LANE, THOMAS R., M.S. Building and Refinement of an *in silico* Homology Model of a Novel G Protein-Coupled Receptor: GPR35. (2011)
Directed by Dr. Patricia H. Reggio. 108pp.

Human GPR35 (hGPR35), a recently deorphanized Class A G-protein coupled receptor, has been shown to exhibit prominent expression in immune and gastrointestinal tissues, with additional expression in pancreatic islets, skeletal muscle, lung tissue, and the dorsal root ganglion. The rat GPR35 (rGPR35) analog, which has 72% sequence identity with human GPR35, has been shown to have expression in similar tissues as with human GPR35. GPR35 has been suggested to be involved in metabolism, heart failure, inflammation, asthma, a mental retardation syndrome associated with the deletion on 2q37.3, type II diabetes, as well as gastric cancer formation, making GPR35 a potential target for the treatment of multiple diseases.

Both zaprinast, the well characterized cGMP-PDE inhibitor, and pamoic acid, a compound which the FDA has classified as an inactive compound, act as agonists at GPR35. However, interesting species differences have been found with these agonists and key mutations have also revealed differences between these two ligands. Pamoic acid is considerably lower in potency in rat GPR35, while zaprinast has increased efficacy in rat GPR35. Further, mutation studies suggest an increase in the potency of zaprinast in a human GPR35 R6.58A mutation. Pamoic acid, on the other hand shows similar potency to wild-type in this same mutant.

To probe the molecular origins of these differences, three separate homology models, an active (R*) hGPR35, an R* hGPR35 R6.58A(240) mutant, and an R* rGPR35 model, were constructed and docking studies were performed with the aforementioned ligands. These studies revealed that the change in residue 5.43 (P5.43 in human; S5.43 in rat) alters the shape of the binding pocket for pamoic acid. In addition, arginines which contribute significantly to the interaction of pamoic acid in hGPR35 (R6.58 and R7.32) become uncharged residues (Q6.58 and

S7.32) in rat GPR35. The increase of the potency of zaprinast in the hGPR35 R6.58A mutant receptor is due to the loss of bulk at position 6.58 (R6.58(240)→ A6.58(240)), that allows for additional interactions with the ligand. The statistically equivalent potencies of pamoic acid for the wild-type and R6.58A(240) mutant hGPR35 receptors is due to the isoenergetic interchange of the direct interaction residue R6.58(240) with R7.32(255) in the R6.58(240)A mutant.

BUILDING AND REFINEMENT OF AN *IN SILICO* HOMOLOGY MODEL OF A NOVEL G PROTEIN-COUPLED RECEPTOR: GPR35

by

Thomas R. Lane

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2011

Approved by

_____
Committee Chair

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____

Committee Members _____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

G-Protein coupled receptors (GPCRs) are the largest class of integral membrane proteins that are both involved in signal transduction and are activated by extracellular signals [1]. The activation of GPCRs is known to invoke a series of downstream effects, often referred to as signal cascades, which are involved in a nearly innumerable number of cellular actions. These actions include, but are not limited to, protein expression patterns, translational regulation, and general cell metabolism. The known endogenous ligands for GPCRs are incredibly diverse, including hormones, peptide neurotransmitters, chemokines, ions, and phospholipids [1]. The direct involvement of GPCRs in multiple key physiological processes, their activation by extracellular signaling, and their ligand diversity, makes GPCRs ideal targets for pharmaceuticals.

All GPCRs are predicted to share a similar topology, which has been exemplified in the published crystal structures of GPCRs Rhodopsin [2-5], Opsin [6], β-2 adrenergic receptor (β2-AR) [7-8], A2A adenosine receptor (A2-AR) [9], dopamine D3 receptor (D3-R) [10], and the CXCR4 chemokine receptor (CXCR4-R) [11]. A GPCR's topology can be separated into four components: an extracellular N-terminus, seven transmembrane α-helices (TMH) creating a closed bundle, intra- (IC) and extracellular (EC) loops connecting the transmembrane α-helices, and an intracellular C-terminus that is initially a short helical segment, that is aligned parallel with the phospholipid head groups of the bilayer, known as helix 8 (Hx 8). The crystal structures published to date suggest that GPCR ligands are bound in pockets formed by the TMHs and that

the residues binding these ligands are found in these pockets and in adjacent loops. The similarities found between different class "A" GPCRs, in both crystal structures and in other biophysical data, suggest that many features may be nearly universal within all similarly classed GPCRs. This doctrine leads to the possibility of creating homology models of other, less understood GPCRs by paralleling some of these universal similarities.

Throughout this document, the amino acid numbering scheme proposed by Ballesteros and Weinstein [12] is used. In this numbering system, the most highly conserved amino acid in each TMH is assigned a locant of .50, and the number preceding this is representative of the TMH number. The amino acids immediately preceding and following the .50 residue are numbered .49 and .51, respectively. An example of numbering, using the β2-AR sequence, is assigning the highly conserved proline on TMH7 a number of 7.50(323), with the number in brackets representing the absolute sequence number. The amino acids preceding and following P7.50(323) are N7.49(322) and L5.51(324) respectively. This numbering scheme applies to the amino acids in the TMH region exclusively; the N- and C-terminus, as well as the loop residues, are numbered using absolute sequence numbers only.

The description of the relative topology of a protein requires the definitions of both the side chain and the backbone dihedral torsional angles. Each dihedral is a rotatable bond, which is defined by the relative position of the atoms directly connected to the atoms of interest. There are three defined backbone dihedrals of amino acid chains: Φ (phi), Ψ (psi) and ω (omega). The Φ, Ψ, and ω dihedral angle rotate about the bond between the N-Cα, Cα –C(O), and C(O)-N atoms, respectively (Figure 1A). Side chain dihedrals are designated with a χ followed by an integer between 1 and 5, which defines the involved atoms. An example of lysine is depicted in figure 1B, which defines χ1, 2, 3, 4, 5 as the rotatable bonds between the atoms $C_\alpha - C_\beta$, $C_\beta - C_\gamma$, $C_\gamma - C_\delta$, and $C_\delta - C_\varepsilon$, respectively. As previously mentioned, the measurement of each side chain

dihedral is defined not by the positions of atoms directly involved in the bond of interest, but by the relative positions of the atoms which are connected to these atoms. The example, shown in Figure 2, depicts the three low energy conformers of the χ1 dihedral: gauche$^+$ (g^+), gauche$^-$ (g^-), and *trans*. The numerical value of the χ1 dihedral is defined by the angle made by the relative positions of an amino acid's backbone nitrogen (N) and its C_γ or Oxygen, depending on the amino acid. The convention is to assign a value for the dihedral angle by calculating the degrees of rotation required to eclipse the two atoms, with clockwise and counterclockwise having positive and negative values, respectively. The ideal value of each rotamer is defined as: $g^+ = -60°$, $g^- = +60°$, and *trans* = 180°. In a α-helix, a χ1 = g^+ or *trans* conformation is the preferred low energy state in the majority of amino acids, with the occasional adoption of the χ1 = g^- rotamer by the residues serine and threonine.

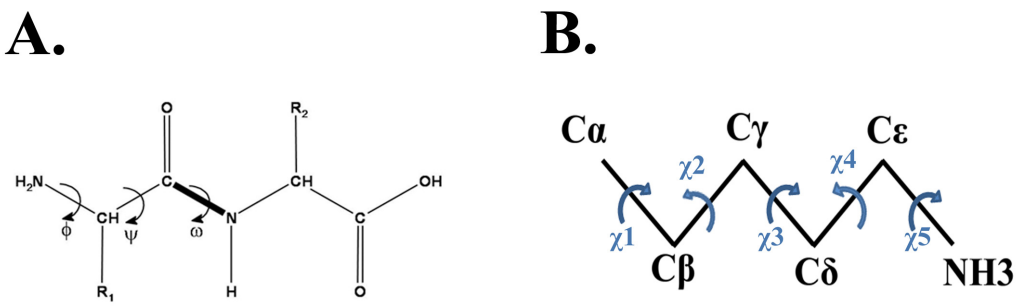


Figure 1. An example of backbone (Φ, Ψ and ω) and sidechain dihedrals. (A) An illustration of the backbone dihedrals Φ, Ψ and ω, using an amino acid skeleton. (B) The sidechain of the amino acid lysine, with labeled heteroatoms and defined sidechain dihedrals χ1 - χ5.

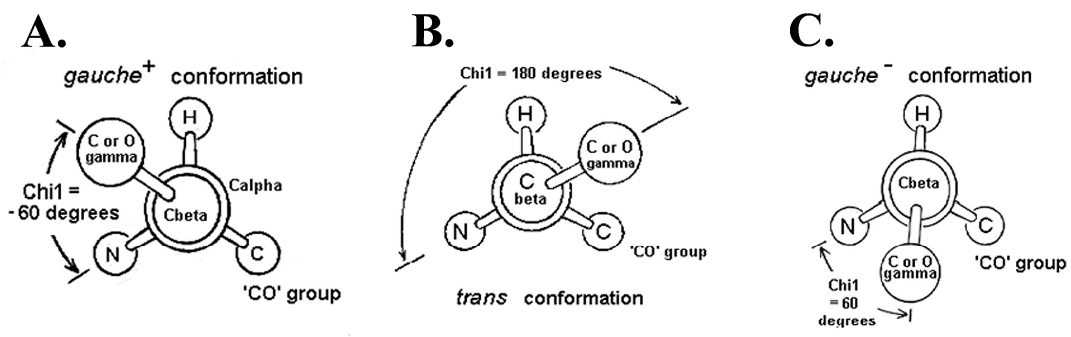


Figure 2. A depiction of the three low energy conformers of the χ1 dihedral. The low energy conformers are (A) gauche$^+$ (g^+), (B) gauche$^-$ (g^-), and (C) *trans.*

Existing knowledge concerning GPCR activation comes mostly from biophysical studies involving the prototypical GPCRs, rhodopsin and the β2-AR. The canonical transition from an inactive (R) to an active (R*) GPCR bundle involves a series of rotameric changes in key, conserved residues, which translate to much larger structural changes in the TMHs. The most dramatic structural change associated with the activation of β2-AR and rhodopsin is the general rotation and straightening of TMH6 about the highly conserved CWXP hinge region [13-21]. In rhodopsin, this change in TMH6 is suggested to cause a net movement of the Cα of E6.30(247) approximately 6 Å away from the intracellular end of TMH3 [18], which is assumed to break the salt bridge between E6.30(247) and D3.50(135). This breaking of the interaction between R3.50 and a polar or negatively charged residue on the intracellular end of TMH6 is presumed to be one of the major features of GPCR activation. This general movement of TMH6 has also been suggested to occur during activation in the M3 muscarinic acetylcholine receptor in *in situ* disulfide crosslinking studies [22-23], further supporting its universality.

A highly conserved tryptophan within the ligand binding pocket on TMH6, W6.48, is suggested to be a toggle switch residue related to activation. The rotameric χ1 change of W6.48 is suggested to occur during the R to R* transition. Multiple spectroscopic studies have

suggested that W6.48 changes both position and interaction partners in rhodopsin [14, 24-25] and in the β2-AR [26] during agonist-dependent activation, suggesting a rotameric change associated with activation. The rhodopsin crystal structures [2-4] shows that the beta-ionone ring of the covalently bound ligand, 11–*cis*-retinal, is proximal to W6.48(265), which constrains it in a χ1 = g+ conformation. When activated by light, the 11-*cis*-retinal isomerizes, moving the beta-ionone ring away from W6.48, relieving the side-chain's torsional constraint [27] which allows for a rotameric χ1 change. In conjunction with spectroscopic studies, this suggests that the χ1 rotamer of W6.48(265) changes from g+ to *trans* during the activation of rhodopsin. Crystal structures of both β2-AR [7] and A2-AR [9] also suggest that 6.48 adopts a χ1 = g+ conformation in its inactive state, suggesting that the χ1 rotamer change from g+ to *trans* is synonymous with GPCR activation. Mutation studies have shown that the mutation of 6.48 to alanine significantly reduces constitutive activity and agonist-induced receptor activation in class A GPCRs GPR119, GPR39, β2-AR, and ghrelin without affecting ligand binding [28], further implicating this residue's importance in activation. Other residues are suggested to be involved with regulating this rotameric change in 6.48. An example is F3.36(200) in CB1, which has been suggested in modeling studies by the Reggio lab to act as a regulatory residue which transitions from a *trans* to a g+ χ1 rotameric state during activation [29]. While in a χ1 = *trans* rotameric state, F3.36(200) constrains W6.48(356) in a χ1= g+ rotamer, impeding activation. F3.36A(200) mutation studies have shown an increase in constitutive activity, suggested that F3.36(200) may be a toggle switch residue in CB1 [30]. Additionally, the aromatic residues often conserved at residue positions 6.52 and 5.47 have been suggested to be part of a rotameric toggle switch. The crystal structures of the β2-AR[31], A2-AR[9], D3-R[10], and CXCR4-R[11] suggest that in an inactive state the aromatic residue at locant 6.52 adopts a χ1 = g^+ conformation, which directly blocks the rotamer change of W6.48. Mutation studies of the β2-AR [26] and the D2 dopamine receptor [32] support

that 6.52 is associated with the rotameric change of 6.48, suggesting that these residues may change rotameric states in tandem. Additionally, the conserved aromatic residue at locant 5.47 is suggested to stabilize the activated, 6.48 $\chi 1$ = *trans* rotamer in the class A GPCRs GPR119, GPR39, β2-AR, and ghrelin [28].

The focus of the thesis research reported here is the class A GPCR, GPR35. GPR35 is a recently identified [33] and de-orphanized [34] GPCR. GPR35 has been suggested to be involved in metabolism [35], heart failure [36], inflammation [37], asthma [38], a mental retardation syndrome associated with the deletion on 2q37.3 [39], type II diabetes [40], as well as gastric cancer formation [41]- making GPR35 a potential target for the treatment of multiple diseases. Due to the extremely broad range of GPR35 physiological involvement, understanding its mechanisms of action and ligand interactions may result in a wide diversity of therapeutic applications, including the development of potential pharmaceutics.

Human GPR35 (hGPR35) has been shown to exhibit prominent expression in immune and gastrointestinal tissues [34], with additional expression in pancreatic islets, skeletal muscle, lung tissue [42], and the dorsal root ganglion [43]. The rat GPR35 (rGPR35) analog, which has 72% sequence identity with hGPR35 [43], has been shown to have expression in similar tissues as with hGPR35, but with the addition of high to moderate expression levels in uterine and neuronal tissues respectively [43]. There have been two isoforms of hGPR35 identified, GPR35a and GPR35b. The GPR35b splice variant was first identified from a human gastric cancer cDNA library [41] and contains an additional 31 amino acids at the N-terminus. HEK293 cells transfected with both hGPR35 and chimeric G-protein α-subunits have suggested that hGPR35 couples to G i/o proteins [43]. This was supported by the inhibition of GPR35 agonist dependent [Ca^{+2}] attenuation by *Bordetella pertussis* toxin in hGPR35-transfected rat sympathetic neurons expressing only native G-proteins [44]. This was recently contrasted in 2011 by data published by

Jenkin *et al*., which suggest that the pertussis toxin-sensitive G proteins were only able to elicit a very modest signal in a [^{35}S]-GTPγS assay following GPR35 activation in a similar HEK293 system [45]. Further experimentation, involving both the chimera and native $G\alpha_{13}$ subunits in a HEK293 cell system, suggest that the $G\alpha_{13}$ subunit couples to the active GPR35 complex [45].

Preliminary unpublished data from a collaborator of the Reggio lab, Mary E. Abood, had originally suggested that GPR35 might be a candidate cannabinoid receptor. GPR55, which has recently been hypothesized and experimentally suggested as a novel cannabinoid receptor, shares 27% sequence identity with GPR35 and is mapped to the same chromosome, 2q37. Preliminary experiments suggested that human GPR35-transfected HEK293 cells demonstrated cannabinoid-stimulated GTPγS binding with two non-classical cannabinoids, CP55940 and Win55212-2- with CP55940 yielding a high potency ($EC_{50} \approx 14$nM). Unfortunately these results were not reproducible and have been contrasted in 2010 by those of Sonoda *et al* [46], whose data suggests that non-classical cannabinoids, CP55940 and the aminoalkylindole,WIN55212-2, are not able to elicit an increase in free intracellular [Ca^{2+}] in hGPR35-transfected HEK293 with ligand concentrations as high as 10 μM. The cannabinoid receptor ligand Δ^9THC, the principal psychoactive component of marijuana, did produce a detectable increase in intracellular [Ca^{2+}], but at concentrations of 5 and 10μM. These data suggest that GPR35 may be in a class separate from the cannabinoid receptors.

The first possible endogenous ligand of GPR35 to be identified, kynurenic acid [37], is one of the major metabolites of the kynurenine pathway; a pathway which is the main route of tryptophan catabolism and has been associated with important physiological roles in the brain. While kynurenic acid (Figure 3) has been shown to be a moderate to a low potency agonist in hGPR35, with an EC_{50} of 36 - 39 μM [34], it has been shown that hGPR35 elicits firm adhesion of leukocytes *in vitro* at kynurenic acid concentrations as low as 300 nM [37]. The baseline

concentration of kynurenic acid in human plasma is approximately 25 nM [47-48], however, during inflammation the catabolism of tryptophan promotes the generation of kynurenic acid, which can increase kynurenic acid's plasma concentration to the μM range [49-50]. These data suggest the conjoined roles of GPR35 and kynurenic acid in regard to their involvement in inflammation, however the potency at which kynurenic acid agonized the diffusely distributed hGPR35 is moderate at best. This suggests the possibility of the existence of a more potent endogenous ligand.

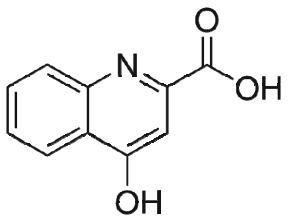


Figure 3. The structure of kynurenic acid.

A significantly more potent possible GPR35 endogenous ligand was recently identified as lysophosphatidic acid (LPA). LPA has assorted physiological properties. These properties range from platelet aggregation, contraction of smooth muscle, stimulation of cell chemotaxis, and a number of cellular G-protein regulated secondary responses. These include, but are not limited to, the stimulation of phospholipases, the mobilization of intracellular Ca^{2+}, the modulation of adenyl cyclase, and the direct involvement in the MAP kinase cascade [51-52]. Of the derivatives of LPA examined, the most potent species identified were 2-acyl LPA. The most potent of the 2-acyl LPA derivatives that were identified was 2-linoleoyl LPA (Figure 4) and 2-oleoyl LPA, each with an EC_{50} ranging from 30 – 50 nM [53]. It has been suggested that 2-acyl LPA activates the

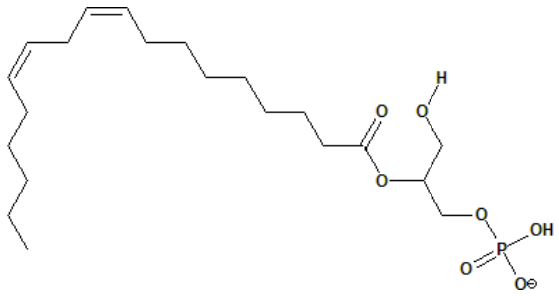


Figure 4. The Structure of 2-linoleoyl lysophosphatidic acid (LPA)

GPCRs P2Y5 [54] and LPA3 [55], with an EC_{50} in the μM or nM range, respectively, which may suggest a similarity between these receptors and GPR35. Several types of phospholipase A1, which selectively generate 2-acyl LPA from the ubiquitous phosphatidic acid (PA), have been identified, suggesting that 2-acyl LPA is a common molecule [46, 56].

There have also been several exogenous ligands for GPR35 that have been identified. The first is the well characterized cyclic guanosine monophosphate-specific phosphodiesterase (cGMP-PDEs) inhibitor, zaprinast [43]. Zaprinast (Figure 5) is a moderately potent agonist for hGPR35, with an EC_{50} ranging from 0.84 – 2.6 μM [43, 57-59] and a significantly more potent agonist in rGPR35 with an EC_{50} ranging from 16-67 nM [43, 58-59]. cGMP-PDEs degrade the intracellular second messenger cGMP, either attenuating or terminating the cellular response associated with this second messenger pathway; therefore, a cGMP-PDEs inhibitor may act to modulate this pathway [60]. Due to zaprinast's promiscuity and its moderate potency, this ligand may not be useful as a tool to understand GPR35 signaling and/or provide a chemical backbone for a potent, specific therapeutic ligand.

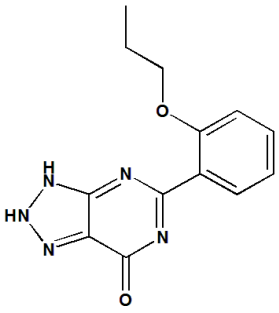


Figure 5. The Structure of Zaprinast

Another recently discovered moderately potent hGPR35 and rGPR35 agonist, with respective EC_{50}s of 4.4-12 μM [58-59] and 16 μM, is the chloride channel blocker 5-nitro-2-(3-phenylpropylamino)benzoic acid (NPPB) [58]. A previous pharmacological study looking at NPPB's (Figure 6) effect on the constriction of rat small pulmonary arteries, has shown activity suggested to be based on the inhibition of N-type calcium channels independent of the blocking of chloride channels [61]. Activation of GPR35 in transfected hGPR35 cells has been shown to activate native $G_{i/o}$, which in rat sympathetic neurons has been suggested to lead to $[Ca^{2+}]$ modulation through the inhibition of N-type calcium channels. These data suggest that GPR35 may naturally be involved in this mechanism. While these data may suggest another potential mechanism of GPR35, NPPB is limited as a tool to study the mechanisms of GPR35 for similar reasons as Zaprinast- promiscuity and moderate potency.

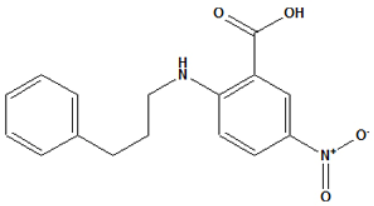


Figure 6. The structure of 5-nitro-2-(3-phenylpropylamino)benzoic acid (NPPB)

Two additional ligands that have been found to be GPR35 agonists are the asthma drugs cromolyn sodium (Intal) [38, 57] and nedocromil sodium (Tilade) [38], each with an EC_{50} measured to be similar to Zaprinast. Cromoglicic acid (Figure 7B) is suggested to have a similar potency in hGPR35 and rGPR35 [38, 57], while nedocromil sodium (Figure 7A) is suggested to be more potent in rGPR35 as compared to the human orthologue [38]. Both of these drugs are thought to act as mast cells stabilizers, preventing the release of inflammatory stimulating chemicals such as histamine, though the target nor the exact mechanism of action of either of these drugs is known definitely [38]. These data suggest that GPR35 may be involved in the disease of asthma, but due to the ambiguity of the target of these drugs, the degree of GPR35 selectivity is uncertain. For this reason and due to their moderate potency for GPR35, these drugs are not ideal candidates to help understand the mechanisms of GPR35.

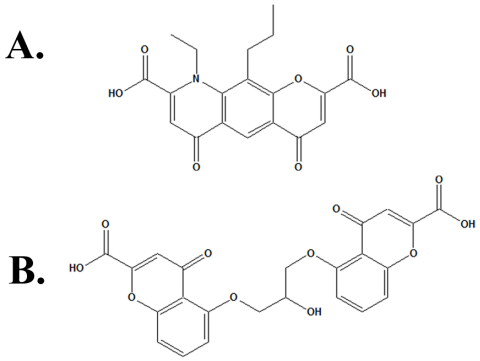


Figure 7. The structures of (A) Nedocromil (Tilade) and (B) Cromoglicic acid (Intal).

A potent hGPR35 agonist, pamoic acid (Figure 8), originally identified by a collaborator with the Reggio lab, Mary Abood [59], is novel compared to the other currently known ligands of GPR35. Pamoic acid is more potent than any other currently known exogenous ligand for GPR35, with calculated EC_{50} values of between 51 - 79 nM [57, 59] and no additional currently known target receptors. Currently, the Food and Drug Administration has pamoic acid classified as an

inactive compound that is packaged with multiple approved drugs to create longer-acting formulations [62]. A truncated list of several drugs packaged with pamoic acid consists of the anthelmintics oxantel pamoate and pyrantel pamoate, the antihistamine hydroxyzine pamoate (Vistaril), and the antidepressant imipramine pamoate (Tofranil-PM) [59]. While pamoic acid is a potent agonist for hGPR35, it is virtually inactive in rGPR35 [57]. Since rGPR35 shares a 72% sequence identity with hGPR35, the divergences in the sequences between the two species may provide important clues to GPR35 structure. Pamoic acid's high potency and selectivity for hGPR35 suggests that it may represent a powerful tool that can lead to the understanding of the receptor's mechanisms of action and may represent a chemical scaffold for future drug therapies.

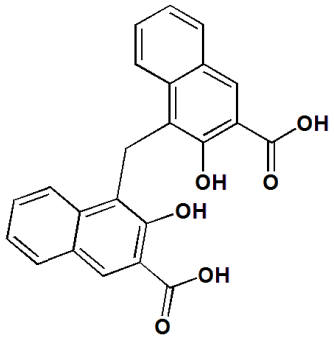


Figure 8. The structure of pamoic acid

A potent antagonist MLS- 0370945 (Figure 9), with an IC_{50} of 72 nM, was recently discovered in a high-throughput drug screen by a collaborator of the Reggio lab, Mary Abood, in conjunction with the Molecular Libraries Probe Production Center at the Sanford-Burnham Institute (AID:2079). This ligand was identified via a dose-dependent β-arrestin-GFP assay with a fixed 10 μM concentration of the agonist zaprinast.

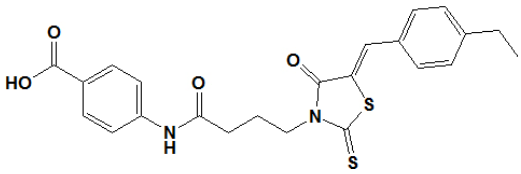


Figure 9. The structure of the compound MLS- 0370945.

Recent mutation data in both the human and rat GPR35 receptors have suggested both binding location and, in conjunction with an accurate model, possible direct interaction residues for the ligands zaprinast and pamoic acid. Mutation data of hGPR35 from the Reggio lab collaborator, Mary Abood, has suggested several residues that may be important in ligand binding and or activation [63]. The residues R6.58(240), R4.60(151), R3.36(100), and K7.40(263) (Figure 10) were all mutated separately to alanine and stably expressed in U2OS cells. The potency of the agonists zaprinast and pamoic acid were assessed for each mutant. Mutant R3.36A(100) was not recruited to the cell surface, so it is speculated that this mutation may have led to the misfolding of the protein. Data published in 2011 by Jenkins *et al* suggests that in stably transfected HEK293T cells, both the human and rat GPR35 R3.36 (human/rat, 100/97) mutant receptors are recruited to the cell surface, but with a trend in the reduction of surface expression as compared to the wild-type receptor, with a significant reduction in the rat GPR35 mutant [45]. Though these mutants showed expression on the cell surface, neither receptor was able to be activated by the agonists pamoic or kynurenic acid. Considered together, both the trend in the reduction of the receptor cell surface expression and the inability to be activated by wild-type agonists suggest that this mutation may indeed cause protein misfolding. The R4.60A(151) mutant expressed on the cell surface, but was unable to be activated by either agonist tested, suggesting that this residue may be involved in ligand binding, protein folding, and/or protein stabilization. The

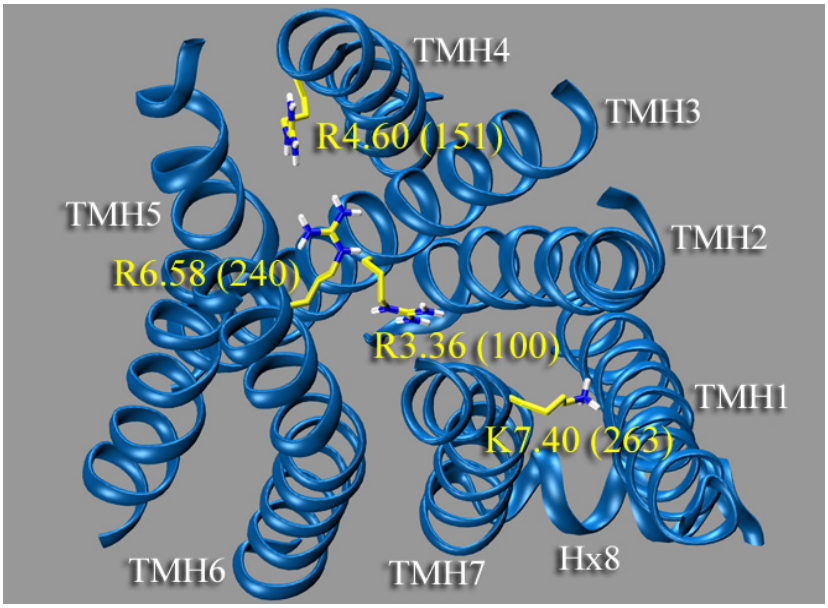


Figure 10. Locations of the mutated residues in hGPR35 by experiment collaborator, Mary Abood.

