborates the extent of active state stabilization and receptor activation induced by ligand (agonist) or G protein binding (Figure 3.3c right). Therefore, if designed microswitches partially mimic the ligand (agonist) effects on receptor structure and activity, then ligand binding to the designed receptor should result in reduced spectral changes compared to WT (Method).

On the extracellular side of TMH6 and TMH7 forming part of the ligand binding site (region R1), we designed 3 microswitches shifting R1 toward the active conformation without substantially perturbing long-range structural couplings (Figure 3.3d; Table 3.2). Consequently, the conformational changes were efficiently propagated over a long-distance, stabilizing the intracellular active conformation and increasing the receptor constitutive activity without preventing the downregulation by the inverse agonist spiperone. Smaller fluorescent changes upon agonist binding compared to WT suggest that the designed microswitches partially mimic the agonist-induced conformational shifts corroborating our structural models (Figure 3.9). Deeply buried in the lipid membrane, the R2 region connects the extracellular and intracellular sides of TMH5 and TMH6, but R2 only partially overlaps with the highly conserved “transmission” region in rhodopsin-like GPCRs[154] (Figure 3.3e). Carefully designed microswitches shifted the region toward the active conformation without substantially perturbing the long-range structural couplings, despite multiple mutations at allosteric residues (positions 5.51, 5.54). Consistent with the predictions, the R2 variant displayed increased constitutive activity from WT but was fully switched off upon spiperone binding. Agonist and Gi binding to the R2 variant displayed a unique combination of fluorescent spectral shifts, consistent with specific residue movements designed at the TMH5-TMH6 helical interface as predicted by our structural models (Figure 3.9). The R3 region undergoes the largest conformational changes upon receptor activation and also comprises the highest number of allosteric residues among the designed regions (Figure 3.3f). The microswitches we engineered in the R3 region induced the largest conformational stability and structural coupling changes of all variants (Table 3.2). Fluorescence spectroscopy, modeling and Gi activation measurements confirmed that the region is locked in an active conformation which partially activates the receptor and does not respond to inverse agonist binding (Figure 3.3f, Figure 3.9). R4 is the region closest to the Gi intracellular binding site and incorporates the highly conserved NPxxY motif. We designed microswitches to rewire the interaction networks between TMH2 and TMH7, shifting the local R4 region toward the active conformation (Figure 3.3g). One of the microswitches at the conserved allosteric residue Y7.53 also decreased the structural coupling

in the inactive state, leading to a loss in inverse agonist potency. As revealed by the reduced spectral changes upon Gi binding and our modeling, the microswitches stabilize R4 in a conformation partially mimicking the Gi-bound WT receptor (Figure 3.3g, Figure 3.9).

We then analyzed the structural coupling properties of the designed R1-4 variants using NMA to understand the origin of their distinct responses to spiperone. We observed a strong correlation between the loss in ligand potency/efficacy and the fraction of allosteric networks affected by the designed microswitches (Figure 3.10).

3.3.5 Allosteric microswitches reprogram D2 ligand sensing

Following our successful allosteric microswitch design at the position 5.54 (Figure 3.2), we designed other predicted allosteric residues with the goal of reprogramming D2 into a biosensor with altered signaling responses to diverse ligands (Figure 3.4a). We focused on sites most distant from both ligand and G-protein binding sites to primarily induce long-range allosteric effects. We scanned all possible amino-acid substitutions and selected those which were predicted to least affect constitutive activity. We selected a representative subset of 15 designed point mutations for experimental validation which covered all TMHs bearing allosteric sites (with the exception of TMH4 and TMH7), and were predicted to differently modulate ligand responses (Table 3.5). Signaling responses were measured for the native potent agonist dopamine and the non-native serotonin which activates D2 WT only very weakly. Except for 2 mutations, the experimentally-observed effects were correctly predicted using our calculations of long-range structural coupling (Methods, Table 3.4, Figure 3.11). Similarly to T5.54M, we identified 3 sites on TMH6 (6.41, 6.44, 6.47) which significantly enhanced the potency of both dopamine and serotonin (Figure 3.4b). These gain of function mutations increased also the signaling responses to serotonin, which reached the same efficacy with D2 F6.44I as with 5HT1AR WT, a GPCR naturally evolved to sense and respond to serotonin (Figure 3.4b). Remarkably, despite the high sensitivity of D2 F6.44I for serotonin, the ligand potency for that D2 variant remained well below that for 5HT1AR WT, indicating that ligand efficacy and potency can be optimized differently. Our calculations predicted also neutral mutations which did not modify the potency of dopamine (Figure 3.4c). Lastly, in addition to T5.54V (Figure 3.2), we identified 3 sites on TMH3, 5 and 6 which significantly decreased the potency and efficacy of dopamine (Figure 3.4d). Mutations I3.40T and I3.40L significantly increased the structural coupling of the agonist to the inactive G-protein binding site, consistent with substantially reduced responses to dopamine (Figure 3.11). Our calculations suggest that the native amino-acid isoleucine at position 3.40, which is highly conserved among class A GPCRs, is optimal for coupling ligand agonists to Gi. Overall, we observed that all the designed D2 with enhanced sensitivity to dopamine had also higher sensitivity to serotonin, suggesting that both ligands share similar allosteric pathways to couple to Gi.

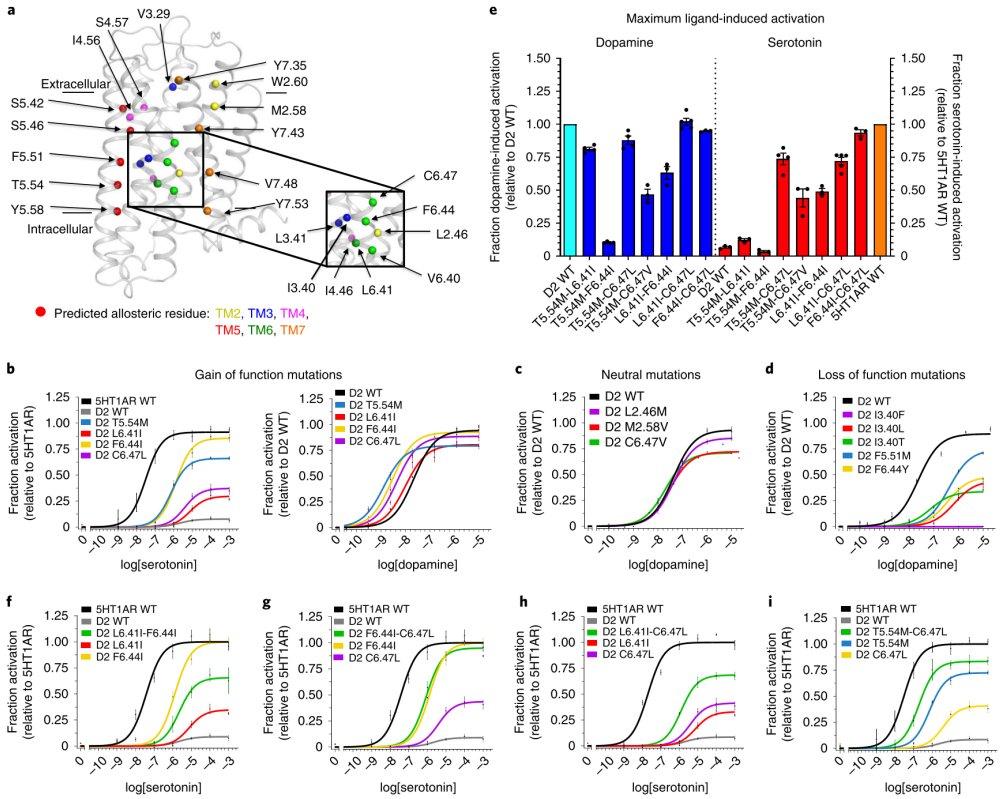


Figure 3.4 – Designed allosteric microswitches reprogram ligand sensing and signaling responses. **a.** Allosteric sites predicted by Molecular dynamics simulations are indicated as spheres on the D2 model structure, and colored based on their TM location. Values shown are from N≥3 independent experiments; data shown are mean ± s.e.m. **b-i.** D2 activation upon agonist dopamine or serotonin binding measured using the cell-based FLIPR assay. Agonist induced (dose response) as percent of maximal dopamine induced D2 WT activity or serotonin induced 5HT1R WT activity. Reported dose responses are from N=3 independent experiments; shown values are mean data shown are mean ± s.e.m. **b-d.** Designed point-mutations are classified based on their signaling effects: gain of function **(b)**; neutral **(c)**; loss of function **(d)**. **e-i.** Double mutants incorporating gain of function mutations. **e.** Dopamine and serotonin induced responses of double mutants incorporating gain of function mutations. **f,h.** Serotonin induced responses by pairs of gain of function mutations that are distant in the structure. **g,i.** Serotonin induced responses by pairs of gain of function mutations that are close in the structure. Agonist induced (dose response) as percent of maximal serotonin induced 5HT1R WT activity. Serotonin induced D2 WT activity is also represented for comparison.

3.3.6 Designed microswitches define multiple allosteric paths

To design biosensors with very high sensitivity to ligands, we combined all gain of function mutations and created double mutant D2 variants. We observed a wide spectrum of effects ranging from a nearly complete loss of signaling responses to additive effects on ligand potency and efficacy (Figure 3.4e). Remarkably, for almost all double mutants, our structural coupling calculations predicted correctly the classes of effects (i.e. gain, neutral and loss of function) (Table 3.4), which strongly depend on the relative location of the designed microswitches.

Loss of function coincided with very close proximity between the allosteric mutations, which were structurally incompatible. Non antagonizing mutations that were only 3 residues apart displayed no additivity and, at best, behaved similarly to the highest gain of function single point mutant (Figure 3.4g,i). Variants incorporating the most distant pairs of mutations (5.54/6.47, 6.41/6.47) displayed significant additivity in serotonin potency and efficacy (Figure 3.4f,h), including the T5.54M-C6.47L variant that responded to serotonin very similarly to the native serotonin 5HT1AR receptor. C6.47 is a key connector between TMH6 and TMH7 while T5.54, L6.41 and F6.44 mainly connect TMH5 with TMH6, implying that individual designed microswitches can modulate the structural coupling between the extracellular and intracellular sites through either the TMH5-6 or the TMH6-7 interfaces. Additive effects occur when distant mutations enhance structural coupling of distinct TMH interfaces and strongly suggest the existence of two allosteric subpathways involving TMH5-6 and TMH6-7 dynamically coupled interactions that can be independently optimized (Figure 3.12).

3.3.7 New model of GPCR regulation and evolution

Overall, our results provide new fundamental insights into the molecular mechanism of GPCR allosteric signaling and evolution. So far, the classical preeminent model of class A GPCR signaling primarily involved a few highly-conserved allosteric microswitches (i.e. DRY ionic lock on TMH3, PIF transmission switch on TMH5,6, W toggle switch on TMH6 and NPxxY motif on TMH7, Figure 3.5a) that trigger consensus structural changes in the receptor during activation [154–156] (Figure 3.5b). Mutation to alanine of most of these sites led to severe loss in ligand-induced receptor signaling, implying that the native switches are close-to-optimal for signal transduction (Figure 3.5c) [173]. However, these motifs are not in direct contact and their connections throughout the receptor structure that define the allosteric signal transduction pathways have remained elusive. Most importantly, a molecular mechanism of signal transduction involving only a few highly-conserved and close-to-optimal signaling switches does not explain the large spectrum of signaling responses displayed by GPCRs.

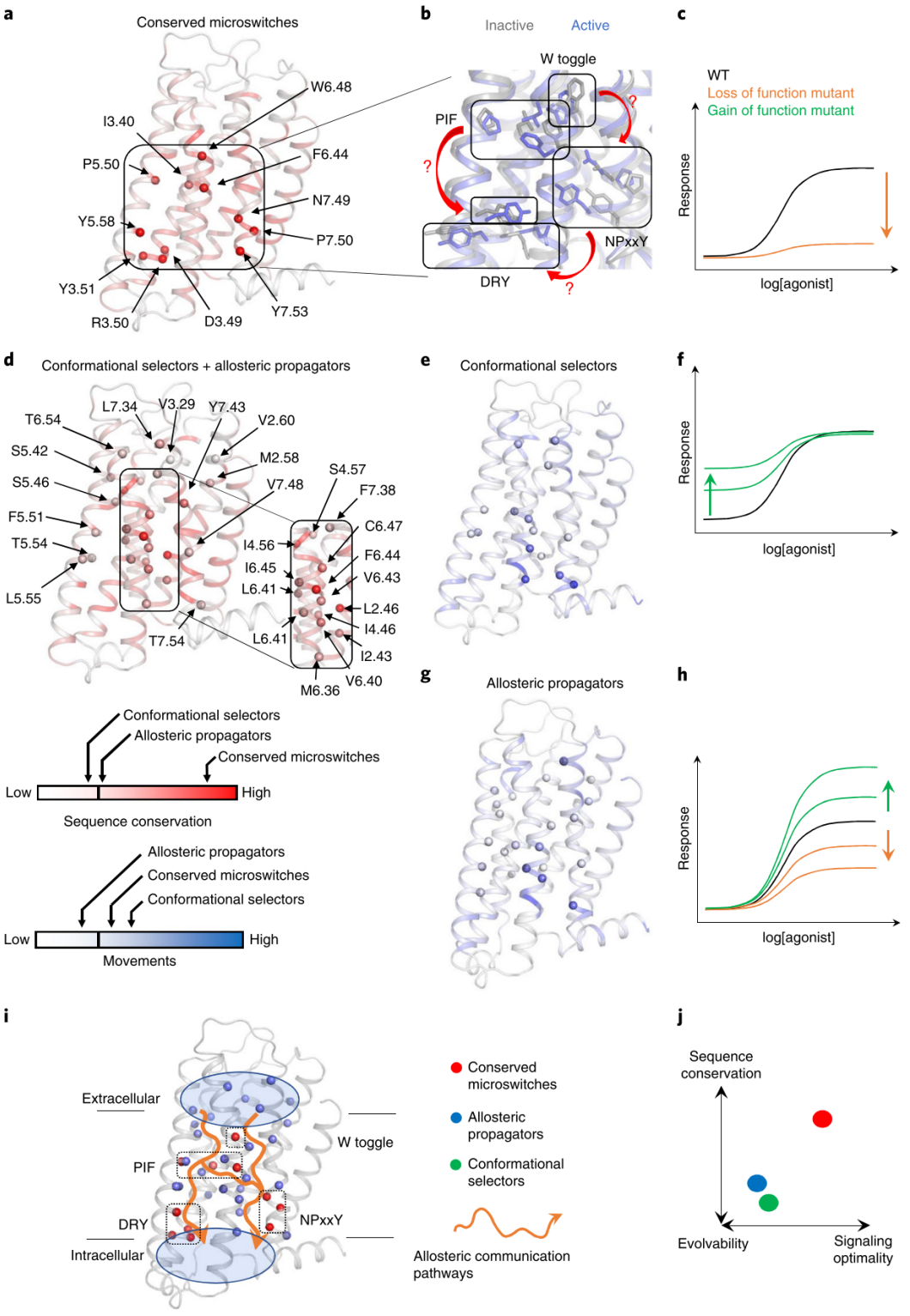

a
Conserved microswitches
W6.48
I3.40
F6.44
P5.50
N7.49
Y5.58
P7.50
Y3.51
R3.50
D3.49
Y7.53
b
Inactive
Active
W toggle
PIF
NPxxY
DRY
c
WT
Loss of function mutant
Gain of function mutant
Response
log[agonist]
d
Conformational selectors + allosteric propagators
L7.34
V3.29
Y7.43
T6.54
V2.60
S5.42
M2.58
S4.57
S5.46
V7.48
F7.38
F5.51
C6.47
T5.54
I4.56
F6.44
I6.45
V6.43
L5.55
L6.41
L2.46
T7.54
I4.46
L6.41
I2.43
M6.36
V6.40
Conformational selectors
Allosteric propagators
Conserved microswitches
Low
High
Sequence conservation
Allosteric propagators
Conserved microswitches
Conformational selectors
Low
High
Movements
e
Conformational selectors
f
Response
log[agonist]
g
Allosteric propagators
h
Response
log[agonist]
i
Extracellular
W toggle
PIF
DRY
NPxxY
Intracellular
Conserved microswitches
Allosteric propagators
Conformational selectors
Allosteric communication pathways
j
Sequence conservation
Evolvability
Signaling optimality

Figure 3.5 – A high-resolution mechanism for the regulation and evolution of GPCR allosteric signaling properties. **a-c.** Properties of conserved allosteric microswitches. **a.** Conserved allosteric microswitches shown in spheres colored by sequence conservation (high: red; low: white). **b.** Conformational changes of the conserved allosteric microswitches represented in sticks upon receptor activation. Red arrows with question marks represent the unknown connectivity linking the conserved microswitches. **c.** Schematic agonist-induced signal response curve representing the deleterious effects of alanine mutations of most conserved microswitches. **d-h.** Properties of conformational selectors and allosteric propagators. **d.** Conformational selectors and allosteric propagators shown in spheres colored by sequence conservation (high: red; low: white). The scale below recapitulates the average sequence conservation for each class of residues in class A GPCRs (the back bar represents the average over all residues). **e.** Conformational changes of the conformational selectors represented as spheres upon receptor activation (blue: high; white: low). **f.** Schematic agonist-induced signal response curve representing the effects of conformational selector mutations stabilizing the receptor active state. **g.** Conformational changes of the allosteric propagators represented as spheres upon receptor activation (blue: high; white: low). The scale below recapitulates the average movement for each class of residues (the back bar represents the average over all residues). **h.** Schematic agonist-induced signal response curve representing the effects of allosteric propagator mutations modulating the coupling between agonist and downstream effectors. **i.** High resolution schematic mapping of GPCR allosteric signal transductions where allosteric propagators connect the conserved microswitches and define several quasi-independent allosteric subpathways. **j.** Schematic diagram representing the relationship between sequence conservation, signaling optimality and evolvability of each class of allosteric residues.

Here, we deconstructed the determinants of signal transduction and uncovered two new classes of microswitches which modulate specific receptor signaling properties. Unlike the above-mentioned consensus microswitches, they exhibit a large diversity in sequence conservation and structural movements and could not be easily detected (Figure 3.5d-f). The class of "stability" microswitches which we now call "conformational selectors" were identified in regions undergoing the largest movements upon receptor activation (Figure 3.5e). By forming distinct contacts in different conformations, they selectively control the time spent by the receptor in each state. Engineering these sites reprograms selectively the receptor conformational stability and constitutive activity (Figure 3.5f). The class of "allosteric" microswitches which we now call "allosteric propagators" usually do not display major structural changes between states (Figure 3.5g) but are dynamically coupled and can propagate signals over long-distances. As such, they represent key regulators of allosteric signaling function. Engineering these sites can modulate selectively the responses to ligands by enhancing or decreasing the receptor allosteric sensing properties (Figure 3.5h).

From this expanded molecular description emerges a novel model of evolution of allosteric regulation where sparse conserved allosteric sites are connected through a network of allosteric propagators that control their coupling and can modulate the strength and direction of the allosteric communication pathways (Figure 3.5i). Conserved microswitches are often close-to-optimal for signal transduction as most amino-acid substitutions at these sites are deleterious. Consistent with a high level of sequence conservation, they correspond to highly-constrained sites with low potential for evolvability and signaling diversity (Figure 3.5j). By contrast, "allosteric propagators" appear farther away from signaling optimality as several designed mutants enhanced signaling functions. Consistent with a lower level of sequence conservation, they provide key sites for evolution of a wide range of allosteric signaling responses in class A GPCRs by modulating the coupling and pathways between the conserved microswitches. Evolutionary coupling between residues in GPCRs were found to maintain ligand responses to those of native receptors[81]. By contrast, our findings indicate that physical coupling can drive the evolution of a wide range of receptor signaling properties for generating functional diversity.

3.3.8 Prediction accuracy and limitations of the method

To assess the broad utility of the computational approach, we calculated its prediction accuracy. Because the designed receptors were selected for experimental validation based on calculated criteria only (Methods), they represent true blind predictions and a stringent dataset for validating our approach. We performed qualitative and quantitative comparisons between calculated and measured properties (Methods). Qualitatively, the direction of the changes in constitutive and ligand-induced activities from WT were correctly predicted for 82% of the designs, respectively (Figure 3.13, Methods). Changes in ligand-free receptor energies and calculated long-range allosteric couplings quantitatively correlated fairly well with changes in constitutive activities (Pearson correlation coefficient r = 0.68) and in ligand potency (r

= 0.54), respectively (Figure 3.14, Methods). To expand the validation of the approach, we predicted the signaling effects of 109 known mutations on several GPCRs. The prediction accuracy on these additional datasets (r varying between 0.55 and 0.69) was comparable to that observed with our designs (Figure 3.15, Methods). These results indicate that the computational approach has strong qualitative and fair quantitative predictive power, similar to structure-based predictions of mutational effects on protein stability by Rosetta [174].

Nevertheless, the predictions are not yet quantitatively very accurate, which currently precludes a complete mechanistic understanding of every design's success or failures. Since the approach relies on the ability to predict the effects of mutations on protein structures, which remains challenging when modeling protein conformations by homology to existing structures, structure prediction inaccuracies are likely a main factor behind the lack of precise quantitative predictions.

Despite its approximation, the two-state model seems to be sufficient to reprogram the receptor constitutive activity and the activity induced by a particular ligand. However, GPCRs can adopt multiple conformational states and more complex design scenarios involving multiple extracellular ligands, allosteric modulators and/or distinct intracellular signaling partners would necessitate a multi-state approach. In addition to accurate structure predictions, the success of our approach relies also on the ability to design residue-residue interactions that are distinct in different states to achieve selective functional effects. Therefore, if no structural information from close homologs is available on these states and/or if their conformations are too similar, the current implementation may not be accurate enough for such complex multi-state design scenarios.

3.4 Discussion

We developed a method for designing allosteric functions in membrane receptors. The method first predicts allosteric sites and pathways in receptor structures using Molecular Dynamics simulations. It then combines multi-state design and normal mode calculations to de novo design residue microswitches modulating receptor conformational stability or long-range coupling between allosteric sites that reprogram specific allosteric signaling properties.

Using this approach, we designed dopamine D2 receptors displaying an unprecedented spectrum of signaling activities with a success rate higher than 80% (Figure 3.13, Methods). Remarkably, we were able to repurpose the D2 receptor to trigger potent signaling responses upon sensing the non-cognate ligand serotonin, which constitute, to our knowledge, the first GPCR-based biosensor engineered by computational structure-based protein design. We also reliably predicted the effects of more than 100 mutations on the signaling activity of diverse rhodopsin class GPCRs (Figure 3.14 and Figure 3.15, Methods), which represents the largest benchmark of its kind. Our method requires only the structures of homolog proteins, and does not rely on evolutionary or functional information [150,175,176]. Therefore, any protein with available structural homologs may be targeted and, in principle, de novo design of allosteric

proteins can be envisioned with the technique. Overall, the computational approach provides an effective mean to design receptor stability and function which largely outperforms random mutagenesis approaches in efficiency and success rate.

Our study reveals fundamental new insights into the mechanisms of regulation and evolution of GPCR signaling with the existence of two unforeseen "conformational selectors" and "allosteric propagators" classes of microswitches which perform largely distinct signaling functions (Figure 3.5). Unlike the highly-conserved allosteric microswitches that have been the focus of most studies so far, these microswitches have higher potential for evolvability and provide the backbone for coding signaling diversity such as the modulation of constitutive activity, ligand efficacy, potency as demonstrated in our study and likely signaling bias as well. Our results uncovered also the existence of topologically distinct allosteric subpathways which function quasi independently in the TM core of the D2 receptor. Their coupling can be modulated separately and they can act synergistically to achieve optimal intracellular effector activation. This particular allosteric network design offers a number of advantages for coding large spectrum of signaling response strengths, fine-grained tuning responses to ligands potentiating distinct signal transduction routes, while providing inherent robustness during evolution. In conclusion, while allosteric pathways in the transmembrane and intracellular regions were largely viewed as conserved routes preserving consensus long-range communications and signaling functions in GPCRs, our study reveals that they actually constitute a critical backbone for driving the evolution of signaling and functional diversity, which redefines our understanding of GPCR allostery.

Overall, our strategy should prove particularly useful for designing novel regulatory pathways and repurposing membrane receptors as biosensors to a vast array of chemical molecules. Foreseeable applications of our method include the de novo design of biosensors and signaling receptors to reprogram cellular functions which is at the heart of current engineering approaches for synthetic biology and cancer immunotherapeutic applications.

3.5 Methods

Mutagenesis and expression of receptors Quickchange PCR mutagenesis was performed on D2 gene (isoform 1 of D2 (http://www.uniprot.org/uniprot/P14416.fasta)) in pcDNA3.1(+) from the cDNA library. Sequenced mutant plasmids were transiently transfected using lipofectamine into HEK293T cells. Briefly, 5x10^6 cells were plated on 10cm tissue culture plates and grown overnight. This was followed by transfection with Lipofectamine 2000 and 15μg of DNA per plate. After 24 hours, transfection medium was replaced with standard growth medium (DMEM supplemented with L-glutamine (2mM), penicillin (100mg/mL), streptomycin (100mg/mL), and fetal bovine serum (10%)) and cells were grown for an additional 24 hours prior to harvesting.

3.5.1 Membrane preparation

Membranes were prepared from transfected cells using sucrose gradient centrifugation as previously described[177]. Briefly, cells from 10cm plates were collected by cell scraper with PBS solution. Cells were pelleted and resuspended in cold hypotonic buffer (1mM Tris-HCL, pH 6.8, 10mM EDTA, protease inhibitor cocktail). Cells were forced through a 26-gauge needle three times. The cell lysate was layered onto a 38% sucrose solution in buffer A (150mM NaCl, 1mM MgCl, 10mM EDTA, 20mM Tris-HCl, pH 6.8, protease inhibitor cocktail) in SW-28 ultracentrifuge tubes. Cells were centrifuged at 15,000 rpm at 4°C for 20 minutes, followed by collection of the interface band with an 18-gauge needle. The collected solution was transferred to Ti-45 ultracentrifuge tubes and the volume brought up to 50mL with buffer A. The sample was spun at 40,000 rpm at 4°C for 30 minutes. The membrane pellets from each 10cm plate was resuspended in 0.5mL buffer A and stored at -80°C in 100uL aliquots.

3.5.2 D2 purification

D2 variants were partially purified from thawed membrane preparations immediately prior to assaying via anti-HA agarose beads. Membrane preparations were solubilized with 1% n-dodecylmaltoside (DDM) for one hour at 4°C, and loaded onto anti-HA agarose beads for one hour at 4°C. The beads were washed with TBS with 0.1% DDM wash buffer three times and HA-tagged receptor variants were eluted with HA peptide (1mg/mL in TBS with 0.1% DDM).

The interpretation of designed mutational effects based on the comparison between receptor activities assumes that the different samples of D2 receptor variants have the same level of purity and contains similar fraction of active receptors. To stringently validate this hypothesis, the levels of purity and activity in all samples of receptor variants after detergent solubilization and purification were measured. Purity of affinity purified D2DR samples was assessed by coomassie-stained SDS-PAGE along with western blots for all receptor variants. All bands observed in the final sample preparations used for the assays were quantified by densitometry. All samples consisted mostly of D2 monomers with very similar levels of purity; i.e. ranging from 92% to 95%. To quantify the fraction of active receptor for each variant, the ability of the receptors to bind the ligand agonist dopamine, which should occur only when the receptor is folded and active, was assessed. Specifically, the maximal amount of bound ligand dopamine per receptor quantity was measured from radioligand binding experiments performed with purified D2 samples and with the same batch of tritium labeled ligand across experiments. The analysis indicated that all purified variants contained a fraction of folded and active receptor that was not significantly different from WT (i.e. maximal difference was 8%).

3.5.3 Expression and purification of the G protein Gi

Gi2 subunits were cloned into pFastbacI followed by transformation into DH10α cells. Recombinant bacmid DNA was isolated and transfected in Sf9 insect cells with Cellfectin II. The transfected cells were grown at 28°C for 72 hours followed by centrifugation in 15mL Falcon tubes to pellet cell debris. The supernatant was saved as P1 viral stock. Virus was amplified by infecting Sf9 cells with P1 viral stock solution at a 2-fold multiplicity-of-infection and cultured for 72 hours prior to harvesting. This amplification process was repeated to generate a high-titer P3, which was used for infection and protein expression. P3 stock was used to infect Sf9 cells at an MOI greater than 4 and harvested 48-60 hours post infection. Cells were washed three times in ice-cold PBS and resuspended in homogenization buffer (10mM Tris-HCl, 25mM NaCl, 10mM MgCl2, 1mM EGTA, 1mM DTT, protease inhibitor cocktail, 10uM GDP, pH 8.0). Gi2 was purified and reconstituted as described previously[178]. Antibodies were purchased from Santa Cruz Biotechnology.

3.5.4 In vitro G protein activation assays

Receptor variants were assayed for their ability to induce guanine nucleotide exchange by the G protein Gi2. To measure constitutive activity the reaction mixture consisted of 4uM Gi2, 20uM of [^{35}S]-GTPγS mix, 50mM Tris-HCl, pH 7.2, 100mM NaCl, 4mM MgCl2, 1mM dithiothreitol. D2DR concentrations were 10nM per sample as estimated from absorbance at 280nm of the anti-HA agarose purified samples. For all samples used for reactions, western blots were performed to further assess receptor quantity using monoclonal anti-HA antibody (Thermo Scientific). Reactions were started by adding 150uL partially purified receptor samples to 300uL of reaction mixture and incubating on ice over a period of time from 0-30 minutes for timecourse measurements. To measure ligand induced D2 activities, increasing amounts of dopamine or spiperone were added to the reaction mixtures. Maximal ligand efficacies were obtained with final ligand concentrations of 20uM and 1uM for dopamine and spiperone, respectively. Reactions were stopped by filter binding onto Millipore nitrocellulose filters or Whatman fiberglass filters pretreated with polyethylenimine. Filters were washed three times with ice-cold TBS prior to incubation with scintillation fluid. Radioactivity counts were measured on a Beckman LS6000 scintillation counter. Mock transfected samples (using empty pcDNA vector) and WT receptor were assayed in every experiments as internal controls. Background binding measured by mock transfected samples was subtracted and activities of receptor variants were normalized relative to WT receptor via densitometry analysis of western blots using ImageJ software. Statistical significance of differences in constitutive or ligand-induced activities was assessed by unpaired two-sided t-tests. Nonlinear curve-fitting was performed for each dose response dataset using Graphpad Prism 6: R2 > 0.9 for all fits, with the exception of D2 WT with spiperone (R2=0.7, Figure 3.3) and D2 R3 variant with spiperone (R2=0.6, Figure 3.3). Measurements were taken from distinct samples.

3.5.5 Trp Channel activation cell-based assay

HEK-293 cells stably expressing TrpC4β channel were kindly provided by Dr. Michael X. Zhu and were transiently transfected with D2 design variants as described below. FMP2 FLIPR Membrane potential assays (Molecular Devices) were performed as previously described[179]. Briefly, the assay utilizes a fluorescent probe that reacts to cations and is coupled with a non-permeable quencher. Activation of TrpC4β causes influx of cations leading to the translocation of the fluorescent probe away from the quencher, resulting in increased fluorescence. TrpC4β has been described to be downstream of Gi activation and serves as a reporter for D2 activation. Stable HEK-293 cells were cultured in DMEM supplemented with 10% FBS and 1% G418 (selection antibiotic) for use with FLIPR Membrane Potential (FMP) assay kits (Molecular Devices). On the day of transfection, cells stably expressing TrpC4β were plated (1.5x105 cells per well) onto 96-well clear, flat-bottom plates precoated with poly-D-lysine (Sigma). Cells were transfected with Lipofectamine 2000 and the appropriate D2 constructs using optimized DNA quantity for the FLIPR assay (1ng DNA per well for D2 WT, 0.5uL Lipofectamine per well) and grown for 24h prior to assaying. A duplicate plate was transfected and grown for assessment of receptor cell surface expression determined by ELISA using anti-HA antibody as described previously[180]. For each variant, DNA titration experiments were performed to determine the DNA quantity required to match the surface expression of D2 WT. Measured activities were normalized to cell surface expression of D2 variants and background signal was removed by subtracting fluorescence signals from mock-transfected cells. Cells were treated with either dopamine, serotonin or spiperone to measure activation and inhibition, respectively. Nonlinear curve-fitting was performed for each dose response dataset with a Hill coefficient of 1 using Graphpad Prism 6: For Figure 4: R2 > 0.9 for all fits, with the exception of M2.58V (R2=0.8) and D2 WT with serotonin (R2=0.7). For Figure 5b-e: R2 > 0.9 for all fits. Measurements were taken from distinct samples.

3.5.6 Characterization of D2 conformations by fluorescence spectroscopy

Following a previously described approach to measure local conformational shifts in the prototypical GPCR β_2AR[181], the monobromobimane fluorescent probe was conjugated to D2 variants via cysteine thiol conjugation. D2 variants were modified to remove exposed cysteine residues close to or inside loop regions (C20A.C132S.C272S*.C281S*) as well as add a cysteine residue to the intracellular helical tips of either TMH6 (Q188C) or TMH5 (L180C) for site specific labeling. Controls were performed to verify that these modified receptors bound ligands and activated Gi2 as the original ones. Receptor variants were expressed and purified as described above. Affinity purified receptor variants were mixed with monobromobimane at the same molarity and incubated on ice for 4-8hrs in the dark. The fluorophore-labeled receptor was separated from unconjugated bimane using a desalting column (HiTrap, GE Health Sciences). Monobromobimane conjugated receptors were incubated with various ligands and effectors (4μM Gi2 and/or 1uM norapomorphine (NPA), and/or 1μM spiperone) for 15 minutes at room temperature. Residual GTP and GDP in all reaction mixtures were

hydrolyzed with apyrase (25 U/mL). Receptor concentrations were 2nM per sample. The fluorescence spectrum was measured on Molecular Devices Flexstation3 plate reader with excitation at 370nm and emission measured from 420-530nm at 1s/nm. The peaks of maximal fluorescence intensities for various D2 variants and conditions were compared using unpaired two-sided t-tests calculated with the GraphPad Prism software. *Note: C272S and C281S are in ICL3 which was not modeled and therefore follow the gene numbering. Measurements were taken from distinct samples. Relationship between fluorescence spectra and local active conformations occupancy

The following model provides the rationale explaining why ligand binding to a designed receptor results in reduced spectral changes compared to WT if the designed microswitches partially mimic the ligand effects on receptor structure and activity (Figure 3.4). Such effects can be formally expressed as follows:

Let's define ΔS^L_{WT}, the change in maximal spectrum intensity due to structural change of the WT receptor reaching 100% active state conformation occupancy upon ligand agonist binding.

$$\Delta S^L_{WT} = S_{WT} - S^L_{WT} \tag{3.1}$$

with S_{WT} and S^L_{WT}, the maximal spectrum intensities of the ligand-free and ligand-bound WT receptors, respectively.

