PYLE, JENNIFER L., M.S. Identification of Agonist and Antagonist Binding Sites at the G-Protein Coupled Receptor, GPR18. (2015)
Directed by Dr. Patricia H. Reggio. 41 pp.

GPR18, a member of the Class A G-Protein Coupled Receptors (GPCRs), is recently a de-orphanized receptor, that upon activation has been found to boost the immune system. Although a GPR18 crystal structure has not been solved, a homology model of the inactive state (R) of GPR18 was built in the Reggio lab based upon the mu-opioid x-ray crystal structure. The present study continues this work by adding loop regions and N- and C-termini to the R state model and by developing a model of GPR18 R* (Active) state, complete with loop regions and N- and C-termini. The complete inactive and active state models were used for docking studies of five known GPR18 antagonists, 1,3-dimethoxy-5-methyl-2-[(1*R*,6*R*)-3-methyl-6-prop-1-en-2-ylcyclohex-2-en-1-yl]benzene; **1**, (2R,3aS,3bR,5aS,11bS,11cS,13aS)-3a,11b,11c-trimethyl-2-(2-methylprop-1-en-1-yl)-2,3,3a,3b,4,5,5a,6,11,11b,11c,12,13,13a-tetradecahydrobenzofuro[4',5':6,7]indeno[1,2-b]indole; **2**, (3S,4S,4aR,6aS,12bS,12cS)-4-((E)-4-methoxy-4-methylpent-2-en-1-yl)-4,12b,12c-trimethyl-1,2,3,4,4a,5,6,6a,7,12,12b,12c-dodecahydrobenzo[6,7]indeno[1,2-b]indol-3-ol; **3**, (Z)-2-(3-((4-chlorobenzyl)oxy)benzylidene)-6,7-dihydro-5H-imidazo[2,1-b][1,3]thiazin-3(2H)-one; **4**, and 2-[(1*R*,6*R*)-6-isopropenyl-3-methylcyclohex-2-en-1-yl]-5-pentylbenzene-1,3-diol; **5** and the GPR18 endogenous ligand, N-arachidonylglycine (NAGly). Conformational searches were performed on all antagonists using a systematic search using the Hartree-Fock method at the 6-31G* level of theory. Because NAGly is highly flexible, its low energy conformations were determined using the Conformational

Memories method. All low energy conformations, less than 3.0 kcal/mol, of each ligand were docked using Glide into their respective bundles based on the state of the receptor each ligand stabilizes. The key interaction site for all antagonists, an R in TMH5, R5.42, anchors each antagonist in the TMH bundle such that rotameric changes in key toggle switch residues, F6.48/H6.52, are prevented, thus preventing the activation of the receptor. The extracellular loop 2 (EC2 loop) residue Y160 further stabilizes each antagonist in the binding site. Identified interactions result in Glide scores consistent with experimental $EC_{50}$ data. The primary interaction site for the agonist, NAGly, was TMH2 residue R2.60. An additional NAGly interaction was identified with the EC2 loop residue K174. Mutation experiments of key residues identified here are underway in a collaborator's lab. These studies will help confirm the importance of these key residues and further our understanding of the receptor.

# IDENTIFICATION OF AGONIST AND ANTAGONIST BINDING SITES AT THE G-PROTEIN COUPLED RECEPTOR, GPR18

by

Jennifer L. Pyle

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
2015

Approved by

Patricia H. Reggio
Committee Chair

APPROVAL PAGE

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

Committee Chair Patricia H. Reggio

Committee Members Mitchell Croatt

Ethan Taylor

November 19, 2015
Date of Acceptance by Committee

November 19, 2015
Date of Final Oral Examination

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER I

# INTRODUCTION

## Background

GPR18, a member of the Class A G-Protein Coupled Receptors (GPCRs), was recently de-orphanized when icosa-5,8,11,14-tetraenoylamino-acetic acid (N-arachidonoylglycine; NAGly) was found to be its endogenous ligand.[1] First isolated from mouse in 1996,[2] GPR18 is thought to be the elusive abnormal cannabinoid receptor (Abn-CBD) identified by Wagner and co-workers.[3] In humans, GPR18 is localized in Chromosome 13q32, in the same region as the Epstein-Barr Virus induced receptor 2 (EBI2)[4,5] and is structurally homologous to it. Both receptors also have almost identical expression rates which have led to the presumption that they might have similar functions as immune response initiators.[6] GPR18 is expressed in several locations in the body including peripheral blood mononuclear cells (PBMC), lymphoid tissues such as spleen and thymus, and moderate expression in brain, testis, ovary, and lung, all of which are associated with the immune system.[7] Recent studies have shown that when channel catfish were infected with viral strains, GPR18 expression was induced in the kidneys.[7] Further studies have shown that GPR18 is significantly expressed in HTLV-1-transformed cell lines (HUT 102, MT-2, and MT-4) as well as lymphoid cell lines (Jurkat, MOLT-4, HUT 78 and Raji) while the non-lymphoid hematopoietic cell lines (U937, HL60, and K562) showed no detectable expression.[1] Due to rapidly evolving

bacteria, the use of antibiotics as the traditional medical first line defense is being called into question. With immune-compromising diseases such as diabetes on the rise, antibiotic resistant strains are a threat to the public health and safety. This pattern of expression suggests that GPR18 may play a role in the immune system and could be a valuable drug target for immune-compromising diseases and antibiotic resistant bacterial strains.