R6.58A(240) mutant showed a statistically significant increase in the potency of zaprinast, while not affecting pamoic acid. The K7.40A(263) mutant showed no statistically significant difference in potency for either zaprinast or pamoic acid, suggesting that K7.40(263) is not a direct interaction residue for these ligands. Since many of the ligands for GPR35 are anions and K7.40(263) is the only positively charged residue in the TMH1-2-7 region of GPR35, a lack of change in potency in the K7.40A(263) mutant suggests that the binding pocket of GPR35 lies rather in the TMH3-4-6 interface, where there are potentially four arginine residues available for direct ligand interaction- residues R3.36(100), R4.60(151), R6.58(240) and R(164) located on the EC2 loop. The recent data published by Jenkins *et al* suggest that Y3.32(human/rat, 96/93) may be a direct interaction site for zaprinast [45]. The Y3.32L(93) mutation resulted in an approximately 60-fold increase in the EC_{50} of zaprinast for rGPR35, suggesting that this may be a direct interaction site. In the hGPR35 Y3.32L(96) mutant receptor, the ability to be activated by

zaprinast was undetectable, though this may be due to the potency difference between rat (16-67 nM) and human GPR35 (0.84 – 2.6 μM). If this mutation caused a similar 60-fold increase in the much higher EC_{50} of zaprinast for hGPR35, than this may increase the concentration of zaprinast needed to elicit a response beyond a reasonable level.

To first assess the ligand binding data and make well-informed mutation recommendations for GPR35, an initial *in silico* homology model based on a rhodopsin template was constructed in the Reggio lab. GPR35 and rhodopsin share most of the highly conserved residues and motifs found in class "A" GPCRs, but have two major sequence variations with the addition and the lack of a proline in TMH1 and TMH2 respectively. The GPR35 model developed here is based on two different crystal structure templates; the A2-AR crystal structure [9] for TMH7 and Hx8 and the β2-AR crystal structure for the remaining TMHs [7]. While the β2-AR shares a relatively low sequence identity of 14% with GPR35, the two receptors share multiple conserved residues across all TMHs including the highly conserved residues N1.50, D2.50, W4.50, P5.50, P6.50 and the conserved E/DRY motif on TMH3. The GPR35 and the β2-AR sequences also share a conserved cysteine at position 3.25, which in the β2-AR crystal structure is involved in a disulfide bridge to a cysteine on the EC2 loop. GPR35 also has a cysteine in the EC2 loop and is presumed here to be involved in a disulfide bridge with C3.25(89). GPR35 and the β2-AR, but not the A2-AR receptor, also shares the conserved residue F5.47, which has been suggested to stabilize the active bundle [28]. The A2-AR receptor was used as a template in lieu of the β2-AR for TMH7 and Hx8 because GPR35 and the A2A receptor both have a two residue elbow region connecting TMH7 and Hx8, while the β2-AR only has one.

While GPR35 shares multiple conserved residues and motifs with other GPCRs that have available crystal structures, there are major sequence variations in several TMHs of GPR35 that do not have a crystal structure counterpart. The GPR35 receptor sequence differs from the

available GPCR crystal structures mainly in the positions of common proline residues in TMH2, 4, 5, and 7 (Table 1). The lack of the proline in the highly conserved NPXXY motif of TMH7 in the GPR35 sequence, which is DAICY at the analogous residue positions, is predicted to have the largest structural variation from the other known GPCR structures. While the addition, subtraction, and shift of prolines in the GPR35 sequence are likely to cause variant helical distortions as compared to other GPCRs structures, there are also multiple non-conserved threonines, glycines, and serines in TMH1, 2, 4, 5, and 6 that have the potential to exert subtle, but potentially important helical deformations. Residues serine and threonine (*i*) in a $\chi 1$ = g- conformation may result in a hydrogen bond between the side chain hydrogen bond donor of the (*i*) residue and carbonyl oxygen of the residue located at approximately one turn towards the N-terminius (position (*i -3*) or (*i-4)*), resulting in the induction and stabilization of a 3-4° bending in the helix [64]

Table 1. GPR35 TMH 2, 4, 5, and 7 sequence position differences of common proline residues.

Helix Number	Change in Proline as compared with GPCRs with available Crystal Structures
TMH2	P2.58 in GPR35 instead of 2.59
TMH4	R4.60 in GPR35 instead of P4.60
TMH5	presence of non-conserved P5.43 in GPR35
TMH7	Lack of NPXXY motif ; this is replaced with DAICY in GPR35

.

CHAPTER II

HYPOTHESES & METHODS

Hypotheses

The ultimate focus of this project was the development of human and rat GPR35 models that are consistent with specific known pharmacology of this important receptor. To this end, three separate homology models, a human GPR35 R, a human GPR35R*, and rat GPR35 R* model, were constructed and refined. Conformational analyses were performed on the agonists pamoic acid and zaprinast and the antagonist MLS-0370945 to calculate plausible conformations available at physiological conditions. Docking studies with each receptor were performed to test the following hypotheses: 1) the increase of the potency of zaprinast in the hGPR35 R6.58A mutant receptor is due to the excision of the bulk of the R6.58 arginine residue, which allows for additional interactions with the ligand, 2) the increase and decrease of the potency of zaprinast and pamoic acid respectively at rGPR35 vs. hGPR35 is primarily due to the conformational differences in TMH5 caused by the lack of the non-conserved proline on P5.43 in rGPR35, and 3) the statistically equivalent potencies of pamoic acid for the wild-type and R6.58A mutant hGPR35 receptors is due to a similar binding affinity of pamoic acid in each of these receptors.

Methods

Modeling of GPR35 using GPCR crystallization data

During the initial development of the model presented here, only a few GPCRs had been crystallized, including Rhodopsin (Rho) [2-4], Opsin [6] the β2-adrenergic receptor (β2-AR) [7, 31, 65], the β1-adrenergic receptor (β1-AR) [66], and the A2-adenosine (A2-AR) receptor [9]. The model of the activated form of GPR35 is based on the 2.4 Å and 2.6 Å resolution crystal structures of the A2-AR [9] and β2-AR [7], respectively. The initial, canonical comparison of GPR35 with other class "A" GPCRs was based on a sequence comparison. The GPR35 sequence was aligned with the sequences of Rho, β2-AR, A2-AR, and the CB1 and CB2 receptors, using highly conserved residues as templates as done previously for the CB1 [67] and CB2 [68] receptors. GPR35 contains several of the highly conserved residues patterns in TMHs 1, 2, 3, 4, 5 (N1.50, D2.50, (E) DRY motif in TMH3, C3.25, W4.50, and P5.50) of other crystallized GPCRs, with notable sequence variations including: 1) a conservative substitution (CFLP) for the TMH6 CWXP motif, 2) a non-conservative substitution (DAICY) for the TMH7 NPXXY motif, and 3) a non-conservative proline addition in TMH5 (P5.43). Additionally, the GPR35 extracellular (EC) EC-1 loop is shorter than most (2 amino acids (aa) vs. 6 in β2-AR and Rho) and the GPR35 EC-2 loop is longer than most (GPR35 11 aa vs. 5 aa in β2-AR, 6 aa in Rho and CB1/CB2, 7 aa in A2-AR).

Conformational Memories (CM)

In order to explore the consequences of a change in the location of a helix deforming residue from the template structure, the Conformational Memories technique was employed. The CM technique uses multiple Monte Carlo/simulated annealing random walks employing the CHARMM all-atom force field [69] in a distance dependent dielectric at 310 K. This method allows for an exploration of the available conformational space of each TMH and is able to converge in a reasonable number of steps. To ensure a reasonable computational time, the Monte Carlo method is preferred over a systematic approach due to the flexible nature of these large molecules. The inclusive rotation of the large number of "degrees of freedom" is approximated through numerous steps in which two dihedrals and one bond angle are rotated at random. Simulated annealing allows for the exploration of all allowed amino acid rotamers and backbone dihedrals by energizing the system to overcome potential energy barriers. The acceptance or rejection of each conformer is based on the Metropolis criterion [70], so the energies of the accepted structures are representative of a Boltzmann energy distribution at the measured temperature. This ensures that the final accepted conformers are reasonable at physiological temperatures. The CM calculations were performed in two phases: the exploratory and the biased phase. In the exploratory phase, the simulated-annealing process has an initially high starting temperature of 3000 K to ensure that the range of each allowed angle is capable of being explored. The accepted conformers from the exploratory phase are stored as "memories" and are used to restrict the conformational search performed in the biased phase. A more detailed description of the CM method is provided below:

In the exploratory phase, the range that each angle is capable of exploring is determined by starting at a temperature of 3000 K and decreasing to 310 K in 18 steps. For each temperature

step, 50,000 Monte Carlo moves are performed in which two dihedrals and one bond angle are varied. Accepted conformers are stored as "memories" and are used to restrict the bond angles and dihedrals that are explored in the biased phase. The "memories" are a representation of a probable distribution of different angles and dihedrals of each TMH as a function of temperature.

The biased phase begins at a temperature of 750 K and decreases to 310 K in 7 steps. As with the exploratory phase, for each temperate step 50,000 Monte Carlo moves were performed and the conformers were accepted or rejected based on the Metropolis criterion. In the biased phase only those dihedral and angle ranges that were suggested to be accessible at 310 K in the exploratory phase were explored. In each CM calculation, a total of 105 TMH conformers were generated.

The sampling ranges for backbone dihedrals Φ (phi), Ψ (psi), ω (omega) were standard values of ±10°, ±10°, ±20°, respectively, and the side chain dihedrals were allowed to vary ±180°. The initial dihedral angles used in CM were transformed from the starting angle of the pre-generated structure to an integer in the range of -180° to 180°; which was restricted to five degree increments (i.e. -5, 0, 5). An example of this is a pre-generated starting structure dihedral angle of -13.4° would be transformed to -15°. The bin ranges were 5°, which gave a minimum range of motion of 10° for explored dihedrals. Bond angle variation ranges were defined as either ±8° for default bond angles or ±15° for sidechain bond angles involving polar hydrogens (i.e. C-O-H: Ser, Thr, Tyr) or the flexible C-S-C bond angle in methionine as described by Whitnell *et al* in 2008 [71]. To avoid the sampling of high energy bond angles, various restrictions were implemented as described by Whitnell *et al* [71]. The bond angles present in aromatic rings and those involving non-polar hydrogens in methylene and methyl groups were not sampled. Based on the examination of all-atom and united-atom topology files [71], a pre-biased "memory" was applied to all bond angles restricting the ranges to 90° – 144° with 1.5° bin ranges.

Analysis of Conformational Memories Transmembrane Helicies

Analyses of helix bends were performed by the program ProKink [72], which is embedded in the Simulaid Conversion program [73]. This method quantifies three attributes of the helix's coordinates as compared to a "hinge" residue; the wobble angle, the face-shift, and the bend angle (Figure 11). As the first point of reference, the α-carbon of the hinge residue is translated to the Cartesian origin and the "pre-hinge" central helical axis (HAx) is aligned parallel with the positive x-axis. The HAx is calculated by the Kahn method [74]. The bend angle (Figure 11B) is measured by first aligning the "post-hinge" HAx axis to the x,y plane. The angle that the "post-hinge" HAx axis makes with the "pre-hinge" HAx axis is defined as the bend angle. At 0°, the hinge is straight and 180° the angle is bent (bent angle). The measurement of the wobble angle (Figure 11A) requires the parallel alignment of the "pre-hinge" HAx to the x-axis and the alignment of the α-carbon of the hinge residue with the y-axis. The wobble angle (-180°,-180°) is the angle that the "pro-hinge" HAx makes with the *y ,z*-plane, with 0° representing the "pro-hinge" HAx moving towards the α-carbon of the hinge residue. The final attribute, the face-shift (Figure 11C, D, and E), is calculated by aligning the "post-hinge" HAx with the negative x-axis. The α-carbon of the hinge residue is aligned with the y-axis and angle between this and the projection of the average of the vectors of the α-carbons of the (i – 3) and (i – 4) represents the face-shift (0°-180°). A value close to 0° is an ideal helix while a positive or negative value represents an under or over wound helix.

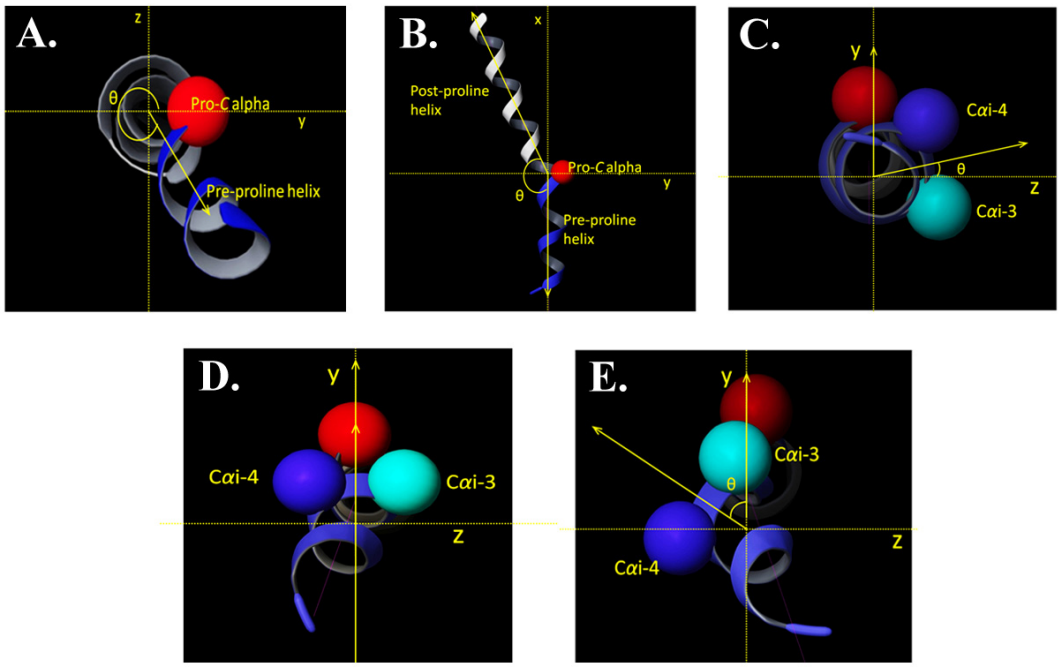


Figure 11. Representation of the three quantifiable attributes of a proline-kink as defined in the program PROKINK. The attributes are defines as A) Wobble Angle, B) Bend Angle, C) Face-shift (over-wound), D) Face-shift ((ideal), E) Face-shift (under-wound).

Implementation of the CM Technique to Calculate the TMH Region of the GPR35 Model

For TMH 1, 2, 4, 5, 6, and 7, a series of possible, sequence-dictated conformations that deviated from the β2-AR or A2-AR template were calculated using the CM simulated annealing technique [71, 75]. In this technique, nearly all dihedrals and bond angles are defined as variable, with various restrictions. The dihedral ranges allowed were based on the sequence of the TMH being explored. If the sequence contained a known helix deforming residue such as proline or glycine, a larger variation range of up to ±50° on the Φ and Ψ was implemented to permit increased flexibility. The proline residue often creates the largest helical distortion. This helical distortion, often referred to as a proline kink, is suggested to affect the Φ and Ψ dihedral angles of the proline (*i*) to the fourth residue preceding it (*i – 4)*. The justifications for the ranges that were explored for TMH 1, 2, 4, 5, 6, and 7 are described below.

TMH 1, 2, 4, and 7

Due to either the lack of a proline or the a lack of a homologous proline locant found in any available crystal structure, TMH1, 2, 4 and 7 used the ideal phi/psi angles of -63°/-41.6° [12] as starting structures.

TMH7**:** TMH7 in GPR35 lacks the conserved P7.50 in the NPXXY motif and instead has the non-conservative substitution of DAXXY. The TMH2 of Rhodopsin, which lacks the conserved proline, contains a sequence of Gly-Gly which acts as a proline analog in terms of its increased flexibility. The Gly-Gly motif in the TMH2 of Rhodopsin creates a helix distortion which emulates the shape of the proline containing TMH2 of the β2-AR and the A2-AR. To determine if the GPR35 TMH7 sequence may dictate a proline-like distortive quality analogous to the Gly-Gly motif of the TMH2 of rhodopsin, a sequence a flexible region (Φ/Ψ ± 50°) was introduced in the virtual *i* to (*i -4*) region (C7.46(269) - A7.50(273)) of TMH7. TMH7 also contains two possible helix distorting residues, S7.39(262) and S7.42(265). These residues were both held in χ1 = g- independently and concurrently to access their ability to distort the helix.

TMH2: GPR35 has a proline that is shifted by one residue extracellularly as compared to the β2AR sequence: P2.59 in the β2AR and P2.58 in GPR35. A flexible region (Φ/Ψ ± 50°) was implemented into the P2.58 (*i*) to L2.54 (*i – 4*) region to mimic the flexibility imposed by the proline.

TMH4: GPR35 does not have the conserved proline located at the extracellular end of the β2AR TMH4. To determine if a sequence dictated conformation might emulate the distortion

typically associated with this conserved proline, the helix deforming potential of the G4.55(146) –R4.60(151) region was analyzed. Two independent conformational searches were performed; the first search explored what effect the increased flexibility of the backbone (Φ/Ψ of ± 50°) in the region of G4.55(164) – R4.60(151) would have on the shape of TMH4 and the second conformational search used the same protocol as the first, but with the addition of S4.56 held in g-.

TMH1: Due to the straightening of the adjacent TMH7, the conformation of TMH1 was explored. TMH1 has three non-consecutive glycines located at positions 1.34, 1.39, and 1.46. Since all of the glycines were non-consecutive, three independent calculations were performed. In each calculation a single glycine was given increased flexibility of the backbone (Φ/Ψ of ± 50°) and assessed independently.

TMH 5 and 6

In TMH5 and 6, the proline locations in GPR35 are inclusively conserved in all currently available GPCR crystal structure sequences. The dihedral angles associated with the proline kink in these helices were examined from Rhodopsin [2], A2-AR [9], and β2-AR[7] and were used to constrain the range of the possible sampling of dihedral angles. The exploratory ranges of the *i* – (*i*-4) backbone dihedral angles of TMH5 and 6 were restricted to those ranges found in their respective crystal structure TMH (Table 2 and Table 3). These constraints allowed for CM to generate conformations of helices with structural similarities to the available crystal structures.

Table 2. Restricted ranges of Conformation memories for TMH5 based on crystal structure ranges. Crystal structure ranges based on the crystal structures values of TMH5 in the β2-AR (2RH1): 2.40 Å, Rhodopsin (1GZM): 2.65 Å, A2-AR(2EML):2.60 Å, and Turkey β1-AR (2VT4): 2.70 Å

Residue Position	Crystal Structure Range: ϕ	CM Range Explored: ϕ	Crystal Structure Range: ψ	CM Range Explored: ψ
i-4	-107.7° to -79.4°	-120° to -70°	-35.5° to -23.1°	-40° to 20°
i-3	-150.6° to -116.6°	-160° to -110°	-70.5° to -54.1°	-70° to -40°
i-2	-63.0° to -57.6°	-80° to -50°	-61.9° to -45.7°	-70° to -40°
i-1	-63.8° to -51.3°	-70° to -40°	-61.5° to -48.7°	-70° to -50°
i (P5.50)	-63.2° to -53.2°	-70° to -40°	-43.5° to -24.8°	-60° to -20°

Table 3. Restricted ranges of Conformation memories for TMH6 based on crystal structure ranges. Crystal structure ranges based on the crystal structures values of TMH6 in the β2-AR (2RH1): 2.40 Å, Rhodopsin (1GZM): 2.65 Å, A2-AR(2EML):2.60 Å, and Turkey β1-AR (2VT4): 2.70 Å

Residue Position	Crystal Structure Range: ϕ	CM Range Explored: ϕ	Crystal Structure Range: ψ	CM Range Explored: ψ
i-4	-65.7° to –54.6°	-70° to -50°	-41.7° to -25.4°	-50° to -20°
i-3	-91.5° to -71.7°	-100° to -60°	-45.1° to -19.4°	-50° to -10°
i-2	-119.0° to -97.3°	-120° to -80°	-29.1° to -26.1°	-30° to -20°
i-1	-55.8° to -51.9°	-60° to -50°	-49.0° to -40.6°	-60° to -30°
i (P5.50)	-57.1° to -47.1°	-70° to -50°	-51.3° to -45.7°	-60° to -40°

TMH5: TMH5 has two proline residues; the conserved proline at position 5.50 and an additional proline at 5.43. The additional proline at 5.43 was assumed to induce a local conformational change that is independent of P5.50(183). The flexibly of the proline kink of the conserved proline (G5.45(178) to P5.50(183)) was restricted to empirically based backbone dihedral ranges and the proline kink of non-conserved proline (S5.39(172) to P5.43(176)) was given the full increased flexibility of the backbone (Φ/Ψ of ± 50°).

TMH6: TMH6 in GPR35 lacks the highly conserved CWXP motif and has a conservative substitution of CFXP. The highly conserved proline in TMH6, P6.50(232), is expected to provide a large amount of flexibility of GPR35's TMH6. GPR35 contains only one other potential helix deforming residue, G6.54(236), which was given an increased backbone flexibility of ± 50° for the Φ and Ψ dihedral angle in all calculations. The importance of exploring the possible conformational space of TMH6 lies in the structural changes associated with the transition from an inactive (R) to an active (R*) GPCR bundle. In β2-AR and rhodopsin, the most dramatic structural change associated with activation is the general rotation and straightening of TMH6 about the highly conserved CWXP hinge region [13-20]. Two independent conformational searches were performed with an identical, ideal helix starting structure to be representative of either an R or R* state. The inactive TMH6 is expected to both have a similar structure to the antagonist bound GPCR crystal structures and to have the suggested toggle switch residue F6.48(230) in a $\chi 1$ = g+ rotameric state. To achieve a set of conformers that met these criteria the flexibly of the i – (*i*-4) region (V6.46 to P6.50) of TMH6 was restricted to the empirically based backbone dihedral ranges (Table 3) and the residue F6.48(230) was held in a $\chi 1$ = g+ rotameric state. To calculate the possible TMH6 conformations for the R* model the standard ± 50° for the Φ and Ψ dihedral angles of the *i* – (*i*-4) region were explored.

Construction of the TMH region of GPR35

The conformers generated by the CM calculations were used to construct the TMH bundle region of the receptor. The receptor was assembled by aligning the intracellular region of TMHs 1, 2, 4, 5 or TMH7 with their respective crystal structure templates of the β2-AR or the A2-AR. For all superimpositions, the C, C_α, and backbone N atoms of the TMH conformers calculated by CM were superimposed with the corresponding atoms of the analogous residues on the crystal structure template. The region chosen for the backbone alignment of each TMH was variable. TMH1, which lacks a proline, was aligned with the backbone of the residues that correspond with the GPR35 sequence G1.46(34) - F1.57(45). This alignment was chosen to ensure that the highly conserved N1.50 was aligned properly, but the potential conformational impact from the highly flexible glycines at residue position G1.34(22), G1.30(18), and G1.46(34) would be explored. For TMHs 2, 4, 5, and 6, which contain a proline in the crystal structure template, the region chosen for superimposition was based on avoiding the region distorted by the proline. The regions chosen for alignment for TMH2, 4, and 5 were as follows: TMH2: E2.38(54) - L2.57(73), TMH4: R4.40(131) - W4.50(141), and TMH5: P5.50(183) - A5.66(199). For TMH6, the residues chosen for superimposition varied between the active and inactive state. For the inactive bundle, the CM calculated helices were superimposed intracellularly with the β2-AR crystal template residues that correspond with the GPR35 sequence T6.30(212) - V6.46(228). For the active bundle, the CM calculated helices were superimposed extracellularly with the backbone of the residues P6.50(232) - V6.61(243) of the final GPR35 inactive (R) bundle. The TMH6 of the GPR35 R bundle was chosen in lieu of the β2-AR due to the biophysical data for rhodopsin, that suggests that the extracellular region of more stationary TMH5 [18] and the dynamic TMH6 are in the same relative orientation in either the R and R* state [76]. As discussed earlier, the template used for the alignment of TMH7 was the A2-AR receptor in lieu of the β2-

AR, due to the more similar elbow length. TMH7, which has the non-conservative substitution DAICY for the highly conserved NPXXY motif, was aligned with the proline kink region due to the energetically unfavorable large gap that occurred between the extracellular tops of TMH7 and TMH6 when the superimposition avoided the proline distortion. The CM calculated conformers were superimposed with the backbone of the residues in the 7.45 to 7.57 region of the A2-AR template.

The TMH3 developed as the starting structure for the GPR35 homology model was based considerably on the β2-AR template. The TMH3 of the β2-AR crystal structure, N3.22(103) to S3.56(137), was mutated to the GPR35 sequence and, after rotameric changes of residues to eliminate overlap, was energy minimized for 500 iterations employing the OPLS2005 all atom force field, extended cutoff (nonbonded: 8.0Å, electrostatic: 20.0Å, hydrogen bonding: 4.0Å), and a distance dependent dielectric in Macromodel 9.1 (Schrodinger Inc., Portand, OR). A harmonic constraint of 1000 kcal/mol was implemented on the backbone dihedrals to maintain the backbone conformation of the crystal structure.

The CM TMH conformers for each explored helix were aligned to their crystal structure template, as described above, and the iterative process of choosing the final transmembrane helices was based on two criteria: 1) eliminating the Van der Waals overlaps with other helices within the bundle and 2) creating an overall topography conducive for the inclusion of the intra- and extracellular loops. Following the choice of the CM conformers, a modification on TMH7 was made. Comparison of the TMH7 sequence of the A2-AR crystal structure and GPR35 showed a homologous sequence motif (A2-AR: S7.42(277), H7.43(278), T7.44(279), N7.45(280), S7.46(281) vs. GPR35: S7.39(262), K7.40(263), L7.41(264), S7.42(265), D7.43(266)) which in the A2-AR created a helical deformation independent of the proline kink initiated by P7.50. To emulate this motif the backbone dihedrals (Φ and Ψ) of the GPR35 TMH7

were changed to match the corresponding dihedrals of the A2-AR crystal structure. Matching the motif found in the A2-AR crystal structure shifted the position of the charged residues K7.40(263) and D7.43(266) (which initially faced towards the lipid) into the likely hydrophilic TMH1-2-7 cavity. This helical manipulation also reduced the unfavorable gap between the intracellular tops of TMH6 and 7. Additionally, the crystal structure of the A2-AR helix 8 and the two residues that precede it were mutated to the GPR35 sequence and subsequently attached to the chosen CM TMH7.

Several waters were added to the bundle based on the positions and abundance of the crystallographic waters found in the A2-AR [9], β2-AR [31], and rhodopsin [2] crystal structures. An average of three water molecules was found in proximity to the highly conserved N1.50, D2.50, and N7.49 interface. These waters were involved in a hydrogen bond network with each other and with D2.50 and either N7.49 or N1.50, depending on the crystal structure. Three waters were placed in this interface of the GPR35 model, with hydrogen bonds formed with each other and the residues at loci 1.50, 2.50, and 7.49. There was also a single water molecule found near the highly conserved R3.50 in the aforementioned crystal structures. To emulate these, a single water molecule was placed in the GPR35 model near R3.50(114).

For the inactive (R) bundle, several highly conserved residue sidechain dihedrals were restrained to crystal structure ranges (Table 4). To emulate the highly conserved salt bridge found between the residues R3.50 and E/D6.30 in multiple GPCRs, a harmonic restraint of 12.5 kj/mol with distance constraints of 2.8 ± 0.4 Å and 7.0 ± 1.0 Å on the hydrogen bond donor/acceptor heteroatoms and Cα respectively of the residues R3.50(114) and T6.30(212) was implemented. Additionally for the R bundle, several residue interaction pairs/motifs, supplemental to the interaction between R3.50 and T6.30, were required to be possible with low energy sidechain rotamers. These interaction pairs/motifs include: 1) the mimicking of the aromatic stacking of

residues 5.47, 5.48, and 6.52 (GPR35:F5.47(180), Y5.48(181), H6.52(234) vs. D3/β2-AR F5.74, Y5.48, F6.52) found in the crystal structures of the β2-AR and human dopamine D3 receptor (D3-DR) [10]; 2) the emulation of an aromatic stack between the highly conserved residues Y7.53(276) and F7.60(283), the first residue on helix 8, found in the rhodopsin crystal structure and is suggested by mutation studies in the 5HT2C receptor [77]; and, 3) the imitation of the highly conserved D3.49(113) and R3.50(114) arginine cage motif of the E/DRY sequence, which is present in all current inactive GPCR crystal structures. Additionally, the χ1 of F6.48 and H6.52 were required to adopt a χ1 = g+ conformation, as suggested by GPCR crystal structures and other biophysical data.