If designed mutations shift the receptor toward X% active state conformation occupancy (X<100) by partially mimicking the effects of ligand binding, this will result in the corresponding spectral shift:

$$\Delta S_{DES} = S_{DES} - S_{WT} \tag{3.2}$$

with S_{DES}, and S_{WT}, the maximal spectrum intensities of the ligand-free designed and WT receptors, respectively. Since the designed mutations only partially mimic ligand binding effects, $0 < \Delta S^L_{DES} < \Delta S^L_{WT}$.

If ligand binding further shifts the designed receptor toward 100% active state conformation occupancy,

$$\Delta S^L_{DES} = S_{DES} - S^L_{DES} \tag{3.3}$$

with S_{DES} and S^L_{DES}, the maximal spectrum intensities of the ligand-free and ligand-bound designed receptors, respectively.

Since both ligand-bound WT and designed receptors reach 100% active state conformation

occupancy, $S^L_{DES} = S^L_{WT}$, then ΔS^L_{DES} can be expressed using equations 3.1 and 3.2:

$$\Delta S^L_{DES} = \Delta S_{DES} + \Delta S^L_{WT} \tag{3.4}$$

Since $\Delta S_{DES} > 0$, $\Delta S^L_{DES} < \Delta S^L_{WT}$, which demonstrates that, if a designed mutation mimics ligand binding, then ligand binding to that designed receptor will result in a smaller spectral change compared to that observed for the WT.

3.5.7 Homology modeling of D2 WT inactive and active conformations

1. Modeling of the D2 WT ligand-free inactive and active conformations

The closest structurally characterized homologs available were used as templates to model the inactive and active state conformational ensembles adopted by the D2 WT sequence isoform 1 of D2 (http://www.uniprot.org/uniprot/P14416.fasta). The inactive state conformations were modeled starting from the X-ray structure of the antagonist bound dopamine D3 receptor (DRD3, PDBID: 3PBL) sharing 48% sequence identity with D2. The active state conformations were modeled from the active state GPCR structures of two distant homologs: opsin (23% sequence identity with D2) bound to Gt (PDBID: 3CAP, 3DQB) and beta-2 adrenergic receptor (26% sequence identity with D2) in complex with ligand agonist and heterotrimeric Gs (PDBID: 3SN6). Models of D2 inactive and active states were generated without bound ligand or G protein using the homology mode of RosettaMembrane as described[166]. From the above-mentioned simulations, ensembles of low energy models were clustered and centers of the most populated clusters were selected as templates for each ligand-free state (R and R*, see Figure 3.1b) in the design calculations of stability microswitches. Specifically, 2 inactive and 6 active state models of D2 WT were selected.

2. Modeling of the D2 WT ligand-bound inactive and active conformations

Ligand-bound D2 models (e.g. AR and AR*, see Figure 3.1b) were generated by performing ligand docking simulations onto the selected inactive and active state ligand-free D2 models obtained as described above. Specifically, ligands were docked onto D2 models as follows. The ligand conformer libraries were generated using Omega (OpenEye Inc.) with default parameters from Rosetta ARLS (automatic RosettaLigand setup; e.g. ewindow 10.0; rms 0.5; Rosetta version 3.2). Ligand docking was performed using a combination of Monte Carlo moves, sidechain repacking, and gradient-based minimization with the Rosetta all-atom potential developed for protein-ligand docking. Receptor backbone and sidechain torsions along with ligand torsions, position, and orientation were optimized simultaneously[165,182]. All ligand-bound receptor models for a given state (i.e. inactive or active) were gathered together and the 5% lowest-energy models were selected and ranked based on the binding energy with the ligand. The selected ligand bound models were then clustered based on structural

similarity of the bound ligand conformation. The representative model of the largest cluster was selected corresponding to the ligand-bound model of the D2 receptor in a given state; i.e. ligand-bound D2 inactive and ligand-bound D2 active state conformations. These models were used to calculate the structural coupling in the ligand-bound inactive and active states and to predict changes in receptor signaling responses to ligand binding.

Modeling of the ligand-free conformations of the D2 local region designed variants To model the apo conformations (i.e. in absence of ligand and Gi) adopted by D2 variants designed in local regions (Figure 3.7), we extensively relaxed fully inactive and active conformations of the designed receptors. We used a structure relaxation protocol that samples all backbone and side-chain degrees of freedom without constraints enabling the polypeptide chain to explore an unrestricted conformational space. This protocol allows each designed local region to identify the lowest energy conformations even if it involves significant conformational changes from the starting template structure. Around 10000 independent simulations were performed for each variant. The 10% lowest energy structures were clustered and the center structure of the largest family of structures was selected as the representative model in Figure 3.7

3.5.8 Computational design of stability microswitches perturbing the apo (ligand-free) equilibrium

1. Multi-state search and selection of stability microswitches

The above-mentioned selected D2 WT ligand-free models were used as a starting ensemble of templates in multi-state design calculations of microswitches altering the relative stability of the ligand-free states. The calculations use a genetic algorithm51 to search and select combinations of amino-acid substitutions introducing novel stability microswitches that maximize the fitness function F (F=x*ΔG_A + y*($\Delta G_A - \Delta G_I$). ΔG_A and ΔG_I are calculated from the energy of the conformations best accommodating the selected sequence and optimizing the number of favorable contacts between protein atoms in the ligand-free active and inactive state, respectively ($\Delta G = G_{design} - G_{WT}$). $\Delta G_A - \Delta G_I = \Delta\Delta G_{apo}$ which represents the change in apo equilibrium energies upon mutations. These energies are determined by relaxing the side-chain conformations of the starting WT templates threaded with the selected designed sequence using an all-atom energy function developed for high-resolution membrane protein modeling and design [105]. The backbone conformation of the starting templates is fixed during these calculations. Iterative rounds of multistate design were performed to optimize the fitness function weights. Designed structures were then refined using an all-atom structure refinement protocol of membrane protein structures to improve the accuracy of the predicted energies (see below). During this refinement step, all designed mutations were also systematically back-mutated to the native residue to identify those that do not contribute significantly to the intended shifts in apo equilibrium. Using the above-mentioned approach, the design of microswitches was performed in 2 steps. In step 1, the TM core of the receptor was scanned

for identifying local regions where individual microswitches could be designed. In step 2, larger combinations of microswitches were designed in these regions.

2. Refinement of designed structures and of stability microswitch selection

Designed structures were refined using an all-atom constrained relax protocol of membrane protein structures to improve the accuracy of the predicted energies calculated by the multi-state design technique. This protocol is based on a Monte Carlo Minimization technique involving multiple cycles of side-chain repacking with fine-grained rotamer libraries followed by side-chain and backbone minimization of the receptor structure in the vicinity of the repacked residues using the all-atom energy function of RosettaMembrane[133,183]. In this protocol, backbone atoms are constrained by a harmonic potential preventing large backbone movements and guiding the sampling toward finding a low energy minimum conformation in the vicinity of the starting structure. The harmonic potential is applied for backbone atom structure deviations larger than 0.2 Angstrom. This approach represents a compromise between full structural relaxation and the necessity to score sequences for a specific functional (i.e. inactive or active) state characterized by a well-defined and restricted conformational space. In these calculations, the protein is modeled explicitly at all-atom and the membrane environment is modeled as an implicit anisotropic solvent membrane object[133]. Specifically, in each variant, designed residues and all residues within 10 Å CA distance from the designed positions were repacked and their backbone and side-chain minimized. The only criterion used to select final refined models is the convergence in energy of the simulations. Specifically, up to hundreds of independent simulations are performed until the energy of the 10% lowest energy relaxed models lie within 0.5 Rosetta Unit (RU) from each other. As part of the refinement process, each designed mutation was systematically back mutated in silico and its structure relaxed as described in the previous paragraph before deciding whether it should be kept in the final variant for experimental testing. Designed mutation contributing at least 0.5 RU of the total energy shift from WT were kept in the final receptor variant. Designed variant mutations leading to the introduction of a polar residue in the TM region were carefully analyzed for optimal satisfaction of the polar groups through hydrogen bonds. Suboptimal polar residues (i.e. not forming hydrogen bonds of at least 0.5 RU in energy) were systematically filtered out from the design solutions.

3.5.9 Computational design of allosteric microswitches perturbing structural couplings

As described below, microswitches altering the structural coupling between a ligand and the G protein binding site in selective states can perturb the allosteric coupling of the ligand to the receptor.

1. Selection of allosteric sites

To most effectively alter the allosteric regulations of the receptor, we reasoned that mutations should be designed at the primary sites encoding the allosteric communication pathways. These sites are called allosteric residues and can be identified through analysis of protein movements extracted from long timescale molecular dynamic simulations7. Allosteric residues constitute highly coupled networks of protein residues connecting the extracellular ligand binding to the intracellular G protein binding sites of the receptor. According to long timescale molecular dynamic simulations of B216,29 and M28, 60% of the allosteric residues in the transmembrane region share the same position in these 2 receptors despite limited sequence and structure homology. Because B2 shares substantially higher sequence and structure similarity with D2 than with M2, we hypothesized that the positions of allosteric residues should remain largely conserved between the B2 and D2 transmembrane domains. Therefore, the D2 residues sharing the same positions with B2 allosteric residues after sequence alignment of the 2 receptors were defined as allosteric residues. The following Ballesteros-Weinstein numbers were assigned to the D2 allosteric residues: L2.46, M2.58, W2.60, V3.29, I3.40, L3.41, I4.46, I4.56, S4.57, S5.42, S5.46, F5.51, T5.54, Y5.58, V6.40, L6.41, F6.44, C6.47, F7.38, Y7.43, V7.48, Y7.53.

Since molecular dynamic simulations are too computationally expensive to be used in protein design and efficiently scan for microswitches at multiple sites, we developed a two-step hybrid approach to design microswitches modulating allosteric responses through changes in structural couplings. The rational and implementation of the approach are described in sections 3-4.

Parameters for running MD simulations

As described[64,78], all simulations on B2AR were performed in explicit POPC lipid membranes including water and ion molecules (NaCl at 0.15 M) using the CHARMM27 force field. Internal water molecules were added with Dowser[184]. Van der Waals and short-range electrostatic interactions were cut off at 9 Å and long-range electrostatic interactions were computed using the Particle Mesh Ewald method[185], using a 64 x 64 x 64 grid with sigma = 2.26 Å and fifth-order B-splines for interpolation. Inactive state simulations were performed for 16 μs on the inactive state inverse agonist carazolol bound B2AR crystal structure (2RH1). Active state simulations were performed for 5.9 μs on the active state B2AR crystal structure (3POG) bound to nanobody and the agonist BI-16701717. Such long timescale simulations may not be necessary as reliable results were also obtained using 5 independent simulations performed each for 200 ns for 1 μs total simulation time[113].

2. Difference in structural couplings as a measure of the allosteric coupling

Allostery occurs when the difference in stability between the active and inactive receptor state is altered upon ligand binding. Therefore, the allosteric coupling due to ligand binding can be expressed as follows:

$$-RTlog(\alpha) = (G_{A,L}–G_{I,L})–(G_A–G_I) \tag{3.5}$$

where α is the allosteric coupling (see Figure 3.1), $G_{A,L} - G_{I,L}$) is the energy difference between active and inactive states of the receptor bound to ligand, and ($G_A - G_I$) is the energy difference between active and inactive states of the ligand-free receptor, i.e. ΔG_{apo}.

As explained previously26, the allosteric effect can be described in terms of structural interactions or coupling between the ligand and the effector binding regions of the receptor. For that purpose, we consider the simple case of a ligand binding to a two-state receptor. This system can be defined as a two-component system where each component has access to 2 states. One component represents the ligand with states bound or unbound. The other component represents the receptor including the effector site, which can adopt an inactive, or an active state. The potential energy of the system in a given state is given by the sum over the one body energy of each component (G^L for the ligand, G^R for the receptor) and their two-body interaction energy or coupling G^C (e.g. $G^C_{A,Lb}$ and $G^C_{A,Lu}$ represent the coupling of the effector active conformation to the ligand binding site conformation with or without bound ligand, respectively). In our model, the two-body interaction energy between ligand and effector sites is calculated from the local thermodynamic coupling energies between the pre-defined allosteric residues or assessed indirectly through long-range dynamic coupling effects (see below, sections 3 and 4). The potential energy of the system for the 4 states can be written as follows:

$$G_A = G^L_{unbound} + G^R_A + G^C_{A,Lu} \tag{3.6}$$

$$G_I = G^L_{unbound} + G^R_I + G^C_{I,Lu} \tag{3.7}$$

$$G_{A,L} = G^L_{bound} + G^R_A + G^C_{A,Lb} \tag{3.8}$$

$$G_{I,L} = G^L_{bound} + G^R_I + G^C_{I,Lb} \tag{3.9}$$

$\Delta G_{apo} = G_A - G_I$ corresponds to equations 3.7 - 3.6. The difference in allosteric coupling of a ligand for two receptor variants x and y with the same ΔGR and ΔG_{apo} (hence the same con-

stitutive activity) can be directly calculated by the difference in the ligand-effector structural coupling (G^C) in the active and inactive state ligand-bound conformations. Hence, equation 3.5 simplifies to equation 3.10.

For a ligand agonist:

$$-RT \cdot log\left(\frac{\alpha_x}{\alpha_y}\right) = \left(G^C_{A,Lb} - G^C_{I,Lb}\right)_x - \left(G^C_{A,Lb} - G^C_{I,Lb}\right)_y = \Delta G^C_x - \Delta G^C_y = \Delta\Delta G^C \tag{3.10}$$

For example, $G^C_{A,Lb}$ represents the coupling of the agonist bound ligand binding site to the effector active conformation.

For a ligand inverse agonist:

$$-RT \cdot log\left(\frac{\alpha_x}{\alpha_y}\right) = \left(G^C_{I,Lb} - G^C_{A,Lb}\right) x - \left(G^C_{I,Lb} - G^C_{A,Lb}\right)_y = -\Delta G^C_x + \Delta G^C_y = \Delta\Delta G^C \tag{3.11}$$

Therefore, microswitches can be designed that alter the allosteric coupling of a ligand to the receptor if they perturb the structural coupling according to equation 3.10.

3. Two-step selection of allosteric microswitches altering the coupling between allosteric sites.

In step 1, potential candidate microswitches were designed at allosteric residues to increase or decrease the interactions with contacting allosteric residues (i.e. local couplings) using the design mode of RosettaMembrane. In step 2, microswitches chosen in step 1 were selected if they also altered the couplings with nonadjacent allosteric residues (i.e. long-range couplings), to ensure long-distance perturbation of the allosteric communication pathways connecting extracellular and intracellular regions of the receptor. These calculations are described in detail in the next section.

4. Calculation of local and long-range structural couplings.

In both steps, the receptor structure of the ligand-bound inactive and active state D2 models were considered for designing microswitches altering the coupling between allosteric residues in the ligand-bound states, i.e. $G^C_{A,Lb}$ and $G^C_{I,Lb}$.

In step 1, the D2 WT models were used as starting templates for design calculations. Specifically, putative allosteric residues identified by homology to B2 were mutated to all possible 20 amino-acids using the structure refinement mode of RosettaMembrane. Perturbations in local structural coupling (G^C) were calculated for all pairs of neighboring allosteric residues constituted by the mutated residue and a directly contacting allosteric residue. The local

structural coupling (G^C) corresponded to the sum of all two-body interactions between the 2 residues. These calculations were performed for each signaling state conformation of the mutant and WT ligand-bound receptors to derive differential structural coupling shifts from WT: ($\Delta\Delta G^C$ = $(\Delta G^C)_{active}$ – $(\Delta G^C)_{inactive}$ = ((G^C (mut) – (G^C (wt))$_{active}$ – (G^C (mut) – (G^C (wt))$_{inactive}$). The total differential structural coupling shift for each mutation was calculated as the sum of shifts over each contacting allosteric residue pair.

In step 2, the predicted inactive and active state structures of the receptor bearing the candidate microswitches displaying large local coupling shifts were analyzed for changes in correlated dynamics between allosteric residues using the software Bio3D[186] calculating normal modes of motions for protein structures55. For each sequence in each state, conformational dynamic fluctuations of all allosteric residues were extracted from the 20 lowest frequency calculated normal modes. Cross-correlation matrices were then calculated to extract correlation in conformational fluctuations (i.e. coupled movements) between two allosteric residues. The correlations have no unit and their magnitude can vary between 0 (no correlation in residue movement) to +1.0 (completely correlated movements). Inter-residue correlations were summed over all pairs of allosteric residues to calculate the total correlation in the movements of allosteric residues for each D2 variant. The difference in total correlation between WT and designed structures (or between two designed structures) defined the long range structural coupling shifts (ΔG^C) upon microswitch substitution in a particular state.

Parameters for running Normal mode calculations.

The software Bio3D performing NMA on protein structures provides the following functions.

1. The function nma performs NMA of the protein structures. The function was run using the recommended default arguments which reproduce the calculation of normal modes (VibrationalModes) in the Molecular Modeling Toolkit (MMTK) package.

Command line example:

```
WT  <- read.pdb("WT_active.pdb")
mode_WT <- nma(WT)
```

2. The function Dccm calculates the cross-correlation matrix from Normal Modes Analysis (NMA) obtained from applying the function “nma” of a protein structure. It returns a matrix of pairwise residue dynamical cross-correlations with elements Cij. The conformational fluctuations of residues i and j are correlated if Cij > 0.

The lowest frequency modes were considered for calculating dynamical cross-correlation since they have been shown to describe the functionally most important global motions of proteins[167]. No selection of specific modes based on their exact frequencies or intensities were performed to avoid any bias in the calculation of the dynamic correlations. In our study,

considering the 10 or 20 lowest frequency modes gave similar results.

Command line example:

```
cijWT <- dccm(mode_WT, nmodes=20, ncore=1)
```

Qualitative prediction accuracy for designed D2 receptors (Figure 3.8)

1. Apo equilibrium shifts.
To calculate the percentage of successfully predicted mutations, we adopted the following criteria. Designed variants predicted to shift the apo equilibrium ($\Delta\Delta G_{apo}$) toward the active state by more than 2 Rosetta Units were defined successful if they displayed higher constitutive activity than WT by more than 2-fold (see Table 3.1, 3.2).

2. Allosteric coupling shifts.
To calculate the percentage of successfully predicted mutations, we adopted the following criteria. We gathered measured mutational effects on ligand-induced signaling responses into 3 categories: enhanced response (i.e. gain of function) if ligand efficacy and/or potency were significantly increased from WT; similar response (i.e. neutral) if ligand efficacy and/or potency were not significantly altered from WT; weakened response (i.e. loss of function) if ligand efficacy and/or potency were significantly decreased from WT. On the prediction side, we gathered calculated mutational effects on the allosteric coupling into 3 categories: enhanced coupling (i.e. gain of function) if the total change in dynamic couplings calculated by NMA according to equation 6 ($\Delta G^{C}_{DESIGN} - \Delta G^{C}_{WT}$, see previous section) were > 0.25; similar coupling (i.e. neutral) if the total change in dynamic couplings were within |0.25|; weakened coupling (i.e. loss of function) if the total change in dynamic couplings were < – 0.25. A prediction was defined successful if the class of effect on the predicted allosteric coupling matched the class of effect on the measured ligand efficacy and potency.

Quantitative relationship between predictions and experiments for both designed D2 receptors and benchmark of known GPCR mutations (Figure 3.9 and Figure 3.10)

1. Apo equilibrium shifts. As described in Figure 3.6, the allosteric two state model (ATSM) provides a mathematical framework for relating the calculated energetic and dynamic properties of the designed receptors to the measured allosteric activities. Without ligand stimulus, the constitutive activity (CA) is directly related to the fraction of ligand free active state receptor (f(R*)) in the active state, which can be calculated from the equilibrium constant of the ligand free receptor in the inactive and active state K_R and therefore from the free energy difference between both states ΔG_{apo} by solving the following equations:

$$CA = \frac{K_R}{(1 + K_R)} \tag{3.12}$$

$$\Delta G_{apo} = -RT \cdot ln\left(\frac{CA}{1-CA}\right) \tag{3.13}$$

However, since Rosetta energies cannot be rigorously converted into physical free energies in kcal/mol, we cannot precisely calculate K_R and the Boltzmann distribution f(R*) from our calculated energy differences. To circumvent that problem, one could in principle fit the data using equation 3.13 describing the exact Boltzmann distribution model but with a free parameter in lieu of RT. However, as shown in Figure 6, datasets covering a large range of CAs (e.g. from 5% to 95% of the maximal receptor activity) are necessary to reliably fit the Boltzmann distribution model but none of our GPCR mutant datasets cover such range in activities (see Figures 3.9 and 3.10). Fortunately, as shown in Figure 3.11, the relationship between CA and $\Delta\Delta G_{apo}$ is quasi linear within the range of measured receptor activities. To a first approximation, such linear relationship enables a direct quantitative comparison between experimental activities and $\Delta\Delta G_{apo}$ energies that is independent of the energy units.

2. Allosteric coupling shifts. As described in Figure 3.6, the allosteric two state model (ATSM) relates the allosteric coupling α to the ligand potency EC50 as follows:

$$EC_{50} = \frac{1+K_R}{K_A(1+\alpha K_R)} \tag{3.14}$$

with K_A being the intrinsic receptor ligand affinity. Our ligand docking simulations and receptor energy calculations indicate that the WT and receptors designed with altered allosteric ligand responses have very similar K_A and K_R. Therefore, the ratio between EC50 between 2 receptor variants X and Y simplifies to:

$$\frac{EC_{50X}}{EC_{50Y}} = \frac{1+\alpha_Y K_{RY}}{1+\alpha_X K_{RX}} \tag{3.15}$$

Since αK_R » 1 for a full agonist such as dopamine[73], 3.15 simplifies to:

$$\frac{EC_{50X}}{EC_{50Y}} = \frac{\alpha_Y}{\alpha_X} \tag{3.16}$$

Then equation 3.10 can be rewritten as follows:

$$RT\left(pEC_{50X} - pEC_{50Y}\right) = -RT\,log\left(\frac{\alpha_x}{\alpha_y}\right) = \Delta\Delta G^C \tag{3.17}$$

Therefore, we expect a linear relationship between changes in structural coupling $\Delta\Delta G^C$ and changes in ligand potency EC_{50}. $\Delta\Delta G^C$ are the changes in dynamic couplings calculated by NMA as described above.

Benchmark of known mutational effects on receptor signaling activities.

Delta opioid, rhodopsin, dopamine and beta 2 adrenergic receptors were selected for the benchmark because large experimental datasets generated with consistent experimental methods were available for these receptors [114,177,187–191].

The following criteria were applied to select the datasets and mutations. Only mutations of residues in the TMH region or helix 8 were considered to focus on long-range allosteric effects and avoid mutations that may directly affect ligand or G-protein binding. Mutations involving proline residues were not considered as they often disrupt TMH structures. For the prediction of allosteric effects affecting ligand potency, mutations inducing large effects on constitutive activities were not considered since equation 12 relating directly EC50 shifts to allosteric coupling shifts would not be valid.

Since the fully induced activities were not available for all mutant receptors, constitutive activities were normalized to the fully induced wildtype activity (at saturating ligand concentration with a full agonist). We further normalized given activities based on receptor expression (either receptor cell surface expression or total expression), where such data were available [114,190]. Activity data for rhodopsin mutations were not normalized.

Our benchmark totalizes 109 mutations across 4 different receptors (dopamine D2, beta2 adrenergic receptor, Delta opioid receptor, rhodopsin) that involves mutations at 40 unique sites and 26 unique combinations of sites.

Available X-ray structures were used for modeling the inactive and active states of opsin (3DQB, 3C9L) and beta2 adrenergic receptor (3SN6, 2RH1) and for modeling the inactive state of the Delta opioid receptor (4N6H). Homology models were generated for the dopamine receptor and for the active state of the Delta opioid receptor as described above. Ligand-free receptor energies and long-range couplings were calculated as described above for the designed receptors, i.e. by RosettaMembrane and NMA, respectively.

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability. The details (i.e. input files and command lines) of the calculations performed in this study as well as the source codes and executables of the modeling and design methods are available from the corresponding author upon request and will be released free of charge for academic users in the software Rosetta.

3.6 Acknowledgements

The authors would like to thank R. Sharma for setting up some NMA calculations, T. Wensel and his laboratory for sharing equipment and experimental protocols and members of the Barth lab for discussion. K-Y.C. and D.K. were partially funded by a training fellowship from

NIH NIGMS T32GM008280. This work was partially supported by a grant from the National Institute of Health (1R01GM097207), by a supercomputer allocation from XSEDE (MCB120101) to P.B. D.K. This work as well as P.B. and D.K. were also partially supported by funding from EPFL and the Ludwig Institute for Cancer Research.

3.7 Supplementary Tables

Variant	ΔΔGapo shifts from WT (Rosetta Energy Units)	Long-range structural coupling shifts from WT calculated by NMA active (inactive)	%CA vs WT dopamine-induced (Fold change from WT)	Ligand potency Dopamine EC50 (uM) Serotonin EC50 (uM) (Fold change from WT)	Maximal ligand induced activity normalized to ligand induced WT (%) Dopamine Serotonin (Fold change from WT)	Successful prediction of CA shifts	Successful prediction of ligand induced shifts
WT	NA	NA	6.7 ± 2.3 (1.0)	2.9 ± 0.1 11 ± 0.6	100 ± 0.6 100 ± 6.6	NA	NA
T5.54L	-1.78	-0.1 (-0.2)	6.5 ± 0.9 (1.0)	3.4 ± 0.6 (1.2) 8.2 ± 0.9 (0.7)	104.7 ± 4.8 (1) 96.4 ± 8.8 (1)	1	1
T5.54I	-5.36	-0.1 (-0.2)	13.7 ± 2.0 (2.0) **	2.5 ± 0.2 (0.9) 9 ± 0.6 (0.9)	97.2 ± 1.8 (1) 95.9 ± 8.8 (1)	1	1
T5.54M	-1.9	0.9 (0.1)	7.7 ± 2.3 (1.2)	2.3 ± 0.1 (0.8) ** 4.3 ± 0.1 (0.4) **	103.1 ± 3.1 (1) 200.6 ± 9.7 (2) **	1	1
T5.54V	-0.24	0 (-1.0)	6.7 ± 1.1 (1.0)	5.3 ± 0.1 (1.8) ** 6.2 ± 0.4 (0.6) **	86.7 ± 4.2 (0.9) ** 68.9 ± 6.3 (0.7) **	1	1

Table 3.1 – Summary of predicted and experimentally measured activation properties of single microswitch designs at position 5.54. The table reports from left to right the predicted energy differences between active and inactive ligand-free states from WT (i.e. apo equilibrium shift, $\Delta\Delta G_{apo}$), change in long-range structural coupling calculated by NMA according to equation 6 (active state: $(G^{C}_{A,Lb})_{design} - (G^{C}_{A,Lb})_{WT}$; inactive state: $(G^{C}_{I,Lb})_{design} - (G^{C}_{I,Lb})_{WT}$), in vitro measured constitutive activity (fold change and percent of fully dopamine induced WT activity), in vitro measured ligand agonist potency (EC50), in vitro measured maximal agonist induced response normalized to WT, successful prediction of changes in constitutive activity (see methods), successful prediction of changes in ligand-induced activity (see methods). 1=successful; 0=not successful. ** unpaired two-sided t-test $p < 0.05$. Experimental values reported as mean ± s.e.m. RU: Rosetta energy Units.

	D2DR Variant	Predicted ΔΔGapo (active-inactive) (RU)	%CA vs WT induced (Fold increase vs WT basal)	Successful prediction of CA shifts
Initial screening designs	WT	NA	6.7 ± 2.3 (1.0)	NA
	V378Y (V6.40)	-15.85	24.8 ± 11.2 (3.7)	1
	T205I.V378Y (T5.54, V6.40)	-19.01	22.7 ± 4.6 (3.4)	1
	M374L.V378Y (M6.36, V6.40)	-10.92	28.1 ± 5.8 (4.2)	1
	T205I.V378Y.F382I (T5.54, V6.40, F6.44)	-21.97	35.2 ± 2.1 (5.3)	1
	F411W (F7.37)	-13.22	7.9 ± 0.4 (1.2)	0
	F202L.L206I.I383L (F5.51, L5.55, I6.45)	-3.15	15.0 ± 1.4 (2.2)	1
	I73Y.T427A (I2.43, T7.54)	-4.95	9.5 ± 0.8 (1.4)	0
	V381L.V421L (V6.43, V7.48)	-4.27	9.4 ± 2.1 (1.4)	0
Local region R1-R4 designs	**R1:** T392S.L407G.F411W (T6.54, L7.34, F7.37)	-3.27	27.6 ± 1.2 (4.1)	1
	R2: F202L.T205I.L206I.I383L (F5.51, T5.54, L5.55, I6.45)	-4.65	17.8 ± 1.3 (2.7)	1
	R3:T205I.M374L.V378Y.V381L.V421I (T5.54, M6.36, V6.40, V6.43, V7.48)	-15.44	26.4 ± 2.0 (3.9)	1
	R4: I73Y.Y426F.T427A (I2.43, Y7.53, T7.54)	-7.23	22.9 ± 1.2 (3.4)	1

Table 3.2 – Summary of predicted and experimentally measured activation properties of designed D2 variants stabilized in the active state. The table reports from left to right the predicted energy differences between active and inactive ligand-free states from WT (i.e. apo equilibrium shift, $\Delta\Delta G_{apo}$), in vitro measured change in constitutive activity (fold change and percent of fully dopamine induced WT activity), successful prediction of changes in constitutive activity (see methods, 1=successful; 0=not successful). Double-lines delineate two categories of designs from top to bottom as follows: 1) initial screening designs (Ballesteros-Weinstein notation34 in parentheses), 2) local region R1-4 designs. RU: Rosetta energy Units. Experimental values reported as mean ± s.e.m.

	D2DR Variant	Long-range structural coupling shifts from WT calculated by NMA inactive (active)	Spiperone IC50 (nM)	Max spiperone induced activity vs WT (%)	Successful prediction of ligand induced shifts
	WT	NA	8.6 ± 0.9	1.4 ± 2.0	NA
Local region R1-R4 designs	**R1:** T392S.L407G.F411W (T6.54, L7.34, F7.37)	-0.9 (-0.3)	29.2 ± 1.1 (3.4) **	7.0 ± 2.9	1
	R2: F202L.T205I.L206I.I383L (F5.51, T5.54, L5.55, I6.45)	-0.8 (-0.2)	48 ± 1.8 (5.6) **	1.3 ± 1.3	1
	R3:T205I.M374L.V378Y.V381L. V421I (T5.54, M6.36, V6.40, V6.43, V7.48)	-2.8 (-0.8)	166.3 ± 38 (19.2) **	18.0 ± 3.8 **	1
	R4: I73Y.Y426F.T427A (I2.43, Y7.53, T7.54)	-1.5 (0.6)	389.1 ± 54.3 (45.1) **	3.7 ± 1.2	1

Table 3.3 – Summary of predicted and experimentally measured inverse agonist regulation of the activity of designed D2 variants stabilized in the active state. The table reports from left to right the predicted change in long-range structural coupling calculated by NMA according to equation 6 (active state: $(G^C_{A,Lb})_{design} - (G^C_{A,Lb})_{WT}$; inactive state: $(G^C_{I,Lb})_{design} - (G^C_{I,Lb})_{WT}$), in vitro measured ligand inverse agonist potency (IC50), in vitro measured maximal inverse agonist induced response normalized to WT, successful prediction of changes in ligand-induced activity (see methods) 1=successful; 0=not successful. Experimental values reported as mean ± s.e.m (except for spiperone induced activity: mean ± stdev). ** unpaired two-sided t-test $p < 0.05$.

Variant	Dopamine Log(EC50) ± S.E.M. (ΔLogEC50 from D2 WT)	Dopamine Maximal Response ± S.E.M.	Serotonin Log(EC50) ± S.E.M. (ΔLogEC50 from 5HT1AR WT)	Serotonin Maximal Response ± S.E.M.	Long-range ΔΔGC coupling shifts from WT calculated by NMA active - inactive	Successful prediction of ligand induced shifts
D2 WT	-7.64 ± 0.04 (0)	1 ± 0	-5.57 ± 0.15 (1.81)	0.07 ± 0.02	N.A.	N.A.
5HT1AR WT	N.D.	N.D.	-7.38 ± 0.05 (0)	1 ± 0	N.A.	N.A.
T205M (T5.54)	-8.73 ± 0.1 (-1.09)	0.77 ± 0.03	-6.07 ± 0.05 (1.32)	0.71 ± 0.02	0.64	1
F382I (F6.44)	-8.58 ± 0.05 (-0.94)	0.94 ± 0.01	-5.87 ± 0.05 (1.51)	0.91 ± 0.04	0.44	1
C385L (C6.47)	-8.37 ± 0.09 (-0.73)	0.97 ± 0.03	-5.47 ± 0.03 (1.92)	0.4 ± 0.04	-0.51	0
L379I (L6.41)	-8.05 ± 0.09 (-0.41)	0.92 ± 0.02	-5.36 ± 0.05 (2.02)	0.36 ± 0.02	0.15	0
C385V (C6.47)	-7.75 ± 0.01 (-0.11)	0.77 ± 0.02	N.D.	0.17 ± 0.02	0.04	1
M90V (M2.58)	-7.29 ± 0.09 (0.35)	0.91 ± 0.04	N.D.	N.D.	0.15	1
L76M (L2.46)	-7.29 ± 0.05 (0.35)	0.92 ± 0.02	N.D.	N.D.	0.15	1
I122T (I3.40)	-7.01 ± 0.06 (0.63)	0.31 ± 0.02	N.D.	N.D.	-0.49	1
Y426F (Y7.53)	-6.87 ± 0.12 (0.77)	0.62 ± 0.07	N.D.	N.D.	-0.4	1
Y209L (Y5.58)	-6.76 ± 0.39 (0.88)	0.31 ± 0.04	N.D.	N.D.	-0.35	1
F382Y (F6.44)	-6.47 ± 0.05 (1.17)	0.53 ± 0.03	N.D.	N.D.	0.2	0
F202M (F5.51)	-6.34 ± 0.03 (1.3)	0.75 ± 0.01	N.D.	N.D.	0.22	0
I122L (I3.40)	-6.19 ± 0.05 (1.45)	0.49 ± 0.04	N.D.	N.D.	-0.65	1
I122F (I3.40)	N.A.	0 ± 0.02	N.D.	N.D.	-0.21	1
T205M.C385L (T5.54, C6.47)	-9.29 ± 0.01 (-1.65)	0.88 ± 0.03	-6.76 ± 0.06 (0.62)	0.74 ± 0.03	0.74	1
L379I.F382I (L6.41, F6.44)	-9.19 ± 0.07 (-1.55)	0.64 ± 0.05	-5.82 ± 0.15 (1.56)	0.49 ± 0.03	0.61	1
L379I.C385L (L6.41, C6.47)	-8.79 ± 0.06 (-1.15)	1.00 ± 0.02	-6.03 ± 0.05 (1.35)	0.74 ± 0.03	-0.33	0
F382I.C385L (F6.44, C6.47)	-8.57 ± 0.11 (-0.93)	0.95 ± 0.02	-6.16 ± 0.05 (1.22)	0.94 ± 0.03	0.74	1
T205M.C385V (T5.54, C6.47)	-8.2 ± 0.06 (-0.56)	0.47 ± 0.05	-5.33 ± 0.07 (2.05)	0.44 ± 0.04	0.62	1
T205M.L379I (T5.54, L6.41)	-8.1 ± 0 (-0.46)	0.81 ± 0.04	-5.09 ± 0.18 (2.29)	0.12 ± 0.01	0.22	1
T205M.F382I (T5.54, F6.44)	N.A.	< 0.1 ± 0	N.A.	< 0.1 ± 0	-1.15	1

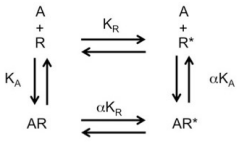
Table 3.4 – Summary of predicted and experimentally measured agonist-induced activity of D2 variants designed with altered allosteric coupling. D2 activities were measured using the cell-based FLIPR assay. The table reports from left to right the experimentally measured dopamine agonist potency (EC50), experimentally measured maximal dopamine induced response normalized to WT, the experimentally measured serotonin agonist potency (EC50), experimentally measured maximal serotonin induced response normalized to WT, the change in long-range structural coupling calculated by NMA according to equation 6 (active - inactive: $((G^{C}_{A,Lb})_{design} - (G^{C}_{A,Lb})_{WT})$ - $((G^{C}_{I,Lb})_{design}$ - $(G^{C}_{I,Lb})_{WT})$), successful prediction of changes in agonist-induced activity (see methods). 1=successful; 0=not successful. Data shown are N=3 independent experiments.

Double mutant	Long-range structural coupling shifts from WT corresponding to ΔΔGC calculated by equation 6	ΔΔGapo shifts from WT calculated by equations (2) - (3)
T205M.C385L (T5.54, C6.47)	0.74	-2.51
L379I.F382I (L6.41, F6.44)	0.61	-2.83
L379I.C385L (L6.41, C6.47)	-0.33	4.51
F382I.C385L (F6.44, C6.47)	0.74	1.99
T205M.C385V (T5.54, C6.47)	0.62	-2.28
T205M.L379I (T5.54, L6.41)	0.22	1.81

Table 3.5 – Allosteric parameters calculated for the D2 variants generated by combining single mutants enhancing agonist ligand sensitivity. $\Delta\Delta G^C$ are calculated from unitless dynamic coupling shifts from WT by NMA and are used to estimate changes in allosteric coupling as described by equation 6. $\Delta\Delta G_{apo}$ (conformational stability shifts from WT calculated by RosettaMembrane in Rosetta Energy Units) correspond to the changes in the ligand-free equilibrium from WT calculated by equation 3.6 - 3.7.

3.8 Supplementary Figures

a A: agonist ligand
R: inactive receptor
R*: active receptor
f_R^*: fraction of active receptor
α: allosteric coupling
K_R: apo equilibrium constant



$K_R = [R^*] / [R]$

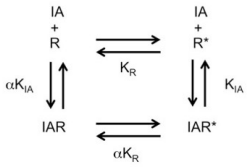
$K_A = [AR] / [A][R]$

$\alpha K_A = [AR^*] / [A][R^*]$

$\alpha K_R = [AR^*] / [AR]$

$f_R^* = ([R^*]+[AR^*]) / ([R]+[R^*]+[AR]+[AR^*]) = (K_R+\alpha K_R K_A[A]) / (1+K_R+K_A[A]+\alpha K_A[A])$
$f_R^*(max)$: maximum receptor activity at [A] -> ∞ $= \alpha K_R / (1+\alpha K_R)$
EC50: [A] at mid-point of the dose response curve $= (1+K_R) / K_A(1+\alpha K_R)$

b IA: inverse agonist ligand
R: inactive receptor
R*: active receptor
f_R^*: fraction of active receptor
f_R: fraction of inactive receptor
α: allosteric coupling
K_R: apo equilibrium constant



$K_R = [R] / [R^*]$

$K_{IA} = [AR^*] / [A][R^*]$

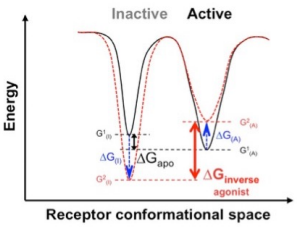
$\alpha K_{IA} = [AR] / [A][R]$

$\alpha K_R = [AR] / [AR^*]$

$f_R = ([R]+[AR]) / ([R]+[R^*]+[AR]+[AR^*]) = (K_R+\alpha K_R K_{IA}[A]) / (1+K_R+K_{IA}[A]+\alpha K_{IA}[A])$
$f_R(max)$: minimal receptor activity at [IA] -> ∞ $= \alpha K_R / (1+\alpha K_R)$
IC50: [IA] at mid-point of the dose response curve $= (1+K_R) / K_{IA}(1+\alpha K_R)$

$f_R^* = 1 - f_R$

c



Inactive state stability: $G_{(I)}$; active state stability: $G_{(A)}$

Ligand-free stability difference: $\Delta G_{apo} = G^1_{(I)} - G^1_{(A)} = -RT\ln(K_R)$

Ligand-bound stability difference: $\Delta G_{inverse\ agonist} = G^2_{(I)} - G^2_{(A)} = -RT(\ln(K_R)+\ln(\alpha))$

Effect of ligand binding on inactive state stability: $\Delta G_{(I)} = G^2_{(I)} - G^1_{(I)}$
Effect of ligand binding on active state stability: $\Delta G_{(A)} = G^2_{(A)} - G^1_{(A)}$

Allosteric effect from ligand binding:
$\Delta\Delta G_{inverse\ agonist} = \Delta G_{(I)} - \Delta G_{(A)} = \Delta G_{inverse\ agonist} - \Delta G_{apo} = -RT\ln(\alpha)$

Stability (apo equilibrium) shift

Allosteric coupling shift

Combination

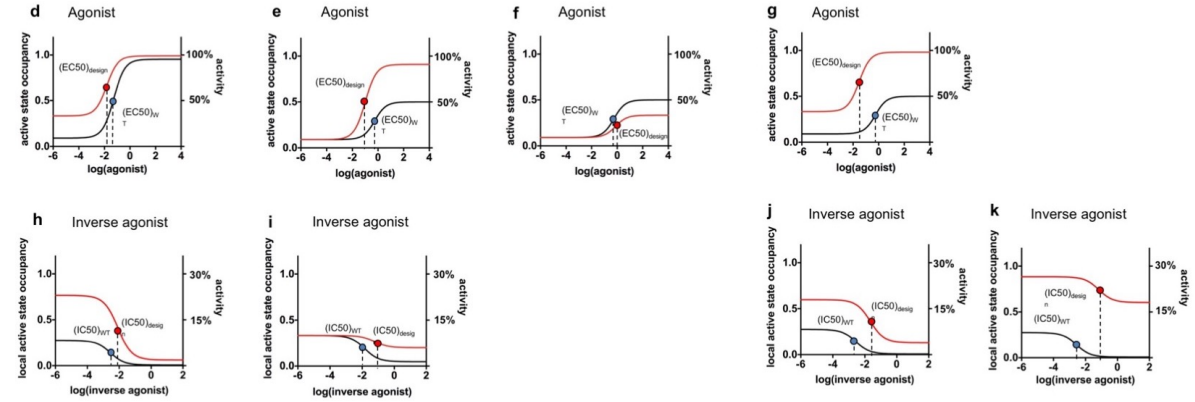


Figure 3.6 – The Allosteric Two State Model (ATSM) predicts the effects on receptor activity of mutation-induced shifts in stability and allosteric coupling. a-b. Thermodynamic cycles depicting ATSMs for agonist (a) or inverse agonist (b) regulation of receptor activity[146]. c. Free energy landscape depicting the ATSM for inverse agonist regulation of receptor activity[146]. d-k. Theoretical dose response curves of receptor activity upon agonist (d-g) or inverse agonist (h-k) titration. The ATSM parameters underlying the simulated curves were selected to best recapitulate structure-based energy, coupling calculations and the experimentally measured dose responses of the different classes of variants designed in the study. Black and red curves stimulate WT D2 and designed D2 variants, respectively. e-g, i-k. Changes in allosteric coupling of a ligand can be achieved by designing mutations altering the structural coupling between the ligand and the G protein binding sites in selective states (Methods). d. Stability (i.e. apo (ligand-free) equilibrium) shift resulting in increased constitutive activity without substantial changes in fully induced agonist activity and ligand potency (valid only for full agonist and allosteric coupling α » 17). This curve stimulates the R1-4 and combined variants responses to the full agonist dopamine. e. Increased allosteric coupling of the ligand agonist resulting in increased induced agonist activity and agonist potency. This curve stimulates for example the T5.54M response to the agonist dopamine. f. Decreased allosteric coupling of the ligand agonist resulting in decreased induced agonist activity and agonist potency. This curve stimulates for example the T5.54V response to the agonist dopamine. g. Theoretical example of combined stability (i.e. apo equilibrium) and allosteric coupling shifts resulting in increased constitutive, agonist induced activities and agonist potency. h-k. Dose responses to inverse agonists are represented for local active states adopted by receptor variants designed in distinct local R regions. h. Stability (i.e. apo equilibrium) shift resulting in increased constitutive activity without substantial change in inverse agonist bound activity and potency. This curve stimulates for example the R2 response to the inverse agonist spiperone. i. Theoretical example of decreased allosteric coupling of the inverse agonist resulting in both decreased inhibition by inverse agonist and decreased inverse agonist potency. j. Moderate stability (i.e. apo equilibrium) shift combined with decreased allosteric coupling of the inverse agonist resulting in increased constitutive activity, slight decreased maximal ligand-induced receptor inhibition and large decreased inverse agonist potency. This curve stimulates the R4 response to the inverse agonist spiperone. k. Large stability (i.e. apo equilibrium) shift combined with decreased allosteric coupling of the inverse agonist resulting in increased constitutive activity, large decrease in maximal ligand-induced receptor inhibition and ligand potency. This curve stimulates the R3 response to the inverse agonist spiperone.

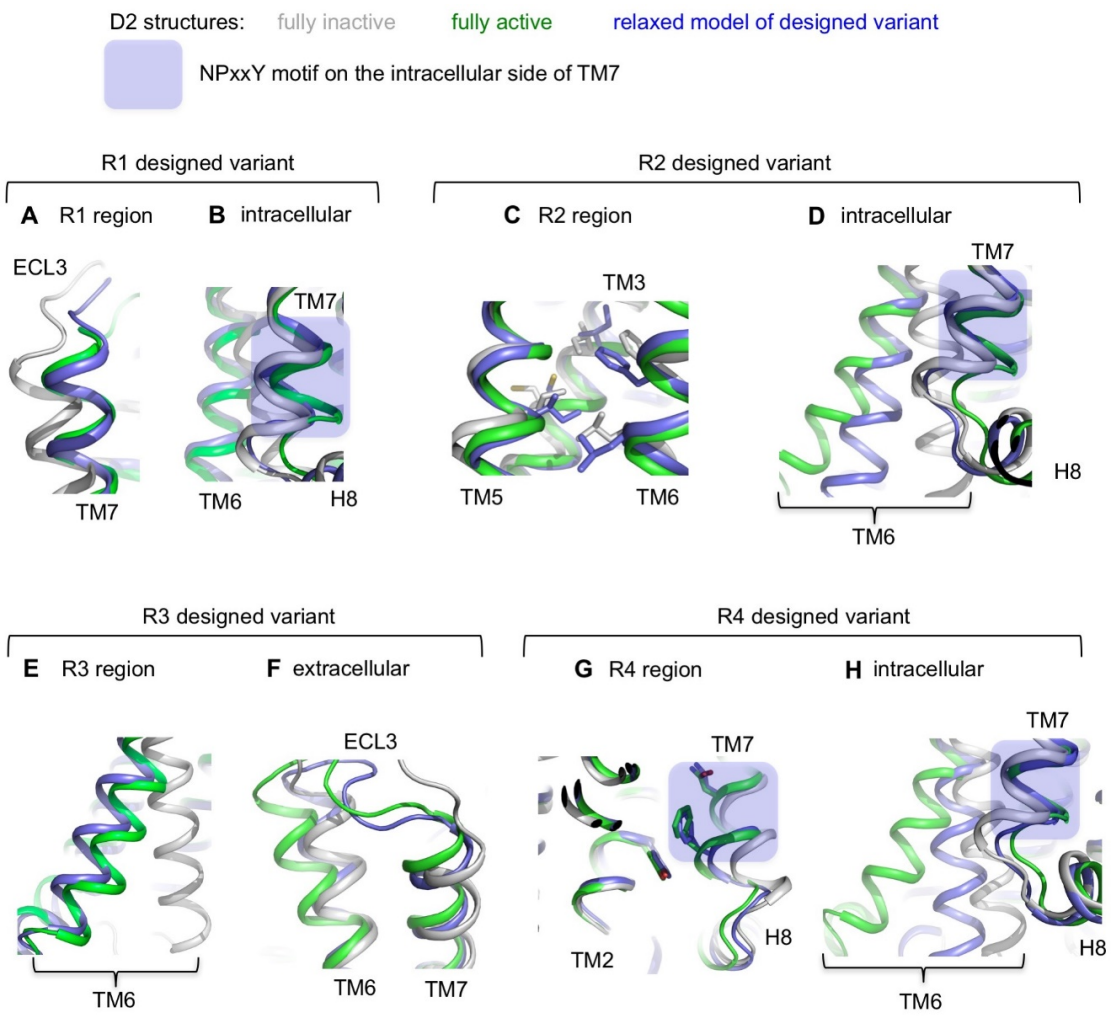

D2 structures: fully inactive fully active relaxed model of designed variant
NPxxY motif on the intracellular side of TM7
R1 designed variant
R2 designed variant
A R1 region
B intracellular
C R2 region
D intracellular
ECL3
TM7
TM6
H8
TM3
TM5
TM6
TM7
H8
TM6
R3 designed variant
R4 designed variant
E R3 region
F extracellular
G R4 region
H intracellular
TM6
ECL3
TM6
TM7
TM7
H8
TM2
TM7
H8
TM6

Figure 3.7 – The Allosteric Two State Model (ATSM) predicts the effects on receptor activity of mutation-induced shifts in stability and allosteric coupling. a-b. Thermodynamic cycles depicting ATSMs for agonist (a) or inverse agonist (b) regulation of receptor activity7. c. Free energy landscape depicting the ATSM for inverse agonist regulation of receptor activity7. d-k. Theoretical dose response curves of receptor activity upon agonist (d-g) or inverse agonist (h-k) titration. The ATSM parameters underlying the simulated curves were selected to best recapitulate structure-based energy, coupling calculations and the experimentally measured dose responses of the different classes of variants designed in the study. Black and red curves stimulate WT D2 and designed D2 variants, respectively. e-g, i-k. Changes in allosteric coupling of a ligand can be achieved by designing mutations altering the structural coupling between the ligand and the G protein binding sites in selective states (Methods). d. Stability (i.e. apo (ligand-free) equilibrium) shift resulting in increased constitutive activity without substantial changes in fully induced agonist activity and ligand potency (valid only for full agonist and allosteric coupling α » 17). This curve stimulates the R1-4 and combined variants responses to the full agonist dopamine. e. Increased allosteric coupling of the ligand agonist resulting in increased induced agonist activity and agonist potency. This curve stimulates for example the T5.54M response to the agonist dopamine. f. Decreased allosteric coupling of the ligand agonist resulting in decreased induced agonist activity and agonist potency. This curve stimulates for example the T5.54V response to the agonist dopamine. g. Theoretical example of combined stability (i.e. apo equilibrium) and allosteric coupling shifts resulting in increased constitutive, agonist induced activities and agonist potency. h-k. Dose responses to inverse agonists are represented for local active states adopted by receptor variants designed in distinct local R regions. h. Stability (i.e. apo equilibrium) shift resulting in increased constitutive activity without substantial change in inverse agonist bound activity and potency. This curve stimulates for example the R2 response to the inverse agonist spiperone. i. Theoretical example of decreased allosteric coupling of the inverse agonist resulting in both decreased inhibition by inverse agonist and decreased inverse agonist potency. j. Moderate stability (i.e. apo equilibrium) shift combined with decreased allosteric coupling of the inverse agonist resulting in increased constitutive activity, slight decreased maximal ligand-induced receptor inhibition and large decreased inverse agonist potency. This curve stimulates the R4 response to the inverse agonist spiperone. k. Large stability (i.e. apo equilibrium) shift combined with decreased allosteric coupling of the inverse agonist resulting in increased constitutive activity, large decrease in maximal ligand-induced receptor inhibition and ligand potency. This curve stimulates the R3 response to the inverse agonist spiperone.

Prediction type	Active versus inactive state stability shift	Allosteric coupling shift
Success rate	81% (13 among 16)	Variant: 82% (23 among 28)

Figure 3.8 – Qualitative success rate of the designed D2 receptor predictions for apo equilibrium shifts and allosteric coupling shifts.

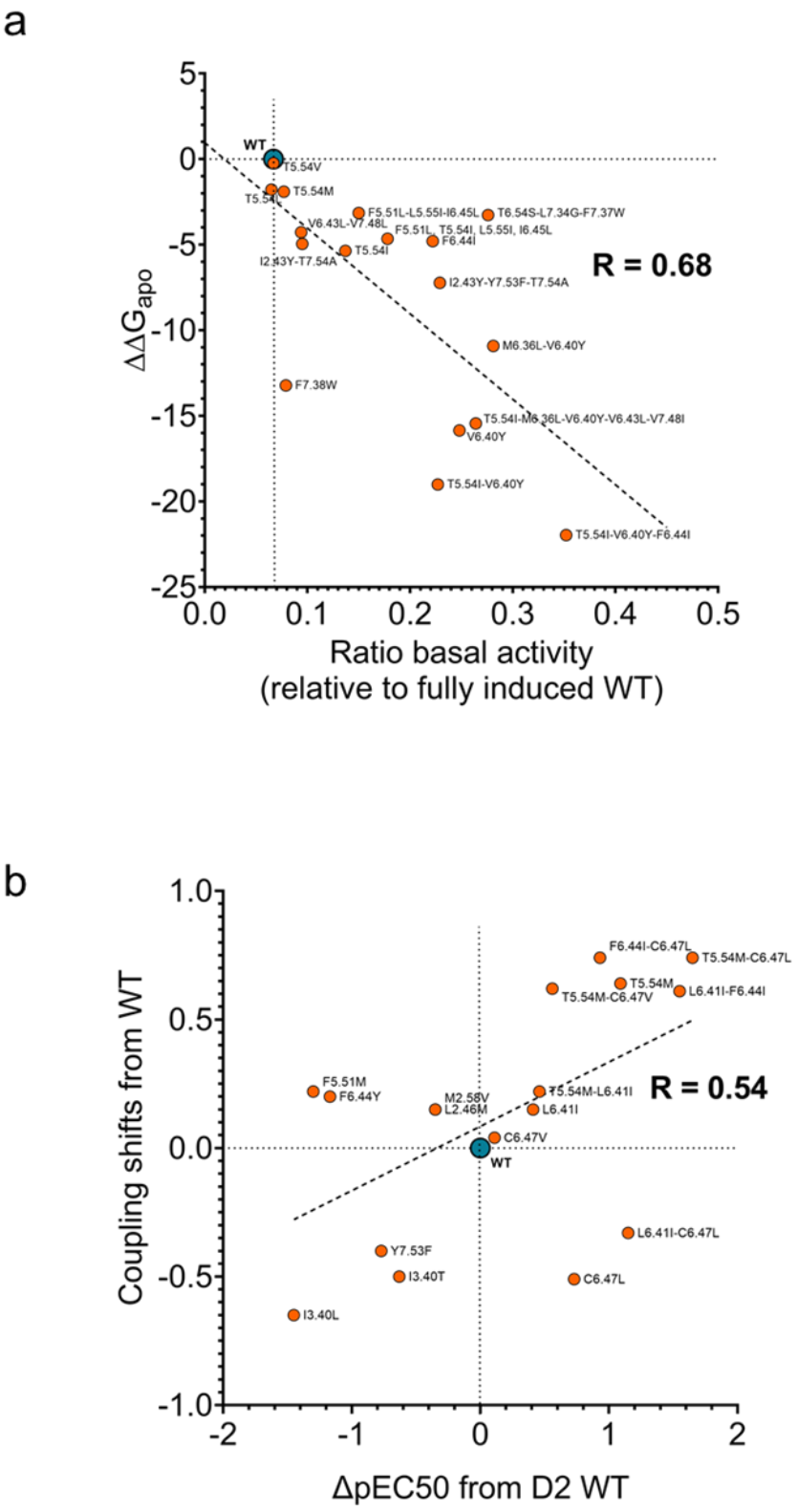

a
WT
T5.54V
T5.54M
T5.54L
F5.51L-L5.55I-I6.45L
T6.54S-L7.34G-F7.37W
V6.43L-V7.48L
F5.51L, T5.54I, L5.55I, I6.45L
F6.44I
T5.54I
I2.43Y-T7.54A
R = 0.68
I2.43Y-Y7.53F-T7.54A
M6.36L-V6.40Y
F7.38W
T5.54I-M6.36L-V6.40Y-V6.43L-V7.48I
V6.40Y
T5.54I-V6.40Y
T5.54I-V6.40Y-F6.44I
ΔΔGapo
Ratio basal activity
(relative to fully induced WT)
b
F6.44I-C6.47L
T5.54M-C6.47L
T5.54M
L6.41I-F6.44I
T5.54M-C6.47V
F5.51M
F6.44Y
M2.58V
L2.46M
T5.54M-L6.41I
L6.41I
R = 0.54
C6.47V
WT
L6.41I-C6.47L
Y7.53F
I3.40T
C6.47L
I3.40L
Coupling shifts from WT
ΔpEC50 from D2 WT

Figure 3.9 – Quantitative relationship between predicted and measured signaling properties for the designed D2 receptors. a. Linear correlation between the ratio of basal activities relative to fully induced WT and predicted shifts in relative ligand-free stability from WT ($\Delta\Delta G_{apo}$ in Rosetta Energy Units) for the designed dopamine D2 receptor variants. n=18; F test p-value=0.0018 b. Linear correlation between measured shifts in dopamine potency (reported as pEC50) from WT and predicted change in structural coupling ($\Delta\Delta G^C$) for dopamine D2 receptor variants designed for altered signaling. The $\Delta\Delta G^C$ are calculated from unitless dynamic correlations by NMA and are used to estimate the changes in allosteric couplings. n=17; F test p-value=0.0263.

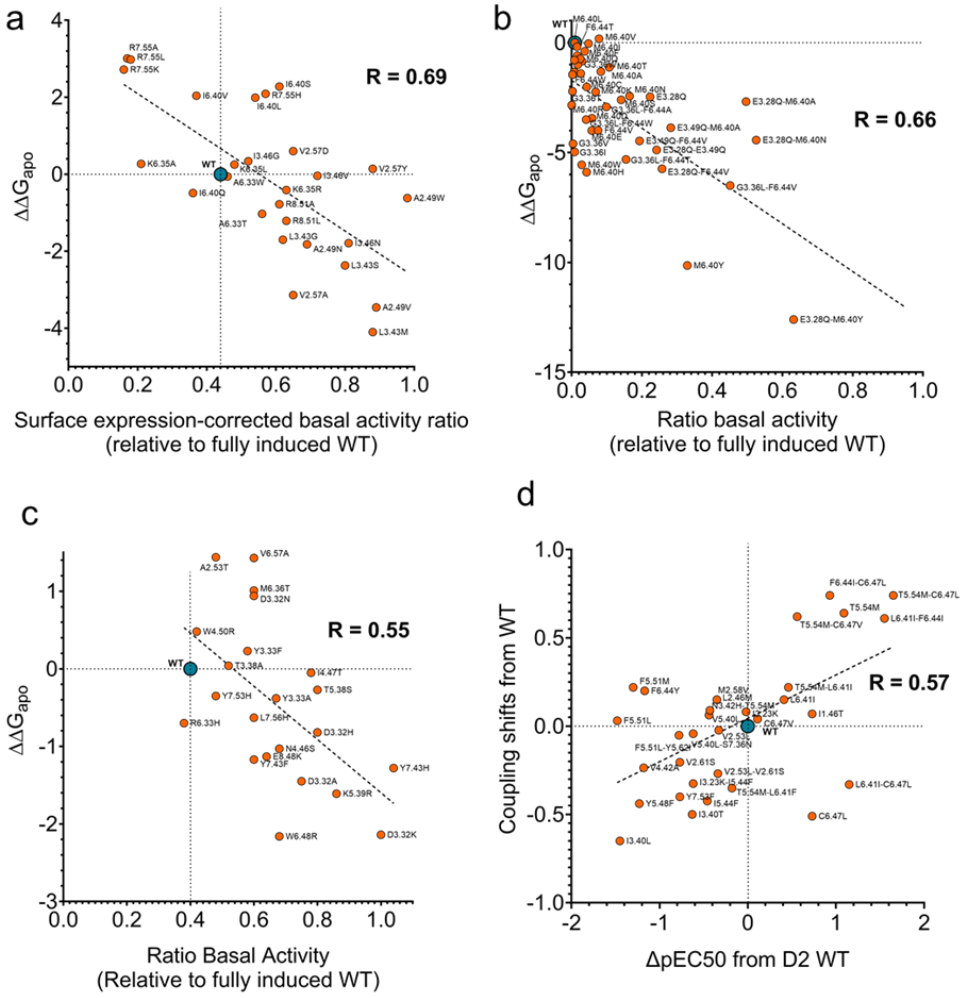


Figure 3.10 – Quantitative relationship between predicted and measured signaling properties for known mutations in diverse GPCRs. **a-c.** Linear correlations between the ratio of basal activities relative to the fully induced WT receptor and predicted shifts in relative ligand-free stability from WT ($\Delta\Delta G_{apo}$ in Rosetta Energy Units) for **(a)** Beta 2 adrenergic receptor (B2AR, n=28, F test p-value<0.0001), **(b)** rhodopsin (n=43, F test p-value<0.0001), **(c)** Delta opioid receptor (n=23, F test p-value=0.0063). **d.** Linear correlation between measured shifts in dopamine potency (reported as pEC50) from WT and predicted change in structural coupling ($\Delta\Delta G^C$) for dopamine D2 receptor variants designed for altered signaling, with predictions for additional available mutational data. The $\Delta\Delta G^C$ are calculated from unitless dynamic correlations by NMA and are used to estimate the changes in allosteric couplings.

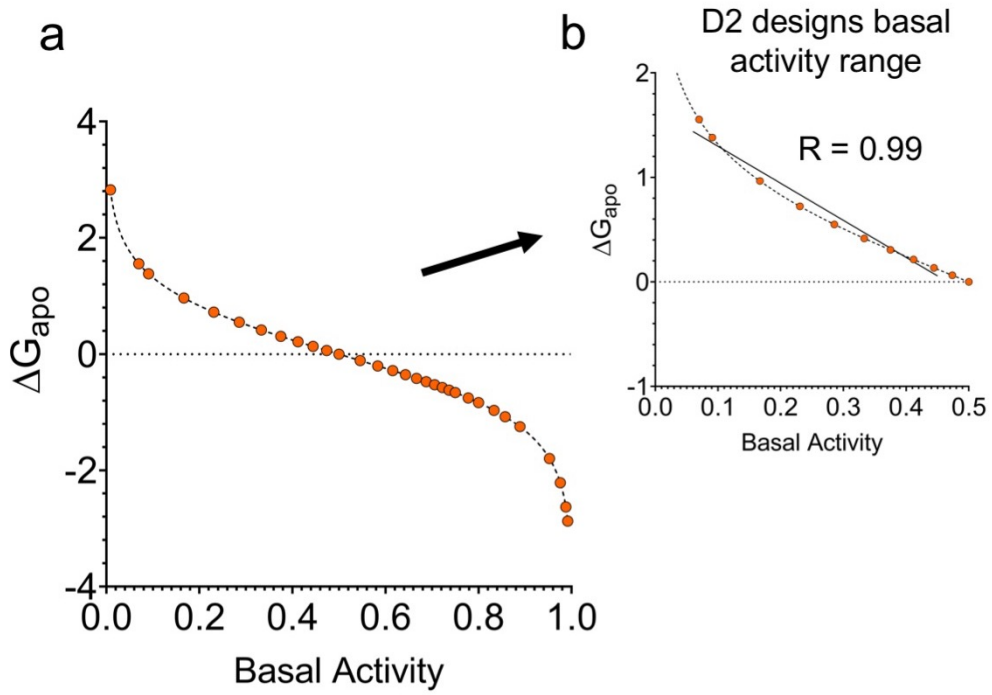


Figure 3.11 – Theoretical relationships between apo equilibrium shifts and constitutive activities according to the allosteric two state model. **a.** Theoretical relationship between the relative stability of the ligand-free active state versus the ligand-free inactive states ($\Delta\Delta G_{apo}$ in kcal/mol) and basal activities (represented as ratios of the maximal receptor activity) calculated by equation 8. **b.** Quasilinear relationship between both quantities for the range of basal activities measured for the D2 receptor.

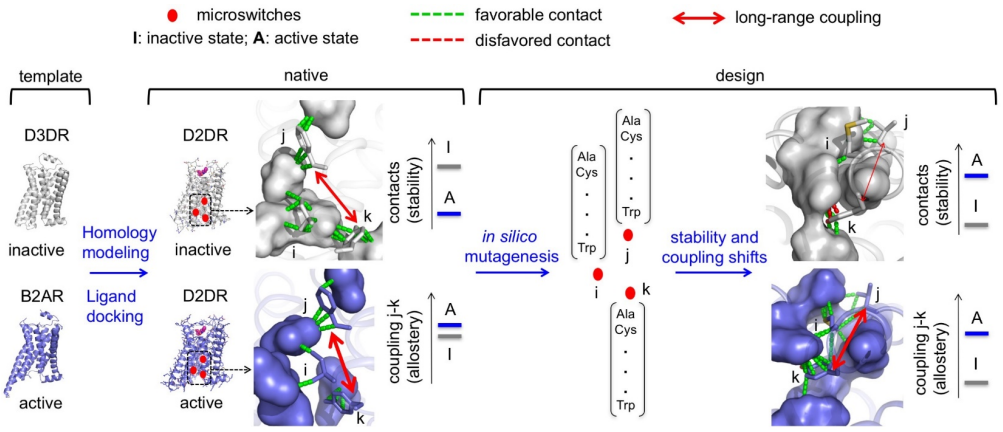


Figure 3.12 – Integrated homology modeling, ligand docking and design framework. Closest homolog structures are selected to model D2 inactive and active state structures. Three native microswitches, including two allosteric residues (positions j and k), are highlighted with short-range contacts (favorable in green and disfavored in red dotted lines) and structural coupling strength by red arrows. The design calculations involve in silico mutagenesis of the 3 positions to all 20 amino acids and selection of the combinations that are predicted to increase (as an example) the stability and structural coupling of the active relative to the inactive state D2 structures. These stability and coupling differences report on the ligand-free equilibrium (K_R) and the allosteric coupling (α).

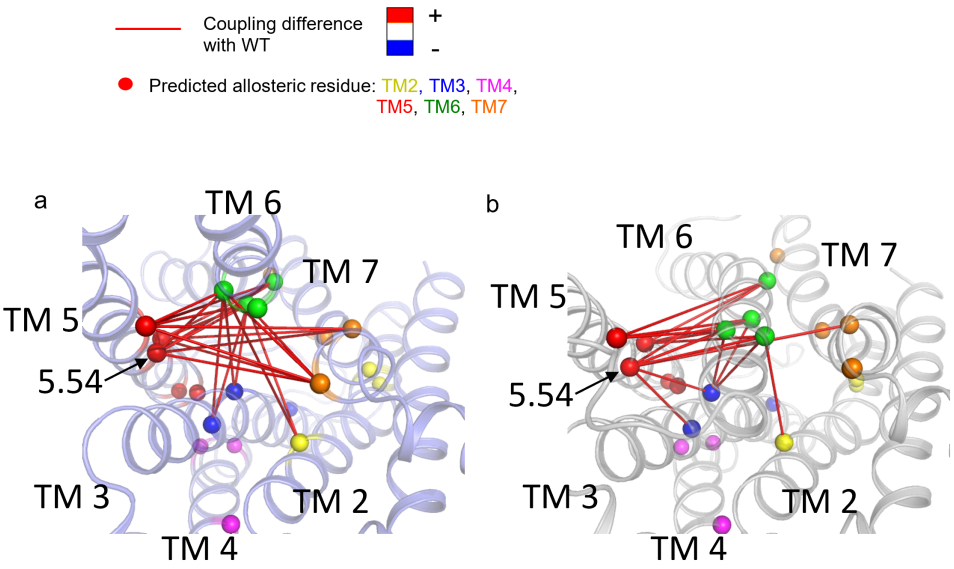


Figure 3.13 – Contrasting long range coupling changes for position 5.54 mutants. a. Met 5.54 strengthened the structural coupling between the intracellular side of TMH7 with TMH5 and TMH6, locking TMH7 in the inward conformation and facilitating the outward motions of TMH6 that are both critical to G protein coupling[146]. b. Conversely, Val 5.54 increased almost exclusively the structural coupling between the intracellular side of TMH6 with TMHs 3 and 5, locking TMH6 in its inactive conformation.