Table 4. Average values of side chain values of conserved residues based on the crystal structure values. ± 1 standard deviation (σ) based on the crystal structure values from β2-AR (2RH1): 2.40 Å, Rhodopsin (1GZM): 2.65 Å, A2-AR(2EML):2.60 Å

Residue	Average χ_1	Average χ_2	Average χ_3	Average χ_4	Average χ_5
N1.50	-71.53° ± 2.18°	-69.43° ± 2.98°	N/A	N/A	N/A
L1.52	-163.8° ± 4.67°	N/A	N/A	N/A	N/A
D2.50	-66.87° ± 5.64°	-32.77° ± 3.51°	N/A	N/A	N/A
W4.50	-78.53° ± 3.51°	92.13° ± 6.84°	N/A	N/A	N/A
W6.48	-72.6° ± 4.47°	100.43° ± 4.16°	N/A	N/A	N/A
D/E3.49	174.43° ± 4.76°	64.87° ± 5.51°	N/A	N/A	N/A
R3.50	-69.83° ± 3.44°	-57.4° ± 2.07°	-77.2° ± 0.60°	-165.27° ± 0.96°	-0.03° ± 0.11°
Y3.51	175.07° ± 3.31°	83° ± 1.67°	N/A	N/A	N/A

The bundle was then pulled apart 3Å from TMH3 and then energy minimized for 3500 iterations to pack the TMHs and relieve side chain clashes. This energy minimization utilized the OPLS2005 all-atom force field in Macromodel 9.1 (Schrodinger Inc., Portland, OR) employing a

distance dependent dielectric with extended cutoffs (nonbonded: 8.0 Å, electrostatic: 20.0 Å, and hydrogen bonding: 4.0 Å) and the Polak-Ribiere conjugate minimization scheme. Prior to energy minimization of the TMH region, all charged residues were neutralized except the residues D2.50(66), R3.50(114), K7.40(263), D7.43(266), and D7.49(272). A harmonic constraint of 1000 kJ/mol was placed on the backbone (Φ, Ψ, and ω) to maintain helicity. Extra- and intracellular loops were generated via the program MODELLER 8.2 [78-80], described below.

To create an active, R* model, multiple changes were implemented to transform the finalized R bundle. After excision of the EC2, EC3 and IC3 loops, the existing TMH6 was replaced with an extracellularly superimposed helix from the TMH6 CM output. In contrast to the restricted flexibility of TMH6 chosen for the R bundle, the TMH6 for the R* bundle was allowed full flexibility of the backbone (Φ/Ψ of ± 50°). To emulate β2-AR biophysical data, which suggests an increase in the water accessibility of C6.47 upon activation [15], TMH6 was rotated 10° counterclockwise from an extracellular perspective. This rotation shifted C6.47 from a lipid environment to the water accessible pocket of the TMH6-7 interface. This rotation also moved the intracellular ends of TMH5 and 6 closer to each other, as suggested by both the opsin crystal structure [6, 17] and disulfide cross-linking studies of the M3 muscarinic acetylcholine receptor [23]. D3.49(113) was protonated to be representative of an active state, as suggested by pH-dependent activation studies in β2-AR [81] and flash photolysis experiments on rhodopsin [82]. In 2007 Knierim *et al.* suggested that the protonation of E3.49(134) occurs subsequent to the movement of TMH6 in rhodopsin [83]. The χ1 rotamers of the toggle switch residues F6.48(230) and H6.52(234) were changed from g+ to *trans* along with the transition of the χ1 of Y7.53(276) from *trans* to g+ as suggested by the opsin crystal structures [6, 17]. Additionally, an offset parallel π-stack between F6.48(230) and F5.47(180) was created, as the interaction between these residues is suggested to stabilize the R* state [28]. With the remaining loops frozen, an energy

minimization with the same force field, cutoffs, and minimization scheme described earlier was performed for 500 iterations, with a distance-dependent dielectric. A harmonic constraint of 1000 kJ/mol was placed on the backbone (Φ. Ψ, and ω) to maintain helicity. An additional harmonic constraint of 1000 kJ/mol was placed on the χ1 of F5.47(180) to maintain the experimentally suggested interaction between F5.47(180) and F6.48(230). The EC2, EC3 and IC3 loops were then generated via the program MODELLER 8.2 [79-80], described below.

Modeling of the rat GPR35 R* Receptor

The construction of the rat GPR35 R* bundle was performed in several steps: 1) TMHs 1, 2, 3, 4, 6, and 7 of the final human GPR35 R* bundle were mutated to the rat sequence; 2) the EC2 and IC3 loops were removed; and 3) TMH5 from the β2-AR crystal structure was mutated to the rat sequence and substituted for the GPR35 human R* TMH5. With the loops and termini frozen, the TMH region was energy minimized for 500 iterations in a distance dependent dielectric, utilizing the same force field, cutoffs, and conjugate method described earlier, but with a harmonic constraint on the backbone (φ, ψ, and ω) of 1000 kJ/mol to maintain helicity. The EC2 loop was then modelled as described above, including the distance restraint to mimic the conserved disulfide bridge. The final loops were then energy minimized, as described above, until a gradient of 0.01 kcal/mol was achieved in a high, constant dielectric of 80.

Generation of N- and C-terminus and EC/IC loops

The program MODELLER 8.2 [78-80] was used to generate the geometries for the non-transmembrane regions of the GPR35 model. The N- and C-termini were defined as Met1-Pro16 and Pro293-Ala309 respectively. The intra- and extracellular loops were defined as: IC1:Arg48-Glu54, IC2: Pro121-Ser129, IC3: Ala199-Thr212, EC1: Asp82-Pro323, EC2: Leu153-Asn169, EC3: Val243-Arg255. During the loop generation the TMH regions were frozen, and in order to mitigate the bias that the position of the typical terminating residues would have on the loop topography, the defined loop length was adjusted. A range extension of ±1 residue was implemented into each loop to lessen this position bias of the terminal residues. Additionally, the length of the EC1 loop was extended due to the presence of the helix altering proline located at loci 3.32(87).

MODELLER generates and optimizes a *de novo* loop/terminus using a pseudo energy function. This energy is a summation of multiple spatial restraints from the CHARMM force field, statistical preferences for backbone and side chain dihedral angles, and statistical preferences for non-bonding atom-atom interactions. The method of conjugate gradients, in conjunction with simulated annealing and molecular dynamics techniques, is then used to optimize the energy function. This program does not create loops/termini as if the receptor bundle was in an explicit bilayer, so each approximately 1000 structures that were generated by each MODELLER calculation were evaluated visually to determine the likelihood of their existence in this environment. If biophysical data was available, this was also used to govern the loop/terminus choice. This is most represented in the EC2 loop where biophysical data suggests there is a disulfide bridge between the cysteine on the EC2 loop and C3.25. To create the disulfide bridge, MODELLER implements a distance, dihedral, and angle constraints on the C_α,

C_β, and the sulfur of C(162) on the EC2 loop and C3.25(89) to force a disulfide bridge. Other biophysical data that was used to direct the EC2 loop choice was based on substituted cysteine accessibility method (SCAM) studies of the dopamine D2 receptor[84], GPCR crystal structure data [2, 7, 9, 11, 31], circular dichroism studies of the Serotonin 5-HT4a [85], and NMR spectroscopic data from the β2-AR [86].

Ligand Conformation Searches

Complete conformational analyses were performed on the compounds zaprinast and pamoic acid using *ab initio* Hatree-Fock calculations at the MMFFs/HF 6-31+G* level in vacuum as encoded by Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612). For each conformational search, local energy minima were identified by the rotation of the subject dihedrals 360°, in 60° increments (6-fold search), followed by a MMFFs/AM1 energy minimization of each rotamer generated. The geometry and energy were then re-evaluated for each unique conformer calculated within 15 kcal/mol of the global conformer at the HF 6-31+G* level. For zaprinast, the energy minimum of each tautomeric form was assessed at the HF 6-31+G*, HF 6-311+G**, DFT 6-31+G*, and DFT 6-311++G** levels of theory. Additionally, pamoic acid was also assessed at HF 6-31G* in a SM8 water solvation model as encoded by Spartan '08 (Wavefunction, Inc.). The conformations of MLS-0370945 were initially assessed at the MMFFs/HF 6-31+G* theory in a vacuum environment. In vacuum the global low energy conformer adopted a folded conformation, which was proposed to be energetically less favourable in a water environment than an elongated conformation. Equilibrium geometries were therefore calculated using a Poisson-Boltzmann water solvation model (Dielectric Constant=80.37, Probe Radius=1.4Å). These calculations were performed at the *ab initio* Hatree-

Fock 6-31+G* level using Jaguar (version 9.0, Schrodinger, LLC, New York, NY) for three MLS-0370945 conformers: the global, the second low energy conformer, and a more elongated conformer. This calculation confirmed the more elongated MLS-0370945 conformer was significantly lower in energy in an aqueous environment.

To calculate the conformational cost of docking, the rotatable dihedrals of the global minimum energy conformer were driven to the corresponding dihedrals of the final docked conformer and then were evaluated. The dihedrals were first driven to the final docked conformers utilizing the force field MMFFs as implemented by Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612). With the rotatable dihedrals angles virtually fixed to the initial values, the equilibrium geometry and the energy of the resultant structure was then calculated at the HF 6-31+G* level and then compared to the global minimum energy calculated at the same level of theory in the same environment. The final docked conformation of the MLS-0370945 compound was evaluated similarly in Jaguar (version 9.0, Schrodinger, LLC, New York, NY), with the inclusion of the Poisson-Boltzmann water solvation model.

Docking of Ligands

A low energy conformer of zaprinast and the global conformers of pamoic acid and MLS-0370945 were used as inputs for the receptor docking studies. The agonists zaprinast and pamoic acid were docked manually in the human wild-type and R6.58A mutant, and the rat GPR35 R* model, while the MLS-0370945 compound was docked solely in the human wild-type R bundle. Each receptor included the intra- and extracellular loops, as well as the N- and C-termini. Due to recent experimental data [45] Y3.32(human/rat, 96/93) was used as the primary

interaction site for zaprinast. Several arginine residues located near the binding pocket of zaprinast (R3.36(100), R4.60(151), R6.58(240) or R7.32(255)) were used as primary interactions for the di-anion pamoic acid. Although there is lack of independent experimental evidence regarding the docking site of pamoic acid alone, supplementary evidence [57] suggests that the binding pockets of pamoic acid and zaprinast share an overlapping binding site. The anionic MLS-0370945 compound, which antagonizes GPR35 activated by zaprinast (AID: 2079), is also suggested to share the same binding site at pamoic acid due to its analogous negative charge.

Ligand/Receptor Minimization

The ligand/receptor complexes were minimized using a Polak-Ribier conjugate gradient minimization and an OPLS2005 all atom force in Macromodel 9.1 (Schrodinger Inc., Portand, OR). An extended cutoff (nonbonded: 8.0Å, electrostatic: 20.0Å, hydrogen bonding: 4.0Å) was used in each stage of the calculation. The minimization was broken into three stages: 1) the extracellular loops EC2, 3 were minimized in a Generalized Born/Surface Area (GB/SA) continuum solvation model for water, while the TMHs and the ligand atoms were frozen, 2) the minimization of the ligand and the TMH region, while the loops and termini residues were frozen, and 3) subsequently an additional iteration of the first step.

The first stage implemented a conjugate gradient minimization scheme as employed in Macromodel, with a GB/SA continuum solvation model for water and a force field defined dielectric. This stage consisted of either 1500 steps or until the 0.05 kJ/mol gradient was reached.

In the second stage, the loops and termini were frozen during the conjugate gradient energy minimization. For the docking studies involving the ligand zaprinast and pamoic acid, the TMH region/ligand minimization, which used the force field and cutoffs described earlier,

consisted of two phases: 1) the TMH region was minimized for 20000 steps or until the bundle reached the 0.05 kJ/mol gradient in a distance-dependent dielectric of 3 ($\varepsilon = 3$) with the hydrogen bond and or cation-π restraints discussed below and 2) the repeat of protocol in the previous phase without the hydrogen bond and or cation-π restraints. The TMH region/ligand minimization for MLS-0370945 compound followed only the second phase of this protocol.

The harmonic constraints placed on the ligand and or the receptor were variable and contingent on the receptor/ligand combination. In all of the docking studies, a harmonic constraint of 1000 kJ/mol were placed on select dihedrals of the ligands to maintain conformational similarity to the low energy conformer calculated at the HF 6-31+G* level of theory. The dihedral restraint for zaprinast maintained the internal hydrogen bond between its rings, the propoxyphenyl and azapurine-6-one functional groups. This restraint was implemented because the lowest energy structure which breaks this internal hydrogen bond was calculated at the HF-6-31+G* level of theory as 3.74 kcal/mol above the global minimum. The dihedral restraints for pamoic acid maintained the planarity of the carboxyl functional group with the naphthalene ring in the 3-hydroxy-2-naphthoic acid portion of pamoic acid. A similar restraint was placed on the dihedral of the MLS-0370945 compound to maintain planarity of the carboxyl functional group and the phenyl ring of the benzoic acid segment of this compound. In certain receptor/ligand combinations, a harmonic constraint of 1000 kJ/mol was also placed the arginines R4.60(151) or R7.32(255) to maintain rotamers that optimize the cation-π interactions not weighted for properly by the OPLS2005 force field. The R4.60(151) restraint was implemented in the pamoic acid docks in both the wild-type and R6.58A(240) human mutants receptor bundles, while the R7.32(255) restraint was employed in the docking of zaprinast in the R6.58A(240) mutant bundle only. A single set of restraints were implemented between the ligand and the human GPR35 receptors to preserve an experimentally suggested [45] hydrogen bond between

Y3.32(96) and zaprinast. Both an angle (150° ± 30°; 100 kJ/mol) and distance (2.8 ± 0.4 Å; 45 kJ/mol) harmonic restraints were implemented to maintain this hydrogen bond during the minimization. Zaprinast docked in the rat GPR35 R* bundle had an analogous restraint between the ligand and Y3.32(93), with an additional restraint to maintain the hydrogen bond between the donor zaprinast and the acceptor S5.43(174), with duplicate constraints on the analogous atoms.

The final stage was essentially an additional iteration of the first stage of the minimization. The TMH region and ligand were frozen and the extracellular loops, EC2 and EC3, were minimized in a GB/SA continuum solvation model for water with the same gradient, force field, and cutoffs described earlier. Contrary to stage 1, the loops were minimized for a maximum of 20000 steps to ensure that the 0.05 kJ/mol gradient was reached.

Calculation of Ligand/Receptor interaction energy

The energies of interaction (EOI) of the ligand/receptor complexes were calculated using the OPLS2005 all atom force field in Macromodel 9.1(Schrodinger Inc., Portand, OR) using the component interactions script. Each ligand/receptor interaction energies are the sums of the individual interaction energy between the ligand and each residue within a 6Å radius. The energies were calculated with an extended cutoff (nonbonded: 8.0Å, electrostatic: 20.0Å, hydrogen bonding: 4.0Å) and a distance dependent dielectric of 3. The non-bonded interaction energies for each residue and the ligand were partitioned into van der Waals and electrostatic contributions.

CHAPTER III

RESULTS

Conformational Memories (CM) Output

Because of the effect that changing one helix conformation has on other helices in the receptor, the helices are presented in the order in which they were fitted.

TMH7

The sequence of TMH7 in GPR35 does not have the conserved P7.50 found within the NPXXY motif, but instead has an A7.50(273). To determine if the sequence of GPR35 may emulate the flexibility of the proline kink caused by the highly conserved P7.50, an increased range of flexibility of ± 50° was given to the backbone in the analogous region of C7.46(269) - A7.50(273). CM calculations suggest that the extracellular end of GPR35 TMH7 diverges from that of the template A2-AR structure (Figure 12). The bend angle of TMH7 shifts from 26.4° in A2-AR to 6.75° ± 3.20° in GPR35, as averaged by 105 CM structures. A7.50(273) was used as the "hinge residue" and i-7 and i+7 residues were the range of superimposition. The wobble angle of the CM TMH7 also differed from the A2-AR crystal structure, with values of -38.66° ± 97.27° for GPR35 and 176.3° for the A2-AR. These values suggest that the divergence of the extracellular end of TMH7is due to the lack of a proline at locant 7.50. Notable is that the superimposition of these CM helices on the A2-AR placed the C_β atoms of the charged residue K7.40 facing toward lipid. As the positions of this charged residue is energetically unfavorable, calculations were performed which forced S7.39(262), S7.42(265), or both S7.39(262) and

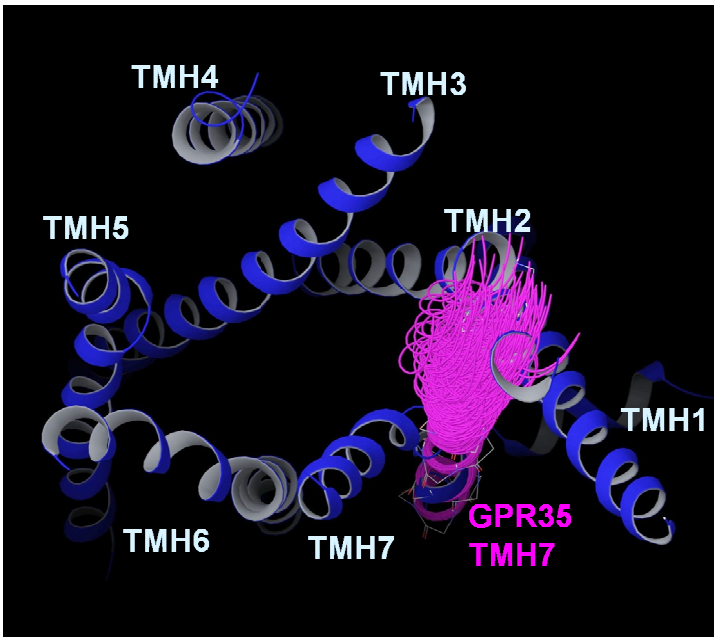


Figure 12. Conformational Memories-calculated hGPR35 TMH7 superimposed on the β2-AR crystal structure. The 105 conformational hGPR35 TMH7, were initially superimposed intracellularly on the TMH7 of the A2-AR crystal structure (PDB ID: 3EML) template structure, with the remaining TMHs from the β2-AR crystal structure (PDB ID: 2RH1).

S7.42(265) in a χ1 = g- (60°±40°) conformation to test if the helix altering properties of a serine in χ1 = g- could change the position of K7.40(263) to a more favorable position. The output showed no conformations which changed the position of these charged residues to a more reasonable position. Additionally, the Prokink analysis of these supplementary CM calculations, with either a single or multiple serines in a χ1 = g+ conformation, did not show a substantial divergence from the initial CM calculations. Each average attribute of bend angle, wobble angle, and face shift fall within one standard deviation of one another. The bend angle, wobble angle, and face shift, with the "hinge residue" K7.40(263) and a superimposition range of and *i*-7 and *i*+7, have coordinated values of: 1) S7.42; 5.64° ± 3.34°, -27.85° ± 96.92°, and 16.80° ± 5.24°, 2) S7.39; 5.79° ± 3.10°, -29.67° ± 95.99°, and 25.90° ± 6.87°, and 3) S7.39 and S7.42 concurrently; 5.48° ± 2.90°, -10.17° ± 113.21°, and 16.36° ± 5.50°. Taken together, all of these results suggest that TMH7 is straighter than TMH7 of the template crystal structure, A2-AR.

TMH1

GPR35's lack of the highly conserved NPXXY motif found in the TMH7 of the crystal structure templates was assumed to cause a significantly straighter TMH7 than the templates, which may influence the conformation of the juxtaposed TMH1. Additionally, the proline position shift in TMH2 (β2-AR & A2-AR: P2.59, GPR35: P2.58(74)) is predicted to cause conformational consequences, which may be translated to the conformation of TMH1. TMH1 was superimposed at its intracellular end on the backbones of the template residues at the positions 1.46 - 1.57 to ensure that the position of the highly conserved N1.50 was aligned properly. GPR35 has three non-consecutive glycines (G1.34(22), G1.40(28), and G1.46(34)) and to assess the potential impact of these flexible residues, three independent CM calculations were performed for each flexible region individually. In each Prokink calculation, the glycine (G1.34(22), G1.40(28), or G1.46(34)) that was assigned as the flexible region in the CM calculation was designated as the "hinge" residue. The largest divergence from the template was in the CM calculation where G1.34(22) was defined as the "hinge residue". This Prokink analysis, with *i*-5 and *i*+7 residues as the range of superimposition, calculates the average face shift, wobble, and bend angles as 12.75° ± 7.93°, -78.35° ± 89.77°, and 7.91° ± 3.96°, as averaged by 105 CM structures, vs. 16.8°, 7.9°, and 4.1° for β2AR. While these results suggest a large structural divergence from the template, visually it is apparent that the conformational divergence from the template is localized to the extracellular end of TMH1 (Figure 13-1). With G1.46(34) identified as the "hinge residue", with *i*-7 and *i*+7 as the range of superimposition, there was a deviation from β2-AR crystal structure in the wobble angle with the measured angle of -32.89° ± 60.93° and -138.6 in GPR35 and the β2-AR crystal structure, respectively. In the additional Prokink calculations, with G1.40(28) and G1.46(34) as "hinge residues", the additional face shift, wobble, and bend angle averages are within one standard deviation of the β2-AR values. Each

CM calculated TMH1 are very similar (Figure 13) to each other and vary from the β2-AR template extracellularly, with a preference to lean towards either TMH7 or TMH2.

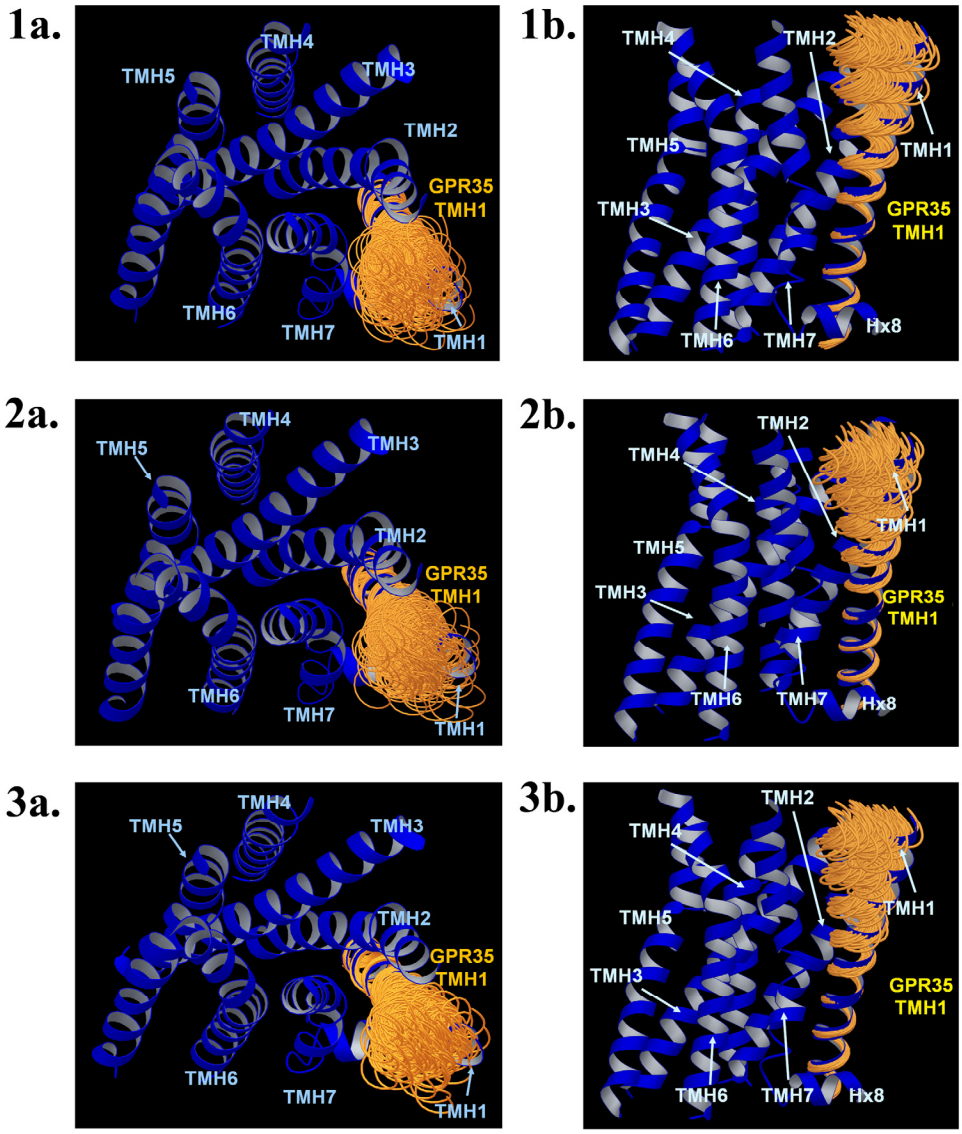


Figure 13. Conformational Memories-calculated hGPR35 TMH1 superimposed on the β2-AR crystal structure. (PDB ID: 2RH1). Each set of 105 conformers (1 = G1.34 (22), 2 = G1.40 (28) or 3 = G1.46 (34)) represents a different flexible region from two perspectives: an extracellular top-down view (a) and a side perspective (b) as seen from the lipid bi-layer interior.

TMH2

GPR35 has a proline toward the extracellular end of the helix, but is shifted by one residue as compared to the β2-AR crystal structure (2.59 to 2.58). The TMH2 proline shift creates a face shift and wobble angle of 53.31° ± 26.90° and -90.53° ± 70.08°, as averaged by 210 CM conformers, compared to 99.6° and -5.0° for the β2-AR TMH2. P2.58(74) was used as the "hinge residue" with *i*-9 and *i*+10 residues as the range of superimposition. This shift allows for TMH2 to move away from the bulk of the straight TMH7 (Figure 14).

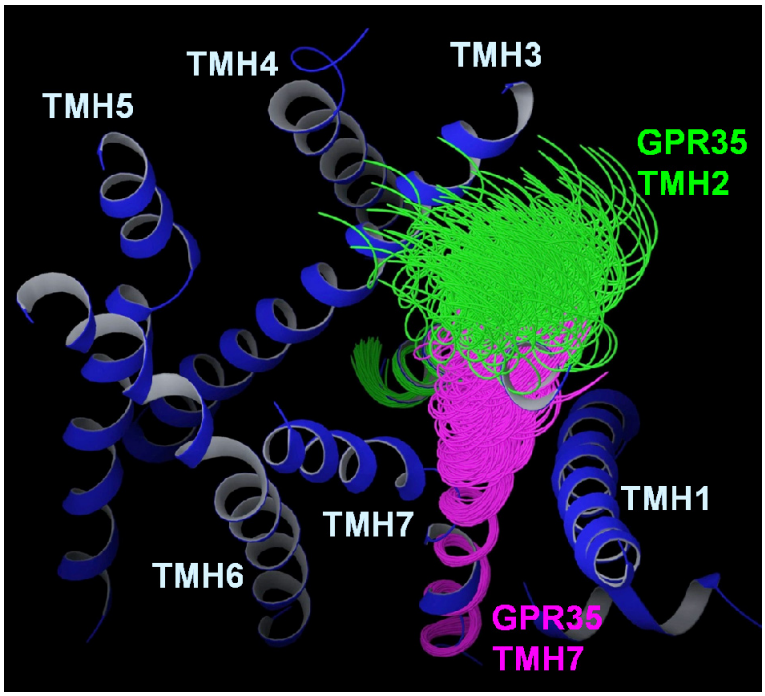


Figure 14. Conformational Memories-calculated hGPR35 TMH7 and TMH2 superimposed on the β2-AR crystal structure. Each sets of 105 CM-calculated helices of hGPR35 TMH7 and TMH2 on their respective helices of the β2-AR crystal structure (PDB ID: 2RH1).

TMH5

TMH5 has two Proline residues: P5.43(176) and the conserved P5.50(183). The range of flexibility of the backbone of the proline kink region *(*P5.50=*i,* flexibly range *i* to *i-4)* of the highly conserved P5.50 was limited to the crystal structure ranges of TMH5 (Table 2). The P5.43 proline kink region *(*P5.43=*i,* flexibly range *i* to *i-4)* had a range of ±50° for the backbone dihedrals. The additional Proline at 5.43 induces a local conformational change (Figure 15). When the "hinge residue" is P5.43, with *i*-7 and *i*+6 residues as the range of superimposition, average face shift, wobble, and bend angles are calculated as 33.81° ± 17.88°, -23.08° ± 141.61°, and -27.36° ± 7.79° vs. 42.5°, 72.4°, and 25.7° for β2AR. When the "hinge residue" is P5.50(183), with *i*-10 and *i*+13 residues as the range of superimposition, the average wobble and bend angle is -86.96° ± 73.58° and 23.33° ± 6.85° for GPR35 vs. -76.3° and 21.0° for β2AR. This difference suggests that TMH5 leans away, as compared to β2AR, from TMH4 (Figure 15).

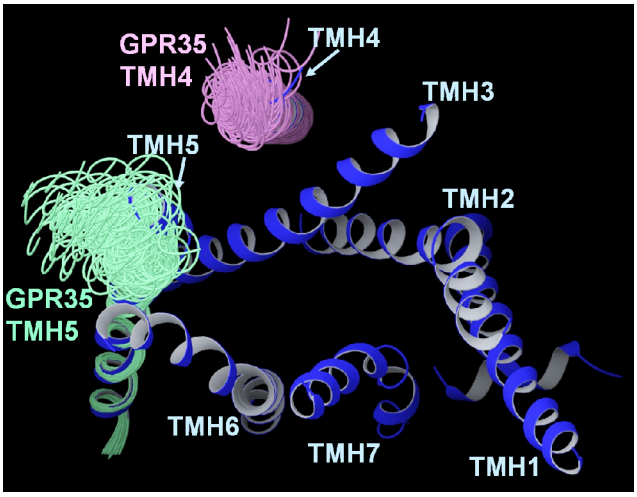


Figure 15. Conformational Memories-calculated hGPR35 TMH4 and TMH5 superimposed on the β2-AR crystal structure. Each sets of 105 CM-calculated helices of TMH4 and TMH5 are superimposed intracellularly on their respective helices of the β2-AR crystal structure (PDB ID: 2RH1).