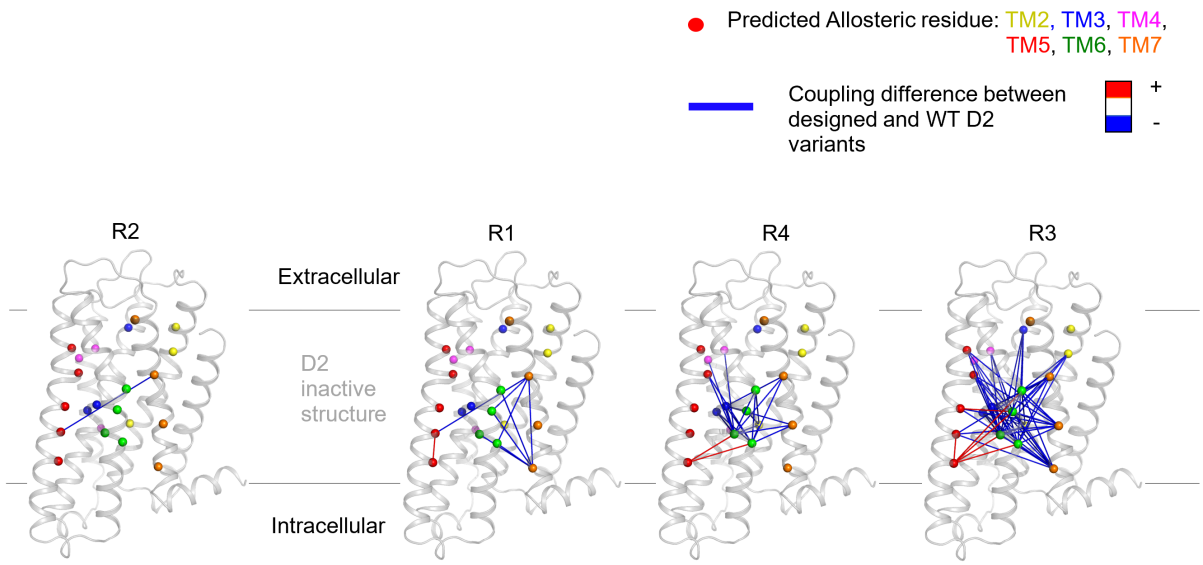


Figure 3.14 – Predicted effects of the designed microswitches on the coupling between spiperone and the closed "inactive" conformation of the G-protein binding site. Large changes in coupling between each variant and WT in the spiperone-bound inactive state are represented by the differences in inter-allosteric residue coupling (i.e. changes in correlated motions of amplitude > 0.02 measured by Normal Mode Analysis, see method). The coupling changes are mapped on the inactive state structure using a colored line (red and blue for coupling increase and decrease, respectively). Allosteric residues are indicated as colored nodes on the structure. The color defines their location on specific TMHs. Variants are ordered from left to right according to the extent of coupling loss from WT. Changes in inter-residue coupling in the active state are minimal. While the structural coupling between allosteric residues was only slightly perturbed by the designed microswitches in R2, the number of allosteric residues losing dynamic correlations gradually increased in the R1, R4 and R3 variants. The largest effects were observed for R3. Our calculations indicated that spiperone mainly lost coupling to the closed “inactive” conformation of the G-protein binding site in these variants while the coupling to the active G-protein bound conformation of the intracellular domain remained largely unaffected (Table 3.3). This selective effect should reduce the ligand’s ability to inhibit receptor activity, consistent with the experimental observations. Except for a few allosteric residues linking TMH5 and TMH6, a large majority of the allosteric networks linking the extracellular ligand to the inactive G-protein binding sites were weakened in the R3 variant. The structural coupling of TMH6 with TMHs 3, 5 and 7, which maintains TMH6 in the inactive conformation was the most affected in the R1, R4 and R3 variants and appears to be a primary determinant for the allosteric inhibition by inverse agonists (Fig.3h).

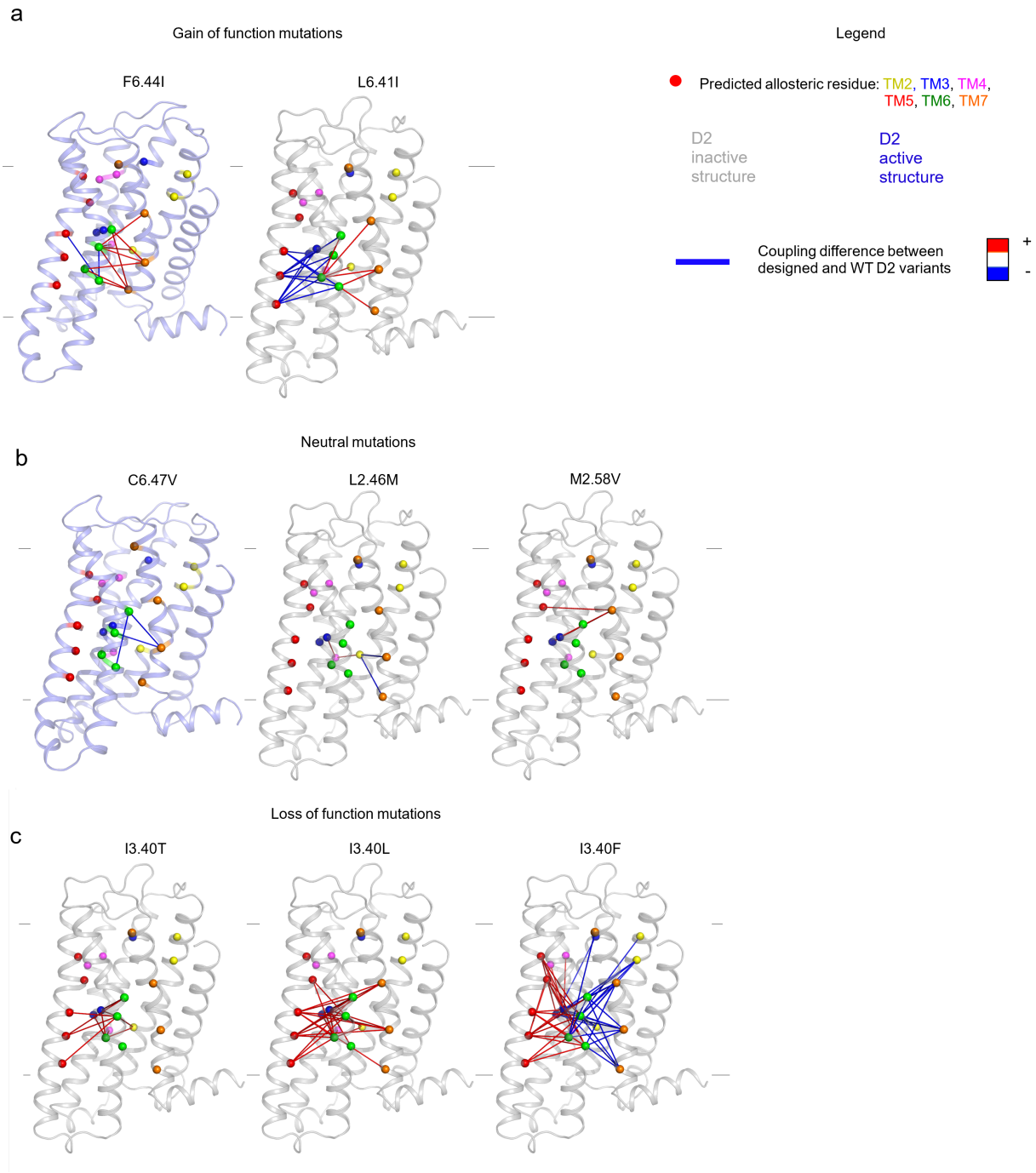


Figure 3.15 – Changes in dynamic coupling in gain-of-function, neutral, and loss-of-function mutations. Changes in dynamic coupling from WT (differences in correlated motions of amplitude > 0.02 measured by Normal Mode Analysis, see method) induced by the designed microswitch are highlighted by a colored line (red and blue for coupling increase and decrease, respectively). The selected state (i.e. inactive, grey or active, blue) is the most perturbed by the mutation.

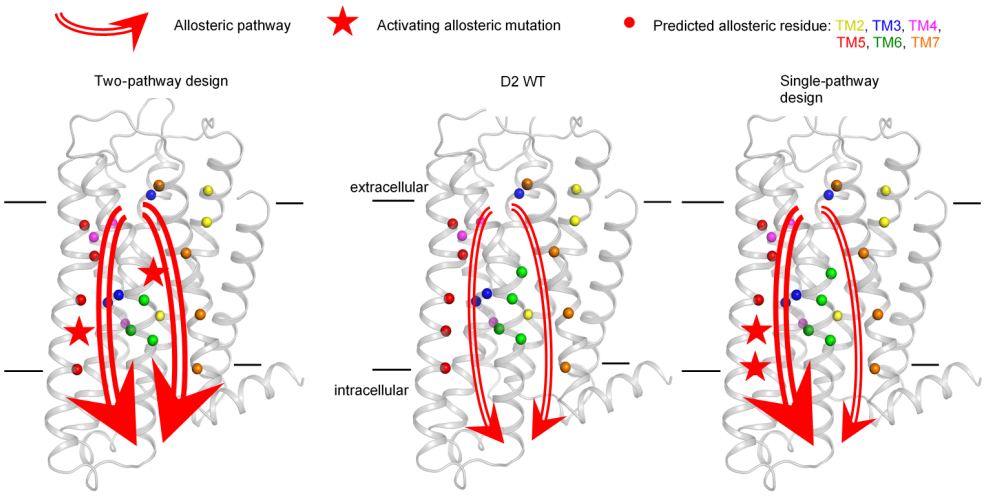


Figure 3.16 – Schematic agonist-induced signal response curve representing the effects of conformational selector mutations stabilizing the receptor active state. Spatially distant of gain-of-function mutation combinations involved in different TMH interactions exhibit additive changes in signaling, whereas nearby gain-of-function mutations either display the effects of the dominant mutation or result in partial or complete loss of function.

4 Determining The Functional, Structural, and Dynamic Effects of Allosteric Mutations on Agonist Signaling in The Dopamine D2 Receptor

D. Keri[1], A.L.J.C. Oggier[1], P. Barth[1]

[1] Swiss Federal Institute of Technology (EPFL), Institute of Bioengineering, Lausanne, Switzerland.

Research in progress

Author Contributions: PB and DK designed study and experiments. DK performed G protein activation and β-arrestin recruitment experiments. AO performed β-arrestin recruitment experiments. PB and DK analyzed the data. DK performed computational work. DK wrote chapter.

4.1 Abstract

We have previously designed a number of highly sensitive dopamine receptors based on the dopamine D2 receptor. These designs not only proved to be roughly up to 100 times more sensitive to dopamine but have also shown increased sensitivity to the non-native ligand serotonin, matching the maximum response seen with the native serotonin receptor 5HT1AR. Here we explored the extent of sensitivity and promiscuity increases in these receptors. We find a general, but not universal, sensitivity increase: interestingly the mutations which serve to increase sensitivity to some agonists result in the loss of sensitivity to others. We are currently in the process of determining the structural features responsible for this divergent effect.

4.2 Introduction

Long-range regulation of distant protein sites, or allostery, is a mechanism widely adopted by biological systems to regulate enzymes and biochemical pathways. This long-range intramolecular communication is thought to be mediated by networks of residues with coupled motions. Cells typically take advantage of such motions through ligand and protein binding or post-translational modifications at one site, resulting in the propagation of the change in local energy through coupled allosteric residues. These long-range correlated motions can be detected using molecular dynamic (MD) simulations and normal mode analysis (NMA) [77,79,192]. Using these tools, it has been possible to modulate the strength of allosteric communication in the dopamine D2 receptor [53].

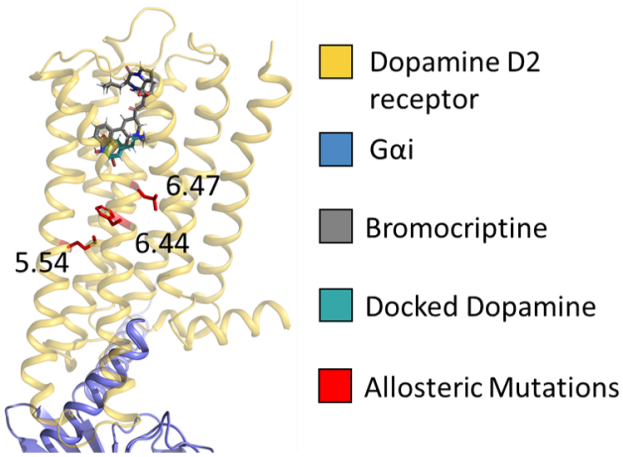


Figure 4.1 – Position of the allosteric gain-of-function mutations in the dopamine D2 receptor structure (6VMS).

The D2 receptor is a member of the G protein-coupled receptor family, which are the premiere signal transducers in cells, relaying external signals to intracellular effectors. These receptor family has evolved to translate diverse extracellular inputs to intracellular effector proteins such as G proteins and arrestins. This allosteric signal transduction across a distance of 30A occurs via allosteric pathways that have been mostly conserved between different receptors. GPCRs act as single-domain, modular units with loose coupling between the ligand-binding module and the effectors binding module [155]. We have recently generated a number of D2 variants (T5.54M, F6.44I, C6.47L, and T5.54M-C6.47L using Ballesteros-Weinstein numbering; shown in Figure 4.1), which are predicted to exhibit an increased level of coupling (i.e. increased correlated motions) between these two modules, experimentally resulting in an increase in sensitivity to not only the endogenous ligand dopamine but also to the non-native ligand serotonin [53]. This indicated an inherent and potentially ligand-independent increase in sensitivity.

The interplay between allosteric mutations and ligand efficacy and potency has not received much attention. Determining the diversity of responses and their structural underpinnings

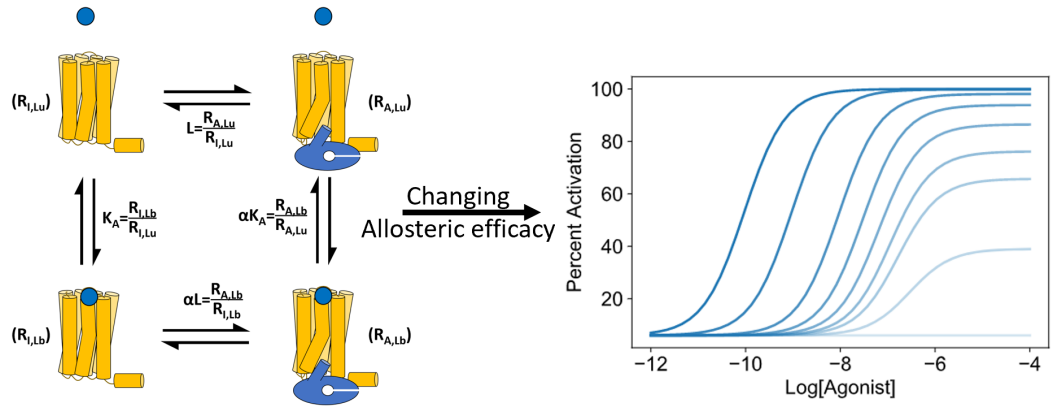


Figure 4.2 – Changes to allosteric efficacy α in the allosteric two-state model. The Allosteric two-state model (left) and theoretical dose response curves with changing values for α (right).

can potentially be useful in future allosteric protein design endeavors. With this in mind, we wanted to explore the scope of enhanced sensitivity and promiscuity exhibited by our designed receptors using a diverse set of ligands. If the allosteric two-state model (ATSM) holds, all agonists should show a proportionate increase in allosteric efficacy (see Allostery introduction for details). Figure 4.2 shows the ATSM (left) and effects on expected dose response curves (right).

4.3 Methods

4.3.1 Computational

Ligand Docking

Simultaneous structure relaxation and coarse grain docking were performed on the dopamine D2 active state structure[193] with heteroatoms, G protein, auxiliary proteins, and ligand removed as previously described[108]. 20,000 decoys were generated in total, keeping only the 10% lowest energy decoys. These decoys were then clustered and the ligand was redocked at full-atom resolution on the 5 largest clusters, generating 20,000 decoys each. For each cluster, the ligand heavy atom coordinates were extracted and used to cluster ligand positions using a DBSCAN-based algorithm. The largest ligand clusters were selected as the putative native ligand-binding modes and the lowest energy models were chosen as the final models.

4.3.2 Experimental

TRP assay

HEK-293 cells stably expressing TrpC4β channel were kindly provided by Dr. Michael X. Zhu and were transiently transfected with A2AR WT and various Gi-s chimeras as described below. FLIPR Membrane potential assays (Molecular Devices) were performed as previously described[179]. Briefly, the assay utilizes a fluorescent probe that reacts to cations and is coupled with a non-permeable quencher. Activation of TrpC4β causes influx of cations leading to the translocation of the fluorescent probe away from the quencher, resulting in increased fluorescence. TrpC4β has been described to be downstream of Gi activation and serves as a reporter for Gi and Gi-s chimeras. Stable HEK-293 cells were cultured in DMEM supplemented with 10% FBS and 1% G418 (selection antibiotic) for use with FLIPR Membrane Potential (FMP) assay kits (Molecular Devices). On the day of transfection, cells stably expressing TrpC4β were plated (1.5x10^5 cells per well) onto 96-well clear, flat-bottom plates pre-coated with poly-D-lysine (Sigma). Cells were transfected with Lipofectamine 2000 and the appropriate D2 constructs using optimized DNA quantity for the FLIPR assay and grown for 24 hours prior to assaying.

β-Arrestin enhanced bystander BRET assay

5x10^4 HEK-293T are seeded into white 96-well plate and are transiently co-transfected with dopamine D2 receptor, RLuc8-β-arrestin2, rGFP-CAAX, and GRK2 in pcDNA 3.1 based on a previous protocol[140]. DNA titration experiments have been previously performed to ensure equalized receptor surface expression. 100μL of DMEM with 10% FBS is added to each well and cells are incubated for 24 hours. The media is removed and wells are washed with 150μL of PBS. 40μL of HPSS buffer supplemented with 0.2% is added to each well. 40μL of coelenterazine 400A is added to each well just before reading for a final concentration of 2.5μM. 40μL of drug is added to wells and plates are incubated for 5 minutes. Donor RLuc8 and acceptor rGFP luminescence values are read at 400nm and 510nm, respectively, on a Mithras2 plate reader and BRET ratios are calculated.

ELISA

HEK-293T cells are transfected with the same conditions used for the assay accompanying the ELISA. 24 hours post-transfection the media is removed from the wells, and cells are fixed to the bottom of the wells with the addition of 40μL of 4% para-formaldehyde (PFA) for 10 minutes. After removal of PFA, a blocking solution of 2% bovine serum albumin (BSA) in PBS is added for 1 hour. Blocking solution is removed and a 1:300 solution of the primary antibody is added in new BSA solution. Plates are then washed 3x with 150μL of PBS. Secondary antibody is added at a 1:2000 ratio in 40μL and cells are again incubated for 1 hour. Cells are once again washed 3x with 150μL of PBS. 90μL of a 1:1 mix of PICO luminescence mix is added to each

well and incubated in the dark for 5 minutes. Finally, the luminescence of each well is read on a Flexstation3 with an integration time of 1 second.

4.4 Results

4.4.1 Altered Signaling of Designed Dopamine D2 Receptors

We tested the previously designed D2 variants using a number of structurally and functionally diverse ligands, looking at the activation of distinct downstream pathways. While we knew these mutations resulted in receptor activation at lower ligand concentrations (i.e. lower EC50s) when observing activation of G proteins, we did not know whether these effects would be mirrored in other commonly-activated GPCR effectors such as β-arrestin. In addition to the endogenous ligands dopamine and serotonin we selected ligands shown in Figure 4.3 due to their unique properties: MLS1547 is reported to be fully biased towards G protein activation[194], whereas UNC9994 is biased towards arrestin activation[195]. Both of these ligands are structurally related to the unbiased partial agonist aripiprazole, the main difference being the linker length between a bicyclic ring and a cyclic amine-connected heterocyclic phenyl group. The bias of these ligands is attributed to this difference in linker length. We also selected the structurally unrelated full agonist bromocriptine. Dose response curves for each ligand are shown in Figures 4.4 and 4.5 for G protein activation and β-arrestin recruitment, respectively.

Similar to the effect seen with dopamine and serotonin, the three structurally similar ligands aripiprazole, MLS1547, and UNC9994 elicited higher G protein activation in the designed receptors, although at slightly different magnitudes depending on the receptor. D2 T5.54M-C6.47L showed the largest shift in EC50 with dopamine, serotonin, and MLS1547; displaying additive shifts in EC50 and maximum responses of the single mutants. Variant F6.44I displayed the highest maximum response to ligands Aripiprazole, MLS1547, UNC9994, and serotonin; while also roughly matching the EC50 of T5.54M-C6.47L. This would indicate F6.44I has a higher allosteric efficacy for certain ligands compared to the double-mutant. Interestingly, the tested partial agonists (MLS1547, UNC9994, aripiprazole) never reached the maximum response seen with dopamine and remained partial agonists. This is in contrast to the behavior seen with serotonin, where we see an increase in maximum response from 5% of WT relative to dopamine to 100% seen with F6.44I.

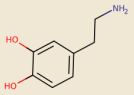
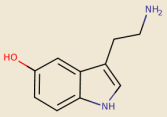
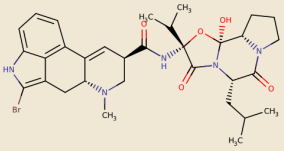
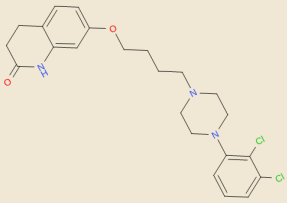
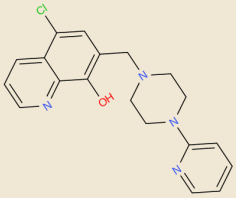
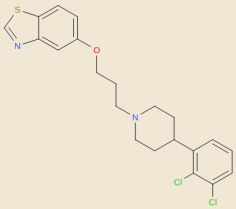
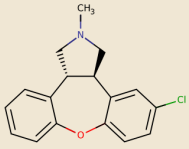
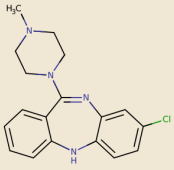
Ligand	Structure	Dopamine D2 WT Action	Bias	Interest
Dopamine		Full agonist	Unbiased	Reference ligand
Serotonin		Very weak partial agonist	N/A	Full agonist for previous D2 designs
Bromocriptine		Full agonist	Unbiased	Unique structure
Aripiprazole		Partial agonist	Unbiased	Unique structure
MLS1547		Partial agonist	Fully G protein biased	Biased ligand
UNC9994		Partial agonist	Fully Arrestin biased	Biased ligand
Asenapine		Antagonist	N/A	Full agonist of serotonin receptors
Clozapine		Antagonist	N/A	Full agonist of some serotonin receptors

Figure 4.3 – Dopamine ligands tested with D2 variants.

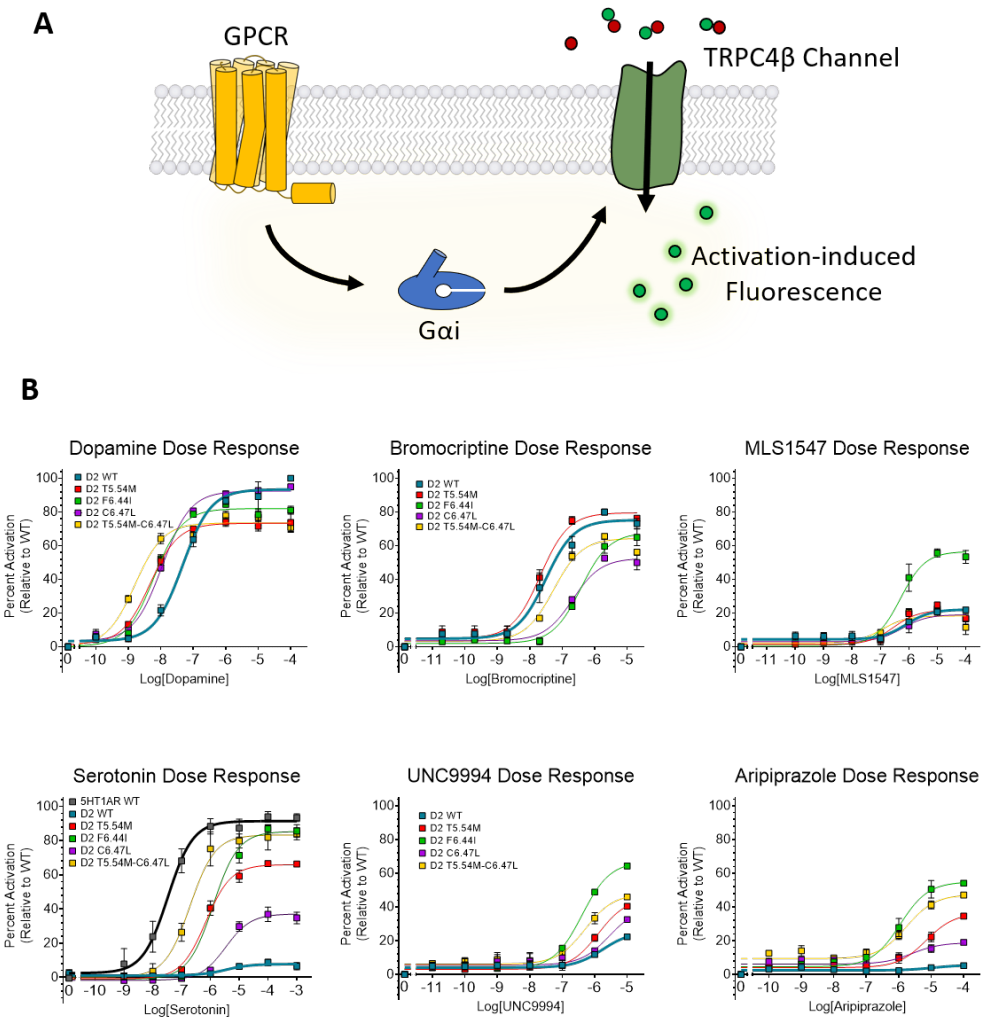


Figure 4.4 – Dopamine D2 dose responses with different ligands looking at G protein activation. **A.** Schematic of the TRP channel activation assay. **B.** Simulated dose response curves are based on experimental Emax and EC50 values. All maximum responses are shown relative to dopamine-induced D2 WT.

G protein Activation Experimental Data

	Dopamine								Bromocriptine							
D2 Variant	Emax	S.E.M.	n	p	EC50	S.E.M.	n	p	Emax	S.E.M.	n	p	EC50	S.E.M.	n	p
WT	1.00	N/A	3		-7.42	0.12	4		0.80	0.02	2		-7.51	0.21	4	
T5.54M	0.82*	0.09	3	<0.0001	-8.70*	0.16	4	<0.0001	0.80	0.02	2	1.0000	-7.68*	0.37	2	0.0470
F6.44I	0.85*	0.05	3	<0.0001	-8.45*	0.09	5	<0.0001	0.76	0.01	2	0.0028	-6.42*	0.07	2	<0.0001
C6.47L	0.92	0.11	2	0.3681	-8.18*	0.11	3	<0.0001	0.50		1	N/A	-6.68*	0.08	2	<0.0001
T5.54M-C6.47L	0.83*	0.05	3	<0.0001	-9.15*	0.13	4	<0.0001	0.63		1	N/A	-7.27*	0.01	3	<0.0001

	Aripiprazole								MLS1547							
	Emax	S.E.M.	n	p	EC50	S.E.M.	n	p	Emax	S.E.M.	n	p	EC50	S.E.M.	n	p
WT	0.05	0.02	3		-6.60	0.21	4		0.16	0.04	3		-6.60	0.19	4	
T5.54M	0.39*	0.04	2	0.0033	-6.00	0.26	4	0.1096	0.22	0.01	3	0.2672	-6.98*	0.42	3	0.0292
F6.44I	0.54*	0.01	3	<0.0001	-5.96	0.24	3	0.1015	0.41*	0.08	3	0.0491	-6.74	0.20	4	0.0834
C6.47L	0.11	0.05	2	0.2775	-7.32	0.52	2	0.1832	0.13	0.05	2	0.6697	-6.55	N/A	1	0.6204
T5.54M-C6.47L	0.45*	0.07	3	0.0053	-6.42	0.18	5	0.5338	0.22	0.03	3	0.2964	-7.37*	0.24	3	0.0055

	UNC9994							
	Emax	S.E.M.	n	p	EC50	S.E.M.	n	p
WT	0.24	0.02	3		-6.01	0.22	3	
T5.54M	0.48*	0.02	3	0.0011	-6.56*	0.33	3	0.2378
F6.44I	0.79*	0.07	3	0.0016	-6.50	0.03	3	<0.0001
C6.47L	0.35*	0.01	2	0.0264	-5.76	0.19	2	0.5333
T5.54M-C6.47L	0.58*	0.08	3	0.0146	-6.67*	0.12	3	0.0037

Table 4.1 – Dopamine D2 experimental dose response E_{max} and Log(EC50) values for G protein activation. * $p < 0.05$ using the F test.

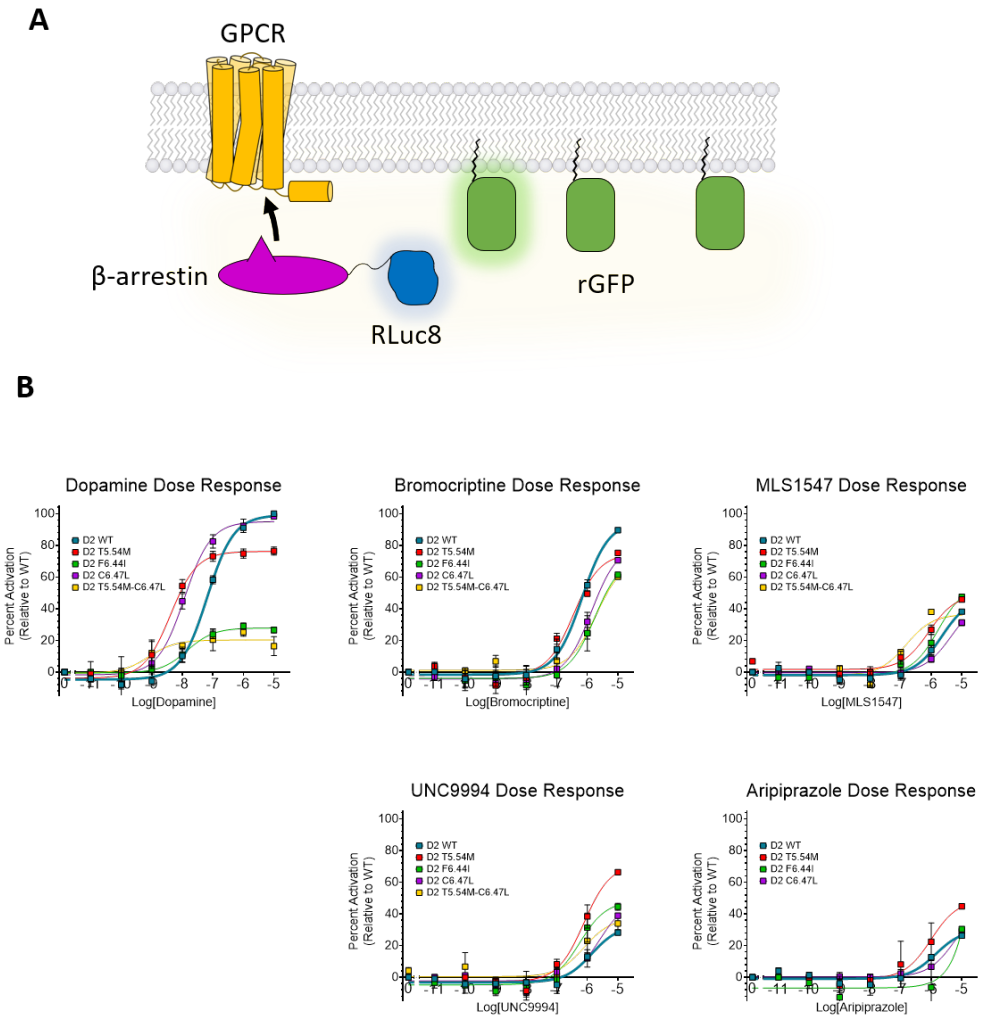


Figure 4.5 – Dopamine D2 dose responses with different ligands looking at β-arrestin2 recruitment. **A.** Schematic of the enhanced bystander BRET β-arrestin assay. **B.** Simulated dose response curves are based on experimental Emax and EC50 values. All maximum responses are shown relative to dopamine-induced D2 WT.

β-Arrestin Recruitment Experimental Data

	Dopamine							Bromocriptine						
D2 Variant	Emax	S.E.M.	p	EC50	S.E.M.	n	p	Emax	S.E.M.	p	EC50	S.E.M.	n	p
WT	0.99	0.02		-7.19	0.05	6		0.95	0.06		-6.18	0.11	6	
T5.54M	0.76*	0.03	<0.0001	-8.39*	0.14	4	<0.0001	0.75*	0.04	0.0006	-6.45	0.12	5	0.063
F6.44I	0.28*	0.01	<0.0001	-7.84*	0.15	3	0.0005	0.73	0.09	0.0820	-5.75*	0.19	4	0.025
C6.47L	0.95	0.02	0.0925	-7.93*	0.08	4	<0.0001	0.81*	0.04	0.0024	-5.87*	0.07	4	3E-04
T5.54M-C6.47L	0.20*	0.03	<0.0001	-9.09*	0.47	2	0.0075	0.72	0.09	0.0939	-5.71*	0.20	2	0.026

	Aripiprazole							MLS1547						
	Emax	S.E.M.	p	EC50	S.E.M.	n	p	Emax	S.E.M.	p	EC50	S.E.M.	n	p
WT	0.30	0.06		-5.86	0.29	2		0.47	0.06		-5.66	0.17	5	
T5.54M	0.49	0.09	N/A	-6.01	N/A	1	N/A	0.49	0.03	0.5158	-6.08*	0.12	4	0.0022
F6.44I	N.D.	N/A	N/A	-2.33	N/A	1	N/A	0.56	0.06	0.1343	-5.76	0.17	3	0.5489
C6.47L	0.47	0.06	N/A	-5.22	N/A	1	N/A	0.45	0.08	0.8411	-5.35	0.20	3	0.0939
T5.54M-C6.47L	N.D.	N/A	N/A	N.D.	N/A		N/A	0.37	0.06	0.1355	-6.82*	0.35	1	0.0207

	UNC9994							Serotonin						
	Emax	S.E.M.	p	EC50	S.E.M.	n	p	Emax	S.E.M.	p	EC50	S.E.M.	n	p
WT	0.33	0.04		-5.86	0.1773	3		0.09	N/A		-5.13	N/A	1	N/A
T5.54M	0.71*	0.05	<0.0001	-6.11*	0.1168	3	0.0456	0.47	N/A		-5.87	N/A	1	N/A
F6.44I	0.48*	0.06	0.0061	-6.25	0.2189	3	0.0522	0.42	N/A		-5.74	N/A	1	N/A
C6.47L	0.49*	0.07	0.0120	-5.61	0.2077	2	0.2204	0.34	N/A		-5.29	N/A	1	N/A
T5.54M-C6.47L	0.37	0.06	0.4324	-6.11	0.3044	3	0.3707	N.D	N/A		N.D.	N/A		N/A

Table 4.2 – Dopamine D2 experimental dose response E_{max} and Log(EC50) values for β-arrestin2 recruitment. * $p < 0.05$ using the F test.

Surprisingly, we see the reverse effect with the ligand bromocriptine for most D2 variants compared to WT. Mutations present on TM6 display a display a higher EC50 than WT and a lower maximum response. C6.47L is the most extreme case, with an increase of 0.8 in Log(EC50), and a 40% drop in maximum response. In contrast, when the same mutant is stimulated with dopamine we see a 0.7 decrease in Log(EC50) and no change in maximum response. Unlike with other ligands, T5.54M exhibits no change in EC50 when stimulated with bromocriptine. This was the only design tested that is not located on TM6.

Shifts in EC50 were fairly consistent between the G protein and β-arrestin activation pathways. Where we see major differences are in the maximum responses, especially for F6.44I and T5.54M-C6.47L. Both of these designs display a poor ability to recruit β-arrestin, with less than 25% activity relative to WT when stimulated with dopamine. Yet with aripiprazole and serotonin F6.44I shows stronger activation than WT, and comparable activation with UNC9994 and MLS1547. Both for G protein activation and β-arrestin recruitment, T5.54M-C6.47L shows more similar patterns of signaling to F6.44I than to its single-mutant constituents. The overall activation pattern of the mutants is fairly complex, with difficult-to-decipher relationships. Figure 4.6 shows shifts in allosteric efficacies derived from TRP channel data. One common factor is the mutants' consistently worse activation and sensitivity to bromocriptine in both pathways. This demonstrates an interesting ability of allosteric mutations to differentially affect the efficacy and potency of distinct drugs.