TMH6

For TMH6, two independent CM calculation were performed in order to create a TMH6 that was representative of either an inactive (R) and active (R*) bundle. For the inactive (R) bundle the backbone dihedral ranges were restricted to ranges found within the crystal structure ranges for TMH6 (Table 3) and the χ1 of F6.48(230) was restrained to a g+ conformation (-60°±40°). The CM calculations for the inactive bundle (R) suggest that the extracellular end of GPR35 TMH6 is very similar to that of the β2-AR. The bend, wobble angle, and face shift of the β2-AR, with P6.50 used as the "hinge residue", were 33.2°, -72.9°, and 85.0°, which correlated well to the average of 105 CM structures for TMH6 of GPR35 (26.68° ± 5.84°, -108.82°± 44.29° and 54.44° ± 18.69°). Additionally, the CM TMH6 helices with increased flexibility also did not diverge from the crystal structure template, although as expected had larger standard deviations. The more flexible TMH6 had a bend, wobble angle, and face shift of 25.46° ± 9.11°, -65.15°± 125.50° and 37.73° ± 28.85°.

TMH4

GPR35 lacks the Proline found at position 4.60 (P4.60 β2-AR, R4.60 GPR35) of the crystal structure template. The CM calculations for TMH4 show a variability in the face shift, bend angle, and wobble angle (1.63° ± 11.83°, 18.04° ± 4.95°, and -123.39° ± 58.70°) of GPR35 conformers vs. the β2-AR (48.0°, 40.0°, and 26.2°) using V4.58 as the "hinge residue" and the range of superimposition of *i*-4 and *i*+10 residues. If W4.50(141) is used as the "hinge residue", the average of GPR35 TMH4 CM calculations (12.72° ± 7.90°, 7.46° ± 5.05°, and -9.26° ± 111.47°) and β2-AR crystal structure (8.4°, 8.5°, and 70.3°) yield very different results. These numbers suggest that the lack of a Proline at 4.60 causes a local change at the extracellular

portion of TMH4 which leans towards the extracellular region of TMH5 compared to the β2-AR crystal structure (Figure 15).

Induced 3-10 helical region in TMH7

Adjusting the backbone dihedrals of the region S7.39(262) - D7.43(266) (7.39 - 7.42: Φ and Ψ, 7.43: Φ only) to match the homologous sequence found in the A2-AR (A2-AR: S7.42(277), H7.43(278), T7.44(279), N7.45(280), S7.46(281) vs. GPR35: S7.39(262), K7.40(263), L7.41(264), S7.42(265), D7.43(266)) induced a 3-10 helical region in the GPR35 TMH7. This alteration in the helix shifted the position of the charged residue K7.40(263) from a lipid environment (Figure 16A), to a more energetically favorable position facing into the hydrophilic channel between helices TMH1, 2, and 3 (Figure 16B). Additionally, this induced 3-10 helix also shifted the extracellular position of TMH7 towards the extracellular top of TMH6 (Figure 17), reducing the energetically unfavorable gap between these helices (Figure 18).

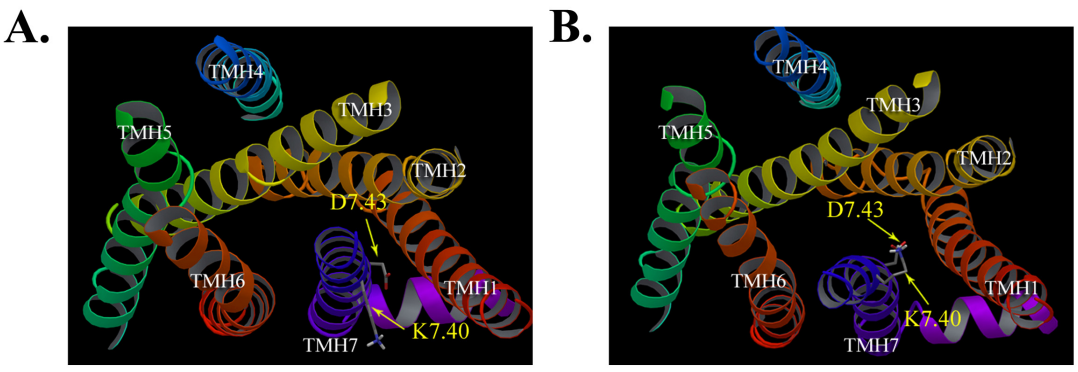


Figure 16. Comparison of a hGPR35 bundles with either a CM-Calculated TMH7 or a TMH7 with an induced 3-10 helical region. A) hGPR35 bundle with CM-calculated TMH7 B) hGPR35 bundle with CM-calculated TMH7 with an induced 3-10 helical region (7.39-7.42: φ and ψ, 7.43: φ only) matching the analogous sequence in the A2-AR crystal structure (PDB ID: 2EML)

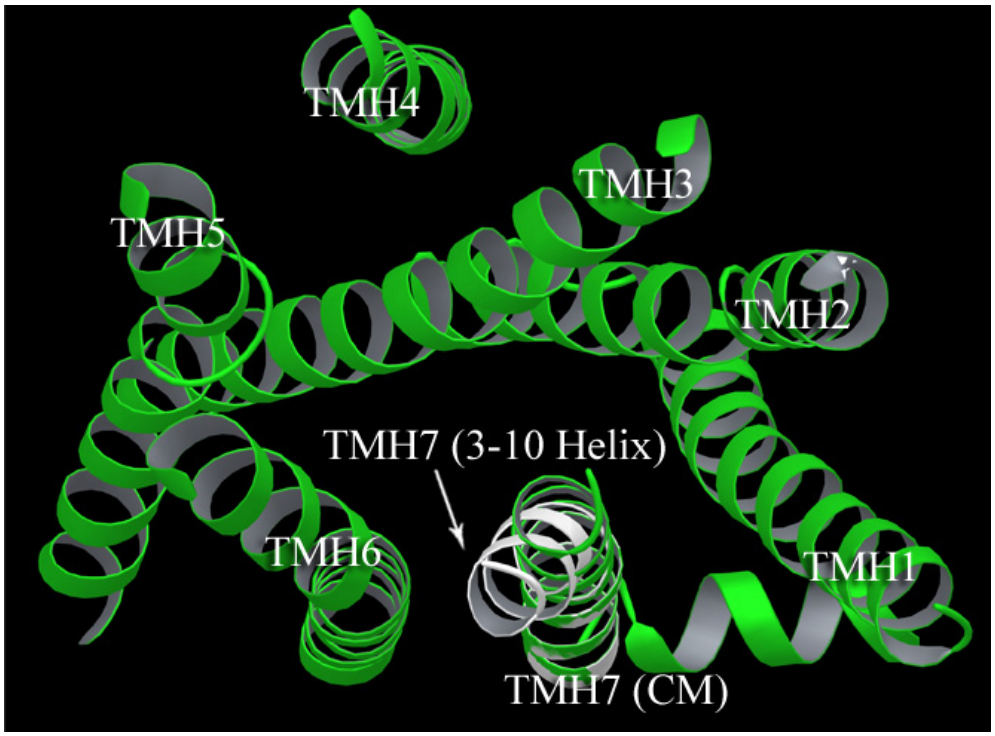


Figure 17. Overlay of the initial CM-calculated TMH7 and the same CM TMH7 with an induced 3-10 helical region (superimposed intracellularly).

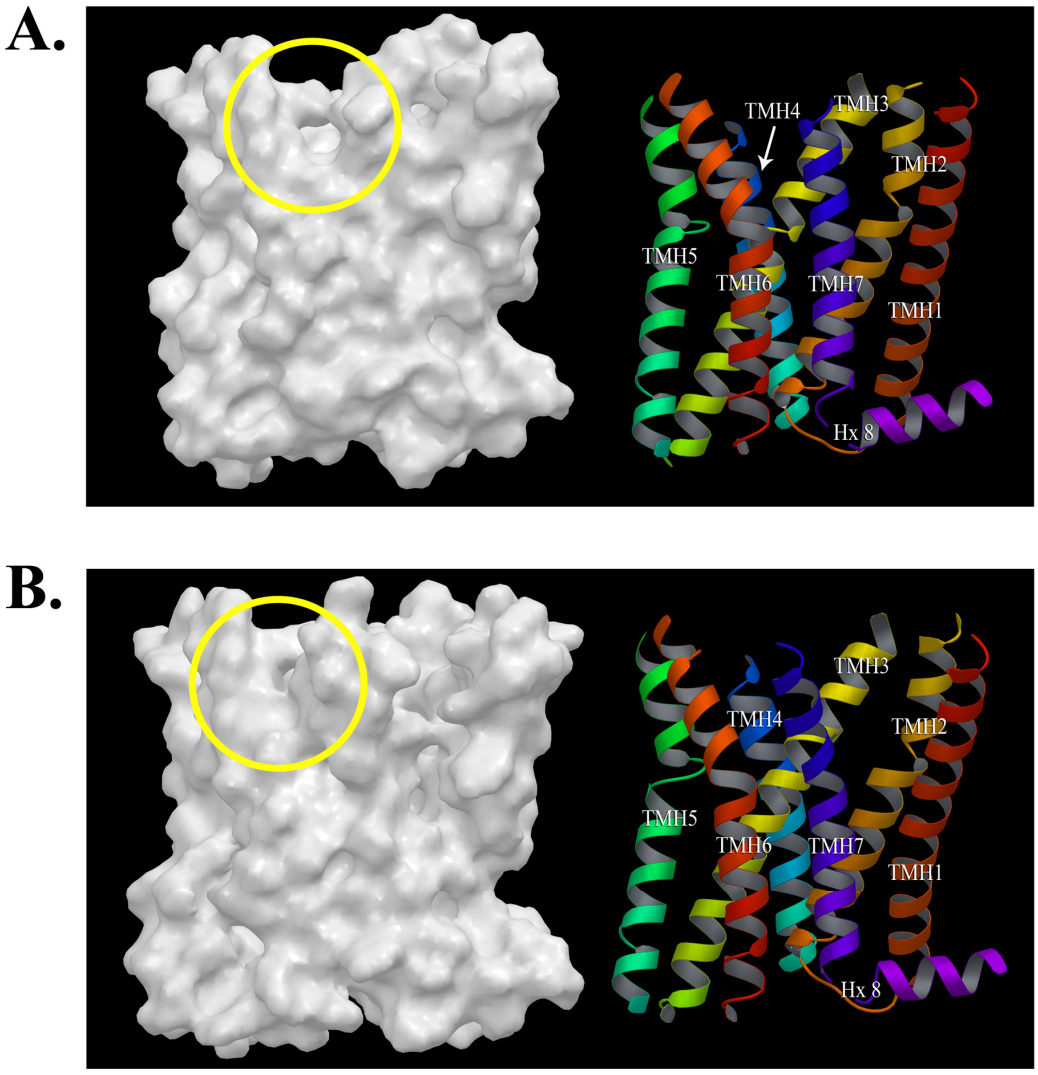


Figure 18. Comparison of the minimized TMH region of GPR35 R bundles with either a CM-Calculated TMH7 or a TMH7 with an induced 3-10 helical region. Both the bundle with an initial TMH7 that was the original CM-calculated TMH7 (A) or the TMH7 with an induced 3-10 helical region (B) are depicted in both a molecular surface and as ribbons. The yellow circles identify the gap in the molecular surface of the protein between TMH6 and 7.

Minimized Human GPR35 R Bundle

Biophysical data suggest two specific characteristics that are nearly universally associated with an inactive (R) GPCR bundle: the aromatic toggle switch residues 6.48 and 6.52 in a $\chi 1$ = g+ conformation and an ionic lock between residues E(D)6.30 and R3.50. The minimized hGPR35 R bundle possesses these representative characteristics, with the toggle switch residues F6.48(230) and H6.52(234) in a $\chi 1$ = g+ conformation (Figure 19B) and a hydrogen bond between T6.30(212) and R3.50(114) (Figure 19A). While GPR35 does not have a negatively charged residues at position 6.30, experimental D6.30N(338) mutation data of the cannabinoid CB1 receptor suggest that the interaction between a hydrogen bond accepting residue and R3.50 is a mimic of the ionic lock between D6.30 and R3.50 [87]. This is further substantiated by mutational studies involving the μ-opioid receptor, which identifies the residues that emulate the ionic lock of the R state of prototypical GPCRs as T6.34(279) and R3.50(165) [88]. In addition, R3.50 is in a highly conserved conformation with specific interactions, known as an arginine-cage motif [89], and is positioned between D3.49 and T6.30. This highly conserved arginine-cage motif is mimicked in the hGPR35 R bundle (Figure 19A).

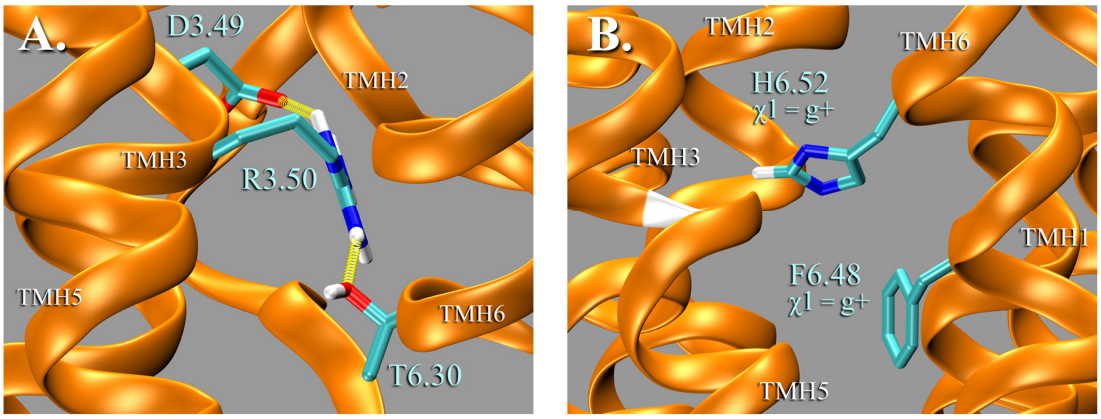


Figure 19. Interactions/Rotameric states of conserved residues of the minimized hGPR35 R Bundle which are suggested to be representative of the inactive state. The representative states are (A) the highly-conserved arginine-cage motif (GPR35: D3.49 (113), R3.50 (114), and T6.30 (212)) and (B)) the $\chi 1 = g^+$ of the toggle switch residues H6.52(234) and F6.48 (230).

A network of aromatic residue interactions found in the crystal structures of the β2-AR [31] and human dopamine D3 receptor [10] between the residues 5.47, 5.48, and 6.52 (GPR35:F5.47, Y5.48, H6.52 vs. D3/A2-AR F5.74, Y5.48, F6.52) was also mimicked in the inactive hGPR35 bundle ((Figure 20). Importantly, the position of H6.52(234) in this mimicked aromatic network reduces the size of the binding pocket, which does not allow the agonists Zaprinast and pamoic acid to fit in the R bundle without major steric overlap, but does not affect the ability to dock the antagonist, MLS-0370945.

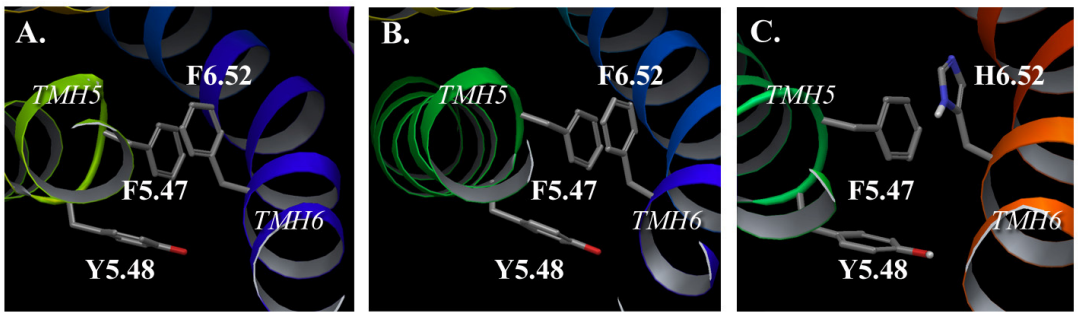


Figure 20. Depiction of the aromatic stacking network between the residues F5.47, Y5.58, and F/H6.52. These depictions are from (A) the dopamine D3 receptor crystal structure (PDB ID: 3PBL), (B) the β2-AR crystal structure (PDB ID: 2RH1) and (C) the minimized hGPR35 R bundle

The final set of residue-residue interactions that were implemented in the hGPR35 R state model was the aromatic stack between Y7.53(276) and F7.60(283), as suggested in the rhodopsin crystal structure [4] and by 5HT2C mutation studies [77]. The aromatic stack in GPR35 is a tilted-T π stack, which differs from the offset parallel interaction between these residues found in the Rhodopsin crystal structure (Figure 21).

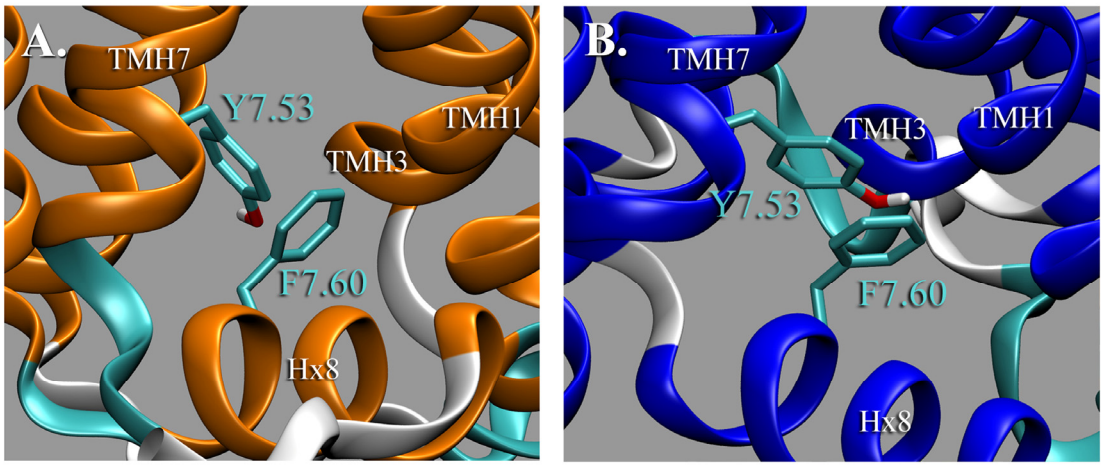


Figure 21. Depiction of the aromatic stack between the residues Y7.53 and F7.60. These are depictions are from the (B) rhodopsin crystal structure (PDB ID: 1GZM) and in the (A) minimized R state of hGPR35

Following the minimization of the hGPR35 R bundle, the residue C(305), located on the C-terminus, was palmitylated. This post translational modification is found experimentally to occur in multiple Class A GPCRs, including rhodopsin [90], β2-AR [91], and the A1-AR [92]. Palmitoylation is the attachment of the 16 carbon-chain, saturated fatty acid palmitic acid to a cysteine via an acyl thioesterification. . The attached palmitic acid is suggested to act as an anchor, holding the C-terminus of the GPCR in the phospholipid bi-layer via the hydrophobic interactions betyween the palmitate and lipid of the fatty acid chains [93] .

Minimized Human GPR35 R* Bundle

There are two characteristics that are associated with an active (R*) vs. an inactive (R) bundle: : 1) the χ1 rotameric change of aromatic toggle switch residues at positions 6.48 and 6.52 and 2) the breaking of the intracellular ionic lock between the residues E(D)6.30 and R3.50. In order to emulate these characteristics, TMH6 of the R bundle was replaced with a different, structurally divergent CM TMH6 that was superimposed extracellularly. After the transition of F6.48(230) from a χ1 = *g*+ to a *trans* conformation, an offset parallel π-stack was formed between F5.47(180) and F6.48(230) (Figure 22A), which is suggested to stabilize the R* state in multiple GPCRs [28]. Figure 22B shows the overlaid images of the inactive bundle, with the toggle switch residues F6.48(230) and H6.52(234) in χ1 = *g*+ conformations (transparent), and the active bundle, with these residues in a χ1 = *trans* conformation.. This overlay demonstrates the large shift in bulk associated with this change. Furthermore, the electrostatic interaction between the residues T6.30(212) and R3.50(114) is broken in the activated hGPR35 bundle (Figure 22C). The distance of the C_α atoms of residue T6.30(212) between the active and inactive bundle was measured as 4.83 Å. This value is analogous to the 5 Å shift of the extracellular end of TMH6 (residues A(241) and R6.35(252)) during the activation of Rhodopsin, as measured by double electron-electron resonance (DEER) spectroscopy [18]. Additionally, D3.49(113) is protonated [81-82] and R3.50(114) is no longer in an arginine-cage motif, , as suggested by crystal structure data for opsin [6], metarhodopsin II [5, 94], and the nanobody-stabilized active state of the β2-AR [8]. Also, the aromatic π stack between Y7.53(276) and F7.60(283) was disrupted in the GPR35 R* bundle (Figure 23A), due to the transition of the χ1 of Y7.53 from *trans* to *g*+ during activation, as suggested in the opsin crystal structure (Figure 23B).

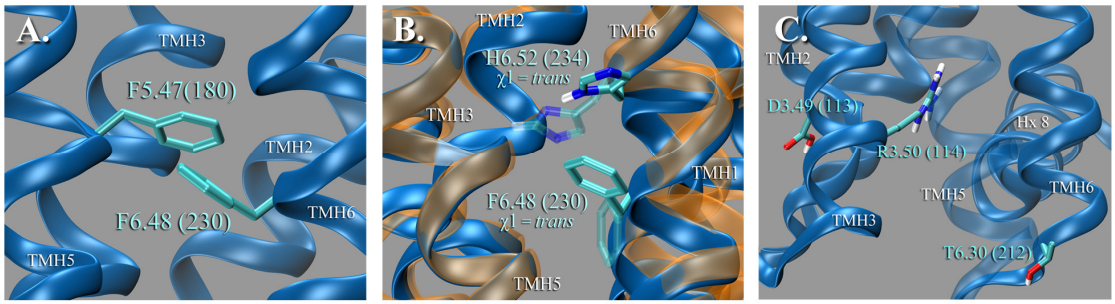


Figure 22. Interactions/Rotameric states of conserved residues of the minimized hGPR35 R* bundle which are suggested to be representative of an active state. The representative states are (A) the aromatic stacking of F5.47 (180) and F6.48 (230), (B)) the χ1 = *trans* of the toggle switch residues H6.52(234) and F6.48 (230) overlaid on the inactive bundle, and (C) the breaking of the electrostatic interaction between R3.50 and T6.30.

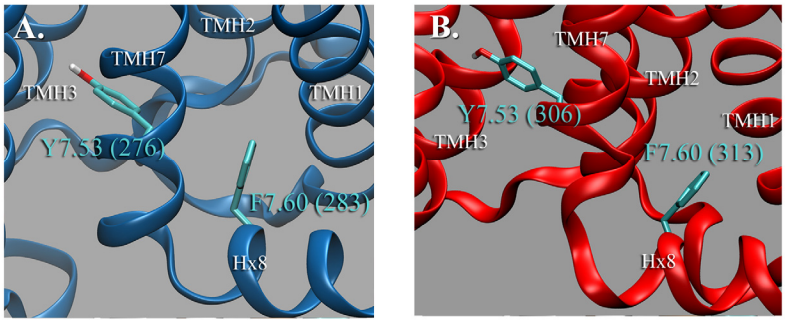


Figure 23. The loss of the aromatic stack between Y7.53 (306) and F7.60 (313) in the Opsin (B) crystal structure (PDB ID: 2CAP) is mimicked in (A) the minimized hGPR35 R* Bundle.

Minimized Human GPR35 R6.58A(240) Mutant Bundle

I propose that the mutation of the residue R6.58(240) to an alanine will cause notable changes in the TMH4-5-6 interface, the proposed binding pocket for the ligands zaprinast and pamoic acid. The residue R6.58(240) in human GPR35 resides on the most extracellular α-helical loop of TMH6, facing the protein interior. In the wild-type human GPR35 bundle absent a ligand, the hydrophobic segment of the R6.58(240) side chain fills the gap between the extracellular-tops of TMH6 and TMH7. This position points the guanidinium group of R6.58(240) towards an area that in a complete cell would contain phospholipid head groups and/or be flooded with water. Mutating R6.58(240) to an alanine creates a large empty space between the extracellular-ends of TMH6 and TMH7. I propose that this unfavorable empty space is to be filled by the

juxtaposed residue R7.32(255). In the absence of the R6.58(240) bulk in the alanine mutant, R7.32(255) is able to adopt an alternative low energy conformation which fills the hole between the tops of the helices, while still exposing the guanidinium group to the extracellular, polar environment.

Minimized rat GPR35 R* Bundle

Rat GPR35 shares a 72% sequence identity with the human orthologue, with high identity in the TMH region, which suggests that their structures may be similar. To create the rat GPR35 model the human GPR35 R* bundle was mutated to match the rat GPR35 sequence with the following two exceptions: 1) first the TMH5 was replaced with the β2-AR TMH5 mutated to the rat sequence and 2) second the original EC2 loop was replaced with a new loop generated by MODELLER. A sequence difference at residue position 5.43, a non-conserved proline in human and serine in rat, causes gross structure difference between rat and human GPR35. This sequence difference causes a 3.91 Å relative shift in the extracellular end of the TMH5 of the minimized rGPR35 wild-type bundle, as compared to the analogous TMH5 of the hGPR35 bundle (Figure 24B). This structure variation causes differences, which will be elaborated on in the following sections, in the docking positions of pamoic acid and zaprinast.

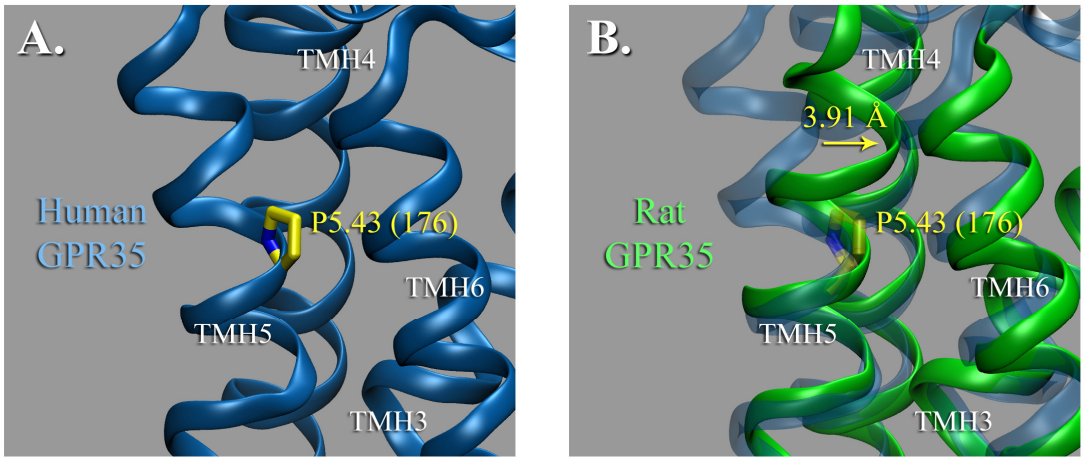
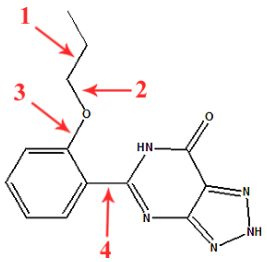


Figure 24. Comparison of the extracellular top of TMH5 in a minimized R* rGPR35 vs. hGPR35 bundle. A) Minimized hGPR35 R* bundle with the carbon atoms (side chain and Cα) of the non-conserved proline, P5.43 (160), colored in yellow B) Minimized rGPR35 R* bundle overlaid on the minimized hGPR35 R* bundle, with the noted relative shift of the position of the extracellular top of the rGPR35 TMH5 of 3.91 Å (distance measured from C_α of residue 5.39)

Conformational Analysis of GPR35 Ligands

Zaprinast

As displayed in Figure 25, zaprinast has 4 rotatable bonds: the bond connecting the phenyl and propoxyl substituents (3), two within the propoxyl substituent (1 and 2), and the bond connecting the propoxyphenyl and azapurine-6-one functional groups (4). Each rotatable bond was explored systematically with a 6-fold rotation, with low energy conformers calculated via a MMFFs force field. Each unique structure within 15 kcal/mol was evaluated at the *ab initio* Hartree Fock (HF) level of theory. These calculations suggested that the rotatable bonds at positions "1","2", and "3" may adopt multiple low energy conformations at physiological temperatures [95].



Zaprinast

Figure 25. The rotatable bonds of zaprinast.