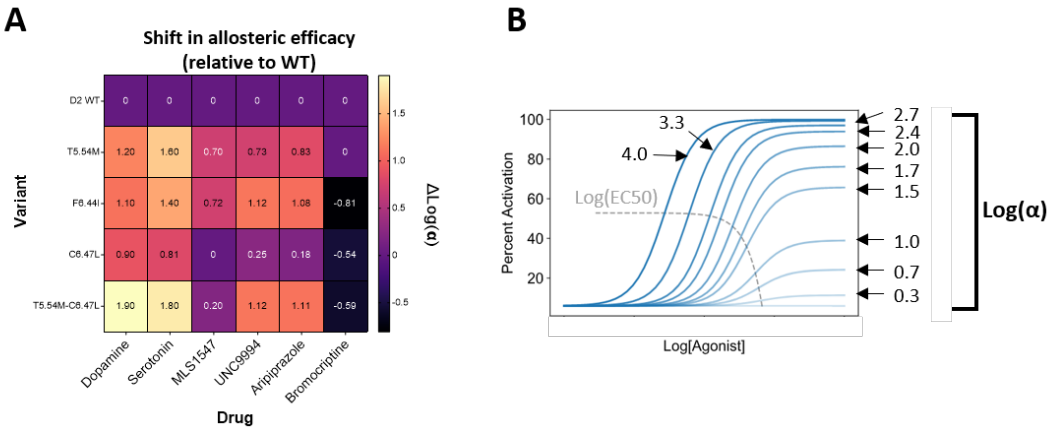


Figure 4.6 – Experimentally determined shifts in allosteric efficacy for each agonist-variant pair from G protein activation experiments. **A** Heatmap of allosteric efficacy α shifts. **B** Theoretical scale of $\log(\alpha)$ values.

We were unable to reproduce the bias seen with the ligands MLS1547 and UNC9994. With D2 WT we see identical efficacies for G protein activation and β-arrestin recruitment for both UNC9994 (25%) and MLS1547 (35%) relative to dopamine. Other groups have also shown contradicting results to the initial publications using different assays [196]. These compounds, it seems, may only show bias under specific conditions. Allen et al. used the cAMP pGloSensor assay to detect ligand-induced decreases in cAMP production, whereas we detected Gi

activation via measuring the activation of the Gi-activated TRPC4B channel. There exists the possbility that somehow Gi is activated in parallel to other effectors, resulting in a net zero change in cAMP production and simultaneous activation of the TRP channel Gi effector.

In addition to testing previously reported agonists of the dopamine D2 receptor, we also tested to reported antagonists which are also considered agonists for serotonin receptors. Given the increase in response we observed with serotonin, we reasoned that such ligands result in such a small shift in ΔG upon ligand binding that its effect might not be noticeable in the WT receptor, but may in the gain-of-function receptors. We selected ligands asenapine and clozapine for this test, both of which remained antagonists for all D2 variants (data not shown).

4.4.2 Differences in contacts between dopamine and bromocriptine

There is a stark difference in the size of dopamine and bromocriptine, and consequently in the contacts the two agonists are able to make with the receptor. A comparison is shown in Figure 4.7. The biggest difference is bromocriptine's ability to make contact with both the typical conserved orthosteric contacts on TM5, TM6, and TM7 and ECL2 simultaneously with the linked heterocyclic moiety. This group creates non-polar contacts with ECL2 and TM7, likely making this region of the receptor more rigid. The protonated form of bromocriptine can form the conserved salt bridge present in all aminergic GPCRs with D3.32 through a tertiary nitrogen. Based on our MD simulations (not shown), deprotonated ligand species unable to form this salt bridge tend to be less stable within the orthosteric site, likely translating to a higher experimental K_{off}. Bromocriptine only has one potential hydrogen bond donor near TM5 with which it can potentially bind to S5.46. Dopamine on the other hand has two and can engage with multiple of the conserved serines S5.42, S5.43, and S5.46. The bromide of bromocriptine also reaches deeper into the binding pocket of D2, making additional stabilizing contacts with I3.40, an allosteric hub. Which of these factors results in the differential shift in signaling between WT and designed receptors is still being investigated.

4.4.3 Altered ligand-induced dynamics between WT and designed receptors

We are in the process of analyzing molecular dynamics trajectories from nanosecond simulations performed on D2 WT and T5.54M-C6.47L, each either bound to dopamine or bromocriptine. The recently-solved active-state structure of D2 bound to bromocriptine[193] and a dopamine-docked model using Iphold[165] were used as the starting structures. A cursory examination of the trajectories reveals that due to the large size and rigidity of bromocriptine, the additional contacts with ECL2 attenuate the fluctuations in movement relative to the trajectories performed with the dopamine-bound models. Allosteric residues have previously been predicted to be present in ECL2[197] and bromocriptine's perturbation of this region may result in the observed effects. Based on the current experimental results it would seem there may be antagonistic allosteric effects between ECL2 and our introduced mutations.

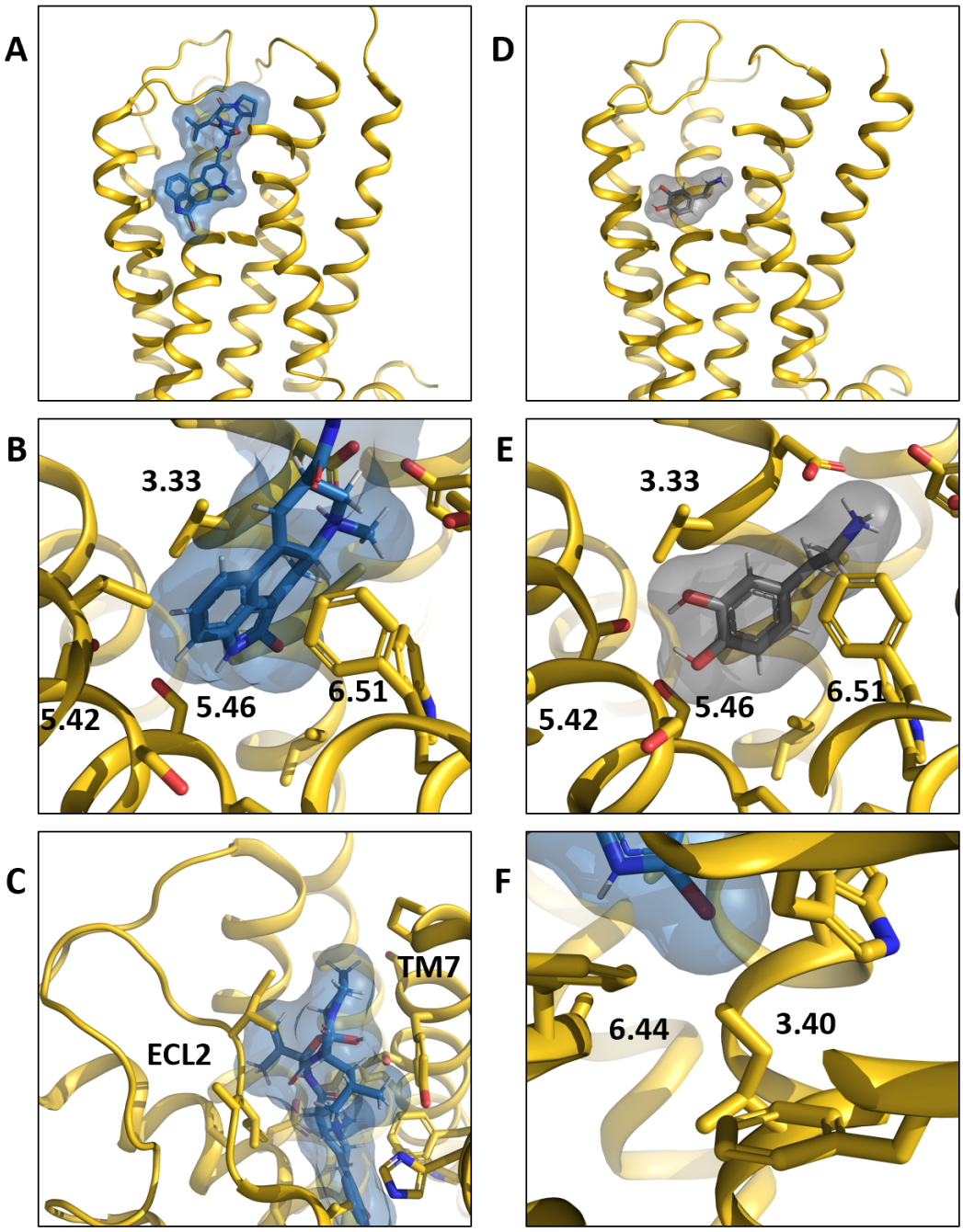


Figure 4.7 – Differences in bromocriptine and dopamine contacts in the dopamine D2 receptor. **A-C. and F.** Contacts made between bromocriptine and the active D2R structure 6VMS[193]. Bromocriptine is unable to form a hydrogen bond with S5.42 **B.**, though most non-polar contacts are maintained with TM3 and TM6. Bromocriptine makes additional non-polar contacts **C.** with ECL2 and TM7 which dopamine can not. The proximity of the bromide group to the allosteric hubs I3.40 and F6.44 (**F.**) may alter dynamics.

4.5 Discussion

Allosteric mutations within the dopamine D2 receptor can result in complex changes in signaling. These changes can manifest either as gain-of-function or loss-of-function signaling properties depending on the agonist. In our case this effect is irrespective of the downstream effector pathway monitored. Despite this complexity, there are a number of interesting patterns: 1) Tested partial agonists remained partial agonists. The highest level of relative activation seen was with UNC9994-stimulated D2 F6.44I, which resulted in a maximum response of 75% relative to dopamine-stimulated D2 WT, compared to the D2 WT E_{max} of 25% for G protein activation. This is somewhat surprising given the E_{max} increase seen with serotonin from an E_{max} of 5% with D2 WT to near 100% with D2 F6.44I. This would indicate that either these ligands stabilizes an alternative and less effective active state than dopamine or serotonin, or they are simply unable to stabilize the canonical active state to the same extent. 2) Unsurprisingly, structurally similar ligands show the same pattern of changes in signaling. Changes in E_{max} (and therefore allosteric efficacies) show similar patterns for MLS1547, UNC9994, and aripiprazole that is different from the other groups/clusters (those being the dopamine-serotonin group and the bromocriptine group). 3) Divergent changes in allosteric efficacies between ligands is possible. We clearly show that our variants increase allosteric efficacy for some ligands, while simultanesouly decreasing it for others. What structural or dynamic changes result in this effect is still unclear, however.

Relation to the allosteric two-state model

Our current results indicate that the simplest form of the allosteric two-state model is insufficient to explain the agonist effects seen with allosteric efficacy-modifying mutations. According to the allosteric two-state model, mutations increasing allosteric efficacy should be independent of agonist used, as long as the agonists bind to the same ligand binding site. Given the size difference between dopamine and bromocriptine, this may not hold. In this case, bromocriptine could be considered to be binding to 2 separate ligand binding sites simultaneously. If this is the case, this second binding site would allosterically antagonize the typical first binding site (i.e. the orthosteric site). However if we consider dopamine and bromocriptine still binding to a single ligand binding site, then it seems the allosteric efficacy α in the ATMS should be split into two different components: a receptor term and a ligand term. The receptor term would describe the inherent allosteric potential of the receptor, and the ligand term the effectiveness of the ligand at triggering this allosteric response. Which of these options is more appropriate is still to be determined by more experimental data.

Next steps

Currently we are considering two experimental approaches to determine what constituent of bromocriptine is responsible for the divergent shift in EC50 in relation to dopamine. The first approach involves the experimental testing of additional ligands analogous to the lysergamide

moiety of bromocriptine which are unable to contact ECL2. If we see the same effect as with bromocriptine, we can rule out ECL2 playing a role. On the other hand, the disappearance of this effect would confirm our suspicions and demonstrate the allosteric effect of ECL2. Another possibility is the involvement of the buried bromide group contacting an allosteric hub and altering signaling. The same aforementioned idea can be applied to confirm or refute this hypothesis. A number of ligands could be used in both of these cases. To rule out the involvement ECL2, we could generate dose response curves with bromolysergamide. If we wanted to rule out the involvement of the bromide, we could test with dihydro-ergotamine.

But the most convenient direction currently is to analyze differences in MD trajectories between dopamine and bromocriptine-bound models. If there are obvious differences in the dynamics of certain regions of the receptor, we can focus on this region. In lieu of the aforementioned ligand testing, a more exciting approach is to discover mutations *in silico* which, when combined with our current receptor designs, reverses the loss in function that is observed with bromocriptine. This would then be followed by experimental validation.

Opportunities in protein design

Introducing allosteric mutations which differentially alter the potency and efficacy of distinct ligands could be a valuable tool once this phenomenon is understood. If two ligands activated a receptor through separate or partially overlapping allosteric pathways, having the ability to selectively tune the allosteric effect of each ligand would increase the design space available. Not only would be possible to modify the ligand binding site, but also distant regions which are coupled to it.

Relation to population diversity and drug effects

GPCRs are the largest class of drug targets. Additionally, thousands of receptor variants have been revealed by genomics data, many of these mutations altering an individual's response to drugs[198]. Here we show how allosteric gain-of-function mutations can result in more promiscuous signaling, responding to previously-inert (or close to inert) ligands. It is interesting to consider what effect such mutations in off-target receptors (i.e. not the intended drug target) may have in clinical applications.

5 Modulating T-cell Activation by Rewiring the Adenosine A2A Receptor Pathway

D. Keri[1], M.J.A. Laval[1], P. Barth[1]

1 Swiss Federal Institute of Technology (EPFL), Institute of Bioengineering, Lausanne, Switzerland.

Research in progress

Author Contributions: PB and DK designed study and experiments. DK created chimeric G proteins. DK performed TRP assay experiments. ML performed cAMP experiments. DK and ML experimentally created designed G protein DNA constructs. DK, ML and PB analyzed the data. DK performed computational work. DK wrote chapter.

5.1 Abstract

Certain solid tumors produce adenosine to create an immunosuppressive tumor microenvironment. One of these immunosuppresinve factors is adenosine, which acts through the adenosine A2A receptor (A2AR) to generate intracellular cyclic AMP (cAMP), leading to T cell inhibition. Here, we attempt to re-wire the adenosine pathway to signal through the G proteins Gi and Gq rather than Gs, resulting in the activation of T cells in the presence of adenosine. When successful, this approach could be integrated into existing chimeric antigen receptor T (CAR T) cell therapies to overcome the artificial inhibitory environment created by tumors. Currently we are in the process of experimentally validating *in silico* designs and optimizing experimental protocols to measure A2AR activity.

5.2 Introduction

A relatively recent and promising immnunotherapy approach to treat certain cancers involves the use of T cells expressing chimeric antigen receptors (CAR T-cells) [199,200]. These CAR T-cells are created from a patient's own extracted T-cells transduced *ex vivo* typically via retroviruses to express a chimeric T-cell receptor capable of binding to a marker the patient's tumor is likely to express. Since their initial inception in the late 1980's CAR T-cell efficacies have been improved through more extensive modifications to the chimeric receptor, essentially fusing additional receptor's cytoplasmic domains to create dual or multi-signaling chimeric receptors. For example, the initial first generation CAR T-cells were ineffective without the simultaneous introduction of interleukin-2 (IL2). Second generation chimeric receptors included a C-term fusion of CD28 or CD137 to stimulate the endogenous production if IL2 upon activation. The current 4th generation CAR T-cell include receptors with a number of such fused receptor domains to co-stimulate numerous cell activation pathways.

CAR T-cell therapy has been successful at treating hematological malignancies such as various forms of leukemia and lymphoma [201], but it has been less successful at treating solid tumors. There are a number of unique environmental factors associated with solid tumors which likely hinder the activation of these cells, including a toxic environment with depleted nutrients, low pH, hypoxia, and high oxidants; suppressive factors; and T regulatory cells [202]. The suppressive factors present in the solid tumor micro-environment can include cytokines, prostaglandin E2 and adenosine; the latter two of which act through Gs-coupled G protein coupled receptors (GPCRs) to activate protein kinase A (PKA) to suppress T cell activation. Blocking PKA localization to the membrane, a requirement for its activity, resulted in the generation of the more efficacious CAR-RIAD T cells [203]. By rationally redesigning these inhibitory and stimulatory pathways, we should be able to further improve CAR T cell efficacy.

G proteins are involved in both the stimulation and inhibition of T cells. In the tumor micro-environment the high adenosine concentration activates the highly-expressed adenosine A2A receptor on T cells [204]. Activation of the Gs adenylyl cyclase-stimulatory G protein results in inhibition of T cells through PKA. Gαq activation, on the other hand, can result in the stimulation of T cells. The direct effect of Gαq activation is calcium release from the endoplasmic reticulum. Downstream effects include the phosphorylation of ERK1/2 and the stimulation of IL2 production [205]. It is these pathways we are planning to take advantage of in our designs.

Our first approach is to rewire the inhibitory signal initiated by adenosine to instead promote T cell activation while simultaneously minimizing or neutralizing this inhibitory signal. This can be achieved by creating chimeric G proteins that contain the Gs C-terminal helix 5 (H5), the primary G protein segment involved in GPCR coupling and selectivity, grafted onto the adenylyl cyclase-inhibiting G protein Gαi. The chimeric G protein, when expressed in cells, would both competitively bind to the active state receptor and non-competitively [206] inhibit adenylyl cyclase. In parallel we plan to create a corresponding Gαq chimera, which will serve as an additional T cell stimulatory signal. The concurrent activation of these two stimulatory

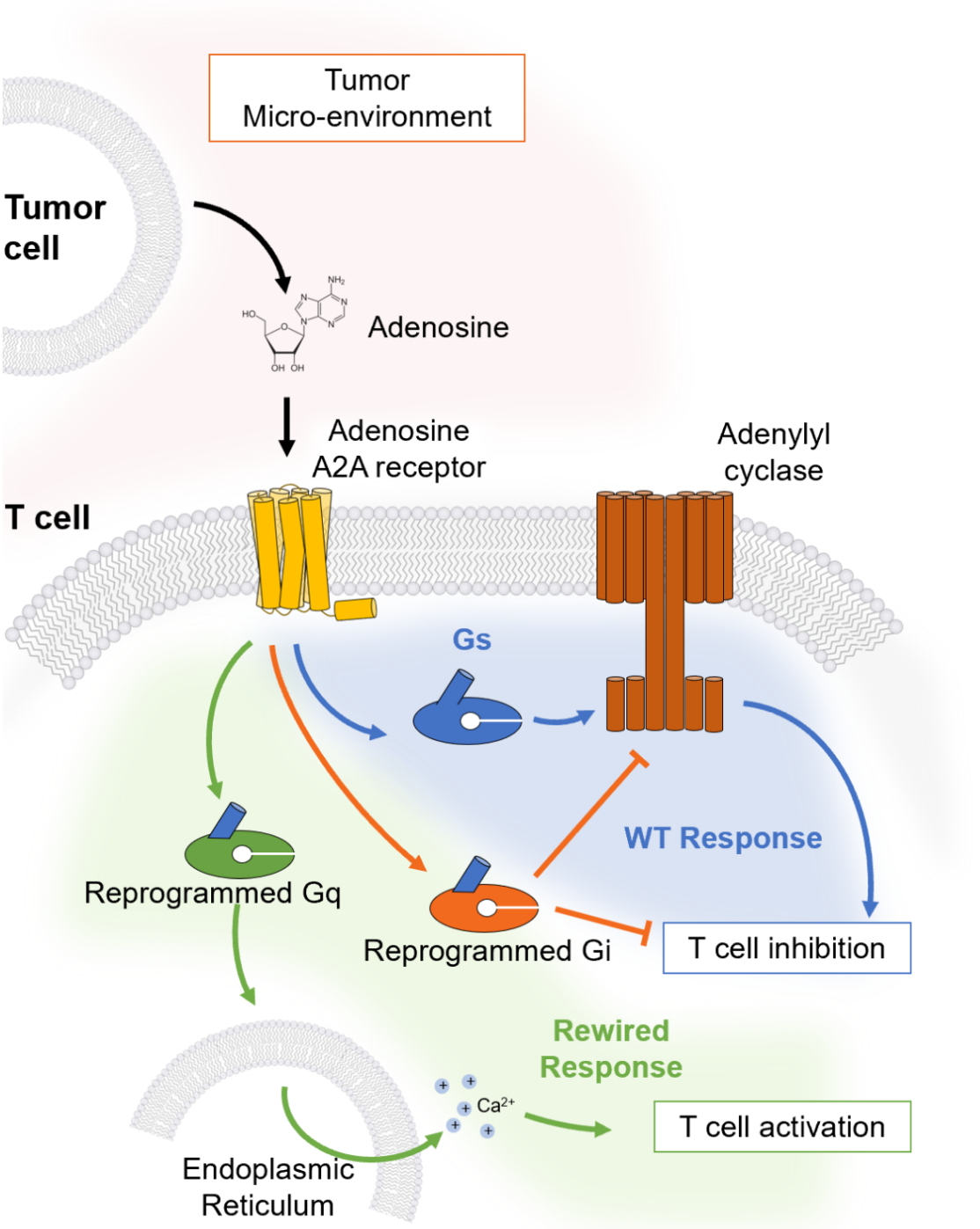


Figure 5.1 – Wildtype (blue) and reprogrammed (green) T-cell response in tumor micro-environment. Normal T cell response to adenosine results in inhibition (blue). The addition of chimeric A2AR-coupled Gi-s and Gq-s G proteins can be used to inhibit WT T cell inhibition in addition to inducing activation (green).

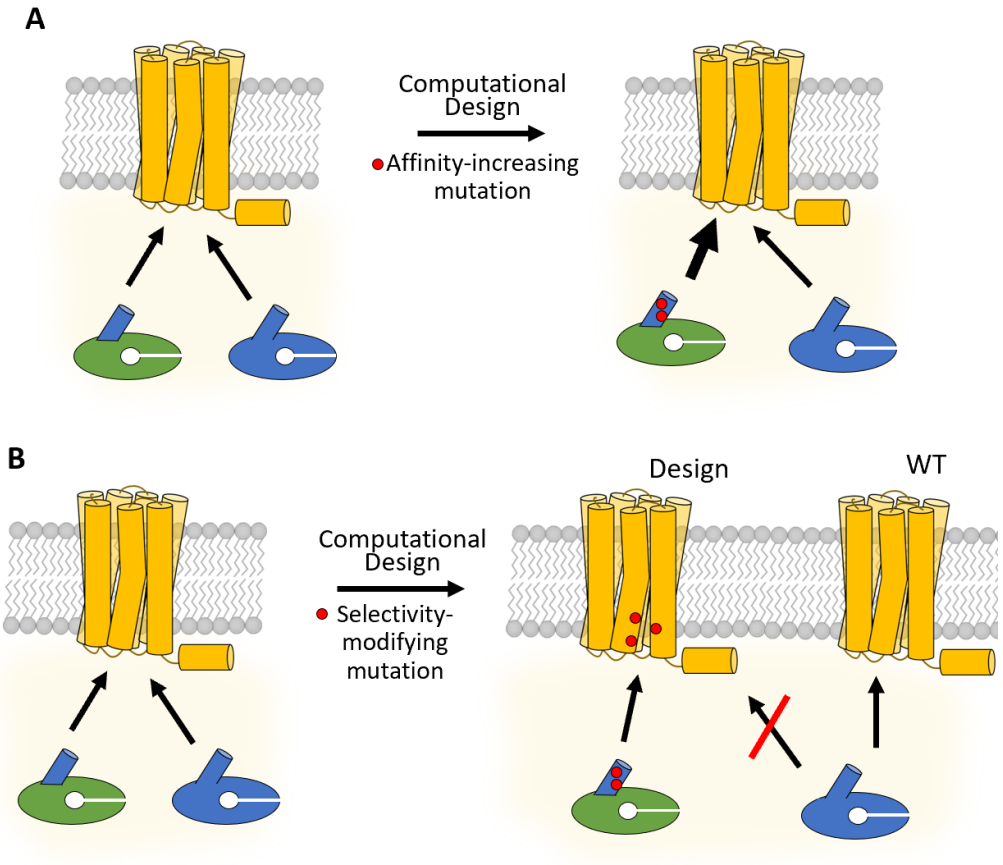


Figure 5.2 – Additional protein engineering techniques to favor reprogrammed T cell pathways. **A.** Use *in silico* mutagenesis to discover mutations in the Gs H5 leading to increased affinity towards A2AR. **B.** Creation of an orthogonal or pseudo-orthogonal pathway to bypass competition from the WT A2AR-Gs pathway. The pseudo-orthogonal option would allow the activation of the designed G protein by WT A2AR, while the fully orthogonal option would not.

pathways upon entry into the normally immunosuppresive, high-adenosine tumor microenvironment will hopefully be sufficient to generate a sustained T cell response. To favor the activation of these chimeras against their WT counterparts we also plan to computationally search for Gαs H5 mutations, via *in silico* mutagenesis, to increase affinity to A2AR. The effect of adenosine on the WT T cell and reprogrammed T cell pathways are shown in Figure 5.1.

The second approach takes advantage of our experience in creating orthogonal pathways for the dopamine D2 receptor[116]. By creating an orthogonal or pseudo-orthogonal pathway, we can ensure minimal interference from endogenous GPCRs and G proteins. By evolving *in silico* A2AR-G protein complexes we can favor mutations which block competitive state binding (e.g. WT receptor to mutant G protein or vice versa) and activation. A schematic is shown in Figure 5.2.

5.3 Methods

5.3.1 Computational

in silico deep-scanning mutagenesis

The initial template A2AR-Gs complex (PDB 6GDG[123] or 5G53[207]) is cleaned by removing all heteroatoms from the structure and is renumbered. The G protein is truncated to only include the last 16 residues within helix 5. The structure is then relaxed using RosettaMembrane with high constraints. An implicit membrane parameter is taken into account using a span file. Due to the low RMSDs, the lowest energy models are not clustered and the lowest energy model is used directly for all following protocols. A separate repack/relax protocol is run for each possible amino acid within helix 5 of the G-protein (16 positions x 20 possibilities). After ensuring all convergence for all simulations, the lowest energy model from each is used to calculate interface energies between the receptor and the G protein helix 5. Total energies for the receptor and G protein helix 5 are also calculated separately. Differences in energies for each chain (receptor, G protein H5), complex, and interface energies are calculated by subtracting the corresponding energies from WT, as demonstrated by these equations:

$$\Delta E_R = E_{R,mut} - E_{R,WT} \tag{5.1}$$

$$\Delta E_G = E_{G,mut} - E_{G,WT} \tag{5.2}$$

$$E_{int,WT} = E_{C,WT} - E_{G,WT} - E_{G,WT} \tag{5.3}$$

$$E_{int,WT} = E_{C,WT} - E_{G,WT} - E_{R,WT} \tag{5.4}$$

$$E_{int,mut} = E_{C,mut} - E_{G,mut} - E_{R,mut} \tag{5.5}$$

$$\Delta E_{int} = E_{int,mut} - E_{int,WT} \tag{5.6}$$

where ΔE_R refers to the chain in energy of the receptor, ΔE_G refers to the change in energy of the G protein H5, $E_{int,WT}$ and $E_{int,mut}$ are the interface energies of the WT and mutant receptor-G protein complex respectively, and ΔE_{int} is the change in the interface energy between WT and mutant complexes.

Mutations to test experimentally were chosen based on their ability to maximize ΔE_{int} while

minimizing changes to ΔE_G and ΔE_G.

Multi-state design of A2AR-Gs orthogonal pairs

Input structures for the Rosetta multistate design (MSD) protocol were prepared as mentioned above. Multistate design protocols were all run with generations of 200 and populations of 100. Crossover frequency was set to 2%. Positions within the G protein H5 were either allowed to mutate individually or in combination with adjacent H5 residues. In either case, receptor residues within 5Åof the H5 design site were allowed to mutate. To discover designs with the desirable properties, different fitness functions were considered.

We considered 5 different states in calculations: WT receptor-WT G protein (WW), WT receptor-mutant G protein (WM), mutant receptor-WT G protein (MW), mutant receptor-mutant G protein (MM, the design), and the inactive state of the receptor. As per the above fitness function, we emphasized the optimization of the interaction energy and stability of the MM state while destabilizing MW and WM competitive states. Parabolic functions were used to deter large changes to stability of the receptor in either active of inactive states.

Promising designs from MSD are then re-scored with RosettaMembrane with alternating repack and backbone minimization steps. Designs that are predicted to be destabilizing with RosettaMembrane are discarded. The rest are slated to be validated experimentally.

5.3.2 Experimental

cAMP LANCE assay

$4*10^4$ HEK-293T cells are seeded into white 96-well plates. are co-transfected with 100ng A2AR and 100ng Gs, both in in pcDNA3.1+ plasmid, using a reverse transfection protocol with 0.25μL Lipofectamine 2000 and DMEM with 10% FBS. 24 hours later the medium is replaced with 100μL DMEM with 5mg/mL BSA and 0.5mM Ro-20-1724 phosphodiesterase (PDE) inhibitor. Cells are then incubated for 30 minutes at room temperature. 100μL of agonist at various concentrations is added to each well. Cell are then incubated on ice for 20 minutes, before 30μL of lysis buffer is added. 6μL of lysate 6μL of anti-cAMP antibody are transferred to a new 96-well plate and incubated for 40 minutes in the dark. 12μL of detection buffer is added to the lysate-antibody mix and incubated again for 1 hour before reading on a FLexstation3 according to the manufacturer's instructions.

TRP assay

HEK-293 cells stably expressing TrpC4β channel were kindly provided by Dr. Michael X. Zhu and were transiently transfected with A2AR WT and various Gi-s chimeras as described below. FLIPR Membrane potential assays (Molecular Devices) were performed as previously

described[179]. Briefly, the assay utilizes a fluorescent probe that reacts to cations and is coupled with a non-permeable quencher. Activation of TrpC4β causes influx of cations leading to the translocation of the fluorescent probe away from the quencher, resulting in increased fluorescence. TrpC4β has been described to be downstream of Gi activation and serves as a reporter for Gi and Gi-s chimeras. Stable HEK-293 cells were cultured in DMEM supplemented with 10% FBS and 1% G418 (selection antibiotic) for use with FLIPR Membrane Potential (FMP) assay kits (Molecular Devices). On the day of transfection, cells stably expressing TrpC4β were plated (1.5x10^5 cells per well) onto 96-well clear, flat-bottom plates precoated with poly-D-lysine (Sigma). Cells were transfected with Lipofectamine 2000 and the appropriate D2 constructs using optimized DNA quantity for the FLIPR assay and grown for 24h prior to assaying. A duplicate plate was transfected and grown for assessment of receptor cell surface expression determined by ELISA using anti-HA antibody as described previously[180]. Eight hours after transfection, wells were treated with pertussis toxin to inhibit endogenous Gi function. Cells were treated with 20μM adenosine to measure Gi-s activation.

ELISA

HEK-293T cells are transfected with the same conditions used for the assay accompanying the ELISA. 24 hours post-transfection the media is removed from the wells, and cells are fixed to the bottom of the wells with the addition of 40μL of 4% para-formaldehyde (PFA) for 10 minutes. After removal of PFA, a blocking solution of 2% bovine serum albumin (BSA) in PBS is added for 1 hour. Blocking solution is removed and a 1:300 solution of the primary antibody is added in new BSA solution. Plates are then washed 3x with 150μL of PBS. Secondary antibody is added at a 1:2000 ratio in 40μL and cells are again incubated for 1 hour. Cells are once again washed 3x with 150μL of PBS. 90μL of a 1:1 mix of PICO luminescence mix is added to each well and incubated in the dark for 5 minutes. Finally, the luminescence of each well is read on a Flexstation3 with an integration time of 1 second.

5.4 Results

5.4.1 Gs designs and validation

Our initial goal was to computationally design a G protein with increased affinity to the adenosine A2AR, and create an orthogonal designed A2AR-G protein pair in parallel. To discover desirable Gs mutations to achieve this goal, we performed *in silico* deep-scanning mutagenesis on the C-term of Gs in an A2AR-Gs active complex, using a solved cryo-EM structure as the starting template[123]. The change in interface energy was used as a predictor in affinity changes. Mutations predicted to change either the active or inactive state stability of the G protein were filtered out.

From our *in silico* deep-scanning most predicted to increase affinity result in additional hydrogen bonds formed between the receptor and G protein. Figure 5.3 shows the effect of

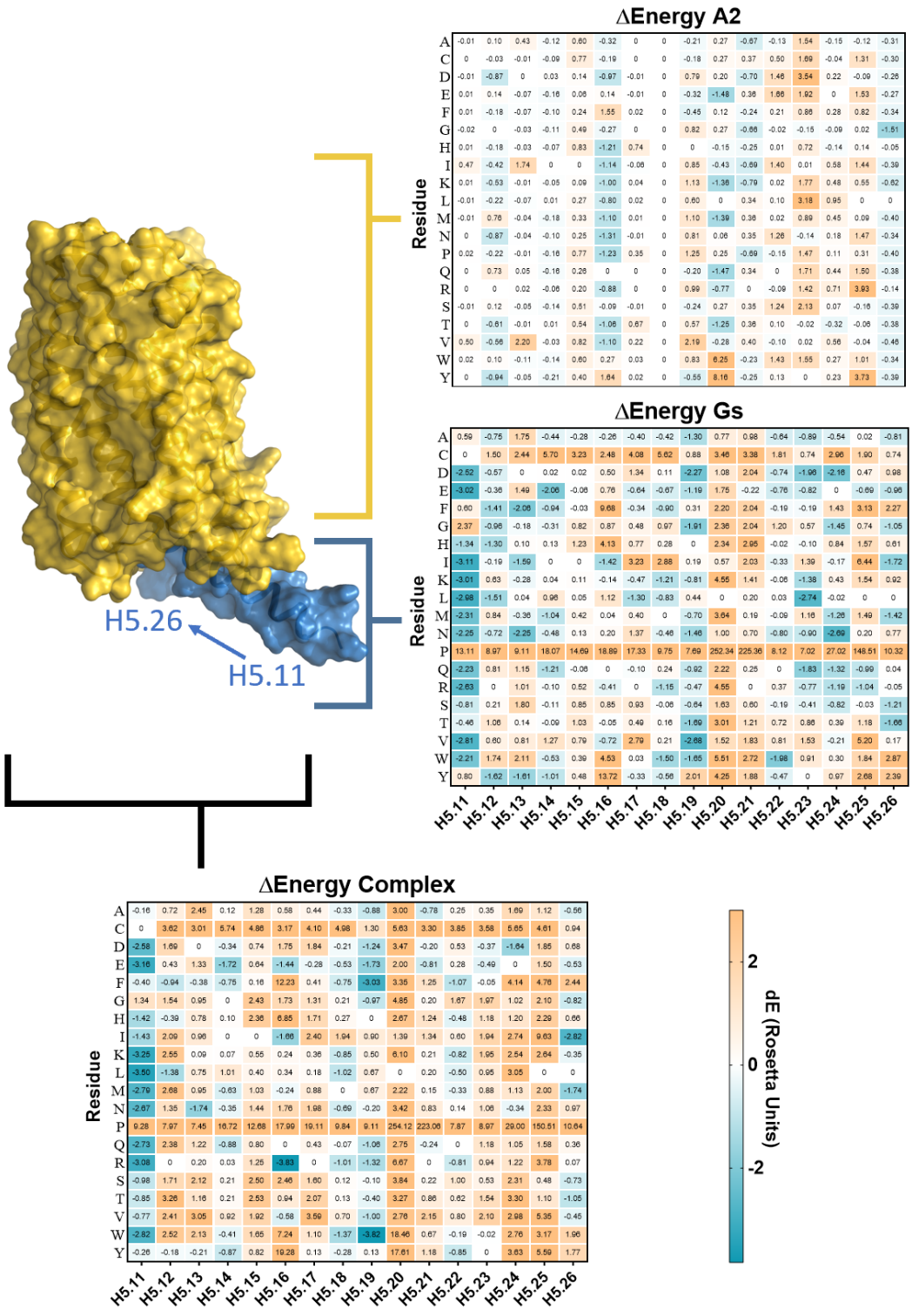


Figure 5.3 – *In silico* deep scanning mutagenesis of G protein C-term. The effect of single-point mutations in the G protein C-term are shown on the complex and receptor and G C-term independently.