The four low energy conformers are all within approximately 2 kcal/mol of the global minimum, with variable low energy rotamers of either *g+* ,*g-,* or *trans* at the rotatable bonds "1","2", and "3" (Figure 26). The global minimum conformer has all three rotatable bonds of the propoxyphenyl group in a *trans* conformation. The second low energy conformer, which has an equivalent "mirror", adopts a gauche rotamer, either g+ or g-, at the rotatable bond "1". The third low energy conformer adopts a g+ rotameric state for the "2" and "3" bonds and a *trans* conformation for bond "2". The fourth conformer has a g+, *g+,* and *trans* rotamer at the "1", "2", and "3" bonds, respectively.

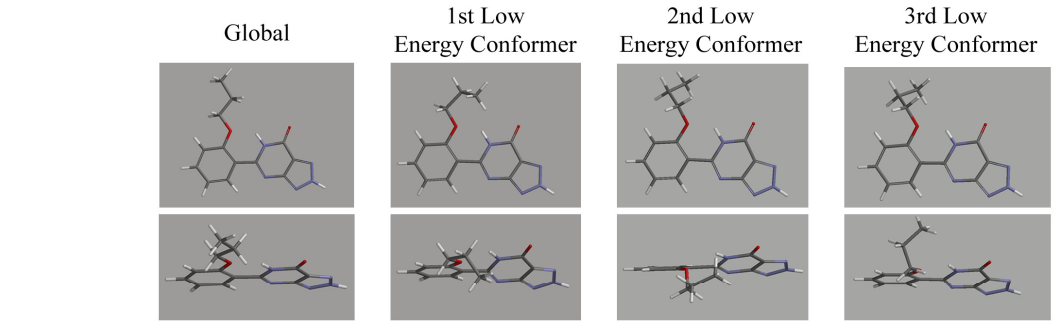
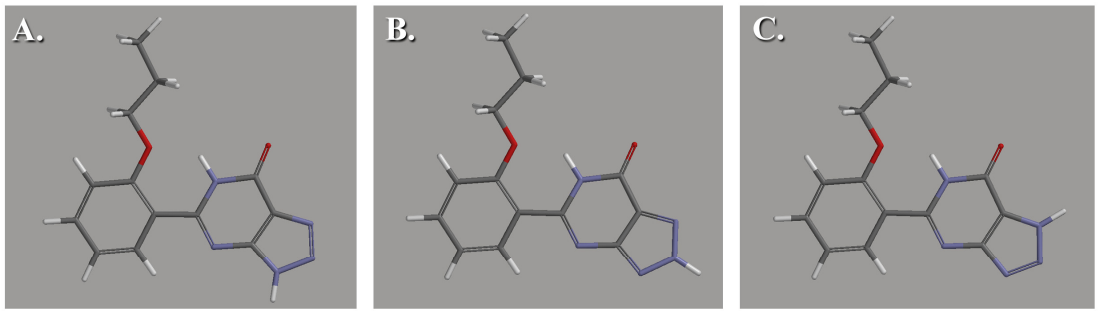


Figure 26. The four low energy conformers of Zaprinast calculated at the HF 6-31+G* level of theory. Values were calculated as implemented in Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612).

Zaprinast has multiple possible tautomeric states and each was evaluated using two different theories: the Hartree Fock and Density function theory (DFT). For the initial calculations both theories utilized the basis set 6-31+G*. To fully explore the tautomeric states of zaprinast a second set of calculations, with a larger basis set of 6-311+G** and 6-311++G** for HF and DFT calculations, respectively, were performed. Both theories identified the same low energy tautomers (Figure 27), though they differed in both the energy-based order and the energetic cost of tautomerization. Each tautomeric state has varied protonation state of the nitrogens on the azapurine-6-one functional group, and are defined as tautomers "A", "B", and "C", with respective IUPAC names of 5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one, 5-(2-Propoxy-phenyl)-2,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one, and 5-(2-Propoxy-phenyl)-1,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one (Figure 27). The HF level of theory at both basis sets identifies the sequence, in order of increasing energies, as tautomer "A", "B", then "C". There is an approximate 1.5 kcal/mol energy difference between each tautomer (Table 5).



5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one

5-(2-Propoxy-phenyl)-2,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one

5-(2-Propoxy-phenyl)-1,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one

Figure 27. The three low energy tautomers of Zaprinast as calculated at the HF/DFT level of theory. Values were calculated as implemented in Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612), with IUPAC names of (A) 5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one, (B) 5-(2-Propoxy-phenyl)-2,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one, and (C) 5-(2-Propoxy-phenyl)-1,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one

Table 5. The energy and energy comparison of multiple, low-energy tautomeric states of Zaprinast at various levels of theory/basis sets. Values were calculated as implemented in Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612).

Theory: HF Basis set: 6-311+G**	Global: **Tautomer "A"** 5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	1st Low Energy Conformer: **Tautomer "B"** 5-(2-Propoxy-phenyl)-2,6--dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	2nd Low Energy Conformer: **Tautomer "C"** 5-(2-Propoxy-phenyl)-1,6--dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one
Energy (kcal/mol)	-578585.28	-578583.73	-578582.28
Δ from global (kcal/mol)	0	1.55	3.01
Theory: HF Basis set: 6-31+G*	Global: **Tautomer "A"** 5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	1st Low Energy Conformer: **Tautomer "B"** 5-(2-Propoxy-phenyl)-2,6--dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	2nd Low Energy Conformer: **Tautomer "C"** 5-(2-Propoxy-phenyl)-1,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one
Energy (kcal/mol)	-578445.71	-578444.08	-578442.42
Δ from global (kcal/mol)	0	1.64	3.3
Theory: DFT Basis set: 6-311++G**	Global: **Tautomer "B"** 5-(2-Propoxy-phenyl)-2,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	1st Low Energy Conformer: **Tautomer "A"** 5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	2nd Low Energy Conformer: **Tautomer "C"** 5-(2-Propoxy-phenyl)-1,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one
Energy (kcal/mol)	-582114.48	-582113.9	-582113.19
Δ from global (kcal/mol)	0	0.58	1.28
Theory:DFT Basis set: 6-31+G*	Global: **Tautomer "B"** 5-(2-Propoxy-phenyl)-2,6--dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	1st Low Energy Conformer: **Tautomer "A"** 5-(2-Propoxy-phenyl)-3,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one	2nd Low Energy Conformer: **Tautomer "C"** 5-(2-Propoxy-phenyl)-1,6-dihydro-[1,2,3]triazolo[4,5-d]pyrimidin-7-one
Tautomer	-581957.79	-581957.49	-581956.44
Δ from global (kcal/mol)	0	0.3	1.35

The energy-based sequence using DFT identifies tautomer "B" as the global minimum, with the higher energy structures identified as "A" then "C". DFT calculates between a 0.3 and 0.58 kcal/mol difference, depending on the basis set, from the global tautomer "B" to the first low energy tautomer "A" (Table 5). Though these theories propose different global tautomers, both suggest the energy difference is small enough between them to expect each at physiological temperatures.

Pamoic Acid

As depicted in Figure 28, pamoic acid has 6 rotatable bonds: the two bonds "1" and "2" connecting the 3-hydroxy-naphthalene-2-carboxyl functional groups, the bonds connecting the carboxyl groups to its respective 3-hydroxy-naphthalene group, bonds "3" and "4", and bonds connecting each hydroxyl group to its respective ring system, bonds "4" and "5". Pamoic acid has a pk_{a2} of approximately 2.1, so it was analyzed as a di-anion to represent its likely physiological protonation state. Conformational analysis of this ligand, utilizing the HF level of theory with a 6-31+G* basis set in vacuum, calculated that the next lowest energy conformer from the global conformation was 7.84 kcal/mol greater in energy. To fully explore this di-anion,

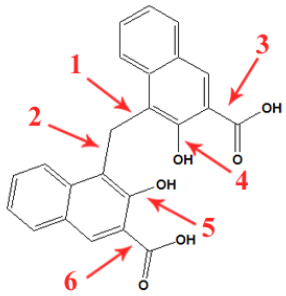
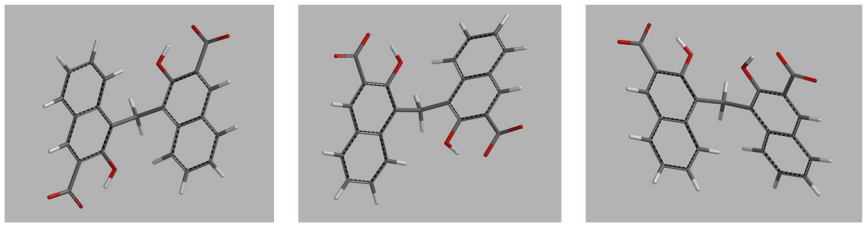


Figure 28. The rotatable bonds of pamoic acid.

a conformational analysis was also performed at HF 6-31G* with a SM8 water solvation model. These calculations identified the identical global conformation as determined in vacuum, but also identified two conformers within 4 kcal/mol from the global minimum (Figure 29).



	Global	1st Low Energy Conformer	2nd Low Energy Conformer
Energy	-833738.53 kcal/mol	-833735.62 kcal/mol	-833734.82 kcal/mol
Δ from Global	N/A	2.91 kcal/mol	3.71 kcal/mol

Figure 29. The three low energy conformers of pamoic acid calculated at the HF 6-31+G* level of theory with the SM8 water solvation model. Values were calculated as implemented in Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612).

MLS-037094

As shown in Figure 30, the compound MLS-0370945 has 9 rotatable bonds located throughout the molecule. The initial quantum mechanical assessment of this compound using HF 6-31+G* in vacuum suggested that the molecule only adopted a folded conformation (Figure 31: conformer A), with the lowest energy elongated conformer being 9.79 kcal/mol higher in energy than the global minimum. The HF 6-31+G* energies recalculated with the addition of solvation energy using a SM5.4/A water solvation model, [96] as implemented by Spartan '08

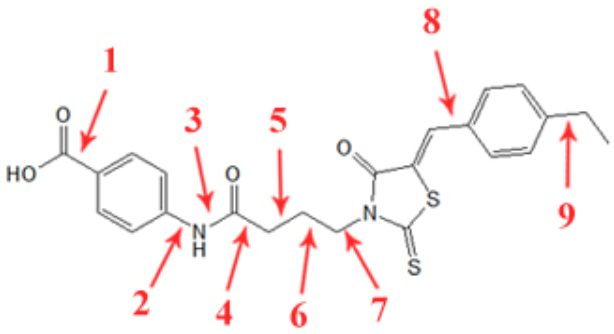


Figure 30. The rotatable bonds of the compound MLS-0370945.

(Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612), suggested that a more elongated structure was only 1.62 kcal/mol higher in energy than the calculated global minimum. This finding provoked a more thorough analysis, in which an equilibrium geometry optimization was calculated at HF 6-31+G* using a Poisson-Boltzmann water solvation model as encoded in Jaguar (version 9.0, Schrodinger, LLC, New York, NY) for the three low energy conformers calculated at HF 6-31+G* using a SM5.4/A water solvation model. The results of the calculations using the Poisson-Boltzmann water solvation model (Figure 32) suggests that the more elongated conformation is substantially lower in energy than the folded conformation. The elongated conformer (Figure 31: Conformer B) after optimization was calculated as the minimum, while the two more folded conformers (Figure 31; A and C) were driven to the same higher energy conformation during the equilibrium geometry optimization. These results (Figure 32) suggest that in water the more elongated conformer is 5.62 kcal/mol lower in energy than the more folded low energy conformer.

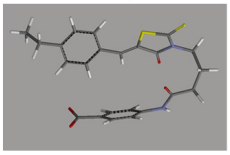
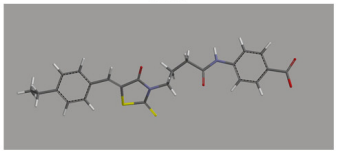
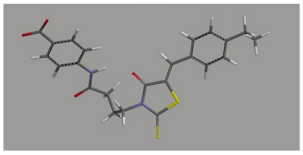
	A.	B.	C.
			
Energy SM5.4/A	-1309474.84 kcal/mol	-1309473.22 kcal/mol	-13094.75.57 kcal/mol
Δ from Global SM5.4/A	N/A	1.62 kcal/mol	2.27 kcal/mol
Energy Vacuum	-1309408.74 kcal/mol	-1309399.55 kcal/mol	-1309400.68 kcal/mol
Δ from Global Vacuum	N/A	9.19 kcal/mol	8.06 kcal/mol

Figure 31. The three low energy conformers of the MLS-0370945 compound calculated at the HF 6-31+G* level of theory (Vacuum vs. SM5.4/A water solvation model). Values were calculated as implemented in Spartan '08 (Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA 92612), with the addition of the individual energies and the comparative energies in vacuum vs. the SM5.4/A water solvation model.

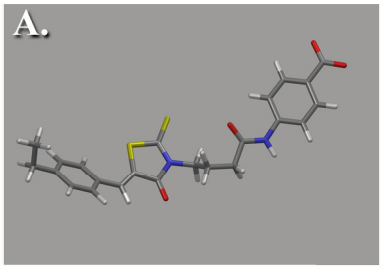
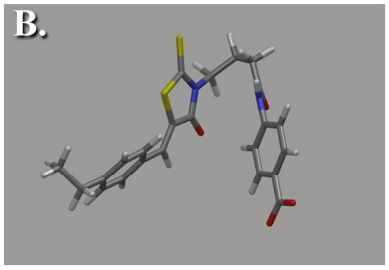
	Global	1st Low Energy Conformer
		
Energy Poisson-Boltzmann	-1309516.50 kcal/mol	-1309510.98 kcal/mol
Δ from Global Poisson-Boltzmann	N/A	5.52 kcal/mol

Figure 32. The two low energy conformers of the MLS-0370945 compound calculated at HF 6-31+G* with the Poisson-Boltzann (PB) water solvation model. Values were calculated as implemented in Jaguar (version 9.0, Schrodinger, LLC, New York, NY), with the addition of the individual energies and the comparative energy.

Docking studies of GPR35 Ligands

For the ligands of pamoic acid, zaprinast, and the compound MLS-0370945, the conformations that were considered for docking purposes did not exceed 5 kcal/mol above the global minimum, as this range is suggested to consider ligand conformations that make up the vast majority of the ligands' biologically relevant conformations [95].

MLS-037094

Human GPR35 R Bundle

The minimum energy conformation of the MLS-0370945 compound, as calculated using a HF 6-31+G* level of theory with a Poisson-Boltzmann water solvation model, was used as the initial manually-docked structure. After the protein-ligand optimization, the MLS-0370945 compound has a large number of interactions with multiple residues (Figure 33). The total protein-ligand "energy of interaction" (EOI) was calculated as -71.32 kcal/mol, with similar EOI contributions from electrostatic and Van der Waals interactions, with approximate values of 44% and 56%, respectively (Table 3A). The largest contributing residue-ligand EOIs are from the arginine residues R4.60(151) and R(164), which form a salt bridge and a hydrogen bond, respectively, with the ligand. Multiple other residues contribute substantially to the total protein-ligand EOI, including the residues Q3.29(93), Y3.32(96), H6.52(234), L6.55(237), R6.58(240), R7.32(255), and the EC2 loop residue F(163) with respective EOI of -6.62, -5.10, -3.34, -4.96, -3.88, -3.73, and -4.01 kcal/mol. Each of the aforementioned residues interacts mainly with the MLS-0370945 compound via Van der Waals interactions. The ligand incurred a small conformational change during the minimization, with a conformational cost of docking of 1.23 kcal/mol (Table 6A). GPCR crystal structures suggest that antagonists directly block the change of the rotameric conformation of either or both of the toggle switch residues at positions 6.48 and 6.52. In the optimized MLS-0370945-hGPR35 R complex, the toggle switch residue H6.52(234)

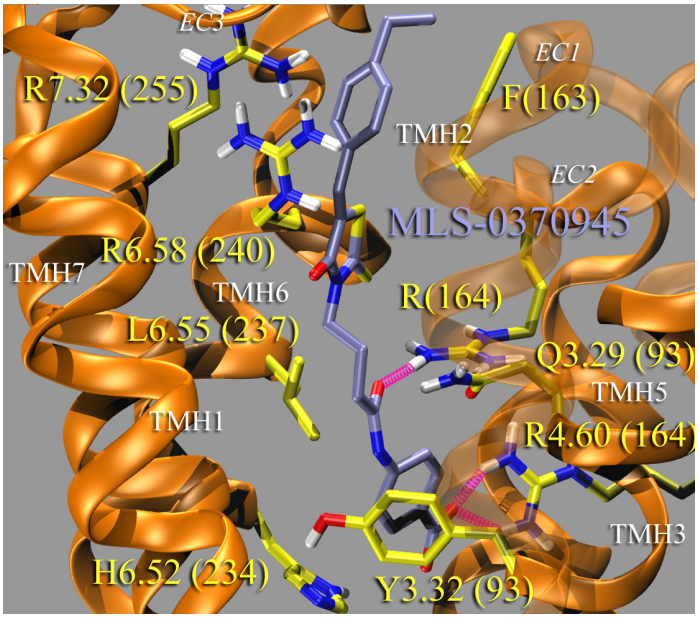


Figure 33. The protein-ligand optimized dock of the MLS-0370945 compound in a minimized hGPR35 wild-type R bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 3 kcal/mol.

Table 6. The protein-receptor EOI of the MLS-03070945-hGPR35 R bundle. A) The total protein-receptor EOI of the MLS-03070945-hGPR35 R bundle, with Van der Waals (VdW) and electrostatic contribution distributions, the ligand conformational cost of docking, and the experimental calculated IC_{50} (Molecular Libraries Probe Production Center at the Sanford-Burnham Institude, AID: 2079). B) An comprehensive list of all residue-ligand EOI that are ≥ 3 kcal/mol.

A.

	Human WT	
Electrostatic Total	-31.19	Electrostatic (kcal/mol)
VdW Total	-40.04	VdW (kcal/mol)
Lig/Rcptr Energy	**-71.23**	**Total Energy (kcal/mol)**
Lig Dock Conformational Cost	1.23	kcal/mol
Total Int. Energy	**-70.01**	**kcal/mol**
IC_{50}	71.8 nM [1]	

B.

Residue	Human WT (kcal/mol)	Interaction
Q3.29 (93)	-6.62	VdW
Y3.32 (96)	-5.10	VdW
R4.60 (151)	-13.73	Salt Bridge
F(163)	-4.01	VdW
R(164)	-8.30	Hydrogen Bond
H6.52 (234)	-3.34	VdW
L6.55 (237)	-4.96	VdW
R6.58 (240)	-3.88	VdW
R7.32 (255)	-3.73	VdW

is locked in an inactive, $\chi 1 = g+$ conformation by the carboxyl functional group of the MLS compound, blocking the $\chi 1$ rotamer change of F6.48(230) (Figure 34). While there is no binding data for the MLS-0370945 compound, the low IC_{50} value of 71.8 nM would suggest that this compound would dock in the GPR35 R bundle with strong interactions, which correlates well with this docking study.

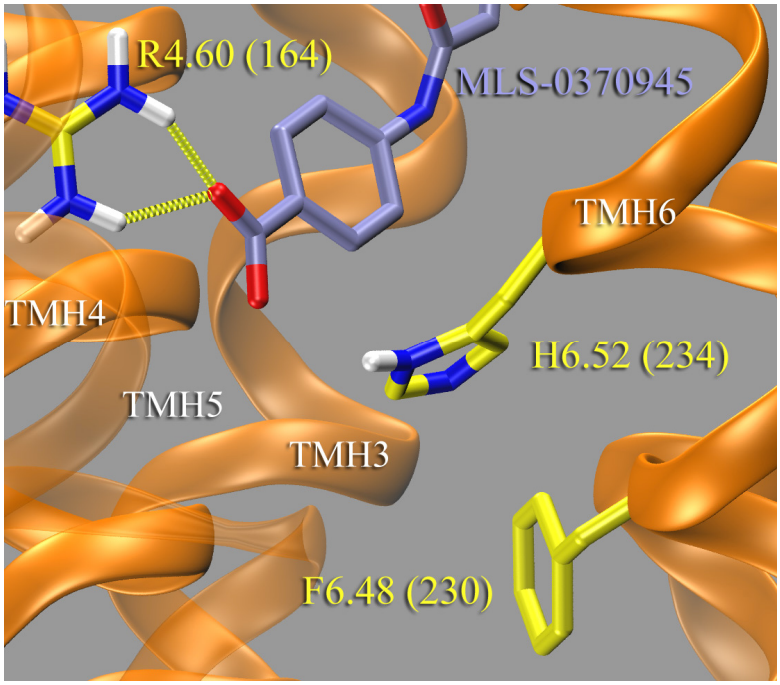


Figure 34. The optimized MLS-0370945-hGPR35 R bundle which highlights that the MLS-0307045 compound locks the hGPR35 in an inactive state. The position of the carboxyl functional group of the MLS-0307045 compound blocks the rotameric change of H6.52 (234), locking the GPCR in an inactive state.

Pamoic acid

The initial conformation of pamoic acid for each manually-docked docking study was performed with the global minimum structure as calculated at the HF 6-31+G* level of theory. The strength of each residue-ligand EOI is depicted as a series of energy categories, with the displayed carbon atoms of each depicted residue identifying its particular category. The category with the largest magnitude has a range of 14.5 -12.0 kcal/mol, with the remaining 3 sequential categories having the same 2.5 kcal/mol range. Each total protein-ligand EOI includes the conformational cost of pamoic acid incurred during docking.

Wild-type Human GPR35 R*

In the optimized dock of pamoic acid in a wild-type GPR35 R* bundle (Figure 35) the total protein-ligand EOI is -78.16 kcal/mol, which has a slightly larger electrostatic component of approximately 60%. There are three individual ligand-residue EOIs, involving pamoic acid and the residues R3.36(100), R4.60(151), and R6.58(240), that contribute the majority of the total protein-ligand EOI, with an approximate combined contribution of 48%. The residues R3.36(100) and R6.58(240), which form salt bridges with the di-anionic pamoic acid, are the largest contributors with a combined approximate 34% of the total ligand-protein EOI, with corresponding EOI values of -13.70 and -12.97 kcal/mol. Residue R4.60(151), which contributes approximately 14% (-10.87 kcal/mol) of the total EOI, forms a cation-π interaction with a naphthalene functional group of pamoic acid. Additional residue-ligand interactions occur between the residues S5.39(172), T(166), R(164), and F(163) and pamoic acid. Both S5.39(172) and T(166) donate hydrogen bonds to pamoic acid, while R(164) and F(163), both on the EC2 loop, interact with pamoic acid mostly through mostly Van der Waals interactions.. Additionally, the ligand incurred a very minor conformational cost of docking of 0.04 kcal/mol (Table 7A). All

abovementioned residues, which all contribute -4.5 kcal/mol or less, are displayed in Table 7B, with both the values of the EOI and the types of interactions listed. Noted is that residue R7.32(255) is unable to interact with pamoic acid due to bulk of the juxtaposed residue, R6.58(240) (Figure 35).

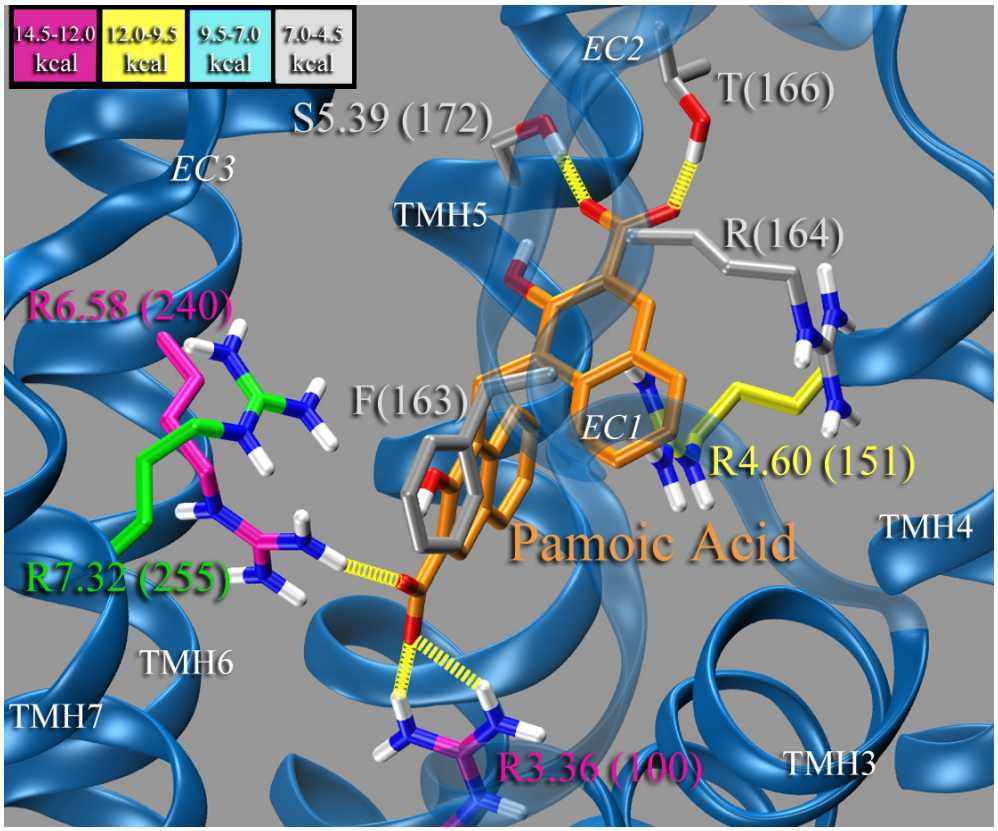


Figure 35. The protein-ligand optimized dock of pamoic acid in a minimized hGPR35 wild-type (WT) R* bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 4.5 kcal/mol. The carbon atoms of each residue are colored in accordance with the magnitude of the residue's EOI with the ligand. The residue R7.32 (255) (green), is not interacting with the ligand, but is displayed to identify its location and conformation in the hGPR35 WT R*bundle optimized with pamoic acid.

Table 7. The protein-receptor EOI of pamoic acid in a hGPR35 wild-type R* bundle. A) The total protein-receptor EOI of pamoic acid in a hGPR35 wild-type R* bundle, with Van der Waals (VdW) and electrostatic component distributions, the ligand conformational cost of docking, and the experimental calculated EC_{50}s from (1) Jenkins et al. (2010) Biochem J, 2) Abood, M. Temple University, Unpublished data 3) Zhao et al. (2010) Mol Pharmacol 78) B) A comprehensive list of all residue-ligand EOI that are ≥ 4.5 kcal/mol.

A.

	Human WT	
Electrostatic Total	-46.99	Electrostatic (kcal/mol)
VdW Total	-31.21	VdW (kcal/mol)
Lig/Rcptr Energy	**-78.20**	**Total Energy (kcal/mol)**
Lig Dock Conformational Cost	0.04	kcal/mol
Total Int. Energy	**-78.16**	**kcal/mol**
EC_{50}	51.3[1],52.0[2], 79[3], 62.3[2] nM	

B.

Residue	Human WT (kcal/mol)	Interaction
R3.36 (100)	-13.70	Salt Bridge
R4.60 (151)	-10.87	Cation-π
F(163)	-5.77	VdW
R(164)	-5.10	VdW
T(166)	-6.30	Hydrogen Bond
S5.39 (172)	-5.44	Hydrogen Bond
R6.58 (240)	-12.97	Salt Bridge

Human R6.58A Mutant GPR35 R*

The optimized dock of pamoic acid in a R6.58A(240) mutant GPR35 R* bundle (Figure 36) has a total protein-ligand EOI of -80.55 kcal/mol, with an approximate 61% electrostatic component (Table 8A). In the R6.58A(240)-pamoic acid dock, the three residues that contribute the majority of the total protein-ligand EOI are the residues R3.36(100), R4.60(151), and R7.32(255), with an approximate contribution of 43% of the total. The residues R3.36(100) and R7.32(255) both have salt-bridges with pamoic acid and contribute -13.48 and -12.28 kcal/mol, respectively. Similar to the wild-type dock, the R6.58A(240) mutant dock has a cation-π interaction between R4.60(151) and pamoic acid, which contributes a residue-ligand EOI of -8.65 kcal/mol. Additional significant residue-ligand EOIs originate from the hydrogen bonds with pamoic acid and the residues S5.39(172) and T(166), and the Van der Waals interactions with the

ligand and the residues S(165) and R(164) on the EC2 loop. The residue-ligand EOIs of the ligand and the residues S5.39(172), T(166), S(165), and R(164) are -5.62, -7.00, -5.14, and -6.58 kcal/mol, respectively. As compared to the total protein-ligand EOI, the conformational cost of docking was nearly insignificant, with a value of 0.08 kcal/mol (Table 8A). All of the aforementioned residues, which all have a minimum residue-ligand EOI magnitude of 4.5 kcal/mol, are displayed in Table 8B, with both the values of the EOI and the types of interactions listed.