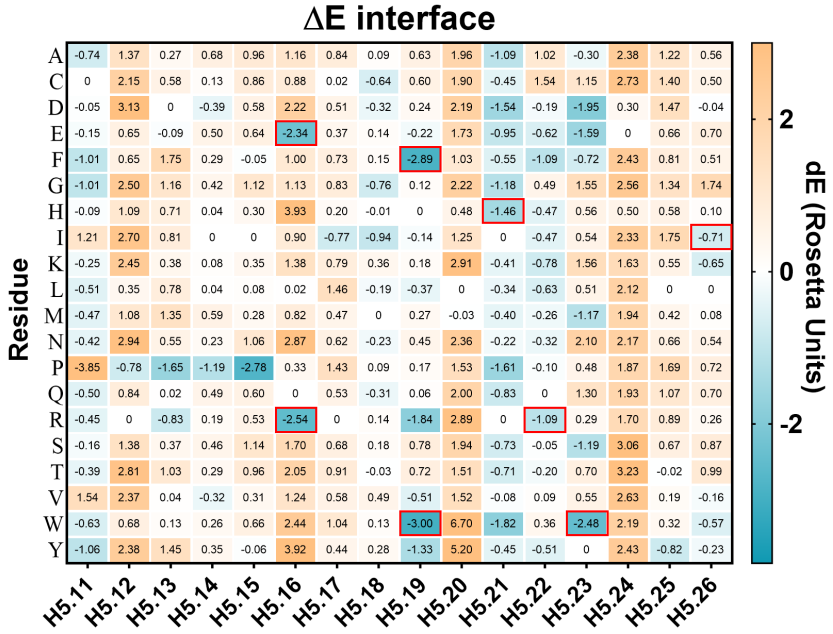


Figure 5.4 – Calculated changes in interface energy (dE_interface) from *In silico* deep scanning mutagenesis of G protein C-term. Values are calculated using equation 5.6.

mutations in the G protein C-term on the A2AR-Gs complex as a whole, and the receptor and G protein C-term separately. As expected, changes in stability of the receptor is relatively small compared to the G protein, usually either caused by losses in hydrogen bonds, loss of packing or additional strain in residue sidechain. G protein residues deeper inside the binding pocket of the receptor have a larger impact on receptor stability, while residues not in contact with the G protein show no changes in energy also as expected (i.e. residue 282).

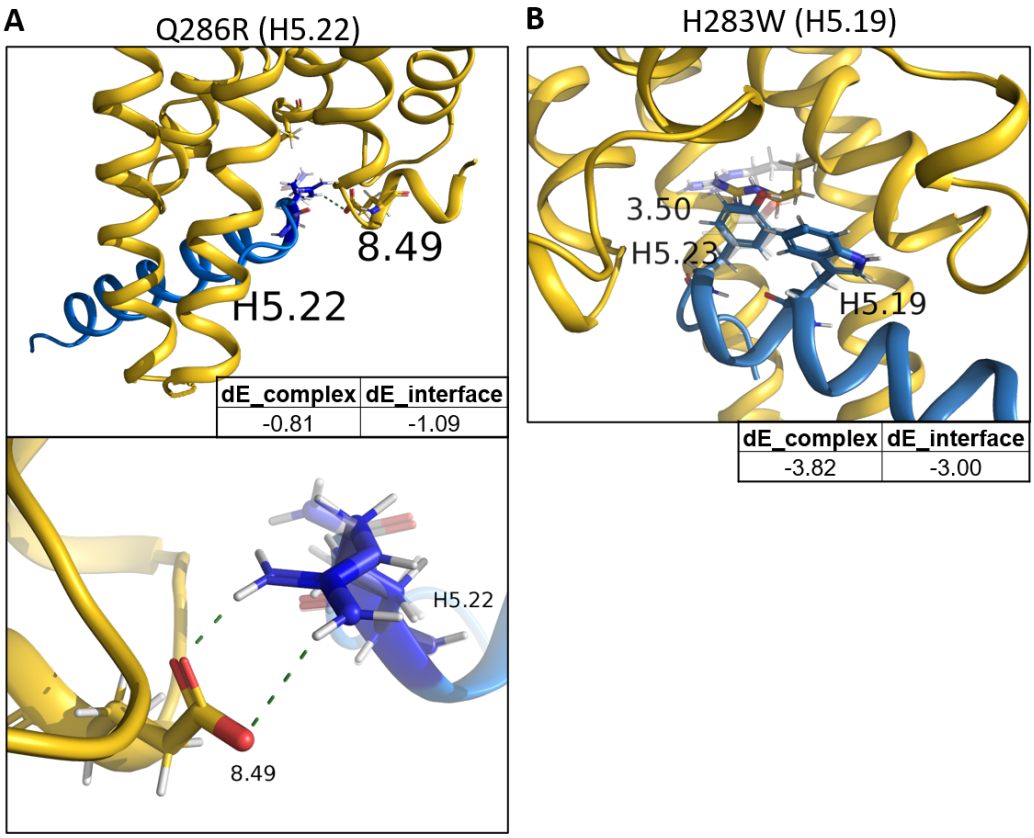


Figure 5.5 – A2AR-Gs designs and experimental validation. **A.** Q286R (H5.22) is predicted to create a salt bridge with D8.49 in helix 8 of A2AR. **B.** H283W (H5.19) pushes on H5.23 (Y) and forces it to make additional contacts with TM3. It also perturbs the conserved R3.50 residue from its typical active state position. dE_complex refers to the change in overall complex stability from WT in Rosetta Units. dE_interface refers to the change in interface energy from WT in Rosetta Units

	Bottom		Top		LogEC50	
Variant	Mean	SEM	Mean	SEM	Mean	SEM
WT	0.0668	0.0064	0.1247	0.0074	-5.5444	0.3119
Q280E **(H5.16)**	**0.0334***	0.0060	0.1224	0.0089	-5.3641	0.2073
Q280R **(H5.16)**	**0.0792***	0.0046	**0.1420***	0.0068	-5.3768	0.2243
H283R **(H5.19)**	**0.0926***	0.0042	**0.1460***	0.0082	-5.1245	0.2782
H283W **(H5.19)**	**0.0955***	0.0037	**0.1515***	0.0075	-5.0823	0.2387
R285H **(H5.21)**	0.0616	0.0068	0.1291	0.0081	-5.5911	0.2925
Q286R **(H5.22)**	0.0656	0.0131	0.1177	0.0063	-6.6454	0.4277
Q286E-Y287Q **(H5.22, H5.23)**	0.0628	0.0081	0.1255	0.0080	-5.7886	0.3221
Y287W **(H5.23)**	0.0779	0.0049	0.1296	0.0051	-5.7433	0.2428
L290I **(H5.26)**	0.0663	0.0063	0.1120	0.0111	-5.2001	0.4603

Table 5.1 – A2AR-Gs designs and experimental validation. Experiments were performed by Marie Laval. N=3 biological replicates with 3 technical replicates each. * indicates p<0.05 difference from WT.

Calculated changes in interface energy scores are shown in Figure 5.4. We chose a number of mutations to make based on these data and changes to overall complex stability based on Figure 5.3 to create experimentally, highlighted with red frames. Most chosen mutations result in hydrogen bonding between the two proteins with the exception of L290I (H5.23), which was predicted to increase packing with leucine 5.69 of the receptor. Of the tested mutations, we observed only one mutation, Q286R (H5.22) which had the desired EC_{50} shift effect. Q286R is predicted to form geometrically non-ideal hydrogen bonds with a glutamate residue (E8.49) in helix 8 of the receptor. Figure 5.5 demonstrates new contacts created between A2AR and designed G proteins Q286R and H283W.

Interestingly, in most (50%)cases we observed an increase in basal complex signaling rather than a decrease in EC50 as initially expected. Results are summarized in Table 5.1. The one exception to this is the Q286R design where we see roughly a tenfold decrease in EC50, although based on current results this shift is not statistically significant. Overall, we see a good correlation between change in A2AR-Gs stability and basal activity. This increase in

basal activity seems to be dependent on the presence of both the receptor and the designed G protein (data not shown), demonstrating this is not an increase in the basal activity of the G protein itself. Although how the expression of the designed G proteins is affected is currently unknown. In contrast, we see no correlation between interface energy scores and receptor-G protein EC_{50} as initially expected.

5.4.2 Gi-s Chimera designs

We created a number of Gi-s chimeric G protein Gi-s-11, Gi-s-13, Gi-s-15, and Gi-s-17. The last 11, 13, 15, or 17 residues of the Gi C-term have been replaced with that of Gs, respectively. We then determined the ability of each chimera to activate the adenosine A2AR receptor, as shown in Figure 5.6. Gi-s-11 exhibits low activity while Gi-s-17 exhibits high receptor-independent basal activity. Neither of these properties is desirable and these chimeras were excluded from further testing. The activities of Gi-s-13 and Gi-s-15 are similar although Gi-s-15 does exhibit a higher level of receptor-independent basal activity akin to Gi-s-17. The increased basal activities stem from progressive destabilization of the native inactive state of Gi as additional bulky and charged Gs residues are substituted into the Gi C-term, resulting in steric clashes and electrostatic repulsions.

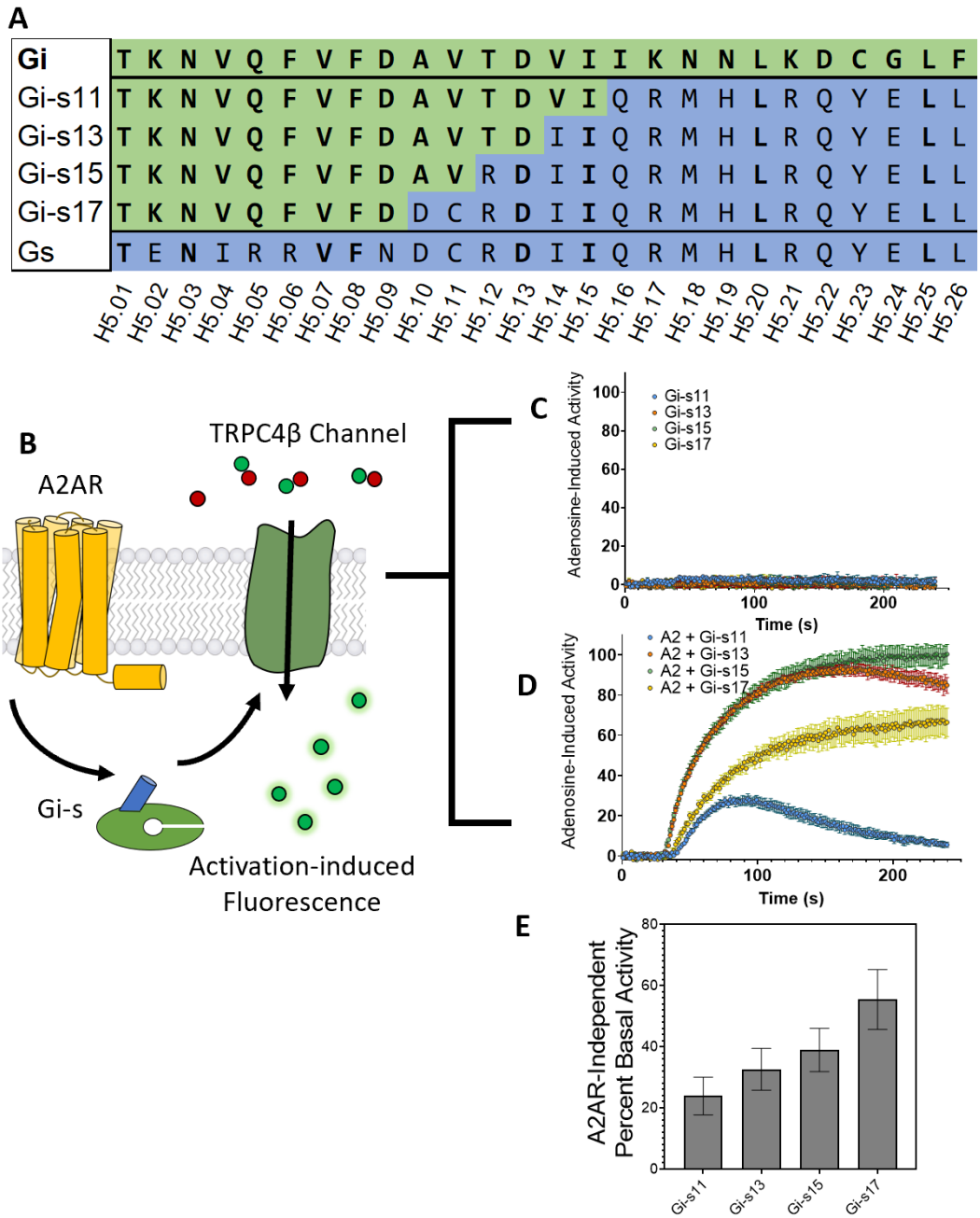


Figure 5.6 – Gi-s sequence comparison and chimera activation of TRP channel. **A.** Sequence comparison of the Gi, Gs, and the chimeric receptor C-term H5. **B.** Schematic of the TRP channel assay used to assess the level of activity of each chimera. **C.** No response is seen without exogenous transiently-transfected A2AR. **D.** Relative signals upon the co-transfection of A2AR and the Gi-s chimeras. **E.** Basal activities seen in the absence of co-transfected A2AR. One representative experiment is shown with three technical replicates.

5.4.3 Multi-state design of A2AR-Gs orthogonal pairs

We have previously successfully created a proof-of-concept receptor-effector orthogonal pair based on the dopamine D_2 receptor and Gi [116]. In the D2-Gi complex we found two hotspot regions which are particularly amenable to design. While we find similar hotspots, they are not as versatile. This proved untenable in most designs, where only the destabilization of only one of the competitive states was possible, as shown in Figure 5.7. Nevertheless, a number of viable designs emerged. The top design is shown in Figure 5.7C. After performing a second cycle of MSD using the most common orthogonal pairs from the first cycle, we obtained two designs predicted to destabilize both competitive states.

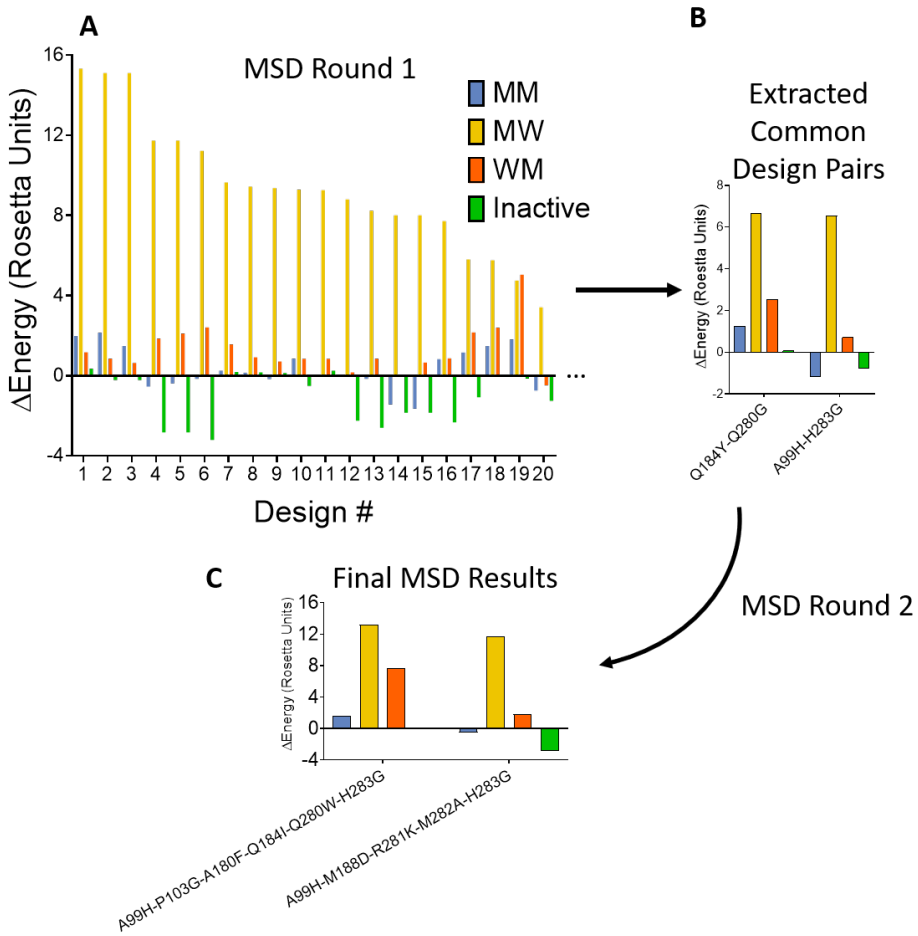


Figure 5.7 – A2AR-Gs orthogonal multistate design selection. **A.** Initial round of multi-state design. The competitive states MW and WM are selected against. Designs which destabilize these states and which do not shift stability of MM away from WT are the most fit. **B.** Most round one designs have one or two pairs of mutations responsible for most of the desired effects. These are extracted based on frequency and fitness and are implemented into the next round of multi-state design. **C.** The second round of multi-state design with implemented common pairs from round one. The weight of WM destabilization is increased in this round to favor WM-destabilizing mutations.

Both design 1 and 2 contain the A99H-H283G (A3.53H, H-H5.19-G) mutation pair which destabilizes the MW competitive state by forcing the two histidines into high energy rotamers which would result in steric clashes otherwise. The second round of MSD emphasized the destabilization of the WM competitive state by using a modified fitness function, increasing the weight of destabilizing WM. The additional mutations introduced in MSD cycle 2 include Q280W (Q-H5.16-W), a bulky residue which results in steric clashes in the WM competitive state with R183 (R5.67) and P103 in ICL3 in the WT receptor. Figure 5.8 shows the differences in binding of the various designs. Two additional mutations, Q184I (5.68) and A180F (5.64), serve to further stabilize Q280W by improving packing and forming π-π interactions. Design 2 only results in a slight destabilization to the WM competitive state, possibly by decreasing packing between R281 and M188. Overall, the success in this design is the additional stabilization of the MM state with the introduction of electrostatic interactions between R281K and M188D.

The experimental validation of th MSD designs are currently on hold while the Gs and Gi-s designs are tested.

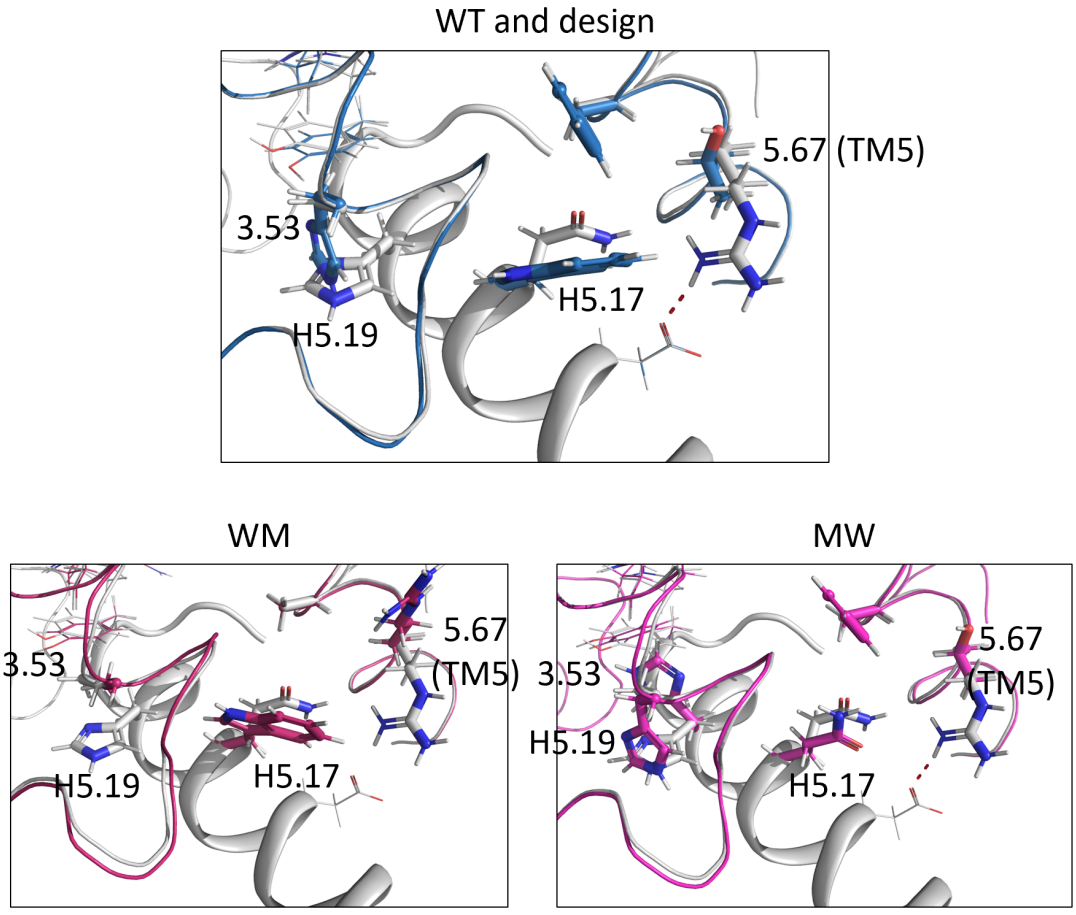


Figure 5.8 – Top A2AR-Gs multistate design structure. Top figure shows WT Gs H5 and receptor interface (white) and the design (MM) counterpart(blue). The histidine at H5.19 of the G protein is effectively swapped with at position 3.53 in the receptor.

5.5 Discussion

So far, we have experimentally tested 10 Gs design mutations predicted to increase G protein affinity towards the adenosine A2AR receptor. Of these 10, only one has shown a decrease in EC50 (Q286R), as seen in figure 5.5. Rather what we tend to see is an increase in basal activity in most cases. As we increase (predicted) affinity of the G protein to the A2AR, we also unavoidably increase the stability of the complex. How this affects the kinetics of G protein binding, activation and eventual dissociation from the receptor is unknown. It may be that receptor affinity to G protein may be evolutionarily well optimized for G protein turnover, which is slowed upon increases in receptor-G protein affinity.

Although technically challenging, direct measurements of G protein C-term affinity to A2AR would be possible via surface plasmon resonance (SPR). This could give us actual K_{on} and K_{off} values and provide useful information as to how C-term affinity actually relates to G protein turnover. The difficulty with this approach would be keeping the purified receptor from misfolding through the procedure and getting data before the receptor sample becomes unusable. To counter this problem, possible solutions are using thermostablized receptors, receptors in nanolipid discs, or a combination of the two.

Another challenge that we will have to deal with is the possibility that the predominant adenylyl cyclase expression in T cells is adenylyl cyclase 7 (AC7)[204,208], as shown by both transcriptomics and proteomics data. AC7 appears to not be inhibited by $G_{\alpha i}$[24,25], partially upending our design plans. AC7 is also activated by $G_{\beta\gamma}$, adding another layer of difficulty. We are currently in the process of designing chimeric Gq-s proteins (Gq core, Gs C-term) that are able to be activated through Gs-coupled receptors and result in Ca^{2+} influx into the cytosol, thus leading to T cell activation. So far we have only performed designs on the A2AR-G protein C-term pair, and not the full G protein (or at least the Gα subunit). Gα makes a number of additional contacts with the receptor which may be useful design positions. Going further, we are also considering creating *de novo* inhibitors of A2AR signaling which selectively bind to the active state receptor and compete with endogenous G proteins. How we can simultaneously inhibit cAMP production and activate T cells still needs to be decided. One possibility is the use of our orthogonal designs signaling through the Gq pathway and in tandem a *de novo* inhibitor can inhibit WT A2AR.

6 Conclusions

In chapter 3, we used the dopamine D2 receptor as a model system to test and validate our *in silico* approach to selectively redesign specific aspects of receptor signaling. Since at the start of the project experimentally-determined structures were not available, we created homology models. Next we took advantage of the metastable nature of GPCRs to selectively modify basal activity of dopamine D2 by using a multi-state design approach to differentially change the active and inactive receptor state stabilities. We reprogrammed specific microswitches in regions R1-R4 distributed throughout the receptor structure (R1 near ligand binding site, R2 in TM core on TM5 and 6, R3 in TM at the TM6 and 7 interface, and R4 near the NPxxY motif). In each case the active state of the receptor was stabilized by increasing the number of packing interactions, while introducing clashes in the inactive state. These 4 designs were tested experimentally to qualitatively match our predictions. The designs ranged in basal activity from 25-35% relative to full induction by ligand, compared to 6% of D2 WT. Moreover, we searched for experimental mutagenesis datasets which reported allosteric mutations that affect basal activity in different GPCRs. We found such datasets for the B2 adrenergic receptor, δ-opioid receptor, and rhodopsin. Collectively they contained 109 mutants. For each of these receptors we showed a significant quantitative correlation between percent basal activity (relative to fully induced activity) and the relative stability change between the active and inactive receptor states upon mutation. In this regard, our findings show a correlation similar to Kellog's[174] *in silico* prediction of thermostability-modifying mutations. Because of their ability to control receptor conformation, we termed the aforementioned microswitches "conformation selectors".

In addition to basal activity, we also reprogrammed the allosteric efficacy of the dopamine D2 receptor. GPCRs have been shown to be a loosely-coupled system, comprised of intra-domain modules such as the ligand binding (i.e. orthosteric) site, the G protein binding site, and the connector region which allows for the communication between the two former regions. As an agonist locks the ligand binding site into an active state, this information is relayed through the connector region to the G protein binding site, influencing its conformation to some extent. We find that we can predict and modify the signaling sensitivity of the receptor

by selectively analyzing the cooperative movements of residues along predicted, functionally-relevant, allosteric pathways. Using the allosteric "pipelines" predicted by Vaidehi[78], we performed *in silico* deep scanning mutagenesis on salient residues to discover mutations which might increase or decrease cooperative movement. We used normal mode analysis (NMA) to achieve this, in conjunction with Rosetta to ensure we do not alter receptor stability. This method has enabled us to discover 11 single-point mutants and 6 double-mutants, which significantly alter receptor sensitivity; in addition to 3 mutations which we correctly predict to have no effect. Overall the range in sensitivites varies from a $\Delta Log(EC_{50})$ of -1.7 to +1.5 relative to WT. According to the allosteric two-state model (ATSM), if we increase the allosteric efficacy of a receptor then there should be an increase in sensitivity to all agonists utilizing the same ligand binding site and allosteric pathway. To test this idea, we tested the most sensitive D2 variants with serotonin, a non-native and very weak D2 agonist. Interestingly, we saw significant increases in maximum responses for these D2 variants with D2 F6.44I reaching the same maximum response seen with the native serotonin receptor 5HT1AR. In comparison D2 WT reaches a relative maximum response of 5%. Further, we looked at how the combination of the single mutants behaved. Typically, combining gain-of-function mutations that are adjacent tends to result in no additive behavior in the best of cases. In other instances, combining such mutations results in complete abrogation of a response (which we successfully predicted qualitatively). Only the combination of two distant mutations resulted in an additive effect to create T5.54M-C6.47L, the most sensitive mutant. As these mutations do not change receptor basal activity but instead have a significant role in allosteric signaling, we termed them "allosteric propagators".

Finally, we looked at the sequence conservation of the two classes of allosteric residues and we find that they are not significantly more conserved than the average TM residue in class A receptors. This is in contrast to the highly conserved NPxxY, PIF, and DRY motifs known to be critical for GPCR signaling. Mutations typically in these regions can result in significant if not total reduction in signaling capability. This led us to propose that allosteric propagators and conformational selectors are responsible for the diversity seen in GPCRs as their effects can be tuned evolutionarily without a complete loss in signaling.

Overall, our work presented in this chapter sheds light on GPCR signaling determinants, which we can accurately tune using our *in silico* approach. Our method can be useful for future structure determination, drug design, and biosensor creation efforts, by tuning receptor sensitivity and/or receptor state stabilization.

In **chapter 4**, we continue to explore the properties of the dopamine D2 receptor variants we designed in chapter 3. We showed that the gain-of-function receptors were more sensitive to both dopamine and serotonin. Here we ask whether this effect extends to the β-arrestin pathway, and to what range of ligands are these receptors more sensitive. To answer these questions, we chose a number of structurally and functionally diverse set of ligands to test via both a G protein activation assay (TRP assay) and a β-arrestin recruitment assay (ebBRET). We decided to test a G protein-biased agonist (MLS1547), a β-arrestin-biased agonist (UNC9994),

and a number of unbiased agonists with unique structures (bromocriptine, aripiprazole). In most cases we see roughly the same pattern of change in sensitivity as we do with dopamine. However, ligands that are partial agonists (MLS1547, UNC9994, aripiprazole) remain partial agonists; unlike what we have seen with serotonin, which essentially becomes a full agonist for D2 F6.44I. This seems to indicate these ligands might actually stabilize an alternate conformation which is less able to bind G protein. Second, the gain-of-function variants actually exhibit a higher EC50 when stimulated with bromocriptine, essentially making them loss-of-function variants for this agonist. To our knowledge such allosteric mutations have not been reported before. We are currently in the process of determining the structural underpinning of the divergent effect of these mutations using MD simulations to guide our next experimental steps. Third, all mutations seem to be biased towards activating the G protein pathway rather than β-arrestin. This is not surprising, since the active state receptor used in our *in silico* predictions was a homology model based off of a G protein-bound template (3SN6). The conclusions so far point to a complex effect of these mutations rather than a simple universal increase or decrease in sensitivity or maximum response to different ligands.

As we increase the sensitivity of the dopamine D2 receptor with the introduction of allosteric mutations, we see a direct correlation with the receptor's sensitivity to other less efficacious agonists. This increase in promiscuity, which is mainly absent in the WT receptor, seems to suggest the dopamine D2 receptor may be evolutionarily tuned not only to maximize sensitivity to the native agonist but also to minimize noise from other potential agonists. How general this is for all GPCRs is not currently clear but is an interesting avenue to explore further.

Ever since its discovery, there has been an unending interest in allostery. Recently a number of photoswitchable allosteric proteins have been reported, increasing collective experience at engineering allosteric pathways[209,210]. Furthermore, the DeGrado group has very recently created a *de novo* allosterically-regulated protein, a phenol oxidase[211]. Our findings contribute to the creation, modulation, and understanding the behavior of allosteric pathways. Stemming from our and others' work, more designed allosteric proteins can be expected in the future.

Chapter 5 takes a different direction from the previous chapters. Here, we attempt to use our knowledge of GPCRs to improve the efficacy of chimeric antigen receptor T-cells (CARTs). One of the limitations of CARTs is their low efficacy against solid tumors due to their toxic and immunosuppresive tumor micro-enviroments (TMEs). One of the immunosuppresive molecules that is usually present in TMEs is adenosine, the endogenous ligand of adenosine receptors which are GPCRs. The predominant adenosine receptor in T cells is the Gs-coupled adenosine A2A receptor. Upon activation, A2AR triggers the downstream production of cAMP, leading to the inhibition of the T cell. Our goal with this project is the rewiring of the A2AR pathway to activate T-cells in the presence of adenosine.

The first step towards this goal is the creation of G protein chimeras. We have created a number of Gi-s chimeras by grafting different lengths of the Gs c-term, major contributor to GPCR selectivity, onto Gi, an adenylyl cyaclase inhibitor. We called these chimeras Gi-s11, Gi-s13,

Gi-s15, and Gi-s17; the number specifying the number of residues we grafted onto Gi from Gs. Gi-s11 showed poor activity, while Gi-s17 showed high basal activity due to destabilizing the inactive state of the chimera. Gi-s13 and Gi-s15 showed promising levels of activity. The next step in regards to chimeras will be to generate analogous Gq-s chimeras. The Gq pathway has also been shown to be involved in T-cell activation.

Next we looked for mutations which would increase the affinity of our chimeras towards A2AR to favor it against Gs WT by performing *in silico* deep scanning mutagenesis on the Gs c-term. We experimentally performed site-directed mutagenesis on initially the Gs WT to test these mutations. Rather than decreasing EC50 of the pathway as we had hoped, most of our designs result in increased basal activity as the A2AR-Gs complex is stabilized.

An alternate strategy we have been working towards is creating an orthogonal pathway to the A2AR-Gs pathway. This involved using a multistate design approach where we only allow the binding of designed G protein to designed receptor and disallow binding of WT counterparts to the designs. This strategy yielded a number of designs have yet to be tested. We are also considering introducing a *de novo* peptide to inhibit WT A2AR signaling, giving our designs sole control over adenosine sensing in the designed T-cells.

Overall, this is a promising approach to increasing CART efficacy, though a lot of experimental testing is still required. After we are confident in the ability of a certain technique to outcompete WT A2AR signaling, we plan to start *in vivo* testing in mice.

As pointed out previously, GPCRs are involved in a myriad of functions in different organs, tissues, and cells. We have shown with the work here and by others in our lab that we are able to computationally reprogram all functional regions of these receptors: the ligand binding site, the effector binding site, and the connector region mediating the allosteric communication between the two. The project presented here is our first foray into rewiring cellular signaling by modifying GPCR pathways and is likely not the last. As cell-based therapeutics mature in general, such cellular reprogramming will play an increasingly large role.

Bibliography

[1] Robert Fredriksson, Malin C Lagerström, Lars-Gustav Lundin, and Helgi B Schiöth. The g-protein-coupled receptors in the human genome form five main families. phylogenetic analysis, paralogon groups, and fingerprints. *Molecular pharmacology*, 63(6): 1256–1272, 2003.

[2] Albert J Kooistra, Stefan Mordalski, Gáspár Pándy-Szekeres, Mauricio Esguerra, Alibek Mamyrbekov, Christian Munk, György M Keserű, and David E Gloriam. Gpcrdb in 2021: integrating gpcr sequence, structure and function. *Nucleic Acids Research*, 49(D1): D335–D343, 2021.