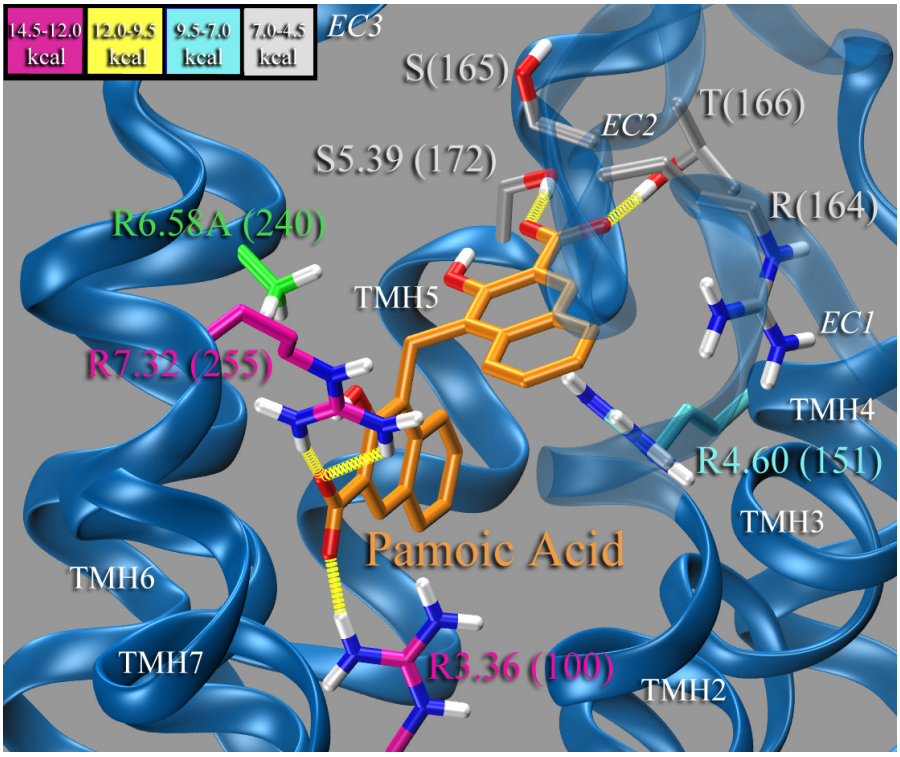


Figure 36. The protein-ligand optimized dock of pamoic acid in a minimized hGPR35 R6.58A (240) mutant R* bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 4.5 kcal/mol. The carbon atoms of each residue are colored in accordance with the magnitude of the residue's EOI with the ligand. The residue R6.58A (green), does not interaction with the ligand, but is displayed to identify its position and bulk.

Table 8. The protein-receptor EOI of pamoic acid in a hGPR35 R6.58A(240) mutant R* bundle. A) The total protein-receptor EOI of pamoic acid in a hGPR35 R6.58A (240) mutant R* bundle, with Van der Waals (VdW) and electrostatic component distributions, the ligand conformational cost of docking, and the experimental calculated EC_{50}s from (1) Abood, M. Temple University, Unpublished data B) A comprehensive list of all residue-ligand EOI that are ≥ 4.5 kcal/mol.

A.

	Human R6.58A Mutant	
Electrostatic Total	-49.19	Electrostatic (kcal/mol)
VdW Total	-31.44	VdW (kcal/mol)
Lig/Rcptr Energy	**-80.63**	**Total Energy (kcal/mol)**
Lig Dock Conformational Cost	0.08	kcal/mol
Total Int. Energy	**-80.55**	**kcal/mol**
EC_{50}	16.0 nM [1]	

B.

Residue	Human R6.58A Mutant (kcal/mol)	Interaction
R3.36 (100)	-13.48	Salt Bridge
R4.60 (151)	-8.65	Cation-π
R(164)	-6.58	VdW
S(165)	-5.14	VdW
T(166)	-7.00	Hydrogen Bond
S5.39 (172)	-5.62	Hydrogen Bond
R7.32 (255)	-12.28	Salt Bridge

Wild-type Rat GPR35 R*

The final, optimized dock of pamoic acid in a wild-type rat GPR35 R* bundle (Figure 37) has a protein-ligand EOI of -64.65 kcal/mol, with an approximate 55% Van der Waals component (Table 9A). The two residue-ligand EOIs that contribute the most to the protein-ligand EOI are from the residues R4.60(148) and R3.36(97), with a 14% and 11% respective contribution. The residue R3.36(97) interacts with the ligand mostly electrostatically, with an EOI of -7.43 kcal/mol. Similar to both of the hGPR35 docks, the residue R4.60(148) in rGPR35 interactions with pamoic acid via a cation-π interaction, with a residue-ligand EOI of -9.17 kcal/mol. Residues Y3.32(93), S5.38(169), and Q(163) all donate hydrogen bonds to the carboxyl functional groups of pamoic acid, while the residue F(160) on the EC2 loop interacts with the ligand via a mostly a Van der Waals interaction. The residue-ligand EOIs of residues Y3.32(93), S5.38(169), Q(163), and F(160) are –6.29, -7.01, -6.62, and –5.54 kcal/mol, respectively. The conformational cost of

docking was more substantial in the rGPR35 dock as compared to the hGPR35 docks, with an energy cost of 0.58 kcal/mol (Table 9A). Table 9B includes an inclusive list of the residue-ligand EOI with a magnitude equal to or greater that 4.5 kcal/mol, as well as the coordinated interaction type. Noted is that there are no residue-ligand EOIs in the rGPR35-pamoic acid dock that meet the criterion of the first two energy categories, which is populated in both the human wild-type and R6.58A(240) mutant docks with pamoic acid. Also noted is that the relative position of pamoic acid is shifted in the wild-type rGPR35-pamoic acid dock, as compared to the analogous hGPR35 wild-type dock, with a relative shift of approximately 3Å (Figure 38).

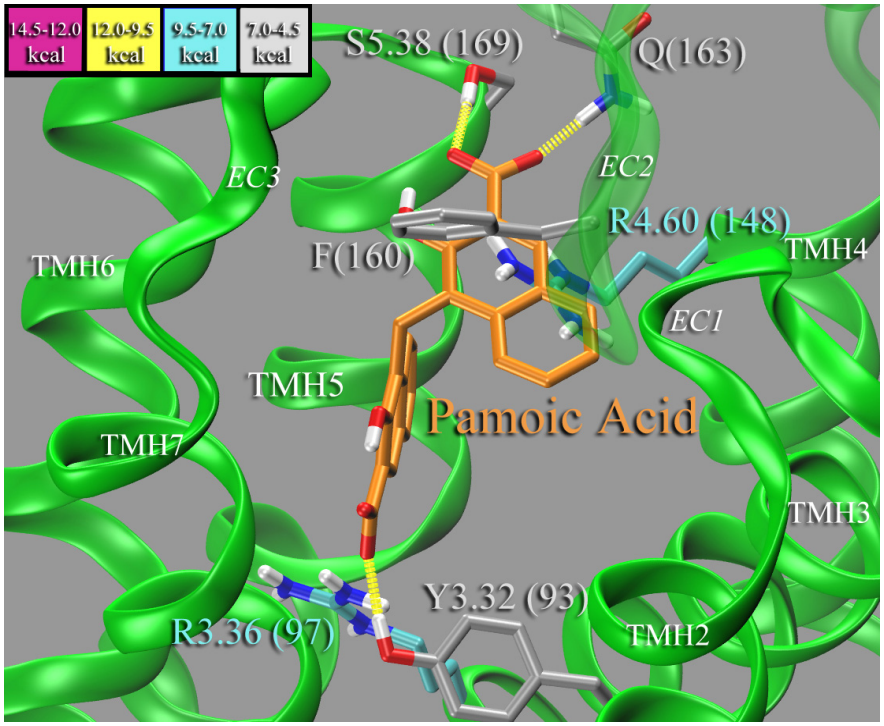


Figure 37. The protein-ligand optimized dock of pamoic acid in a minimized rGPR35 wild-type R* bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 4.5 kcal/mol. The carbon atoms of each residue are colored in accordance with the magnitude of the residue's EOI with the ligand.

Table 9. The protein-receptor EOI of pamoic acid in a rGPR35 wild-type R* bundle. A) The total protein-receptor EOI of pamoic acid in a rGPR35 wild-type R* bundle, with Van der Waals (VdW) and electrostatic component distributions, the ligand conformational cost of docking, and the experimental calculated EC_{50} from (1) Jenkin *et al.* (2010) Biochem J B) A comprehensive list of all residue-ligand EOI that are ≥ 4.5 kcal/mol.

A.

	Rat WT	
Electrostatic Total	-29.42	Electrostatic (kcal/mol)
VdW Total	-35.23	VdW (kcal/mol)
Lig/Rcptr Energy	**-64.65**	**Total Energy (kcal/mol)**
Lig Dock Conformational Cost	0.58	kcal/mol
Total Int. Energy	**-64.07**	**kcal/mol**
EC_{50}	> 100 μM[1]	

B.

Residue	Rat WT (kcal/mol)	Interaction
Y3.32 (93)	-6.29	Hydrogen Bond
R3.36 (97)	-7.43	Electrostatic
R4.60 (148)	-9.17	Cation-π
F(160)	-5.54	VdW
Q(163)	-6.62	Hydrogen Bond
Ser5.38 (169)	-7.01	Hydrogen Bond

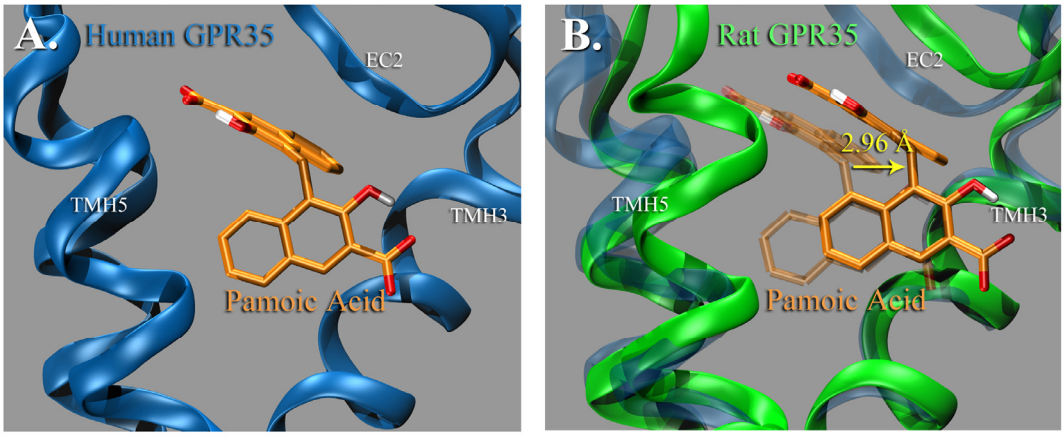


Figure 38. The comparison of the position of pamoic acid in a optimized hGPR35 vs. rGPR35 wild-type R* bundle. A) Pamoic acid docked (optimized) in a minimized hGPR35 wild-type R* bundle (blue) B) Minimized rGPR35 R* bundle (green) overlaid on the minimized hGPR35 R* bundle (transparent), with the noted relative shift of the position of pamoic acid in the rGPR35 (opaque) of 2.96 Å (distance measured from the carbon atoms linking the rings of pamoic acid)

Zaprinast

The initial conformation of zaprinast for each manually-docked, docking study was performed with the first low energy conformational and tautomeric state of zaprinast, as calculated at the HF 6-31+G* level of theory. The strength of each residue-ligand EOI is represented as described previously, using a series of energy categories. The category with the largest magnitude has a range of 6.0 - 4.8 kcal/mol, with the remaining 3 sequential categories having the same 1.2 kcal/mol range. Each total protein-ligand EOI includes the conformational and tautomeric cost of docking the ligand.

Wild-type Human GPR35 R*

Zaprinast docked in a human GPR35 R* bundle with optimized protein-ligand interactions (Figure 39) had a measured total protein-ligand EOI of -24.24 kcal/mol (Table 10A). The vast majority of the protein-ligand EOI between Zaprinast and the wild-type hGPR35 are from Van der Waals interactions, with an approximate 94% contribution. Noted is that there are no individual residue-ligand interactions that populate the first energy category. The interaction with the largest magnitude is the experimentally suggested [57] residue-ligand interaction between Y3.32(96) and Zaprinast, with a contribution of -3.92 kcal/mol, which is approximately 15% of the protein-ligand EOI. Another significant residue-ligand EOI is the tilted-T aromatic stack between F5.47(180) and Zaprinast, with a value of -3.25 kcal/mol. There are a series of leucines and a glutamine, L3.33((97), L6.51(233), L6.55(237), and Q3.29(93), that together contribute approximately 44% of the protein-ligand EOI with respective EOI values of -2.76, -3.07, -3.19, and -2.76 kcal/mol. The abovementioned residues proximity and EOI are visualized and listed in Figure 39 and Table 10B, respectively. Additional residue-ligand interactions include the mostly Van der Waals interactions between zaprinast and the residue R6.58(240) and the residue C(162),

located on the EC2 loop, with respective EOI values of -2.06 and -1.81 kcal/mol. The comformational and tautomeric cost of docking was significant, with an energy of 2.77 kcal/mol (comformational: 1.13 kcal/mol, tautomeric: 1.64 kcal/mol) Noted is that the residues R7.32(255) and S7.39(262) (Figure 39) have little to no interaction with zaprinast, but their positions and interactions are integral to be able to explain the ligand binding difference in the R6.58A(240) mutant- specifically that in the wild-type dock S7.39(262) is hydrogen bonded to Y3.32(96).

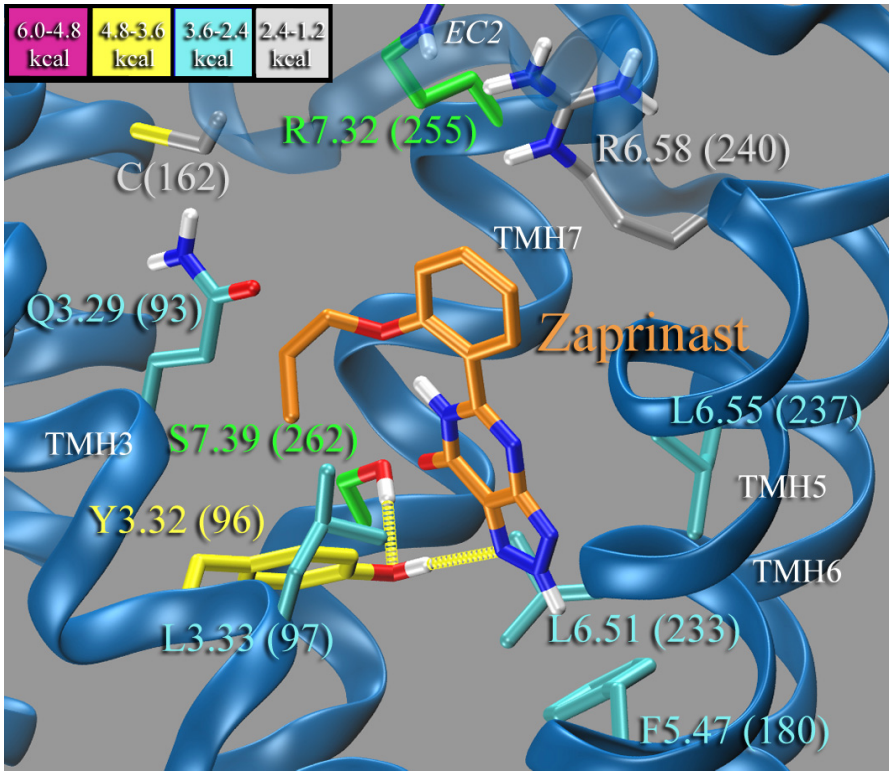


Figure 39. The protein-ligand optimized dock of Zaprinast in a minimized hGPR35 wild-type R* bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 1.2 kcal/mol. The carbon atoms of each residue are colored in accordance with the magnitude of the residue's EOI with the ligand. The residues R7.32 (255) and S7.39 (262) (green) have little to no interaction with the ligand, but are displayed to identify their location and or interaction partners.

Table 10. The protein-receptor EOI of Zaprinast in a hGPR35 wild-type R* bundle. A) The total protein-receptor EOI of Zaprinast in a hGPR35 wild-type R* bundle, with Van der Waals (VdW) and electrostatic component distributions, the ligand's conformational/tautomeric cost of docking, and the experimental calculated EC_{50}s from (1)Taniguchi et al. (2006) Febs Lett 580(21), (2) Zhao et al. (2010) Mol Pharmacol 78(4), (3) Abood, M. Temple University, Unpublished data, (4) Taniguchi et al (2008) Pharmacology 82(4), and (5)Jenkins et al. (2010) Biochem J B) A comprehensive list of all residue-ligand EOI that are ≥ 1.2 kcal/mol.

A.

	Human WT	
Electrostatic Total	-1.74	Electrostatic (kcal/mol)
VdW Total	-25.27	VdW (kcal/mol)
Lig/Rcptr Energy	-27.01	**Total Energy (kcal/mol)**
Lig Dock Conformational /Tautomerization Cost	2.77	kcal/mol
Total Int. Energy	**-24.24**	**kcal/mol**
EC_{50}	0.84[1], 1.0[2,3], 1.62[4], 2.95[5], 2.8[3] µM	

B.

Residue	Human WT (kcal/mol)	Interaction
Q3.29 (93)	-2.76	VdW
Y3.32 (96)	-3.92	Hydrogen Bond
L3.33 (97)	-2.76	VdW
C(162)	-1.81	Tilted-T
F5.47 (180)	-3.25	VdW
L6.51 (233)	3.07	VdW
L6.55 (237)	-3.19	VdW
R6.58 (240)	-2.06	VdW

Human R6.58A Mutant GPR35 R*

The final optimized dock of zaprinast in a human R6.58A(240) mutant GPR35 R* bundle (Figure 40) had a calculated final protein-ligand EOI of -28.65 kcal/mol (Table 11A). Similar to Zaprinast docked in the human wild-type, the majority of the protein-ligand EOI in the R6.58A(240) mutant dock is from Van der Waals interactions, with an approximate 90% contribution. The residue-ligand EOI that has the largest magnitude is Y3.32(96), which is hydrogen bonded to Zaprinast, with an EOI of -5.47 kcal/mol. In addition to an electrostatic component, the Y3.32(96)-Zaprinast EOI also includes a largely contributing Van der Waals component. This residue-ligand EOI populates the highest energy category and contributes approximately 18% of the total protein-ligand EOI. A significant energy contribution also comes

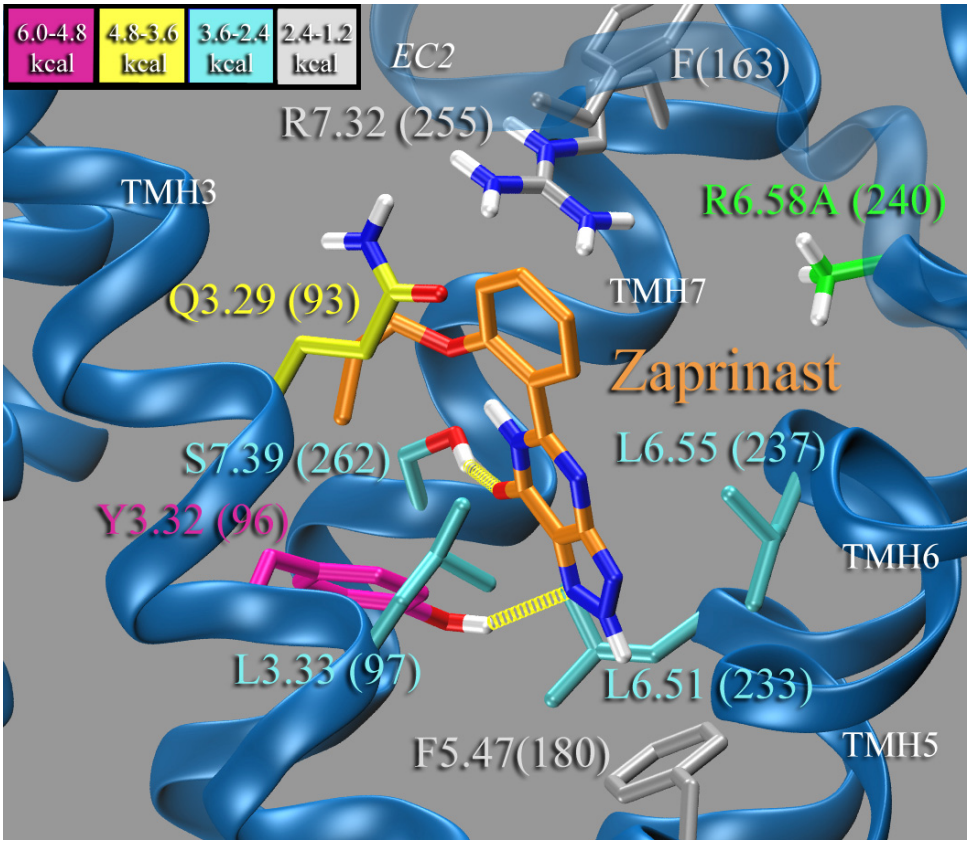


Figure 40. The protein-ligand optimized dock of Zaprinast in a minimized hGPR35 R6.58A (240) mutant R* bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 1.2 kcal/mol. The carbon atoms of each residue are colored in accordance with the magnitude of the residue's EOI with the ligand. The residue R6.58A (green) does not interaction with the ligand, but is displayed to identify its position and bulk.

Table 11. The protein-receptor EOI of Zaprinast in a hGPR35 R6.58A (240) R* bundle. A) The total protein-receptor EOI of Zaprinast in a hGPR35 R6.58A (240) R* bundle, with Van der Waals (VdW) and electrostatic component distributions, the ligand's conformational/tautomeric cost of docking, and the experimental calculated EC_{50} from (1) Abood, M. Temple University, Unpublished data B) A comprehensive list of all residue-ligand EOI that are ≥ 1.2 kcal/mol.

A.

	Human R6.58A Mutant	
Electrostatic Total	-3.02	Electrostatic (kcal/mol)
VdW Total	-27.90	VdW (kcal/mol)
Lig/Rcptr Energy	-30.92	Total Energy (kcal/mol)
Lig Dock Conformational /Tautomerization Cost	2.27	kcal/mol
Total Int. Energy	**-28.65**	**kcal/mol**
EC_{50}	0.08 [1]μM	

B.

Residue	Human R6.58A Mutant (kcal/mol)	Interaction
Q3.29 (93)	-4.01	VdW
Y3.32 (96)	-5.47	Hydrogen Bond
L3.33 (97)	-2.63	VdW
F(163)	-1.24	VdW
F5.47 (180)	-2.37	Tilted-T
L6.51 (233)	-2.94	VdW
L6.55 (237)	-3.33	VdW
R7.32 (255)	-1.24	Cation-π
S7.39 (262)	-2.69	Hydrogen Bond

from the mostly Van der Waals interaction between the ligand and the residues Q3.29(93), L3.33(97), L6.51(233), L6.55(237), and F(163) on the EC2 loop with respective EOI values of -4.01, -2.63, -2.94, -3.33, and -1.24 kcal/mol. The additional residue-ligand interactions that have a EOI magnitude of greater than 1.2 kcal/mol are a hydrogen bond with S7.39(262), a tilted-T aromatic stack with F5.47(180), and a cation-π interaction with R7.32(255), with the corresponding EOI of -2.69, -2.37, -1.24 kcal/mol. The energy cost of docking Zaprinast in the hGPR35 R6.58A(240) mutant was 2.27 kcal/mol (conformational: 0.63 kcal/mol, tautomeric: 1.64 kcal/mol). Table 11 displays a quantitative synopsis of this dock, including the protein-ligand EOI, the conformational and tautomeric cost of docking, and an inclusive list of the values of all residue-ligand interactions with a magnitude above 1.2 kcal/mol.

The cation-π interaction between zaprinast and R7.32(255) occurs between the propoxy-benzyl functional group of zaprinast and the guanidinium group of R7.32(255) (Figure 41). The central carbon on the guanidinium group has a measured distance of 4.152 Å from the plane of the benzene, which is similar to the computationally-determined ideal distance of 3.96 Å for analogous molecules [97]. The angle of the involved N-H atoms of the guanidinium group of R7.32 and the benzene centroid of zaprinast was measured as 162.69°, which is comparable to the computationally-derived ideal angle of 180° [97].

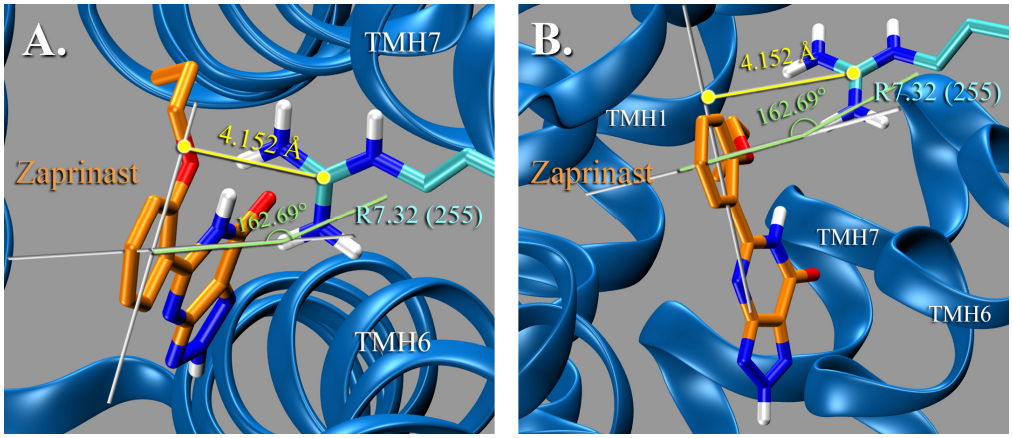


Figure 41. Depiction of the quantitative assessment of the cation-π interaction between R7.32(255) and Zaprinast in the optimized hGPR35 R6.58A mutant R* dock. This interaction is depicted from an extracellular top (A) and side perspective (B). The central carbon of the guanidinium functional group of R7.32 (255) has a distance of 4.152 Å from the plane of propoxy-benzyl functional group of Zaprinast and the angle of the involved N-H atoms of the guanidinium group of R7.32 and the benzene centroid of Zaprinast was measured as 162.69°.

Wild-type Rat GPR35 R*

The finalized, optimized zaprinast dock in a rat GPR35 R* bundle (Figure 42) has a measured total protein-ligand EOI of -31.58 kcal/mol. Similar to zaprinast docked in human GPR35, the type of non-bonded protein-ligand interaction that contributed the most towards the protein-ligand EOI in the rat dock was determined as Van der Waals interactions, with an

approximate 90% contribution. In the optimized rGPR35-zaprinast dock there are two residue-ligand EOIs that populate the energy category with the largest magnitude, Y3.32(93) and L7.35(256). These two residues both individually contribute approximately 5 kcal/mol, and there summed residue-ligand EOIs contribute about 30% of the protein-ligand EOI. Y3.32(93) donates a hydrogen bond to zaprinast, but the majority of the EOI for both the residues Y3.32(93) and L7.35(256) is due to Van der Waals interactions. S5.43(174) accepts a hydrogen bond from Zaprinast and has an residue-ligand EOI of -2.04 kcal/mol. Additionally, the residues Q3.29(90), L3.33(94), L6.55(235), Q6.58(238), L7.35(256), T7.36(257), and F(160) on the EC2 loop contribute approximately 62% of the protein-ligand EOI, mostly via Van der Waals interactions, with the respective residue-ligand EOIs of -3.09, -2.15, -3.35, -3.47, -5.08, -2.04, and -1.92 kcal/mol. The conformational and tautomeric cost of docking of zaprinast were similar in the wild-type rat dock as compared to the human docks, with an energy cost of 2.31 kcal/mol (conformational: 0.67, tautomeric: 1.64 kcal/mol) (Table 12A). Table 12 lists the total protein-ligand EOI, the ligand conformational and tautomerization costs, and the EOI value and the type of interaction of all residue-ligand interactions with an EOI magnitude of greater than or equal to 1.2kcal/mol.

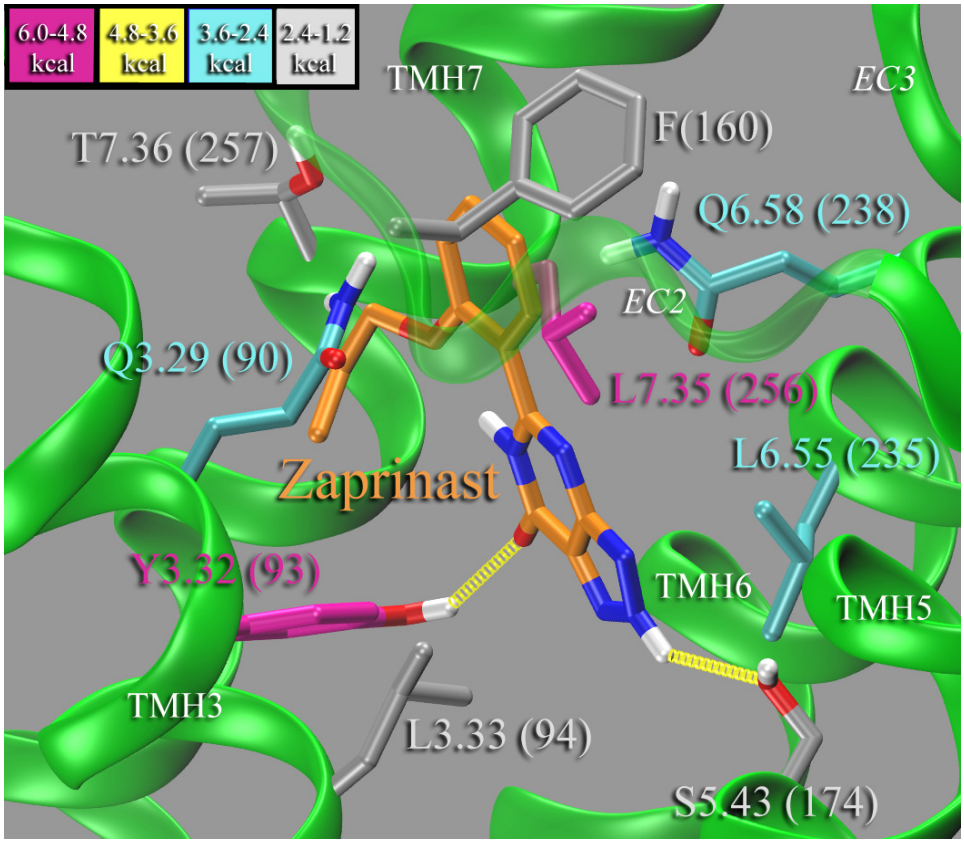


Figure 42. The protein-ligand optimized dock of Zaprinast in a minimized rGPR35 wild-type R* bundle. The residues displayed represent an inclusive list of the residues that have an EOI magnitude of ≥ 1.2 kcal/mol. The carbon atoms of each residue are colored in accordance with the magnitude of the residue's EOI with the ligand.

Table 12. The protein-receptor EOI of Zaprinast in a rGPR35 wild-type R* bundle. A) The total protein-receptor EOI of Zaprinast in a rGPR35 wild-type R* bundle, with Van der Waals (VdW) and electrostatic component distributions, the ligand's conformational/tautomeric cost of docking, and the experimental calculated EC_{50}s from (1)Taniguchi et al. (2006) Febs Lett 580(21), (2) Ohshiro et al. (2008) Biochem Biophys res commun 365(2), (3)Jenkins et al. (2010) Biochem J, and(4) Taniguchi et al (2008) Pharmacology 82(4) B) A comprehensive list of all residue-ligand EOI that are ≥ 1.2 kcal/mol.

A.

	Rat WT	
Electrostatic Total	-3.32	Electrostatic (kcal/mol)
VdW Total	-30.57	VdW (kcal/mol)
Lig/Rcptr Energy	**-33.89**	**Total Energy (kcal/mol)**
Lig Dock Conformational /Tautomerization Cost	2.31	kcal/mol
Total Int. Energy	**-31.58**	**kcal/mol**
EC_{50}	0.016[1], 0.290[2], 0.047[3], 0.676[4] µM	

B.

Residue	Rat WT (kcal/mol)	Interaction
Q3.29 (90)	-3.09	VdW
Y3.32 (93)	-4.85	Hydrogen Bond
L3.33 (94)	-2.15	VdW
F(160)	-1.92	VdW
S5.43 (174)	-2.04	Hydrogen Bond
L6.55 (235)	-3.35	VdW
Q6.58 (238)	-3.47	VdW
L7.35 (256)	-5.08	VdW
T7.36 (257)	-2.04	VdW

CHAPTER IV

DISCUSSION

The goal of this project was to build accurate GPR35 models to be able to suggest viable reasons for the pharmacological differences or similarities between human wild-type, rat wild-type, and a R6.58A(240) human GPR35 mutant for the ligands pamoic acid and zaprinast. As no crystal structure is currently available for GPR35, a homology model was built using the crystal structures templates of β2-AR and A2-AR for the TMH region. Though GPR35 and these template GPCRs share many conserved residues/motifs, GPR35 has several non-conservative sequence differences. These non-conservative sequence differences, such as the lack of a proline in TMH7 in the highly conserved NPXXY motif, are suggested to cause gross differences in the TMH region of GPR35 vs. the crystal structure template. Important to note is that in GPCRs a sequence divergence, such a lack of a conserved proline near the center of a helical structure, is probable to not only cause major structure changes in the orientation of the extracellular positions of the divergent TMH, but will also affect the orientation of the juxtaposed helices. The implications of these sequence divergences were assessed using the Monte Carlo, simulated annealing technique, Conformational Memories (CM). This technique has corroborated that many of these non-conservative sequence divergences cause structural difference between GPR35 and the crystal structure templates.

As mentioned earlier, the TMH7 of GPR35 lacks the NPXXY motif and instead has a sequence of DAICY. TMH7 was initially suggested to diverge the most severely from the template structures and therefore was assessed first. For GPR35, the lack of an NPXXY motif in appears create a much straighter TMH7 than the template structures. The shift in the extracellular

position of TMH7 is suggested by CM to be compensated for in the structure of GPR35 by the additional variation of a proline at 2.58 instead of at the locant 2.59 in TMH2. These changes work together to yield a conformational difference in the TMH1-2-7 region of GPR35, but creates an overall cohesive receptor conformation. The lack of a conserved proline in TMH4 (P4.60) causes a conformational change in TMH4, moving the extracellular end of TMH4 closer to TMH5. Such a conformational change would allow for a shorter EC2 loop, which is accommodated by the GPR35 sequence, as GPR35 has a shortened EC2 loop. These studies suggest that the influence of a single conformational change may be pervasive, but is ultimately compensated for by additional, sequence-dictated conformational changes in the approximal helices.

After the initial development and minimization of the active forms of a human wild-type, human R6.58A mutant, and rat GPR35 models, docking studies were performed with the ligands pamoic acid and zaprinast. The assessment of each dock was based on the magnitude of each individual protein-ligand "energy of interaction" (EOI). The calculated protein-ligand EOI may be representative of the actually binding affinity in physiological conditions, though several variables are unable to be considered in this experimental system. The impact that these variables, described below, would have on the EOI is assumed to be similar in each bundle with the same docked ligand. This hypothesis suggests that while the measured protein-ligand EOI may not be a direct reflection of the experimental binding energy of the ligand, the differences in energy of the comparative docks with the same ligand should be accurate representations of the differences in ligand-binding energy. Essentially by comparing the protein-ligand EOI, in lieu of assessing them separately, the results are essentially normalized. Each set of comparisons described will evaluate the differences and similarities of the human wild-type docks vs. the rat wild-type or human R6.58A(240) mutant GPR35 docks.

The probable most-influential variable not able to be directly considered is the effect of the solvent. It is assumed that water permeates the TMH4-5-6 channel, though the number of water molecules is unknown and most likely transient. The concentration of water in the binding pocket would have multiple effects on the accurate measurement of the protein-ligand EOI. One of the most prominent effects that water has on the strength of protein-ligand interactions lies in the ability of water to shield the electrostatic component of protein-ligand interaction. The amount of shielding of a solvent can be represented by a dielectric constant. The dielectric constant varies with different solvents, with examples of tetrachloro-methane, ethyl acetate, ethanol, and water having approximate respective dielectrics of 2.23, 5.99, 24.85, and 78.0 [98]. These values are canonical representations of a real solvent, as they represent stationary, completely homogonous bulk solvent. There are multiple, more accurate ways of representing solvent effects, such as the Generalized Born with solvent accessible surface area solvation model (GBSA) [99]. The GBSA model considers not only the shielding affect of the solvent, but also considers the energy cost of creating a hole in an equilibrated solvent, the energy of the polarization of the solvent, and the surface accessibility of the molecule of interest to the solvent. Unfortunately implicit solvation models, such as the GBSA model, require that the molecule of interest be in a bulk solvent, which would make this model ineffectual for determining the solvent effects on ligands docked in the current GPR35 model. For these reasons, the method used to emulate the effects of water on the protein-ligand EOI only considered the shielding effect. A distance-dependent dielectric, which increases with distance, was used to represent the electrostatic dampening effects of water on the protein-ligand interactions. The fundamental concept of why a distance-dependent dielectric can emulate the shielding effects of water is that once the distance between two atoms reaches a threshold distance, the dielectric will be equal to the high-dielectric of bulk water. The dielectric constant of the distance-dependent dielectric

chosen for the protein-ligand minimizations and the EOI calculations was three. There have been a variety of dielectric constants chosen to represent the interior of GPCRs, including the dielectric constants of 1 [100], 2 [101] 4 [102-103] , and 5 [104]. The value of three was chosen for the analysis of GPR35, as it is representative of the literature values.

The first sets of docks to be compared are the human wild-type docks vs. the rat wild-type GPR35 docks. There is pharmacological experimental data for both human and rat GPR35 that measure the potency of both the ligands pamoic acid and zaprinast. These potencies, which are from experiments performed by collaborators and other research groups, show a strong continuity in the measured potency between groups (Table 13). The potencies of pamoic acid for human wild-type vs. rat wild-type GPR35 are significantly different from one another (Table 13A), with an average EC_{50} value of approximately 61 nM for human wild-type and an EC_{50} of greater than 100 μM in rat wild-type GPR35. Important to note is that though pamoic acid had a significantly lower potency in rat GPR35, there was detected activation [57]. Unpublished results (Mary Abood, Temple Univeristy) also suggested a similar trend with pamoic acid at mouse GPR35, with an approximate EC_{50} of 5μM. Mouse and rat GPR35 have an 85% identity, with an identical match of all of the suggested primary interaction residues between pamoic acid and rGPR35, including a serine at 5.43. The dose-dependent antinociceptive effect of pamoic acid was also detected *in vivo* in mouse, but with the large A_{50} of 40.5 mg/kg [59]. The large sequence identity between mouse and rat GPR35, particularly the fact that the suggested direct interaction residues in rat are identical between species, strongly suggests these results in mouse would parallel those in rat. The potency of zaprinast is also variable between human and rat GPR35. The EC_{50} values for zaprinast have been calculated by multiple experimental groups, with a statistically significant difference ($p < 0.001$) of EC_{50}s between species, with an approximately 62 fold increase in rat over human GPR35 [57].

Table 13. EC_{50}s identified for pamoic acid and Zaprinast in human and rat wild-type GPR35. A) EC_{50}s identified for pamoic acid in human and rat wild-type GPR35 from (1) Jenkins et al. (2010) Biochem J, (2) Abood, M. Temple University, Unpublished data, and (3) Zhao et al. (2010) Mol Pharmacol 78(4) B) EC_{50}s identified for Zaprinast in human and rat wild-type GPR35 from (1)Taniguchi et al. (2006) Febs Lett 580(21), (2) Zhao et al. (2010) Mol Pharmacol 78(4), (3) Abood, M. Temple University, Unpublished data, (4) Taniguchi et al (2008) Pharmacology 82(4), (5) Jenkins et al. (2010) Biochem J, and (6) Ohshiro et al. (2008) Biochem Biophys res commun 365(2
***Denotes statistically significant difference ($p < 0.0001$)**

A.

	Human WT	Rat WT
EC_{50}	51.3[1], 52.0[2], 62.3[2], 79[3], nM	> 100[1] µM

B.

	Human WT	Rat WT
EC_{50}	0.84 [1], 1.0 [2,3], 1.62 [4], 2.95*[5], 2.8 [3] µM	0.016 [1], 0.290 [6], 0.047*[5], 0.676 [4] µM

Comparison of the EOIs of the human and rat wild-type GPR35 docks with pamoic acid offer an explanation for these substantial interspecies potency differences of pamoic acid. Pamoic acid has an increase of 14.09 kcal/mol in rat GPR35 as compared to human (Table 14). As a larger, or less negative, EOI value represents a weaker protein-ligand interaction, the increase of approximately 14 kcal/mol represents a substantially reduced protein-ligand interaction of pamoic acid in rat as compared to human GPR35. This divergence can be mostly attributed to a small number of residue-ligand interaction differences, involving residues at positions 3.36, 6.58, and the residue on the EC2 loop that is two positions past the cysteine that forms a conserved disulfide bridge with C3.25. As noted in earlier, rat GPR35 does not have any residue-ligand EOI that are greater than 9.5 kcal/mol (Table 9B), while the human wild-type GPR35 dock has three (Table 7B and Table 8B). A justifiable reason that the rat dock lacks residue-ligand interactions at these energy levels, while human GPR35 does, is that rat sequence lacks the positively charged

Table 14. The listed comparisons of the EOIs of pamoic acid docked (optimized) in the human and rat GPR35 minimized R* bundles. The displayed residues are an inclusive list of residues that have a ≥ 2 kcal/mol residue-ligand EOI difference.

Residue (Human/Rat)	Human WT (kcal/mol)	Rat WT (kcal/mol)	Δ from WT (kcal/mol)
Y3.32 (96/93)	-1.16	-6.29	-5.13
R3.36 (100/97)	-13.70	-7.43	+6.27
R(164) / S(161)	-5.10	-0.19	+4.91
N/S 5.38 (171/169)	-0.72	-7.01	-6.29
R/Q 6.58 (240/238)	-12.97	-2.03	+10.94
Salt Bridge	**VdW**		
Total Int. Energy	**-78.16**	**-64.07**	Kcal/mol
ΔE from Human WT	**0**	**+14.09**	Kcal/mol

residues at loci 6.58 and 7.32 found in the human sequence. Rat GPR35 has a glutamine and serine at the respective positions of 6.58 and 7.32, and therefore cannot form a salt bridge with pamoic acid at these positions. Additionally, the EOI of R3.36 and pamoic acid in the rat GPR35 dock has a substantially lower magnitude as compared to the human docks, with values of -7.43 kcal/mol for rat and approximately -13.5 kcal/mol for the human docks. This difference is suggested to be caused by the shift of the final docked position of pamoic acid in rat GPR35 as compared to human GPR35 (Figure 38B). As described earlier, rat GPR35 lacks the non-conserved proline found in hGPR35 at position 5.43, which causes an approximate 4 Å relative shift in the extracellular position of TMH5 (Figure 24B). This change in the position of TMH5 is proposed to be the cause of the relative docking position shift of pamoic acid by approximately 3 Å. This shift increased the relative distance between R3.36 and pamoic acid in rat GPR35 as compared to hGPR35, which mitigated the residue-ligand EOI.

A comparison of the interactions of R3.36 and pamoic acid between the human wild-type and rat wild-type GPR35 docks (Figure 43) shows that the position of pamoic acid in rat does not

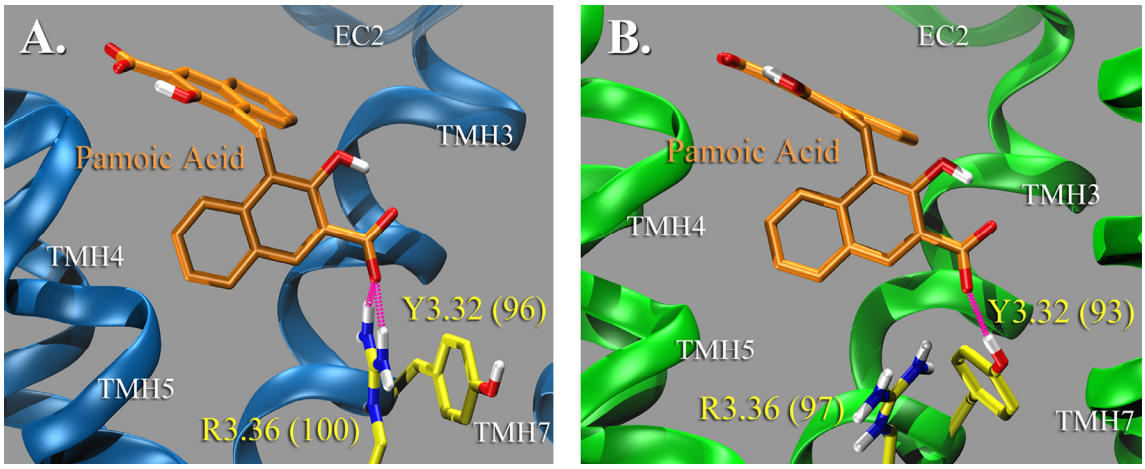


Figure 43. The comparison of the R3.36-pamoic acid residue-ligand interactions in both the minimized human (A) and rat (B) wild-type GPR35 R* bundles optimized with the ligand pamoic acid.

allow for as strong of an interaction with R3.36(97), but is instead replaced with an energetically weaker hydrogen bond with the juxtaposed residue Y3.32(93). The final substantial residue-ligand energy difference lies with the residue on the EC2 loop that is two positions past the conserved cysteine that forms a disulfide bridge with C3.25. The residue at this position, which is experimentally suggested to face into the TMH4-5-6 interface by SCAM studies in the dopamine D2 receptor [84] and by multiple crystal structures [9-11, 31], is an arginine in the human (R(164)) and a serine in rat GPR35 (S(161)). The residue-ligand interaction at this position is approximately 5 kcal/mol better in human GPR35 as compared to rat (Table 14).

In summary, the suggested reason for the difference in the EC_{50} of pamoic acid in rat vs. human GPR35 is due to three residue sequence variations. The non-conserved proline found in human GPR35, P5.43(176), is absent in the rat sequence. This difference is suggested to cause the relative shift of the extracellular end of TMH5 approximately 4Å, which in turn shifts the relative position of pamoic acid a similar distance. This shift increased the distance between R3.36 and pamoic acid in rat GPR35 as compared to human, which mitigates their relative EOI.

Additionally, one of the residue-ligand EOI with the largest magnitude in hGPR35 is a salt-bridge between the residue R6.58 and pamoic acid. This residue is a glutamine in rGPR35, and therefore is not able to create the same strong interaction. Finally, an arginine on the EC2 loop in human, R(164), contributes, through mostly Van der Waals interactions, an approximately 5 kcal/mol better interaction with pamoic acid than the analogous residue, S(161), found in the rat sequence.

There have been several experimental groups which have assessed the potency of the agonist zaprinast in both human and rat wild-type GPR35. These experimental studies (Table 13B) have yielded a consistent trend of zaprinast having a higher potency in rat vs. human GPR35, with a statistically significant difference found in 2010 by Jenkins *et al,* with an approximately 60 fold increase in potency, with respective EC_{50} values of 0.047μM and 2.95μM. These differences in potency correlate with the differences in the calculated protein-ligand EOI, with an improved total EOI in the rat dock of -7.34 kcal/mol (Table 15). The suggested reasons for the interspecies difference in total protein-ligand EOI of zaprinast, which lacks a formal charge in physiological conditions, are more subtle that those found for the di-anion pamoic acid. The comparison of all residue-ligand EOIs that exceed a difference of 0.75 kcal/mol (Table 15) suggests that the majority of the approximate 8 kcal/mol difference originate from the EOI differences of the residues at positions 7.35, 7.36, and 5.43.

Table 15. The listed comparisons of the EOIs of Zaprinast docked (optimized) in the human and rat GPR35 minimized R* bundles. The displayed residues are an inclusive list of residues that have a ≥ 0.75 kcal/mol residue-ligand EOI difference.

Residue (Human/Rat)	Human WT (kcal/mol)	Rat WT (kcal/mol)	Δ from Human WT (kcal/mol)
Y3.32 (96/93)	-3.92	-4.86	-0.93
C(162/159)	-1.81	-0.54	+1.27
F(163/160)	-0.99	-1.92	-0.93
Pro/Ser 5.43 (176/174)	-0.22	-2.04	-1.82
F5.47 (178/180)	-3.25	-0.65	+2.60
L6.51 (233/231)	-3.07	-0.91	+2.16
Arg/Gln 6.58 (240/238)	-2.06	-3.47	-1.41
L7.35 (258/256)	-0.09	-5.08	-4.99
Tyr/Thr 7.36 (259/257)	0	-2.04	-2.04
H-Bond	VdW		
Total Int. Energy	**-24.24**	**-31.58**	Kcal/mol
ΔE from Human WT	**0**	-7.34	Kcal/mol

The most substantial interspecies EOI difference of -4.99 kcal/mol occurs at the residue L7.35 (human/rat, 258/256) (Table 15). While the residue at this position is a leucine in both rat and human GPR35, the cause of this significant EOI difference is a sequence difference at a proximal residue position. The residue at position 6.54, which in juxtaposition to L7.35, is the beta-branched amino acid isoleucine in rGPR35 and a glycine in hPR35. In hGPR35 the lack of the bulk of G6.54(236) allows L7.35(258) to adopt the $\chi 1 = g^+$ conformation, which allows for the Van der Waals interactions with the nearby residues I(254), V6.57(239), and T7.38(261) (Figure 44A). In rGPR35, the bulk of I6.54(234) forces L7.35(256) into a $\chi 1$ = *trans* conformation (Figure 44B), which places its bulk into the binding pocket, allowing for a direct interaction with zaprinast.

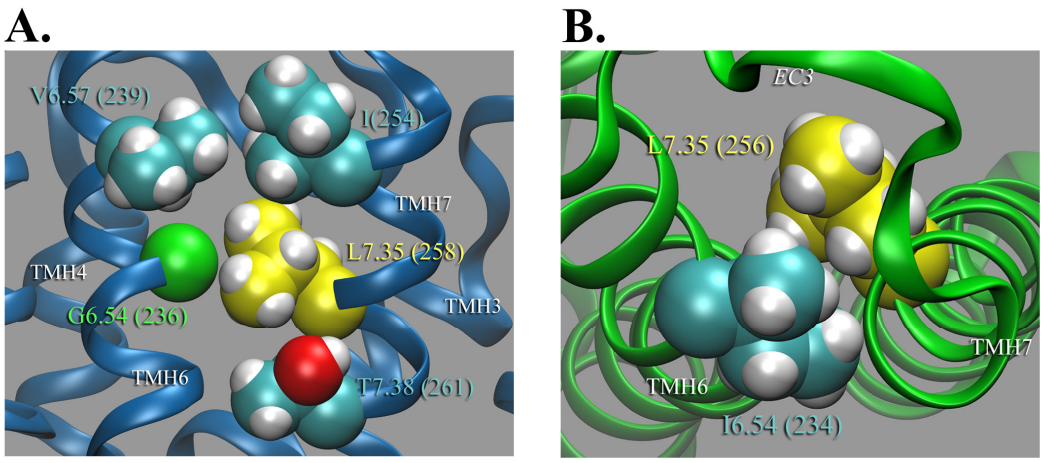


Figure 44. Comparison of the χ1 rotamer of L7.35 in hGPR35 vs. rGPR35. Residues are displayed in accordance to their Van der Waals radius. In hGPR35 (A) the lack of bulk of G6.54 (236) allows for L7.35 (258) to be in a χ1 = g+ conformation, allowing for favorable interactions with V6.57 (239), I(254), and T7.38 (261). This is opposed to rGPR35 (B), where the bulk of I6.54 (234) forces L7.35 (256) into a χ1 = *trans* conformation, putting its bulk into the binding pocket.

The basis for the interspecies EOI differences involving the residues at positions 7.36 and 5.43 is more unambiguous, as the residues at these positions vary between species. In hGPR35, the residues at 7.36 and 5.43 are a tyrosine and non-conserved proline, respectively. In rGPR35, which has the more favorable residue-ligand EOIs of -2.04 and -1.82 kcal/mol at the coordinating residues of 7.36 and 5.43, has a threonine and serine at these positions. In hGPR35 Y7.36(259) is only able to adopt a χ1 = *trans* conformation, due to the bulk of the surrounding residues, which negates the possibility of interacting directly with the ligand. The smaller T7.36(257) in rGPR35 is in a $\chi 1 = g^+$ conformation, which places the hydrophobic portion of the side-chain towards lipid and the hydrophilic segment towards zaprinast. In rGPR35, the serine at position 5.43 accepts a hydrogen bond from zaprinast. This hydrogen bond between zaprinast and S5.43(174) in rGPR35 partially directs the docked position of the ligand, which accounts for difference in the orientation of zaprinast in the human GPR35 dock as compared to the rat dock.

Additional, significantly different interspecies residue-zaprinast EOIs that favor rGPR35 occur at the positions 6.58, 3.32, and the position on the EC2 loop that immediately follows the cysteine involved in a disulfide bridge with C3.25. At position 6.58 in GPR35 the sequences are different between species, with this residue being an arginine in human and a glutamine in rat. In the hGPR35 dock, the guanidinium group of R6.58(240) is facing extracellularly, towards an area that in a complete cell would contain phospholipid head groups and be flooded with water. Though the charged portion of R6.58(240) is not interacting with zaprinast, the more hydrophobic region does contribute an EOI of -2.06 kcal/mol via Van der Waals interactions. The smaller, less polar Q6.58(238) is able to adopt a low energy conformation that is able to interact more substantially with zaprinast than the arginine found in hGPR35 at the analogous position, with an improved interaction of -1.41 kcal/mol. The elevated EOI in rat vs. human GPR35 of -0.93 kcal/mol occur at both the residues of Y3.32 (96/93, human/rat) and F (163/160, human/rat) on the EC2 loop. The altered position of zaprinast in the rGPR35 dock, due to the unique hydrogen bond between the ligand and S5.43, allows for the comparatively increased Van der Waals interaction at both Y3.32(93) and F(160).

While the overall change in protein-ligand EOI difference between human and rat GPR35 for zaprinast is approximately -8 kcal/mol, not all of the residue-ligand EOIs are better in rat. The residues as locants 5.47 and 6.51, which are a phenylalanine and a leucine, respectively, in both species, have EOI differences of -2.60 and -2.16 kcal/mol that favor the human dock. These differences are attributed to the relative change in position of zaprinast. In the human wild-type dock, zaprinast forms a tilted-T aromatic stack with F5.47(178), while in the rat dock this interaction is replaced by the hydrogen bond with S5.43(174). The change in the orientation of zaprinast to be able to retain the hydrogen bond with S5.43(174) also increases the relative distance of L6.51 in rGPR35 as compared to hGPR35, reducing the Van der Waals interaction.

An additional residue-ligand EOI that favors the human dock occurs at the cysteine on the EC2 loop that is assumed to have a highly-conserved disulfide bridge with C3.25. This relative EOI difference of 1.27 kcal/mol is similarly attributed to the relative change in position of zaprinast in rat, due to the hydrogen bond with S5.43 (174).

In summary, these docking studies suggest that zaprinast has an increased binding affinity in rGPR35 as compared to hGPR35, with a calculated protein-ligand EOI of -7.34 kcal/mol favoring rGPR35, which parallels the experimental data of other groups, which suggest that zaprinast has an increased potency of between 25 and 60 fold in rat as compared to human GPR35. The difference in the protein-ligand EOI is due to a multitude of differences in residue-ligand EOIs including: 1) the increased Van der Waals interactions in rGPR35 at residues Y3.32(93) and F(160) due to the relative change in position of zaprinast, 2) a hydrogen bond between zaprinast and S5.43(174), which is only possible in rGPR35 due to a sequence difference, 3) an increase in the Van der Waals interaction at between zaprinast and the residues at the positions 6.58 and 7.36, due to a direct sequence difference, 4) the increased Van der Waals interactions of zaprinast and L7.35 in rGPR35, due to L7.35(256) forced rotamer of $\chi 1$ = *trans* caused by the bulk of the juxtaposed I6.54(234), and 5) the comparatively mitigated EOIs in rGPR35 with the residues F5.47(178), L6.51(231), and the highly-conserved cysteine on the EC2 due to the relative position change of zaprinast.

The final series of docking studies to be compared are between the wild-type and the R6.58A(240) mutant of human GPR35. Pamoic acid has shown to have a high potency in the hGPR35 wild-type, with similar EC_{50}s from multiple groups and publications (Table 16A), with an average EC_{50} of approximately 60 nM. The EC_{50} of pamoic acid for the R6.58A(240) mutant was calculated as 62.3 nM (30.5 – 127) by collaborator Mary Abood (unpublished data), which showed a trend toward an increased potency from the wild-type, but was not statistically

significantly different. Zaprinast is moderately potent at the hGPR35 wild-type, with an approximate average EC_{50}, as calculated by multiple groups (Table 16), of 1.8 μM. Contrasting pamoic acid, zaprinast in the R6.58A(240) hGPR35 mutant showed a statistically significant increased potency ($p < 0.05$) as compared with its wild-type counterpart, with a calculated EC_{50} of 0.08 μM (Unpublished results, Mary Abood).

Table 16. The EC_{50}s identified for pamoic acid and Zaprinast in the human wild-type and the R6.58A(240) GPR35 mutant. A) The EC_{50}s identified for pamoic acid in the human wild-type and the R6.58A(240) GPR35 mutant by (1) Jenkins et al. (2010) Biochem J, (2) Abood, M. Temple University, Unpublished data, and (3) Zhao et al. (2010) Mol Pharmacol 78(4). B) The EC_{50}s identified for Zaprinast in the human wild-type and R6.58A (240) mutant by (1)Taniguchi et al. (2006) Febs Lett 580(21), (2) Zhao et al. (2010) Mol Pharmacol 78(4), (3) Abood, M. Temple University, Unpublished data, (4) Taniguchi et al (2008) Pharmacology 82(4), and (5)Jenkins et al. (2010) Biochem J)

*** Denotes statistically significant difference ($p < 0.05$)**

A.

	Human WT	Human R6.58A
EC_{50}	51.3[1],52.0[2], 62.3[2], 79[3] nM	16.0[2] nM

B.