[3] Michael C Lorenz, Xuewen Pan, Toshiaki Harashima, Maria E Cardenas, Yong Xue, Jeanne P Hirsch, and Joseph Heitman. The g protein-coupled receptor gpr1 is a nutrient sensor that regulates pseudohyphal differentiation in saccharomyces cerevisiae. *Genetics*, 154(2):609–622, 2000.

[4] W. I. Weis and B. K. Kobilka. The molecular basis of g protein-coupled receptor activation. *Annu Rev Biochem*, 87:897–919, 2018. ISSN 1545-4509 (Electronic) 0066-4154 (Linking). doi: 10.1146/annurev-biochem-060614-033910.

[5] L. Martinez-Munoz, J. M. Rodriguez-Frade, R. Barroso, C. O. S. Sorzano, J. A. Torreno-Pina, C. A. Santiago, C. Manzo, P. Lucas, E. M. Garcia-Cuesta, E. Gutierrez, L. Barrio, J. Vargas, G. Cascio, Y. R. Carrasco, F. Sanchez-Madrid, M. F. Garcia-Parajo, and M. Mellado. Separating actin-dependent chemokine receptor nanoclustering from dimerization indicates a role for clustering in cxcr4 signaling and function. *Mol Cell*, 70(1):106–119 e10, 2018. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2018.02.034.

[6] A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat, H. Hubner, X. P. Huang, M. F. Sassano, P. M. Giguere, S. Lober, D. Duan, G. Scherrer, B. K. Kobilka, P. Gmeiner, B. L. Roth, and B. K. Shoichet. Structure-based discovery of opioid analgesics with reduced side effects. *Nature*, 537(7619):185–+, 2016. ISSN 0028-0836. doi: 10.1038/nature19112.

[7] RW Butcher, GA Robison, JG Hardman, and EW Sutherland. The role of cyclic amp in hormone actions. *Advances in enzyme regulation*, 6:357–389, 1968.

[8] Richard AF Dixon, Brian K Kobilka, David J Strader, Jeffrey L Benovic, Henrik G Dohlman, Thomas Frielle, Mark A Bolanowski, Carl D Bennett, Elaine Rands, Ronald E Diehl, et al. Cloning of the gene and cdna for mammalian β-adrenergic receptor and homology with rhodopsin. *Nature*, 321(6065):75–79, 1986.

[9] Brian K Kobilka, RA Dixon, Thomas Frielle, Henrik G Dohlman, Mark A Bolanowski, Irving S Sigal, Teresa L Yang-Feng, Uta Francke, Marc G Caron, and Robert J Lefkowitz. cdna for the human beta 2-adrenergic receptor: a protein with multiple membrane-spanning domains and encoded by a gene whose chromosomal location is shared with that of the receptor for platelet-derived growth factor. *Proceedings of the National Academy of Sciences*, 84(1):46–50, 1987.

[10] B. K. Kobilka, T. S. Kobilka, K. Daniel, J. W. Regan, M. G. Caron, and R. J. Lefkowitz. Chimeric alpha 2-,beta 2-adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. *Science*, 240(4857):1310–6, 1988. ISSN 0036-8075 (Print) 0036-8075 (Linking).

[11] Joël Bockaert and Jean Philippe Pin. Molecular tinkering of g protein-coupled receptors: an evolutionary success. *The EMBO journal*, 18(7):1723–1729, 1999.

[12] G Glenn Gregorio, Matthieu Masureel, Daniel Hilger, Daniel S Terry, Manuel Juette, Hong Zhao, Zhou Zhou, Jose Manuel Perez-Aguilar, Maria Hauge, Signe Mathiasen, et al. Single-molecule analysis of ligand efficacy in β 2 ar–g-protein activation. *Nature*, 547 (7661):68–73, 2017.

[13] Laura M Wingler, Matthias Elgeti, Daniel Hilger, Naomi R Latorraca, Michael T Lerch, Dean P Staus, Ron O Dror, Brian K Kobilka, Wayne L Hubbell, and Robert J Lefkowitz. Angiotensin analogs with divergent bias stabilize distinct receptor conformations. *Cell*, 176(3):468–478, 2019.

[14] Shunsuke Imai, Tomoki Yokomizo, Yutaka Kofuku, Yutaro Shiraishi, Takumi Ueda, and Ichio Shimada. Structural equilibrium underlying ligand-dependent activation of β 2-adrenoreceptor. *Nature Chemical Biology*, pages 1–10, 2020.

[15] Asuka Inoue, Francesco Raimondi, Francois Marie Ngako Kadji, Gurdeep Singh, Takayuki Kishi, Akiharu Uwamizu, Yuki Ono, Yuji Shinjo, Satoru Ishida, Nadia Arang, et al. Illuminating g-protein-coupling selectivity of gpcrs. *Cell*, 177(7):1933–1947, 2019.

[16] Charlotte Avet, Arturo Mancini, Billy Breton, Christian Le Gouill, Alexander S. Hauser, Claire Normand, Hiroyuki Kobayashi, Florence Gross, Mireille Hogue, Viktoriya Lukasheva, Sandra Morissette, Eric Fauman, Jean-Philippe Fortin, Stephan Schann, Xavier Leroy, David E. Gloriam, and Michel Bouvier. Selectivity landscape of 100 therapeutically relevant gpcr profiled by an effector translocation-based bret platform. *bioRxiv*, 2020. doi: 10.1101/2020.04.20.052027. URL https://www.biorxiv.org/content/early/2020/04/24/2020.04.20.052027.

[17] M. Hillenbrand, C. Schori, J. Schoppe, and A. Pluckthun. Comprehensive analysis of heterotrimeric g-protein complex diversity and their interactions with gpcrs in solution. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11): E1181–E1190, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1417573112.

[18] S. M. Khan, J. Y. Sung, and T. E. Hebert. Gbetagamma subunits-different spaces, different faces. *Pharmacol Res*, 111:434–441, 2016. ISSN 1096-1186 (Electronic) 1043-6618 (Linking). doi: 10.1016/j.phrs.2016.06.026.

[19] N. Okashah, Q. Wan, S. Ghosh, M. Sandhu, A. Inoue, N. Vaidehi, and N. A. Lambert. Variable g protein determinants of gpcr coupling selectivity. *Proc Natl Acad Sci U S A*, 116(24):12054–12059, 2019. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1905993116.

[20] M. Sandhu, A. M. Touma, M. Dysthe, F. Sadler, S. Sivaramakrishnan, and N. Vaidehi. Conformational plasticity of the intracellular cavity of gpcr-g-protein complexes leads to g-protein promiscuity and selectivity. *Proc Natl Acad Sci U S A*, 116(24):11956–11965, 2019. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1820944116.

[21] L. J. Santiago and R. Abrol. Understanding g protein selectivity of muscarinic acetylcholine receptors using computational methods. *Int J Mol Sci*, 20(21), 2019. ISSN 1422-0067 (Electronic) 1422-0067 (Linking). doi: 10.3390/ijms20215290.

[22] M. Seyedabadi, M. H. Ghahremani, and P. R. Albert. Biased signaling of g protein coupled receptors (gpcrs): Molecular determinants of gpcr/transducer selectivity and therapeutic potential. *Pharmacol Ther*, 200:148–178, 2019. ISSN 1879-016X (Electronic) 0163-7258 (Linking). doi: 10.1016/j.pharmthera.2019.05.006.

[23] A. Inoue, F. Raimondi, F. M. N. Kadji, G. Singh, T. Kishi, A. Uwamizu, Y. Ono, Y. Shinjo, S. Ishida, N. Arang, K. Kawakami, J. S. Gutkind, J. Aoki, and R. B. Russell. Illuminating g-protein-coupling selectivity of gpcrs. *Cell*, 177(7):1933–1947 e25, 2019. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2019.04.044.

[24] Jacques Hanoune and Nicole Defer. Regulation and role of adenylyl cyclase isoforms. *Annual review of pharmacology and toxicology*, 41(1):145–174, 2001.

[25] Roger K Sunahara and Ron Taussig. Isoforms of mammalian adenylyl cyclase: multiplicities of signaling. *Molecular interventions*, 2(3):168, 2002.

[26] Gerald W Zamponi, Emmanuel Bourinet, Donald Nelson, Joel Nargeot, and Terry P Snutch. Crosstalk between g proteins and protein kinase c mediated by the calcium channel α 1 subunit. *Nature*, 385(6615):442–446, 1997.

[27] Shahriar M Khan, Adam Min, Sarah Gora, Geeda M Houranieh, Rhiannon Campden, Mélanie Robitaille, Phan Trieu, Darlaine Pétrin, Ashley M Jacobi, Mark A Behlke, et al. Gβ4γ1 as a modulator of m3 muscarinic receptor signalling and novel roles of gβ1

subunits in the modulation of cellular signalling. *Cellular signalling*, 27(8):1597–1608, 2015.

[28] Rhiannon Campden, Darlaine Pétrin, Mélanie Robitaille, Nicolas Audet, Sarah Gora, Stéphane Angers, and Terence E Hébert. Tandem affinity purification to identify cytosolic and nuclear g$\beta\gamma$-interacting proteins. In *Nuclear G-Protein Coupled Receptors*, pages 161–184. Springer, 2015.

[29] C. Martin, S. L. C. Moors, M. Danielsen, C. Betti, C. Fabris, D. Sejer Pedersen, E. Pardon, M. Peyressatre, K. Feher, J. C. Martins, J. Mosolff Mathiesen, M. C. Morris, N. Devoogdt, V. Caveliers, F. De Proft, J. Steyaert, and S. Ballet. Rational design of nanobody80 loop peptidomimetics: Towards biased beta2 adrenergic receptor ligands. *Chemistry*, 23(40): 9632–9640, 2017. ISSN 1521-3765 (Electronic) 0947-6539 (Linking). doi: 10.1002/chem.201701321.

[30] P. Donthamsetti, E. F. Gallo, D. C. Buck, E. L. Stahl, Y. Zhu, J. R. Lane, L. M. Bohn, K. A. Neve, C. Kellendonk, and J. A. Javitch. Arrestin recruitment to dopamine d2 receptor mediates locomotion but not incentive motivation. *Mol Psychiatry*, 2018. ISSN 1476-5578 (Electronic) 1359-4184 (Linking). doi: 10.1038/s41380-018-0212-4.

[31] D. Keri and P. Barth. Reprogramming g protein coupled receptor structure and function. *Curr Opin Struct Biol*, 51:187–194, 2018. ISSN 1879-033X (Electronic) 0959-440X (Linking). doi: 10.1016/j.sbi.2018.07.008.

[32] Laura M. Bohn, Robert J. Lefkowitz, Raul R. Gainetdinov, Karsten Peppel, Marc G. Caron, and Fang-Tsyr Lin. Enhanced morphine analgesia in mice lacking $beta$-arrestin 2. *Science*, 286(5449):2495–2498, 1999. ISSN 0036-8075. doi: 10.1126/science.286.5449.2495. URL https://science.sciencemag.org/content/286/5449/2495.

[33] Keshava Rajagopal, Erin J. Whalen, Jonathan D. Violin, Jonathan A. Stiber, Paul B. Rosenberg, Richard T. Premont, Thomas M. Coffman, Howard A. Rockman, and Robert J. Lefkowitz. $beta$-arrestin2-mediated inotropic effects of the angiotensin ii type 1a receptor in isolated cardiac myocytes. *Proceedings of the National Academy of Sciences*, 103(44):16284–16289, 2006. ISSN 0027-8424. doi: 10.1073/pnas.0607583103. URL https://www.pnas.org/content/103/44/16284.

[34] Jonathan D. Violin, Scott M. DeWire, Dennis Yamashita, David H. Rominger, Lisa Nguyen, Kevin Schiller, Erin J. Whalen, Maxine Gowen, and Michael W. Lark. Selectively engaging $beta$-arrestins at the angiotensin ii type 1 receptor reduces blood pressure and increases cardiac performance. *Journal of Pharmacology and Experimental Therapeutics*, 335(3):572–579, 2010. ISSN 0022-3565. doi: 10.1124/jpet.110.173005. URL https://jpet.aspetjournals.org/content/335/3/572.

[35] Danijela Markovic and RA John Challiss. Alternative splicing of g protein-coupled receptors: physiology and pathophysiology. *Cellular and molecular life sciences*, 66(20): 3337–3352, 2009.

[36] Frederick J Monsma, Loris D McVittie, Charles R Gerfen, Lawrence C Mahan, and David R Sibley. Multiple d2 dopamine receptors produced by alternative rna splicing. *Nature*, 342(6252):926–929, 1989.

[37] Patrizia Bovolin, Serena Bovetti, Aldo Fasolo, Z Katarova, G Szabo, MT Shipley, FL Margolis, and AC Puche. Developmental regulation of metabotropic glutamate receptor 1 splice variants in olfactory bulb mitral cells. *Journal of neuroscience research*, 87(2): 369–379, 2009.

[38] Zoltan Rekasi, Tamas Czompoly, Andrew V Schally, and Gabor Halmos. Isolation and sequencing of cdnas for splice variants of growth hormone-releasing hormone receptors from human cancers. *Proceedings of the National Academy of Sciences*, 97(19):10561–10566, 2000.

[39] Adolfo Rivero-Müller, Yen-Yin Chou, Inhae Ji, Svetlana Lajic, Aylin C Hanyaloglu, Kim Jonas, Nafis Rahman, Tae H Ji, and Ilpo Huhtaniemi. Rescue of defective g protein–coupled receptor function in vivo by intermolecular cooperation. *Proceedings of the National Academy of Sciences*, 107(5):2319–2324, 2010.

[40] Lee F Kolakowski Jr. Gcrdb: a g-protein-coupled receptor database. *Receptors & channels*, 2(1):1, 1994.

[41] Juan A Ballesteros and Harel Weinstein. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in g protein-coupled receptors. In *Methods in neurosciences*, volume 25, pages 366–428. Elsevier, 1995.

[42] Christoffer K Goth, Ulla E Pet aj a-Repo, and Mette M Rosenkilde. G protein-coupled receptors in the sweet spot: glycosylation and other post-translational modifications. *ACS pharmacology & translational science*, 3(2):237–245, 2020.

[43] A. Sente, R. Peer, A. Srivastava, M. Baidya, A. M. Lesk, S. Balaji, A. K. Shukla, M. M. Babu, and T. Flock. Molecular mechanism of modulating arrestin conformation by gpcr phosphorylation. *Nat Struct Mol Biol*, 25(6):538–545, 2018. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). doi: 10.1038/s41594-018-0071-3.

[44] X. E. Zhou, Y. He, P. W. de Waal, X. Gao, Y. Kang, N. Van Eps, Y. Yin, K. Pal, D. Goswami, T. A. White, A. Barty, N. R. Latorraca, H. N. Chapman, W. L. Hubbell, R. O. Dror, R. C. Stevens, V. Cherezov, V. V. Gurevich, P. R. Griffin, O. P. Ernst, K. Melcher, and H. E. Xu. Identification of phosphorylation codes for arrestin recruitment by g protein-coupled receptors. *Cell*, 170(3):457–469 e13, 2017. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2017.07.002.

[45] K. Palczewski, 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*, 289(5480):739–45, 2000. ISSN 0036-8075 (Print) 0036-8075 (Linking).

[46] S. G. Rasmussen, 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*, 450(7168):383–7, 2007. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature06325.

[47] V. Cherezov, 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*, 318(5854):1258–65, 2007. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1150577.

[48] Chong Wang, Yi Jiang, Jinming Ma, Huixian Wu, Daniel Wacker, Vsevolod Katritch, Gye Won Han, Wei Liu, Xi-Ping Huang, Eyal Vardy, John D. McCorvy, Xiang Gao, X. Edward Zhou, Karsten Melcher, Chenghai Zhang, Fang Bai, Huaiyu Yang, Linlin Yang, Hualiang Jiang, Bryan L. Roth, Vadim Cherezov, Raymond C. Stevens, and H. Eric Xu. Structural basis for molecular recognition at serotonin receptors. *Science*, 340(6132):610–614, 2013. ISSN 0036-8075. doi: 10.1126/science.1232807. URL https://science.sciencemag.org/content/340/6132/610.

[49] Aaron A Thompson, Wei Liu, Eugene Chun, Vsevolod Katritch, Huixian Wu, Eyal Vardy, Xi-Ping Huang, Claudio Trapella, Remo Guerrini, Girolamo Calo, et al. Structure of the nociceptin/orphanin fq receptor in complex with a peptide mimetic. *Nature*, 485(7398): 395–399, 2012.

[50] F. Magnani, M. J. Serrano-Vega, Y. Shibata, S. Abdul-Hussein, G. Lebon, J. Miller-Gallacher, A. Singhal, A. Strege, J. A. Thomas, and C. G. Tate. A mutagenesis and screening strategy to generate optimally thermostabilized membrane proteins for structural studies. *Nat Protoc*, 11(8):1554–71, 2016. ISSN 1750-2799 (Electronic) 1750-2799 (Linking). doi: 10.1038/nprot.2016.088.

[51] D. J. Scott, L. Kummer, P. Egloff, R. A. D. Bathgate, and A. Pluckthun. Improving the apo-state detergent stability of nts1 with chess for pharmacological and structural studies. *Biochimica Et Biophysica Acta-Biomembranes*, 1838(11):2817–2824, 2014. ISSN 0005-2736. doi: 10.1016/j.bbamem.2014.07.015.

[52] Marco Schütz, Jendrik Schöppe, Erik Sedlák, Matthias Hillenbrand, Gabriela Nagy-Davidescu, Janosch Ehrenmann, Christoph Klenk, Pascal Egloff, Lutz Kummer, and Andreas Plückthun. Directed evolution of g protein-coupled receptors in yeast for higher functional production in eukaryotic expression hosts. *Scientific reports*, 6(1):1–16, 2016.

[53] K. M. Chen, D. Keri, and P. Barth. Computational design of g protein-coupled receptor allosteric signal transductions. *Nat Chem Biol*, 2019. ISSN 1552-4469 (Electronic) 1552-4450 (Linking). doi: 10.1038/s41589-019-0407-2.

[54] Soumadwip Ghosh, Tobias Bierig, Sangbae Lee, Suvamay Jana, Adelheid Löhle, Gisela Schnapp, Christofer S Tautermann, and Nagarajan Vaidehi. Engineering salt bridge networks between transmembrane helices confers thermostability in g-protein-coupled receptors. *Journal of chemical theory and computation*, 14(12):6574–6585, 2018.

[55] S. G. Rasmussen, 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*, 477(7366):549–55, 2011. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature10361.

[56] Patrick Scheerer, Jung Hee Park, Peter W Hildebrand, Yong Ju Kim, Norbert Krauß, Hui-Woog Choe, Klaus Peter Hofmann, and Oliver P Ernst. Crystal structure of opsin in its g-protein-interacting conformation. *Nature*, 455(7212):497–502, 2008.

[57] Jung Hee Park, Patrick Scheerer, Klaus Peter Hofmann, Hui-Woog Choe, and Oliver Peter Ernst. Crystal structure of the ligand-free g-protein-coupled receptor opsin. *Nature*, 454 (7201):183–187, 2008.

[58] G. Pandy-Szekeres, C. Munk, T. M. Tsonkov, S. Mordalski, K. Harpsoe, A. S. Hauser, A. J. Bojarski, and D. E. Gloriam. Gpcrdb in 2018: adding gpcr structure models and ligands. *Nucleic Acids Research*, 46(D1):D440–D446, 2018. ISSN 0305-1048. doi: 10.1093/nar/gkx1109.

[59] Begonia Y Ho, Andreas Karschin, Theresa Branchek, Norman Davidson, and Henry A Lester. The role of conserved aspartate and serine residues in ligand binding and in function of the 5-ht1a receptor: A site-directed mutation study. *FEBS letters*, 312(2-3): 259–262, 1992.

[60] Martijn Bruysters, Heinz H Pertz, Aloys Teunissen, Remko A Bakker, Michel Gillard, Pierre Chatelain, Walter Schunack, Henk Timmerman, and Rob Leurs. Mutational analysis of the histamine h1-receptor binding pocket of histaprodifens. *European journal of pharmacology*, 487(1-3):55–63, 2004.

[61] A. M. Ring, A. Manglik, A. C. Kruse, M. D. Enos, W. I. Weis, K. C. Garcia, and B. K. Kobilka. Adrenaline-activated structure of beta2-adrenoceptor stabilized by an engineered nanobody. *Nature*, 502(7472):575–579, 2013. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature12572.

[62] Ron O Dror, Daniel H Arlow, David W Borhani, Morten Ø Jensen, Stefano Piana, and David E Shaw. Identification of two distinct inactive conformations of the β2-adrenergic receptor reconciles structural and biochemical observations. *Proceedings of the National Academy of Sciences*, 106(12):4689–4694, 2009.

[63] S.G.F. Rasmussen, 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.A. 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 $beta$ 2 adrenergic receptor-gs protein complex. *Nature*, 477(7366): 549–557, 2011. doi: 10.1038/nature10361.

[64] R. Nygaard, Y. Z. Zou, R. O. Dror, T. J. Mildorf, D. H. Arlow, A. Manglik, A. C. Pan, C. W. Liu, J. J. Fung, M. P. Bokoch, F. S. Thian, T. S. Kobilka, D. E. Shaw, L. Mueller, R. S. Prosser, and B. K. Kobilka. The dynamic process of beta(2)-adrenergic receptor activation. *Cell*, 152(3):532–542, 2013. ISSN 0092-8674. doi: 10.1016/j.cell.2013.01.008.

[65] Rita Rahmeh, Marjorie Damian, Martin Cottet, Hélène Orcel, Christiane Mendre, Thierry Durroux, K. Shivaji Sharma, Grégory Durand, Bernard Pucci, Eric Trinquet, Jurriaan M. Zwier, Xavier Deupi, Patrick Bron, Jean-Louis Banères, Bernard Mouillac, and Sébastien Granier. Structural insights into biased g protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proceedings of the National Academy of Sciences*, 109 (17):6733–6738, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1201093109. URL https://www.pnas.org/content/109/17/6733.

[66] Matthew T Drake, Jonathan D Violin, Erin J Whalen, James W Wisler, Sudha K Shenoy, and Robert J Lefkowitz. β-arrestin-biased agonism at the β2-adrenergic receptor. *Journal of Biological Chemistry*, 283(9):5669–5676, 2008.

[67] Hemlata Dwivedi-Agnihotri, Madhu Chaturvedi, Mithu Baidya, Tomasz Maciej Stepniewski, Shubhi Pandey, Jagannath Maharana, Ashish Srivastava, Natarin Caengprasath, Aylin C Hanyaloglu, Jana Selent, et al. Distinct phosphorylation sites in a prototypical gpcr differently orchestrate β-arrestin interaction, trafficking, and signaling. *Science advances*, 6(37):eabb8368, 2020.

[68] Max F Perutz, Michael G Rossmann, Ann F Cullis, Hilary Muirhead, Georg Will, and ACT North. Structure of hæmoglobin: a three-dimensional fourier synthesis at 5.5-å. resolution, obtained by x-ray analysis. *Nature*, 185(4711):416–422, 1960.

[69] J. Monod, J. Wyman, and J.-P. Changeux. On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology*, 12(1):88–118, 1965. doi: 10.1016/S0022-2836(65)80285-6.

[70] DE Koshland Jr, G Nemethy, and D_ Filmer. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry*, 5(1):365–385, 1966.

[71] Vincent J Hilser, James O Wrabl, and Hesam N Motlagh. Structural and energetic basis of allostery. *Annual review of biophysics*, 41:585–609, 2012.

[72] Paul Leff. The two-state model of receptor activation. *Trends in pharmacological sciences*, 16(3):89–97, 1995.

[73] C.-J. Tsai and R. Nussinov. A unified view of "how allostery works". *PLoS Computational Biology*, 10(2), 2014. doi: 10.1371/journal.pcbi.1003394.

[74] K Gunasekaran, Buyong Ma, and Ruth Nussinov. Is allostery an intrinsic property of all dynamic proteins? *Proteins: Structure, Function, and Bioinformatics*, 57(3):433–443, 2004.

[75] Jian Dai and Huan-Xiang Zhou. Reduced curvature of ligand-binding domain free-energy surface underlies partial agonism at nmda receptors. *Structure*, 23(1):228–236, 2015.

[76] Christopher L McClendon, Gregory Friedland, David L Mobley, Homeira Amirkhani, and Matthew P Jacobson. Quantifying correlations between allosteric sites in thermodynamic ensembles. *Journal of chemical theory and computation*, 5(9):2486–2502, 2009.

[77] Zachariah H Foda, Yibing Shan, Eric T Kim, David E Shaw, and Markus A Seeliger. A dynamically coupled allosteric network underlies binding cooperativity in src kinase. *Nature communications*, 6(1):1–10, 2015.

[78] S. Bhattacharya and N. Vaidehi. Differences in allosteric communication pipelines in the inactive and active states of a gpcr. *Biophysical Journal*, 107(2):422–434, 2014. ISSN 0006-3495. doi: 10.1016/j.bpj.2014.06.015.

[79] N. Vaidehi and S. Bhattacharya. Allosteric communication pipelines in g-protein-coupled receptors. *Current Opinion in Pharmacology*, 30:76–83, 2016. ISSN 1471-4892. doi: 10.1016/j.coph.2016.07.010.

[80] G. M. Suel, S. W. Lockless, M. A. Wall, and R. Ranganathan. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Biol*, 10 (1):59–69, 2003. ISSN 1072-8368 (Print) 1072-8368 (Linking). doi: 10.1038/nsb881.

[81] Y. M. Sung, A. D. Wilkins, G. J. Rodriguez, T. G. Wensel, and O. Lichtarge. Intramolecular allosteric communication in dopamine d2 receptor revealed by evolutionary amino acid covariation. *Proc Natl Acad Sci U S A*, 113(13):3539–44, 2016. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1516579113.

[82] A. J. Venkatakrishnan, X. Deupi, G. Lebon, F. M. Heydenreich, T. Flock, T. Miljus, S. Balaji, M. Bouvier, D. B. Veprintsev, C. G. Tate, G. F. Schertler, and M. M. Babu. Diverse activation pathways in class a gpcrs converge near the g-protein-coupling region. *Nature*, 536 (7617):484–7, 2016. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature19107.

[83] K. Sokolina, S. Kittanakom, J. Snider, M. Kotlyar, P. Maurice, J. Gandia, A. Benleulmi-Chaachoua, K. Tadagaki, A. Oishi, V. Wong, R. H. Malty, V. Deineko, H. Aoki, S. Amin, Z. Yao, X. Morato, D. Otasek, H. Kobayashi, J. Menendez, D. Auerbach, S. Angers, N. Przulj, M. Bouvier, M. Babu, F. Ciruela, R. Jockers, I. Jurisica, and I. Stagljar. Systematic protein-protein interaction mapping for clinically relevant human gpcrs. *Molecular Systems Biology*, 13(3), 2017. ISSN 1744-4292. doi: ARTN91810.15252/msb.20167430.

[84] Zhou Qingtong, Dehua Yang, Meng Wu, Yu Guo, Guo Wangjing, Zhong Li, Cai Xiaoqing, Dai Antao, Jang Wonjo, Eugene I Shakhnovich, et al. Common activation mechanism of class a gpcrs. *eLife*, 8, 2019.

[85] D. M. Rosenbaum, S. G. F. Rasmussen, and B. K. Kobilka. The structure and function of g-protein-coupled receptors. *Nature*, 459(7245):356–363, 2009. ISSN 0028-0836. doi: 10.1038/nature08144.

[86] K. J. Kohlhoff, D. Shukla, M. Lawrenz, G. R. Bowman, D. E. Konerding, D. Belov, R. B. Altman, and V. S. Pande. Cloud-based simulations on google exacycle reveal ligand modulation of gpcr activation pathways. *Nature Chemistry*, 6(1):15–21, 2014. ISSN 1755-4330. doi: 10.1038/Nchem.1821.

[87] D. M. Thal, Z. Vuckovic, C. J. Draper-Joyce, Y. L. Liang, A. Glukhova, A. Christopoulos, and P. M. Sexton. Recent advances in the determination of g protein-coupled receptor structures. *Curr Opin Struct Biol*, 51:28–34, 2018. ISSN 1879-033X (Electronic) 0959-440X (Linking). doi: 10.1016/j.sbi.2018.03.002.

[88] E. Ghosh, P. Kumari, D. Jaiman, and A. K. Shukla. Methodological advances: the unsung heroes of the gpcr structural revolution. *Nature Reviews Molecular Cell Biology*, 16(2): 69–81, 2015. ISSN 1471-0072. doi: 10.1038/nrm3933.

[89] Sheng Wang, Tao Che, Anat Levit, Brian K Shoichet, Daniel Wacker, and Bryan L Roth. Structure of the d2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature*, 555(7695):269–273, 2018.

[90] A. Jazayeri, M. Rappas, A. J. H. Brown, J. Kean, J. C. Errey, N. J. Robertson, C. Fiez-Vandal, S. P. Andrews, M. Congreve, A. Bortolato, J. S. Mason, A. H. Baig, I. Teobald, A. S. Dore, M. Weir, R. M. Cooke, and F. H. Marshall. Crystal structure of the glp-1 receptor bound to a peptide agonist. *Nature*, 546(7657):254–+, 2017. ISSN 0028-0836. doi: 10.1038/nature22800.

[91] S. P. Andrews, G. A. Brown, and J. A. Christopher. Structure-based and fragment-based gpcr drug discovery. *ChemMedChem*, 9(2):256–75, 2014. ISSN 1860-7187 (Electronic) 1860-7179 (Linking). doi: 10.1002/cmdc.201300382.

[92] N. Vaidehi, R. Grisshammer, and C. G. Tate. How can mutations thermostabilize g-protein-coupled receptors? *Trends Pharmacol Sci*, 37(1):37–46, 2016. ISSN 1873-3735 (Electronic) 0165-6147 (Linking). doi: 10.1016/j.tips.2015.09.005.

[93] T. Aristotelous, S. Ahn, A. K. Shukla, S. Gawron, M. F. Sassano, A. W. Kahsai, L. M. Wingler, X. Zhu, P. Tripathi-Shukla, X. P. Huang, J. Riley, J. Besnard, K. D. Read, B. L. Roth, I. H. Gilbert, A. L. Hopkins, R. J. Lefkowitz, and I. Navratilova. Discovery of beta 2 adrenergic receptor ligands using biosensor fragment screening of tagged wild-type receptor. *Acs Medicinal Chemistry Letters*, 4(10):1005–1010, 2013. ISSN 1948-5875. doi: 10.1021/ml400312j.

[94] E. R. Siuda, B. A. Copits, M. J. Schmidt, M. A. Baird, R. Al-Hasani, W. J. Planer, S. C. Funderburk, J. G. McCall, R. W. Gereau, and M. R. Bruchas. Spatiotemporal control of opioid signaling and behavior. *Neuron*, 86(4):923–935, 2015. ISSN 0896-6273. doi: 10.1016/j.neuron.2015.03.066.

[95] S. M. Spangler and M. R. Bruchas. Optogenetic approaches for dissecting neuromodulation and gpcr signaling in neural circuits. *Curr Opin Pharmacol*, 32:56–70, 2017. ISSN 1471-4973 (Electronic) 1471-4892 (Linking). doi: 10.1016/j.coph.2016.11.001.

[96] B. L. Roth. Dreadds for neuroscientists. *Neuron*, 89(4):683–94, 2016. ISSN 1097-4199 (Electronic) 0896-6273 (Linking). doi: 10.1016/j.neuron.2016.01.040.

[97] G. Barnea, W. Strapps, G. Herrada, Y. Berman, J. Ong, B. Kloss, R. Axel, and K. J. Lee. The genetic design of signaling cascades to record receptor activation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(1):64–69, 2008. ISSN 0027-8424. doi: 10.1073/pnas.0710487105.

[98] NH Kipniss, PCDP Dingal, TR Abbott, YC Gao, HF Wang, AA Dominguez, L Labanieh, and LS Qi. Engineering cell sensing and responses using a gpcr-coupled crispr-cas system. *Nature Communications*, 8, 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-02075-1.

[99] WK Kroeze, MF Sassano, XP Huang, K Lansu, JD McCorvy, PM Giguere, N Sciaky, and BL Roth. Presto-tango as an open-source resource for interrogation of the druggable human gpcrome. *Nature Structural & Molecular Biology*, 22(5):362–U28, 2015. ISSN 1545-9993. doi: 10.1038/nsmb.3014.

[100] Katharine Herrick-Davis, Graeme Milligan, and Giuseppe Di Giovanni. *G-protein-coupled Receptor Dimers*. Springer, 2017. ISBN 3319601725.

[101] B. Farran. An update on the physiological and therapeutic relevance of gpcr oligomers. *Pharmacological Research*, 117:303–327, 2017. ISSN 1043-6618. doi: 10.1016/j.phrs.2017.01.008.

[102] E. Vardy, J. E. Robinson, C. Li, R. H. J. Olsen, J. F. DiBerto, P. M. Giguere, F. M. Sassano, X. P. Huang, H. Zhu, D. J. Urban, K. L. White, J. E. Rittiner, N. A. Crowley, K. E. Pleil, C. M. Mazzone, P. D. Mosier, J. Song, T. L. Kash, C. J. Malanga, M. J. Krashes, and B. L. Roth. A new dreadd facilitates the multiplexed chemogenetic interrogation of behavior. *Neuron*, 86(4):936–946, 2015. ISSN 0896-6273. doi: 10.1016/j.neuron.2015.03.065.

[103] N. H. Joh, T. Wang, M. P. Bhate, R. Acharya, Y. B. Wu, M. Grabe, M. Hong, G. Grigoryan, and W. F. DeGrado. De novo design of a transmembrane zn2+-transporting four-helix bundle. *Science*, 346(6216):1520–1524, 2014. ISSN 0036-8075. doi: 10.1126/science.1261172.

[104] P. Lu, D. Min, F. DiMaio, K. Y. Wei, M. D. Vahey, S. E. Boyken, Z. Chen, J. A. Fallas, G. Ueda, W. Sheffler, V. K. Mulligan, W. Xu, J. U. Bowie, and D. Baker. Accurate computational

design of multipass transmembrane proteins. *Science*, 359(6379):1042–1046, 2018. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aaq1739.

[105] K. Y. Chen, F. Zhou, B. G. Fryszczyn, and P. Barth. Naturally evolved g protein-coupled receptors adopt metastable conformations. *Proc Natl Acad Sci U S A*, 109(33):13284–9, 2012. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1205512109.

[106] K. Y. Chen, J. Sun, J. S. Salvo, D. Baker, and P. Barth. High-resolution modeling of transmembrane helical protein structures from distant homologues. *PLoS Comput Biol*, 10(5):e1003636, 2014. ISSN 1553-7358 (Electronic) 1553-734X (Linking). doi: 10.1371/journal.pcbi.1003636.

[107] Y. Peng, J. D. McCorvy, K. Harpsoe, K. Lansu, S. Yuan, P. Popov, L. Qu, M. Pu, T. Che, L. F. Nikolajsen, X. P. Huang, Y. Wu, L. Shen, W. E. Bjorn-Yoshimoto, K. Ding, D. Wacker, G. W. Han, J. Cheng, V. Katritch, A. A. Jensen, M. A. Hanson, S. Zhao, D. E. Gloriam, B. L. Roth, R. C. Stevens, and Z. J. Liu. 5-ht2c receptor structures reveal the structural basis of gpcr polypharmacology. *Cell*, 172(4):719–730 e14, 2018. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2018.01.001.

[108] X. Feng, J. Ambia, K. M. Chen, M. Young, and P. Barth. Computational design of ligand-binding membrane receptors with high selectivity. *Nat Chem Biol*, 13(7):715–723, 2017. ISSN 1552-4469 (Electronic) 1552-4450 (Linking). doi: 10.1038/nchembio.2371.

[109] P. Barth and A. Senes. Toward high-resolution computational design of the structure and function of helical membrane proteins. *Nat Struct Mol Biol*, 23(6):475–80, 2016. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). doi: 10.1038/nsmb.3231.

[110] Francesca Magnani, Yoko Shibata, Maria J Serrano-Vega, and Christopher G Tate. Co-evolving stability and conformational homogeneity of the human adenosine a2a receptor. *Proceedings of the National Academy of Sciences*, 105(31):10744–10749, 2008.

[111] Kuang-Yui Michael Chen, Daniel Keri, and Patrick Barth. Computational design of g protein-coupled receptor allosteric signal transductions. *Nature Chemical Biology*, 16 (1):77–86, 2020.

[112] C. Arber, M. Young, and P. Barth. Reprogramming cellular functions with engineered membrane proteins. *Curr Opin Biotechnol*, 47:92–101, 2017. ISSN 1879-0429 (Electronic) 0958-1669 (Linking). doi: 10.1016/j.copbio.2017.06.009.

[113] A. K. Nivedha, C. S. Tautermann, S. Bhattacharya, S. Lee, P. Casarosa, I. Kollak, T. Kiechle, and N. Vaidehi. Identifying functional hotspot residues for biased ligand design in g-protein-coupled receptors. *Mol Pharmacol*, 93(4):288–296, 2018. ISSN 1521-0111 (Electronic) 0026-895X (Linking). doi: 10.1124/mol.117.110395.

[114] A. M. Schonegge, J. Gallion, L. P. Picard, A. D. Wilkins, C. Le Gouill, M. Audet, W. Stallaert, M. J. Lohse, M. Kimmel, O. Lichtarge, and M. Bouvier. Evolutionary action and

structural basis of the allosteric switch controlling beta2ar functional selectivity. *Nat Commun*, 8(1):2169, 2017. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/s41467-017-02257-x.

[115] T. Flock, A. S. Hauser, N. Lund, D. E. Gloriam, S. Balaji, and M. M. Babu. Selectivity determinants of gpcr-g-protein binding. *Nature*, 545(7654):317–322, 2017. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature22070.

[116] M. Young, T. Dahoun, B. Sokrat, C. Arber, K. M. Chen, M. Bouvier, and P. Barth. Computational design of orthogonal membrane receptor-effector switches for rewiring signaling pathways. *Proc Natl Acad Sci U S A*, 2018. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1718489115.

[117] B. R. Conklin, Z. Farfel, K. D. Lustig, D. Julius, and H. R. Bourne. Substitution of three amino acids switches receptor specificity of gq alpha to that of gi alpha. *Nature*, 363 (6426):274–6, 1993. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/363274a0.

[118] I. Gomes, M. A. Ayoub, W. Fujita, W. C. Jaeger, K. D. Pfleger, and L. A. Devi. G protein-coupled receptor heteromers. *Annu Rev Pharmacol Toxicol*, 56:403–25, 2016. ISSN 1545-4304 (Electronic) 0362-1642 (Linking). doi: 10.1146/annurev-pharmtox-011613-135952.

[119] B. Stephens and T. M. Handel. Chemokine receptor oligomerization and allostery. *Oligomerization and Allosteric Modulation in G-Protein Coupled Receptors*, 115:375–420, 2013. ISSN 1877-1173. doi: 10.1016/B978-0-12-394587-7.00009-9.

[120] J. K. Lai, J. Ambia, Y. Wang, and P. Barth. Enhancing structure prediction and design of soluble and membrane proteins with explicit solvent-protein interactions. *Structure*, 25 (11):1758–1770.e8, 2017. ISSN 1878-4186. doi: 10.1016/j.str.2017.09.002.

[121] A. Koehl, H. Hu, S. Maeda, Y. Zhang, Q. Qu, J. M. Paggi, N. R. Latorraca, D. Hilger, R. Dawson, H. Matile, G. F. X. Schertler, S. Granier, W. I. Weis, R. O. Dror, A. Manglik, G. Skiniotis, and B. K. Kobilka. Structure of the micro-opioid receptor-gi protein complex. *Nature*, 2018. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/s41586-018-0219-7.

[122] Y. Kang, O. Kuybeda, P. W. de Waal, S. Mukherjee, N. Van Eps, P. Dutka, X. E. Zhou, A. Bartesaghi, S. Erramilli, T. Morizumi, X. Gu, Y. Yin, P. Liu, Y. Jiang, X. Meng, G. Zhao, K. Melcher, O. P. Ernst, A. A. Kossiakoff, S. Subramaniam, and H. E. Xu. Cryo-em structure of human rhodopsin bound to an inhibitory g protein. *Nature*, 2018. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/s41586-018-0215-y.

[123] J. Garcia-Nafria, Y. Lee, X. Bai, B. Carpenter, and C. G. Tate. Cryo-em structure of the adenosine a2a receptor coupled to an engineered heterotrimeric g protein. *Elife*, 7, 2018. ISSN 2050-084X (Electronic) 2050-084X (Linking). doi: 10.7554/eLife.35946.

[124] Y. Zhang, B. Sun, D. Feng, H. Hu, M. Chu, Q. Qu, J. T. Tarrasch, S. Li, T. Sun Kobilka, B. K. Kobilka, and G. Skiniotis. Cryo-em structure of the activated glp-1 receptor in

complex with a g protein. *Nature*, 546(7657):248–253, 2017. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature22394.

[125] Rebecca F Alford, Andrew Leaver-Fay, Jeliazko R Jeliazkov, Matthew J O'Meara, Frank P DiMaio, Hahnbeom Park, Maxim V Shapovalov, P Douglas Renfrew, Vikram K Mulligan, Kalli Kappel, et al. The rosetta all-atom energy function for macromolecular modeling and design. *Journal of chemical theory and computation*, 13(6):3031–3048, 2017.

[126] Carol A Rohl, Charlie EM Strauss, Kira MS Misura, and David Baker. Protein structure prediction using rosetta. In *Methods in enzymology*, volume 383, pages 66–93. Elsevier, 2004.

[127] Rhiju Das and David Baker. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.*, 77:363–382, 2008.

[128] Maxim V Shapovalov and Roland L Dunbrack Jr. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure*, 19(6):844–858, 2011.

[129] Brian Kuhlman, Gautam Dantas, Gregory C Ireton, Gabriele Varani, Barry L Stoddard, and David Baker. Design of a novel globular protein fold with atomic-level accuracy. *science*, 302(5649):1364–1368, 2003.

[130] Lin Jiang, Eric A Althoff, Fernando R Clemente, Lindsey Doyle, Daniela Röthlisberger, Alexandre Zanghellini, Jasmine L Gallaher, Jamie L Betker, Fujie Tanaka, Carlos F Barbas, et al. De novo computational design of retro-aldol enzymes. *science*, 319(5868):1387–1391, 2008.

[131] Daniel-Adriano Silva, Shawn Yu, Umut Y Ulge, Jamie B Spangler, Kevin M Jude, Carlos Labão-Almeida, Lestat R Ali, Alfredo Quijano-Rubio, Mikel Ruterbusch, Isabel Leung, et al. De novo design of potent and selective mimics of il-2 and il-15. *Nature*, 565(7738):186–191, 2019.

[132] Peilong Lu, Duyoung Min, Frank DiMaio, Kathy Y Wei, Michael D Vahey, Scott E Boyken, Zibo Chen, Jorge A Fallas, George Ueda, William Sheffler, et al. Accurate computational design of multipass transmembrane proteins. *Science*, 359(6379):1042–1046, 2018.

[133] P. Barth, J. Schonbrun, and D. Baker. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proceedings of the National Academy of Sciences of the United States of America*, 104(40):15682–15687, 2007. doi: 10.1073/pnas.0702515104.

[134] Andrew Leaver-Fay, Ron Jacak, P Benjamin Stranges, and Brian Kuhlman. A generic program for multistate protein design. *PloS one*, 6(7):e20937, 2011.

[135] Ivet Bahar, Timothy R Lezon, Ahmet Bakan, and Indira H Shrivastava. Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chemical reviews*, 110(3):1463–1497, 2010.

[136] Y. Zhang, P. Doruker, B. Kaynak, S. Zhang, J. Krieger, H. Li, and I. Bahar. Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. *Curr Opin Struct Biol*, 62:14–21, 2019. ISSN 1879-033X (Electronic) 0959-440X (Linking). doi: 10.1016/j.sbi.2019.11.002.

[137] Basak Isin, AJ Rader, Harpreet Kaur Dhiman, Judith Klein-Seetharaman, and Ivet Bahar. Predisposition of the dark state of rhodopsin to functional changes in structure. *PROTEINS: Structure, Function, and Bioinformatics*, 65(4):970–983, 2006.

[138] T. Perica, Y. Kondo, S.P. Tiwari, S.H. McLaughlin, K.R. Kemplen, X. Zhang, A. Steward, N. Reuter, J. Clarke, and S.A. Teichmann. Evolution of oligomeric state through allosteric pathways that mimic ligand binding. *Science*, 346(6216), 2014. doi: 10.1126/science.1254346.

[139] J. Luo, Y. Zhu, M.X. Zhu, and H. Hu. Cell-based calcium assay for medium to high throughput screening of trp channel functions using flexstation 3. *Journal of Visualized Experiments*, (54), 2011. doi: 10.3791/3149.

[140] Y. Namkung, C. Le Gouill, V. Lukashova, H. Kobayashi, M. Hogue, E. Khoury, M. Song, M. Bouvier, and S. A. Laporte. Monitoring g protein-coupled receptor and beta-arrestin trafficking in live cells using enhanced bystander bret. *Nat Commun*, 7:12178, 2016. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/ncomms12178.

[141] G.M. Süel, S.W. Lockless, M.A. Wall, and R. Ranganathan. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature Structural Biology*, 10(1):59–69, 2003. doi: 10.1038/nsb881.

[142] K.A. Reynolds, R.N. McLaughlin, and R. Ranganathan. Hot spots for allosteric regulation on protein surfaces. *Cell*, 147(7):1564–1575, 2011. doi: 10.1016/j.cell.2011.10.049.

[143] Q. Cui and M. Karplus. Allostery and cooperativity revisited. *Protein Science*, 17(8):1295–1307, 2008. doi: 10.1110/ps.03259908.

[144] I. Luque, S.A. Leavitt, and E. Freire. The linkage between protein folding and functional cooperativity: Two sides of the same coin? *Annual Review of Biophysics and Biomolecular Structure*, 31:235–256, 2002. doi: 10.1146/annurev.biophys.31.082901.134215.

[145] M.W. Clarkson, S.A. Gilmore, M.H. Edgell, and A.L. Lee. Dynamic coupling and allosteric behavior in a nonallosteric protein. *Biochemistry*, 45(25):7693–7699, 2006. doi: 10.1021/bi060652l.

[146] S. Bhattacharya and N. Vaidehi. Differences in allosteric communication pipelines in the inactive and active states of a gpcr. *Biophysical Journal*, 107(2):422–434, 2014. doi: 10.1016/j.bpj.2014.06.015.

[147] Y. Miao, S.E. Nichols, P.M. Gasper, V.T. Metzger, and J.A. McCammon. Activation and dynamic network of the m2 muscarinic receptor. *Proceedings of the National Academy*

of Sciences of the United States of America, 110(27):10982–10987, 2013. doi: 10.1073/pnas.1309755110.

[148] K.J. Kohlhoff, D. Shukla, M. Lawrenz, G.R. Bowman, D.E. Konerding, D. Belov, R.B. Altman, and V.S. Pande. Cloud-based simulations on google exacycle reveal ligand modulation of gpcr activation pathways. *Nature Chemistry*, 6(1):15–21, 2014. doi: 10.1038/nchem.1821.

[149] R.N. McLaughlin Jr., F.J. Poelwijk, A. Raman, W.S. Gosal, and R. Ranganathan. The spatial architecture of protein function and adaptation. *Nature*, 491(7422):138–142, 2012. doi: 10.1038/nature11500.

[150] Y.-M. Sung, A.D. Wilkins, G.J. Rodriguez, T.G. Wensel, and O. Lichtarge. Intramolecular allosteric communication in dopamine d2 receptor revealed by evolutionary amino acid covariation. *Proceedings of the National Academy of Sciences of the United States of America*, 113(13):3539–3544, 2016. doi: 10.1073/pnas.1516579113.

[151] D.M. Rosenbaum, S.G.F. Rasmussen, and B.K. Kobilka. The structure and function of g-protein-coupled receptors. *Nature*, 459(7245):356–363, 2009. doi: 10.1038/nature08144.

[152] H.N. Motlagh, J.O. Wrabl, J. Li, and V.J. Hilser. The ensemble nature of allostery. *Nature*, 508(7496):331–339, 2014. doi: 10.1038/nature13001.

[153] J.W. Wisler, K. Xiao, A.R.B. Thomsen, and R.J. Lefkowitz. Recent developments in biased agonism. *Current Opinion in Cell Biology*, 27(1):18–24, 2014. doi: 10.1016/j.ceb.2013.10.008.

[154] A.J. Venkatakrishnan, X. Deupi, G. Lebon, C.G. Tate, G.F. Schertler, and M. Madan Babu. Molecular signatures of g-protein-coupled receptors. *Nature*, 494(7436):185–194, 2013. doi: 10.1038/nature11896.

[155] R.O. Dror, D.H. Arlow, P. Maragakis, T.J. Mildorf, A.C. Pan, H. Xu, D.W. Borhani, and D.E. Shaw. Activation mechanism of the β 2-adrenergic receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 108(46):18684–18689, 2011. doi: 10.1073/pnas.1110499108.

[156] S. Ahuja and S.O. Smith. Multiple switches in g protein-coupled receptor activation. *Trends in Pharmacological Sciences*, 30(9):494–502, 2009. doi: 10.1016/j.tips.2009.06.003.

[157] V. Katritch, V. Cherezov, and R.C. Stevens. Diversity and modularity of g protein-coupled receptor structures. *Trends in Pharmacological Sciences*, 33(1):17–27, 2012. doi: 10.1016/j.tips.2011.09.003.

[158] T. Flock, A.S. Hauser, N. Lund, D.E. Gloriam, S. Balaji, and M.M. Babu. Selectivity determinants of gpcr-g-protein binding. *Nature*, 545(7654):317–322, 2017. doi: 10.1038/nature22070.

[159] D.J. Scott, L. Kummer, D. Tremmel, and A. Plückthun. Stabilizing membrane proteins through protein engineering. *Current Opinion in Chemical Biology*, 17(3):427–435, 2013. doi: 10.1016/j.cbpa.2013.04.002.

[160] F. Magnani, Y. Shibata, M.J. Serrano-Vega, and C.G. Tate. Co-evolving stability and conformational homogeneity of the human adenosine a2a receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 105(31):10744–10749, 2008. doi: 10.1073/pnas.0804396105.

[161] P. Egloff, M. Hillenbrand, C. Klenk, A. Batyuk, P. Heine, S. Balada, K.M. Schlinkmann, D.J. Scott, M. Schütz, and A. Plückthun. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6):E655–E662, 2014. doi: 10.1073/pnas.1317903111.

[162] C.A. Sarkar, I. Dodevski, M. Kenig, S. Dudli, A. Mohr, E. Hermans, and A. Plückthun. Directed evolution of a g protein-coupled receptor for expression, stability, and binding selectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39):14808–14813, 2008. doi: 10.1073/pnas.0803103105.

[163] L. Ye, N. Van Eps, M. Zimmer, O.P. Ernst, and R. Scott Prosser. Activation of the a 2a adenosine g-protein-coupled receptor by conformational selection. *Nature*, 533: 265–268, 2016. doi: 10.1038/nature17668.

[164] Michael V LeVine and Harel Weinstein. Aim for allostery: using the ising model to understand information processing and transmission in allosteric biomolecular systems. *Entropy*, 17(5):2895–2918, 2015.

[165] X. Feng, J. Ambia, K.-Y.M. Chen, M. Young, and P. Barth. Computational design of ligand-binding membrane receptors with high selectivity. *Nature Chemical Biology*, 13 (7):715–723, 2017. doi: 10.1038/nchembio.2371.

[166] K.-Y.M. Chen, J. Sun, J.S. Salvo, D. Baker, and P. Barth. High-resolution modeling of transmembrane helical protein structures from distant homologues. *PLoS Computational Biology*, 10(5), 2014. doi: 10.1371/journal.pcbi.1003636.

[167] I. Bahar, T.R. Lezon, L.-W. Yang, and E. Eyal. Global dynamics of proteins: Bridging between structure and function. *Annual Review of Biophysics*, 39(1):23–42, 2010. doi: 10.1146/annurev.biophys.093008.131258.

[168] J.A. Ballesteros and H. Weinstein. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in g protein-coupled receptors. *Methods in Neurosciences*, 25(C):366–428, 1995. doi: 10.1016/S1043-9471(05)80049-7.

[169] Y. Kang, X.E. Zhou, X. Gao, Y. He, W. Liu, A. Ishchenko, A. Barty, T.A. White, O. Yefanov, G.W. Han, Q. Xu, P.W. De Waal, J. Ke, M.H.E. Tan, C. Zhang, A. Moeller, G.M. West, B.D.

Pascal, N. Van Eps, L.N. Caro, S.A. Vishnivetskiy, R.J. Lee, K.M. Suino-Powell, X. Gu, K. Pal, J. Ma, X. Zhi, S. Boutet, G.J. Williams, M. Messerschmidt, C. Gati, N.A. Zatsepin, D. Wang, D. James, S. Basu, S. Roy-Chowdhury, C.E. Conrad, J. Coe, H. Liu, S. Lisova, C. Kupitz, I. Grotjohann, R. Fromme, Y. Jiang, M. Tan, H. Yang, J. Li, M. Wang, Z. Zheng, D. Li, N. Howe, Y. Zhao, J. Standfuss, K. Diederichs, Y. Dong, C.S. Potter, B. Carragher, M. Caffrey, H. Jiang, H.N. Chapman, J.C.H. Spence, P. Fromme, U. Weierstall, O.P. Ernst, V. Katritch, V.V. Gurevich, P.R. Griffin, W.L. Hubbell, R.C. Stevens, V. Cherezov, K. Melcher, and H.E. Xu. Crystal structure of rhodopsin bound to arrestin by femtosecond x-ray laser. *Nature*, 523(7562):561–567, 2015. doi: 10.1038/nature14656.

[170] A. Koehl, H. Hu, S. Maeda, Y. Zhang, Q. Qu, J.M. Paggi, N.R. Latorraca, D. Hilger, R. Dawson, H. Matile, G.F.X. Schertler, S. Granier, W.I. Weis, R.O. Dror, A. Manglik, G. Skiniotis, and B.K. Kobilka. Structure of the *mu*-opioid receptor-gi protein complex. *Nature*, 558 (7711):547–552, 2018. doi: 10.1038/s41586-018-0219-7.

[171] B. Carpenter, R. Nehmé, T. Warne, A.G.W. Leslie, and C.G. Tate. Structure of the adenosine a2a receptor bound to an engineered g protein. *Nature*, 536(7614):104–107, 2016. doi: 10.1038/nature18966.

[172] Søren GF Rasmussen, Brian T DeVree, Yaozhong Zou, Andrew C Kruse, Ka Young Chung, Tong Sun Kobilka, Foon Sun Thian, Pil Seok Chae, Els Pardon, Diane Calinski, et al. Crystal structure of the β 2 adrenergic receptor-gs protein complex. *Nature*, 477(7366): 549–555, 2011.

[173] F. Horn, J. Weare, M.W. Beukers, S. Hörsch, A. Bairoch, W. Chen, Ø. Edvardsen, F. Campagne, and G. Vriend. Gpcrdb: An information system for g protein-coupled receptors. *Nucleic Acids Res.*, 26:277–281, 1998.

[174] E.H. Kellogg, A. Leaver-Fay, and D. Baker. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins: Structure, Function and Bioinformatics*, 79(3):830–838, 2011. doi: 10.1002/prot.22921.

[175] G.J. Rodriguez, R. Yao, O. Lichtarge, and T.G. Wensel. Evolution-guided discovery and recoding of allosteric pathway specificity determinants in psychoactive bioamine receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 107(17):7787–7792, 2010. doi: 10.1073/pnas.0914877107.

[176] A.-M. Schönegge, J. Gallion, L.-P. Picard, A.D. Wilkins, C. Le Gouill, M. Audet, W. Stallaert, M.J. Lohse, M. Kimmel, O. Lichtarge, and M. Bouvier. Evolutionary action and structural basis of the allosteric switch controlling β2ar functional selectivity. *Nature Communications*, 8(1), 2017. doi: 10.1038/s41467-017-02257-x.

[177] M. Han, S.O. Smith, and T.P. Sakmar. Constitutive activation of opsin by mutation of methionine 257 on transmembrane helix 6. *Biochemistry*, 37(22):8253–8261, 1998. doi: 10.1021/bi980147r.

[178] S.G. Graber, R.A. Figler, and J.C. Garrison. Expression and purification of functional g protein α subunits using a baculovirus expression system. *Journal of Biological Chemistry*, 267(2):1271–1278, 1992.

[179] Melissa Miller, Meng Wu, Jia Xu, David Weaver, Min Li, and Michael X Zhu. High-throughput screening of trpc channel ligands using cell-based assays. 2011.

[180] M. Young, T. Dahoun, B. Sokrat, C. Arber, K.M. Chen, M. Bouvier, and P. Barth. Computational design of orthogonal membrane receptor-effector switches for rewiring signaling pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 115(27):7051–7056, 2018. doi: 10.1073/pnas.1718489115.

[181] J.Y. Xiao, G.V. Ruiz, M.R. Whorton, S.G.F. 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*, 106(23):9501–9506, 2009. doi: 10.1073/pnas.0811437106.

[182] I.W. Davis and D. Baker. Rosettaligand docking with full ligand and receptor flexibility. *Journal of Molecular Biology*, 385(2):381–392, 2009. doi: 10.1016/j.jmb.2008.11.010.

[183] K.-Y.M. Chen, F. Zhou, B.G. Fryszczyn, and P. Barth. Naturally evolved g protein-coupled receptors adopt metastable conformations. *Proceedings of the National Academy of Sciences of the United States of America*, 109(33):13284–13289, 2012. doi: 10.1073/pnas.1205512109.

[184] Li Zhang and Jan Hermans. Hydrophilicity of cavities in proteins. *Proteins: Structure, Function, and Bioinformatics*, 24(4):433–438, 1996.

[185] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n log (n) method for ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092, 1993.

[186] Lars Skjærven, Xin-Qiu Yao, Guido Scarabelli, and Barry J Grant. Integrating protein structural dynamics and evolutionary analysis with bio3d. *BMC bioinformatics*, 15(1): 399, 2014.

[187] M. Han, S.W. Lin, M. Minkova, S.O. Smith, and T.P. Sakmar. Functional interaction of transmembrane helices 3 and 6 in rhodopsin. replacement of phenylalanine 261 by alanine causes reversion of phenotype of a glycine 121 replacement mutant. *Journal of Biological Chemistry*, 271(50):32337–32342, 1996. doi: 10.1074/jbc.271.50.32337.

[188] M. Han, S.W. Lin, S.O. Smith, and T.P. Sakmar. The effects of amino acid replacements of glycine 121 on transmembrane helix 3 of rhodopsin. *Journal of Biological Chemistry*, 271(50):32330–32336, 1996. doi: 10.1074/jbc.271.50.32330.

[189] Jin Li, Peng Huang, Chongguang Chen, J Kim de Riel, Harel Weinstein, and Lee-Yuan Liu-Chen. Constitutive activation of the μ opioid receptor by mutation of d3. 49 (164),

but not d3. 32 (147): D3. 49 (164) is critical for stabilization of the inactive form of the receptor and for its expression. *Biochemistry*, 40(40):12039–12050, 2001.

[190] A. Cavalli, A.-M. Babey, and H.H. Loh. Altered adenylyl cyclase responsiveness subsequent to point mutations of asp 128 in the third transmembrane domain of the δ-opioid receptor. *Neuroscience*, 93(3):1025–1031, 1999. doi: 10.1016/S0306-4522(99)00280-8.

[191] F.M. Decaillot, K. Befort, D. Filliol, S.Y. Yue, P. Walker, and B.L. Kieffer. Opioid receptor random mutagenesis reveals a mechanism for g protein-coupled receptor activation. *Nature Structural Biology*, 10(8):629–636, 2003. doi: 10.1038/nsb950.

[192] Joanna F Swain and Lila M Gierasch. The changing landscape of protein allostery. *Current opinion in structural biology*, 16(1):102–108, 2006.

[193] Jie Yin, Kuang-Yui M Chen, Mary J Clark, Mahdi Hijazi, Punita Kumari, Xiao-chen Bai, Roger K Sunahara, Patrick Barth, and Daniel M Rosenbaum. Structure of a d2 dopamine receptor–g-protein complex in a lipid membrane. *Nature*, pages 1–5, 2020.

[194] Xin Chen, John D McCorvy, Matthew G Fischer, Kyle V Butler, Yudao Shen, Bryan L Roth, and Jian Jin. Discovery of g protein-biased d2 dopamine receptor partial agonists. *Journal of medicinal chemistry*, 59(23):10601–10618, 2016.

[195] John A Allen, Julianne M Yost, Vincent Setola, Xin Chen, Maria F Sassano, Meng Chen, Sean Peterson, Prem N Yadav, Xi-ping Huang, Bo Feng, et al. Discovery of β-arrestin–biased dopamine d2 ligands for probing signal transduction pathways essential for antipsychotic efficacy. *Proceedings of the National Academy of Sciences*, 108(45):18488–18493, 2011.

[196] Richard Ågren, Peter Århem, Johanna Nilsson, and Kristoffer Sahlholm. The beta-arrestin-biased dopamine d2 receptor ligand, unc9994, is a partial agonist at g-protein-mediated potassium channel activation. *International Journal of Neuropsychopharmacology*, 21(12):1102–1108, 2018.

[197] S. Bhattacharya and N. Vaidehi. Mechanism of allosteric communication in gpcr activation from microsecond scale molecular dynamics simulations. *Biophysical Journal*, 112 (3):498a–499a, 2017. ISSN 0006-3495. doi: DOI10.1016/j.bpj.2016.11.2697.

[198] Alexander S Hauser, Sreenivas Chavali, Ikuo Masuho, Leonie J Jahn, Kirill A Martemyanov, David E Gloriam, and M Madan Babu. Pharmacogenomics of gpcr drug targets. *Cell*, 172(1-2):41–54, 2018.

[199] Hollie J Jackson, Sarwish Rafiq, and Renier J Brentjens. Driving car t-cells forward. *Nature reviews Clinical oncology*, 13(6):370, 2016.

[200] Cheng Zhang, Jun Liu, Jiang F Zhong, and Xi Zhang. Engineering car-t cells. *Biomarker research*, 5(1):1–6, 2017.

[201] Sattva S Neelapu, Sudhakar Tummala, Partow Kebriaei, William Wierda, Cristina Gutierrez, Frederick L Locke, Krishna V Komanduri, Yi Lin, Nitin Jain, Naval Daver, et al. Chimeric antigen receptor t-cell therapy—assessment and management of toxicities. *Nature reviews Clinical oncology*, 15(1):47, 2018.

[202] Kheng Newick, Edmund Moon, and Steven M Albelda. Chimeric antigen receptor t-cell therapy for solid tumors. *Molecular Therapy-Oncolytics*, 3:16006, 2016.

[203] Kheng Newick, Shaun O'Brien, Jing Sun, Veena Kapoor, Steven Maceyko, Albert Lo, Ellen Puré, Edmund Moon, and Steven M Albelda. Augmentation of car t-cell trafficking and antitumor efficacy by blocking protein kinase a localization. *Cancer immunology research*, 4(6):541–551, 2016.

[204] Gianni Monaco, Bernett Lee, Weili Xu, Seri Mustafah, You Yi Hwang, Christophe Carre, Nicolas Burdin, Lucian Visan, Michele Ceccarelli, Michael Poidinger, et al. Rna-seq signatures normalized by mrna abundance allow absolute deconvolution of human immune cell types. *Cell reports*, 26(6):1627–1640, 2019.

[205] Pierre Launay, Henrique Cheng, Subhashini Srivatsan, Reinhold Penner, Andrea Fleig, and Jean-Pierre Kinet. Trpm4 regulates calcium oscillations after t cell activation. *Science*, 306(5700):1374–1377, 2004.

[206] John JG Tesmer and Stephen R Sprang. The structure, catalytic mechanism and regulation of adenylyl cyclase. *Current opinion in structural biology*, 8(6):713–719, 1998.

[207] B. Carpenter, R. Nehme, T. Warne, A. G. Leslie, and C. G. Tate. Structure of the adenosine a(2a) receptor bound to an engineered g protein. *Nature*, 536(7614):104–7, 2016. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature18966.

[208] Benjamin J Schmiedel, Divya Singh, Ariel Madrigal, Alan G Valdovino-Gonzalez, Brandie M White, Jose Zapardiel-Gonzalo, Brendan Ha, Gokmen Altay, Jason A Greenbaum, Graham McVicker, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell*, 175(6):1701–1715, 2018.

[209] Onur Dagliyan, Nikolay V Dokholyan, and Klaus M Hahn. Engineering proteins for allosteric control by light or ligands. *Nature protocols*, 14(6):1863–1883, 2019.

[210] Olga Bozovic, Brankica Jankovic, and Peter Hamm. Sensing the allosteric force. *Nature communications*, 11(1):1–7, 2020.

[211] Fabio Pirro, Nathan Schmidt, James Lincoff, Zachary X Widel, Nicholas F Polizzi, Lijun Liu, Michael J Therien, Michael Grabe, Marco Chino, Angela Lombardi, et al. Allosteric cooperation in a de novo-designed two-domain protein. *Proceedings of the National Academy of Sciences*, 117(52):33246–33253, 2020.

Daniel Keri

Contact Information

https://www.linkedin.com/in/daniel-keri-18ba98a2/

Education

École Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland

Ph.D. Candidate, Bioengineering and Biotechnology, October 2017 - Present

- Thesis Topic: "Reprogramming G Protein-Coupled Receptor Structure, Function, and Signaling by Computational Design" (award-nominated)
- Advisor: Patrick Barth
- Assisted in the setup of the new lab at EPFL
- Developed new experimental protocols for the new lab
- Trained new members of the lab

Baylor College of Medicine, Houston, Texas USA

Predoctoral Fellow, Biochemistry and Molecular Biology, August 2014 - October 2017

- Thesis Topic: "Reprogramming G Protein-Coupled Receptor Structure, Function, and Signaling by Computational Design"
- Advisor: Patrick Barth

California State University Channel Islands, Camarillo, California USA

B.S., Chemistry, May, 2012

California State University Channel Islands, Camarillo, California USA

B.S., Biology (Molecular Biology), May, 2012

Honors and Awards

Keck Fellow 2016 - 2018

The Keck Fellowship is provided by the Gulf Coast Consortia for Quantitative Biomedical Sciences, as part of the Houston Area Molecular Biophysics Training Program, to select doctoral students who utilize a combination of computational and experimental tools to carry out their research.

California State University Channel Islands Alzheimer's Institute Scholarship, 2010

The Alzheimer's Scholarship is a highly competitive and desired scholarship, given only to two students annually.

Academic Experience

École Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland
Graduate Student **August, 2014 - October 2017**

Baylor College of Medicine, Houston, Texas USA
Predoctoral Fellow **August, 2014 - October 2017**
Includes current Ph.D. research, and Ph.D. level coursework.

California State University Channel Islands, Camarillo, California USA
Student Researcher - Biology **June 2010 - December, 2011**
Maintained and propagated human neural stem cell (hNSC) lines. Performed cytotox-

icity studies and antioxidation experiments on hNSC's utilizing LDH and MTS assays, both of which rely on UV/Vis spectroscopy.

Student Researcher - Chemistry **December 2010 - June, 2011**
Performed chemical syntheses with the goal of creating analogues of the organic molecule curcumin.

Tutor **May 2012 - May, 2014**
Tutored high school students in various mathematics courses (geometry, algebra).

Publications

Chen, K. Y. M.*, **Keri, D.***, Barth, P. (2020). Computational design of G Protein-Coupled Receptor allosteric signal transductions. *Nature Chemical Biology*, 16(1), 77-86.

Keri, D., Barth, P. (2018). Reprogramming G protein coupled receptor structure and function. C*urrent opinion in structural biology*, 51, 187-194.

***Contributed Equally**

Conference Presentations

D. Keri, P. Barth. 2018. Engineering novel allosteric responses in the Dopamine D2 Receptor. 2018 RosettaCon.

D. Keri, P. Barth. 2018. Engineering novel allosteric responses in the Dopamine D2 Receptor. 2018 Gordon Research Conference on Membrane Protein Folding.

D. Keri, P. Barth. 2017. Engineering novel allosteric responses in the Dopamine D2 Receptor. 2017 RosettaCon.

D. Keri, P. Barth. 2017. Validation Of A Novel *in silico* Method To Predict Allosteric Changes In The Dopamine D2 Receptor. 2017 Sealy Structural Biology and Molecular Biophysics Symposium.

D. Keri, P. Barth. 2016 Validation Of A Novel *in silico* Method To Predict Allosteric Changes In The Dopamine D2 Receptor. 2016 Keck Center Annual Research Conference.

J. Cortez, W. Ferguson, D. Keri, P. Keshavarzian, V. Vijayaraghvan, G. M. Rishton, C. Wang. 2011. The Study of Oxidative Stress on Human Neural Stem Cell Induced by tert-butyl Hydroperoxide. 23rd CSUPERB Symposium.

Professional Experience

Baxter Biosciences, Thousand Oaks, California USA

Manufaturing Associate **June, 2013 - May, 2014**
Support cell culture operations for the production of a bulk drug substance. Set up, operate, maintain, and tear down equipment such as bioreactors, tanks, filtration systems, and production lines. Operate general production equipment such as pH meters, conductivity meters, and autoclaves. Monitor and record critical process parameters.

Experimental Skills

- GTPyS binding assay
- TRP Channel Assay
- Cell culture (CHO, HEK, HeLa, hNSC)
- PCR
- Western Blotting
- 1D NMR spectroscopy

- IR spectrsoscopy
- Organic small molecule synthesis

Computer Skills

- Statistical Software: SPSS, Python Packages (Pandas)
- Languages: Python, Perl, Java, C++, BASH
- Applications: Microsoft Office, LaTeX, Mathematica, Graphpad
- Operating Systems: Unix/Linux, Windows.