	Human WT	Human R6.58A
EC_{50}	0.84[1], 1.0[2,3], 1.62[4], 2.95[5], 2.8[3]* μM	0.08[3]* μM

Table 17. The listed comparisons of the EOIs of pamoic acid docked (optimized) in the human wild-type and R6.58A (240) human GPR35 minimized R* bundles. The displayed residues are an inclusive list of residues that have a ≥ 2 kcal/mol residue-ligand EOI

Residue	Human WT (kcal/mol)	Human R6.58A (kcal/mol)	Δ from WT (kcal/mol)
F(163)	-5.77	-3.04	+2.73
S(165)	-0.67	-5.14	-4.47
R6.58(A) (240)	-12.97	-0.81	+12.15
Arg7.32 (255)	N/A	-12.28	-12.28
Salt Bridge			
Total Int. Energy	**-78.16**	**-80.55**	Kcal/mol
ΔE from Human WT	**0**	**-2.39**	Kcal/mol

The potency of the ligand pamoic acid is not statistically significantly different between the GPR35 wild-type and R6.58A(240) mutation, though a trend towards a mild increased potency in the mutant is suggested. The hGPR35 docks suggest a slight increase in the binding affinity for pamoic acid in the R6.58A(240) mutant, with a difference in the total protein-ligand EOI of -2.39 kcal/mol (Table 17). The GPR35 R6.58A(240) mutant and wild-type docks suggests these proteins have similar residue-protein EOIs for pamoic acid, with the major exception of the exchange of the primary interaction residue R6.58(240) for R7.32(255). In the wild-type dock, R6.58(240) is a primary interaction residue which forms a salt bridge with pamoic acid, with the resulting residue-ligand EOI of -12.97 kcal/mol. The mutation of arginine at position 6.58 to an alanine removes a primary interaction residue, though this removal allows for additional residue-ligand interactions. In the R6.58A(240) mutant, the removal of the bulk of R6.58(240) allows for the juxtaposed residue R7.32(255) to form a nearly equivalent residue-ligand EOI with pamoic acid of -12.28 kcal/mol. To achieve this near isoenergetic interchange of the residue R6.58(240)

with R7.32(255) in the mutant vs. wild-type dock, the position of pamoic acid shifted approximately 2Å (measured from carbon linking the napthalene rings).

The comparison of these docks suggests that the shift of pamoic acid changes multiple EOIs between the ligand and several residues on the EC2 loop. The most significant residue-ligand EOI differences between pamoic acid and residues found in the EC2 loop occur at the residues F(163) and S(165). The residue-ligand EOI differences (Table 17) of these residues in the R6.58A(240) mutant as compared to the wild-type shows a diminished interaction with F(163) and a improved EOI with S(165), with respective differences of +2.73 kcal/mol and -4.47 kcal/mol. The net difference of approximately -2 kcal/mol of these residue-ligand EOI account for the majority of the total protein-ligand difference of -2.39 kcal/mol favoring the R6.58A(240) mutant.

In summary, the potency of pamoic acid is not statistically significantly different between the hGPR35 wild-type and R6.58A(240) mutant, though there is a trend which suggest the mutant may have a slight increased potency to pamoic acid. These results parallel the differences in the protein-ligand EOI, which shows a -2.39 kcal/mol difference in the mutant as compared to the wild-type. These results suggest that the loss of the bulk of the primary wild-type interaction residue R6.58(240) in the R6.58A(240) mutant allows for the nearly isoenergetically exchange of R6.58(240) with the juxtaposed residue R7.32(255). This interchange of primary interaction residues causes a shift of pamoic acid of approximately 2 Å, affecting the residue-ligand EOIs between the ligand and several EC2 loop residues. The variance in the EOIs of the EC2 loop residues accounts for the majority of the difference of -2.39 kcal/mol in the mutant as compared to the wild-type.

The ligand zaprinast has shown to have a statistically significant difference in potency in hGPR35 wild-type vs. R6.58A mutant, with respective EC_{50} values of 2.8 μM (0.5 – 15) and 0.08

μM (0.02 – 0.3) (Mary Abood, Unpublished Data). These potency differences correlate with the differences in the measured total protein-ligand EOI, which has an improved, lower EOI of -4.41 kcal/mol in the R6.58A mutant dock. In the GPR35 wild-type dock the residue R6.58(240) interacts positively via Van der Waal forces with zaprinast, with a -2.06 kcal/mol residue-ligand EOI. Though this is a positive interaction, the position of R6.58(240) creates a barrier for zaprinast, therefore dictating its location and consequently limiting its possible interactions. In the R6.58A(240) mutant the positive R6.58(240) EOI is lost, but the removal of the bulk of this residue allows for a position change of zaprinast, leading to additional/improved residue-ligand interactions.

The total EOI divergence can be mostly attributed to a small number of residue-ligand EOI differences, involving the residues Q3.29(93), Y3.32(96), F5.47(180), R7.32(255), and S7.39(262). The position change of zaprinast in the R6.58A(240) mutation improves the Van der Waal interaction of Q3.29(93) and Y3.32(96) in the R6.58A(240) mutant as compared to wild-type, with respective EOI value differences of -1.25 and -1.55 kcal/mol. The shift in the ligand position also mitigates the residue-ligand EOI in the mutant between zaprinast and F5.47(180), with an EOI difference of 0.88 kcal/mol. Additionally, a hydrogen bond and a cation-π interaction are gained in the mutant at the respective residues S7.39(262) and R7.32(255), with the coordinated EOI value differences of -1.84 and -1.24 kcal/mol. Noted is that the distance and angle of the cation-π interaction between R7.32(262) and the propoxy-benzyl functional group of Zaprinast are similar to the suggested ideal values. The EOI at the ideal distance/angle, as calculated at the MP2/6-31+G* level of theory, for the analogous molecules of benzene and N-methyl guanidine is -13.96 kcal/mol [97]. The EOI between the residue R7.32(255) and zaprinast, as calculated by the OPLS2005 force field, was -1.24 kcal/mol, which diverges substantially from

the -13.96 kcal/mol calculated for the abovementioned analogous molecules, suggesting that this interaction may be more significant than the current parameters indicate.

In summary, the statistically significant, approximately 35 fold increase in ligand potency of zaprinast in the hGPR35 R6.58A(240) mutant as compared to wild-type is paralleled by the difference in the measured protein-ligand EOI, which favors the mutant by -4.41 kcal/mol (Table 18). The removal of the bulk of R6.58(240) in the R6.58A(240) mutation allows for improved and additional residue interactions with zaprinast. There are numerous differences in the residue-ligand interactions that caused the improved protein-ligand EOI in the mutant dock, with the most notable differences being: the hydrogen bond with zaprinast and S7.39(262), the cation-π interaction with the ligand and R7.32(255), and the improved Van der Waals interactions with zaprinast and the residues Q3.29(93) and Y3.32(96).

Table 18. The comparison of the EOIs of Zaprinast docked in the human wild-type and the R6.58A (240) mutant R* GPR35 bundles. The displayed residues are an inclusive list of residues that have a ≥ 0.75 kcal/mol residue-ligand EOI difference.

Residue	Human WT (kcal/mol)	Human R6.58A (kcal/mol)	Δ from WT (kcal/mol)
Q3.29 (93)	-2.76	-4.01	-1.25
Y3.32 (96)	-3.92	-5.47	-1.55
F5.47 (180)	-3.25	-2.37	+0.88
R6.58(A) (240)	-2.06	-0.1	+1.96
R7.32 (255)	N/A	-1.24	-1.24
Ser7.39 (262)	-0.85	-2.69	-1.84
H-Bond	**Cation-π**	**VdW**	
Total Int. Energy	**-24.11**	**-28.65**	Kcal/mol
ΔE from Human WT	**0**	**-4.41**	Kcal/mol

As discussed earlier, the mutation of R4.60(151) to alanine yielded ambiguous results. In this mutation, though it showed surface expression, both the agonists zaprinast and pamoic acid, were unable to elicit a response via a β2-arrestin assay or an ERK activity assay, (Abood, Mary.

Unpublished data). These data suggest that R4.60(151) is involved with protein folding, ligand binding, and or receptor activation. In the hGPR35 R*-pamoic acid docks, R4.60(151) forms a cation-π interaction with the ligand, which suggests that it is a primary interaction residue for this ligand. While this supports the inactivity of pamoic acid in the R4.60A(151) mutant, this does not offer a explanation for the effects on zaprinast. In the hGPR35-zaprinast docks R4.60A(151) is not suggested to have any interaction with the ligand, yet this mutation eliminates all zaprinast-induced signaling. The GPR35 models offer two practical explanations for this possible inconsistency. In the GPR35 models the residue R4.60(151) sits directly in the TMH3-4-5 interface, and is surrounded by multiple hydrophobic residues, including F5.42(175), L3.33(97), L4.57(148), and P5.43(176). The mutation of this large, charged residue to an alanine would create an energetically-unfavorable, large hole which would be filled by the surrounding residues or by water. As there are no adjacent residues in which a rotamer change would fill this gap, it is possible that the intracellular tops of TMH3, 4, and or 5 may adjust to maximize the hydrophobic interactions of the aforementioned residues- eliminating the hole, but collapsing the intracellular portion of the binding pocket. Additional support for the collapse of the binding pocket lies in the recruitment of water. With the charged R4.60(151) supplanted with the neutral alanine, the binding pocket would be become less hydrophilic, which would likely reduce the average number of waters in this channel. The lack of water in this area will lead to an even larger hole that is suggested by just the removal of the arginine, making the collapsing of the binding pocket more energetically favorable.

The fundamental goal of this project was to be able to ascertain viable reasons for the differences in the pharmacology of the ligands pamoic acid and zaprinast in rat wild-type, human wild-type, and in engineered human GPR35 mutants. The current models offer feasible explanations for the current known pharmacology of these ligands and additional mutational

studies are forthcoming to test the validity of these predictions. Currently, the suggested direct interaction residues with the ligand pamoic acid are being used to steer the development of novel molecules. There are multiple goals actuating the engineering these novel molecules. The development of molecule with a high enough affinity for GPR35 can act as a radiolabel which would be instrumental in the understanding of this misunderstood GPCR. There are a multitude of ligands that are thought to act at GPR35, but the development of a radioligand would allow for the reassessment of many of these ligands, potentially leading to a better understanding of the currently convoluted purpose of GPR35 *in vivo*. The other major goal would be develop patentable pharmaceutics, both agonists and antagonists that are specific to GPR35. The extensive range of GPR35's proposed involvement *in vivo* make this a potentially invaluable receptor to target, as it has associations with a nearly innumerable number of diseases and disorders. Additionally, the knowledge gained during the development of this novel receptor, which has the unusual quality of having multiple charged residues within the binding pocket, will hopefully translate to other similar GPCRs.

REFERENCES

1. Musnier, A., et al., *GPCR signalling to the translation machinery.* Cell Signal, 2010. **22**(5): p. 707-16.

2. Li, J., et al., *Structure of bovine rhodopsin in a trigonal crystal form.* J Mol Biol, 2004. **343**(5): p. 1409-38.

3. Okada, T., et al., *Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5982-7.

4. Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. **289**(5480): p. 739-45.

5. Choe, H.W., et al., *Crystal structure of metarhodopsin II.* Nature, 2011. **471**(7340): p. 651-655.

6. Park, J.H., et al., *Crystal structure of the ligand-free G-protein-coupled receptor opsin.* Nature, 2008. **454**(7201): p. 183-7.

7. Cherezov, V., et al., *High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor.* Science, 2007. **318**(5854): p. 1258-65.

8. Rasmussen, S.G., et al., *Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor.* Nature, 2011. **469**(7329): p. 175-80.

9. Jaakola, V.P., et al., *The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist.* Science, 2008. **322**(5905): p. 1211-7.

10. Chien, E.Y., et al., *Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist.* Science, 2010. **330**(6007): p. 1091-5.

11. Wu, B., et al., *Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists.* Science, 2010. **330**(6007): p. 1066-71.

12. Weinstein, J.A.B.a.H., *Integrated methods for the construction of three dimensional models and computational probing of structure function relations in G protein coupled receptors* Methods in Neurosciences, 1995. **25**: p. 366-428.

13. Nakanishi, J., et al., *FRET-based monitoring of conformational change of the beta2 adrenergic receptor in living cells.* Biochem Biophys Res Commun, 2006. **343**(4): p. 1191-6.

14. Lin, S.W. and T.P. Sakmar, *Specific tryptophan UV-absorbance changes are probes of the transition of rhodopsin to its active state.* Biochemistry, 1996. **35**(34): p. 11149-59.

15. Javitch, J.A., et al., *Constitutive activation of the beta2 adrenergic receptor alters the orientation of its sixth membrane-spanning segment.* J Biol Chem, 1997. **272**(30): p. 18546-9.

16. Jensen, A.D., et al., *Agonist-induced conformational changes at the cytoplasmic side of transmembrane segment 6 in the beta 2 adrenergic receptor mapped by site-selective fluorescent labeling.* J Biol Chem, 2001. **276**(12): p. 9279-90.

17. Scheerer, P., et al., *Crystal structure of opsin in its G-protein-interacting conformation.* Nature, 2008. **455**(7212): p. 497-502.

18. Altenbach, C., et al., *High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation.* Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7439-44.

19. Ghanouni, P., et al., *Agonist-induced conformational changes in the G-protein-coupling domain of the beta 2 adrenergic receptor.* Proc Natl Acad Sci U S A, 2001. **98**(11): p. 5997-6002.

20. Farrens, D.L., et al., *Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin.* Science, 1996. **274**(5288): p. 768-70.

21. Ballesteros, J.A., et al., *Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6.* J Biol Chem, 2001. **276**(31): p. 29171-7.

22. Li, J.H., et al., *Ligand-specific changes in M3 muscarinic acetylcholine receptor structure detected by a disulfide scanning strategy.* Biochemistry, 2008. **47**(9): p. 2776-88.

23. Ward, S.D., et al., *Use of an in situ disulfide cross-linking strategy to study the dynamic properties of the cytoplasmic end of transmembrane domain VI of the M3 muscarinic acetylcholine receptor.* Biochemistry, 2006. **45**(3): p. 676-85.

24. Klein-Seetharaman, J., et al., *Differential dynamics in the G protein-coupled receptor rhodopsin revealed by solution NMR.* Proc Natl Acad Sci U S A, 2004. **101**(10): p. 3409-13.

25. Ruprecht, J.J., et al., *Electron crystallography reveals the structure of metarhodopsin I.* EMBO J, 2004. **23**(18): p. 3609-20.

26. Shi, L., et al., *Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch.* J Biol Chem, 2002. **277**(43): p. 40989-96.

27. Borhan, B., et al., *Movement of retinal along the visual transduction path.* Science, 2000. **288**(5474): p. 2209-12.

28. Holst, B., et al., *A conserved aromatic lock for the tryptophan rotameric switch in TM-VI of seven-transmembrane receptors.* J Biol Chem, 2010. **285**(6): p. 3973-85.

29. Singh, R., et al., *Activation of the cannabinoid CB1 receptor may involve a W6 48/F3 36 rotamer toggle switch.* J Pept Res, 2002. **60**(6): p. 357-70.

30. McAllister, S.D., et al., *Structural mimicry in class A G protein-coupled receptor rotamer toggle switches: the importance of the F3.36(201)/W6.48(357) interaction in cannabinoid CB1 receptor activation.* J Biol Chem, 2004. **279**(46): p. 48024-37.

31. Rasmussen, S.G., et al., *Crystal structure of the human beta2 adrenergic G-protein-coupled receptor.* Nature, 2007. **450**(7168): p. 383-7.

32. Javitch, J.A., et al., *A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice.* Biochemistry, 1998. **37**(4): p. 998-1006.

33. O'Dowd, B.F., et al., *Discovery of three novel G-protein-coupled receptor genes.* Genomics, 1998. **47**(2): p. 310-3.

34. Wang, J., et al., *Kynurenic acid as a ligand for orphan G protein-coupled receptor GPR35.* J Biol Chem, 2006. **281**(31): p. 22021-8.

35. Leonard, J., Chu, Z.L., Unett, D.J., Gatlin, J.E., Gaidarov, I., Qui J., Skinner P.J., and Boatman, P.D. , *GPR35 and modulators thereof for the treatment of metabolic-related disorders*, in *World Intellectual Property Organization*. 2005. p. .

36. Min, K.D., et al., *Identification of genes related to heart failure using global gene expression profiling of human failing myocardium.* Biochem Biophys Res Commun, 2010. **393**(1): p. 55-60.

37. Barth, M.C., et al., *Kynurenic acid triggers firm arrest of leukocytes to vascular endothelium under flow conditions.* J Biol Chem, 2009. **284**(29): p. 19189-95.

38. Yang, Y., et al., *G-protein-coupled receptor 35 is a target of the asthma drugs cromolyn disodium and nedocromil sodium.* Pharmacology, 2010. **86**(1): p. 1-5.

39. Shrimpton, A.E., et al., *Molecular delineation of deletions on 2q37.3 in three cases with an Albright hereditary osteodystrophy-like phenotype.* Clin Genet, 2004. **66**(6): p. 537-44.

40. Vander Molen, J., et al., *Population genetics of CAPN10 and GPR35: implications for the evolution of type 2 diabetes variants.* Am J Hum Genet, 2005. **76**(4): p. 548-60.

41. Okumura, S., et al., *Cloning of a G-protein-coupled receptor that shows an activity to transform NIH3T3 cells and is expressed in gastric cancer cells.* Cancer Sci, 2004. **95**(2): p. 131-5.

42. Horikawa, Y., et al., *Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus.* Nat Genet, 2000. **26**(2): p. 163-75.

43. Taniguchi, Y., H. Tonai-Kachi, and K. Shinjo, *Zaprinast, a well-known cyclic guanosine monophosphate-specific phosphodiesterase inhibitor, is an agonist for GPR35.* FEBS Lett, 2006. **580**(21): p. 5003-8.

44. Guo, J., et al., *Inhibition of N-type calcium channels by activation of GPR35, an orphan receptor, heterologously expressed in rat sympathetic neurons.* J Pharmacol Exp Ther, 2008. **324**(1): p. 342-51.

45. Jenkins, L., et al., *Agonist activation of the G protein-coupled receptor GPR35 involves transmembrane domain III and is transduced via Galpha and beta-arrestin-2.* Br J Pharmacol, 2011. **162**(3): p. 733-48.

46. Sonoda, H., et al., *A novel phosphatidic acid-selective phospholipase A1 that produces lysophosphatidic acid.* J Biol Chem, 2002. **277**(37): p. 34254-63.

47. Forrest, C.M., et al., *Kynurenine pathway metabolism in patients with osteoporosis after 2 years of drug treatment.* Clin Exp Pharmacol Physiol, 2006. **33**(11): p. 1078-87.

48. Hartai, Z., et al., *Decreased serum and red blood cell kynurenic acid levels in Alzheimer's disease.* Neurochem Int, 2007. **50**(2): p. 308-13.

49. Forrest, C.M., et al., *Levels of purine, kynurenine and lipid peroxidation products in patients with inflammatory bowel disease.* Adv Exp Med Biol, 2003. **527**: p. 395-400.

50. Forrest, C.M., et al., *Purine, kynurenine, neopterin and lipid peroxidation levels in inflammatory bowel disease.* J Biomed Sci, 2002. **9**(5): p. 436-42.

51. Moolenaar, W.H., *Lysophosphatidic acid, a multifunctional phospholipid messenger.* J Biol Chem, 1995. **270**(22): p. 12949-52.

52. Moolenaar, W.H., et al., *Lysophosphatidic acid: G-protein signalling and cellular responses.* Curr Opin Cell Biol, 1997. **9**(2): p. 168-73.

53. Oka, S., et al., *GPR35 is a novel lysophosphatidic acid receptor.* Biochem Biophys Res Commun, 2010. **395**(2): p. 232-7.

54. Yanagida, K., et al., *Identification and characterization of a novel lysophosphatidic acid receptor, p2y5/LPA6.* J Biol Chem, 2009. **284**(26): p. 17731-41.

55. Bandoh, K., et al., *Lysophosphatidic acid (LPA) receptors of the EDG family are differentially activated by LPA species. Structure-activity relationship of cloned LPA receptors.* FEBS Lett, 2000. **478**(1-2): p. 159-65.

56. Higgs, H.N., et al., *Cloning of a phosphatidic acid-preferring phospholipase A1 from bovine testis.* J Biol Chem, 1998. **273**(10): p. 5468-77.

57. Jenkins, L., et al., *Identification of novel, species selective agonists of the G protein-coupled receptor GPR35 that promote recruitment of beta-arrestin-2 and activate Galpha13.* Biochem J, 2010.

58. Taniguchi, Y., H. Tonai-Kachi, and K. Shinjo, *5-Nitro-2-(3-phenylpropylamino)benzoic acid is a GPR35 agonist.* Pharmacology, 2008. **82**(4): p. 245-9.

59. Zhao, P., et al., *Targeting of the orphan receptor GPR35 by pamoic acid: a potent activator of extracellular signal-regulated kinase and beta-arrestin2 with antinociceptive activity.* Mol Pharmacol, 2010. **78**(4): p. 560-8.

60. Francis, S.H., I.V. Turko, and J.D. Corbin, *Cyclic nucleotide phosphodiesterases: relating structure and function.* Prog Nucleic Acid Res Mol Biol, 2001. **65**: p. 1-52.

61. Kato, K., A.M. Evans, and R.Z. Kozlowski, *Relaxation of endothelin-1-induced pulmonary arterial constriction by niflumic acid and NPPB: mechanism(s) independent of chloride channel block.* J Pharmacol Exp Ther, 1999. **288**(3): p. 1242-50.

62. Neubig, R.R., *Mind your salts: when the inactive constituent isn't.* Mol Pharmacol, 2010. **78**(4): p. 558-9.

63. Zhao, P., et al., *Crucial positively charged residues for ligand activation of the GPR35 receptor*, in *Poster session presented at the annual Society for Neuroscience meeting,* . 2010: San Francisco, CA.

64. Ballesteros, J.A., et al., *Serine and threonine residues bend alpha-helices in the chi(1) = g(-) conformation.* Biophys J, 2000. **79**(5): p. 2754-60.

65. Rosenbaum, D.M., et al., *GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function.* Science, 2007. **318**(5854): p. 1266-73.

66. Warne, T., et al., *Structure of a beta1-adrenergic G-protein-coupled receptor.* Nature, 2008. **454**(7203): p. 486-91.

67. McAllister, S.D., et al., *An aromatic microdomain at the cannabinoid CB(1) receptor constitutes an agonist/inverse agonist binding region.* J Med Chem, 2003. **46**(24): p. 5139-52.

68. Zhang, R., et al., *Cysteine 2.59(89) in the second transmembrane domain of human CB2 receptor is accessible within the ligand binding crevice: evidence for possible CB2 deviation from a rhodopsin template.* Mol Pharmacol, 2005. **68**(1): p. 69-83.

69. MacKerell, A.D., Jr., N. Banavali, and N. Foloppe, *Development and current status of the CHARMM force field for nucleic acids.* Biopolymers, 2000. **56**(4): p. 257-65.

70. Metropolis, N., et al., *Equation of State Calculations by Fast Computing Machines.* Journal of Chemical Physics, 1953. **21**(1087).

71. Whitnell, R.M., et al., *Conformational memories with variable bond angles.* J Comput Chem, 2008. **29**(5): p. 741-52.

72. Visiers, I., B.B. Braunheim, and H. Weinstein, *Prokink: a protocol for numerical evaluation of helix distortions by proline.* Protein Eng, 2000. **13**(9): p. 603-6.

73. Mezei, M., *Simulaid: a simulation facilitator and analysis program.* J Comput Chem, 2010. **31**(14): p. 2658-68.

74. Christopher, J.A., R. Swanson, and T.O. Baldwin, *Algorithms for finding the axis of a helix: fast rotational and parametric least-squares methods.* Comput Chem, 1996. **20**(3): p. 339-45.

75. Guarnieri, F. and H. Weinstein, *Conformational Memories and the Exploration of Biologically Relevant Peptide Conformations: An Illustration for the Gonadotropin-Releasing Hormone.* J. Am. Chem. Soc., 1996. **118**(24): p. 5580-5589.

76. Struthers, M., et al., *Tertiary interactions between the fifth and sixth transmembrane segments of rhodopsin.* Biochemistry, 1999. **38**(20): p. 6597-603.

77. Prioleau, C., et al., *Conserved helix 7 tyrosine acts as a multistate conformational switch in the 5HT2C receptor. Identification of a novel "locked-on" phenotype and double revertant mutations.* J Biol Chem, 2002. **277**(39): p. 36577-84.

78. Marti-Renom, M.A., et al., *Comparative protein structure modeling of genes and genomes.* Annu Rev Biophys Biomol Struct, 2000. **29**: p. 291-325.

79. Fiser, A., R.K. Do, and A. Sali, *Modeling of loops in protein structures.* Protein Sci, 2000. **9**(9): p. 1753-73.

80. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.

81. Ghanouni, P., et al., *The effect of pH on beta(2) adrenoceptor function. Evidence for protonation-dependent activation.* J Biol Chem, 2000. **275**(5): p. 3121-7.

82. Arnis, S. and K.P. Hofmann, *Two different forms of metarhodopsin II: Schiff base deprotonation precedes proton uptake and signaling state.* Proc Natl Acad Sci U S A, 1993. **90**(16): p. 7849-53.

83. Knierim, B., et al., *Sequence of late molecular events in the activation of rhodopsin.* Proc Natl Acad Sci U S A, 2007. **104**(51): p. 20290-5.

84. Shi, L. and J.A. Javitch, *The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice.* Proc Natl Acad Sci U S A, 2004. **101**(2): p. 440-5.

85. Baneres, J.L., et al., *Molecular characterization of a purified 5-HT4 receptor: a structural basis for drug efficacy.* J Biol Chem, 2005. **280**(21): p. 20253-60.

86. Bokoch, M.P., et al., *Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor.* Nature, 2010. **463**(7277): p. 108-12.

87. Nebane, N.M., B. Kellie, and Z.H. Song, *The effects of charge-neutralizing mutation D6.30N on the functions of CB1 and CB2 cannabinoid receptors.* FEBS Lett, 2006. **580**(22): p. 5392-8.

88. Huang, P., et al., *The local environment at the cytoplasmic end of TM6 of the mu opioid receptor differs from those of rhodopsin and monoamine receptors: introduction of an ionic lock between the cytoplasmic ends of helices 3 and 6 by a L6.30(275)E mutation inactivates the mu opioid receptor and reduces the constitutive activity of its T6.34(279)K mutant.* Biochemistry, 2002. **41**(40): p. 11972-80.

89. Ballesteros, J., et al., *Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor.* J Biol Chem, 1998. **273**(17): p. 10445-53.

90. Ovchinnikov Yu, A., N.G. Abdulaev, and A.S. Bogachuk, *Two adjacent cysteine residues in the C-terminal cytoplasmic fragment of bovine rhodopsin are palmitylated.* FEBS Lett, 1988. **230**(1-2): p. 1-5.

91. O'Dowd, B.F., et al., *Palmitoylation of the human beta 2-adrenergic receptor. Mutation of Cys341 in the carboxyl tail leads to an uncoupled nonpalmitoylated form of the receptor.* J Biol Chem, 1989. **264**(13): p. 7564-9.

92. Gao, Z., et al., *Palmitoylation of the recombinant human A1 adenosine receptor: enhanced proteolysis of palmitoylation-deficient mutant receptors.* Biochem J, 1999. **342 (Pt 2)**: p. 387-95.

93. Qanbar, R. and M. Bouvier, *Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function.* Pharmacol Ther, 2003. **97**(1): p. 1-33.

94. Standfuss, J., et al., *The structural basis of agonist-induced activation in constitutively active rhodopsin.* Nature, 2011. **471**(7340): p. 656-60.

95. Perola, E. and P.S. Charifson, *Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding.* J Med Chem, 2004. **47**(10): p. 2499-510.

96. Chambers, C.C., et al., *Model for Aqueous Solvation Based on Class IV Atomic Charges and First Solvation Shell Effects.* J. Phys. Chem, 1996. **100** (40): p. 16041-16446.

97. Singh, N.J., et al., *Comprehensive Energy Analysis for Various Types of π-Interaction.* J. Chem. Theory Comput, 2009. **5**(3): p. 515-529.

98. Gallivan, J.P. and D.A. Dougherty, *A Computational Study of Cation−π Interactions vs Salt Bridges in Aqueous Media: Implications for Protein Engineering.* J. Am. Chem. Soc., 2000. **122**(5): p. 870-874.

99. Qiu, D., et al., *The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii.* J. Phys. Chem. A, 1997. **101**(16): p. 3005–3014.

100. Roth, B.L., et al., *Salvinorin A: a potent naturally occurring nonnitrogenous kappa opioid selective agonist.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11934-9.

101. Janovick, J.A., et al., *Molecular mechanism of action of pharmacoperone rescue of misrouted GPCR mutants: the GnRH receptor.* Mol Endocrinol, 2009. **23**(2): p. 157-68.

102. Tuccinardi, T., et al., *Cannabinoid CB2/CB1 selectivity. Receptor modeling and automated docking analysis.* J Med Chem, 2006. **49**(3): p. 984-94.

103. Ravna, A.W., I. Sylte, and S.G. Dahl, *Molecular mechanism of citalopram and cocaine interactions with neurotransmitter transporters.* J Pharmacol Exp Ther, 2003. **307**(1): p. 34-41.

104. Chen, J.Z., J. Wang, and X.Q. Xie, *GPCR structure-based virtual screening approach for CB2 antagonist search.* J Chem Inf Model, 2007. **47**(4): p. 1626-37.