Other serious concerns for public health are neurodegenerative diseases such as multiple sclerosis and Alzheimer's, which are caused by deregulation of receptor cell migration to affected areas in the central nervous system (CNS).[8] The migration is regulated by microglia, which play a huge role in active immune response in the brain against pro-inflammatory and cytotoxic responses. In quantitative PCR studies, GPR18 has been shown to be expressed in the mouse microglia cell line BV-2, [9] as well as, human endometrial cell line HEC-1B.[10] This suggests that the endocannabinoid system has a part in regulating microglial migration.[8] This pattern of expression, in turn, suggests that GPR18 could also be a potential pharmaceutical target for various inflammatory diseases. Endometriosis is another condition caused by abnormal cell migration in the wall of the uterus causing intense pain and in some cases infertility.[11] Recent studies have shown that GPR18 plays some part in the migration of endometrial cells. Through in silico analysis, GPR18 was found to be expressed in primary melanoma cells.[12] Further, quantitative PCR studies by Qin and co-workers have shown that GPR18 is also expressed in melanoma metastases.[13] Both of these studies show that GPR18 is a potential therapeutic target for aggressive skin cancer melanoma. Therefore,

while the exact role of GPR18 in the human body an area of research, its involvement and mere presence in so many major areas of the body make it undeniably important for further research.

GPCRs make up the largest family of transmembrane eukaryotic receptors, making them a focus group among receptor studies. GPCRs are responsible for regulating a large portion of cell signaling through signal transduction over cell membranes and are included in multiple systems within an organism. GPCRs are important pharmacological targets because they respond to a vast number of factors such as light and hormones.[3]

The structures of Class A GPCRs include seven transmembrane α-helices (TMHs), an extracellular N-terminus, intracellular C-terminus, extracellular (EC) and intracellular (IC) loops that connect the TMHs, and a C-terminal helix helical section, parallel to membrane, Helix 8 (H8) (see Figure 1). With GPCRs varying in the number of amino acids in their sequences, the Ballesteros-Weinstein numbering and was created in 1995 to allow direct comparisons between receptor sequences.[14] In this system, the most highly conserved residue in each TMH is denoted as x.50 where x is the number of the helix. Sequentially left or right of that residue, the number after x will represent the next or previous amino acid in the sequence. As an example, the most conserved residue in TMH2 is aspartic acid D2.50. The residue immediately before this is V2.49 and immediately after is L2.51 in the GPR18 sequence. GPCRs sequences are characterized by conserved motifs within each TMH. These motifs seem to either play a role in the function of the receptor or can be used in studying the homology between receptors.

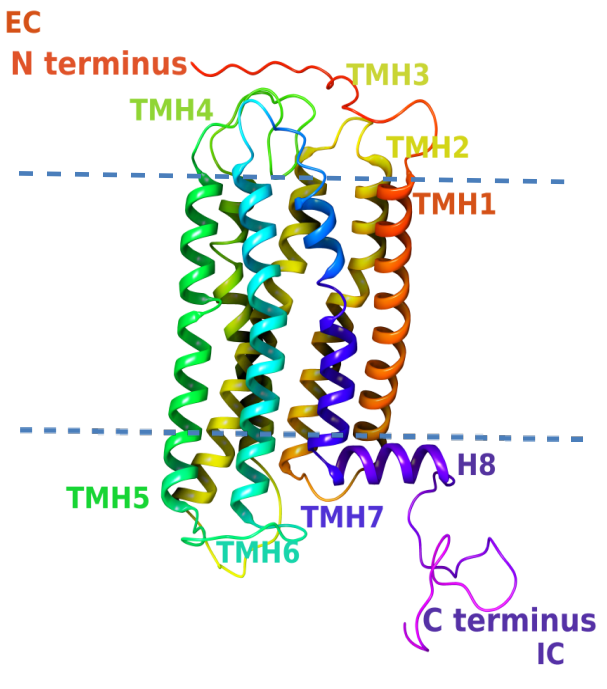



Figure 1. Example of Typical GPCR.

Rhodopsin was the first GPCR to be crystallized.[15] For this reason, rhodopsin formed the basis for homology models of many GPCRs. These homology models have been used to study the binding of ligands and to study receptor function. Although an increasing amount of x-ray crystal structures of GPCRs are becoming available, the method of homology modeling is still a useful tool for researchers working on a receptor still waiting to be crystallized. GPR18 is one of many receptors that for various reasons has not been crystallized. Previous work in this lab resulted in the creation of a homology model of GPR18 based on the μ-opioid receptor structure.[16] The μ-opioid receptor exhibits 54% sequence homology to GPR18 within its TMH regions.

## GPCR Activation

Within each GPCR binding pocket is a set of residues that functions as a "toggle switch" that changes conformation upon agonist binding. The resultant conformational change within the binding pocket, leads to a change in the receptor intracellular domains that leads to signal transduction. These binding pocket "toggle switch" residues typically include residue 6.48 of the TMH6 CWXP hinge motif and another residue that interacts with 6.48, limiting its conformation. In order for an agonist to activate a GPCR bundle, the toggle switch must be "tripped." The key conformational change in the toggle switch involves the χ1 dihedral on W6.48 which starts on residue 6.48 which is at -60° ($g_+$) in the inactive state but moves to a χ1 of 180° (trans) in the activated state.[15]

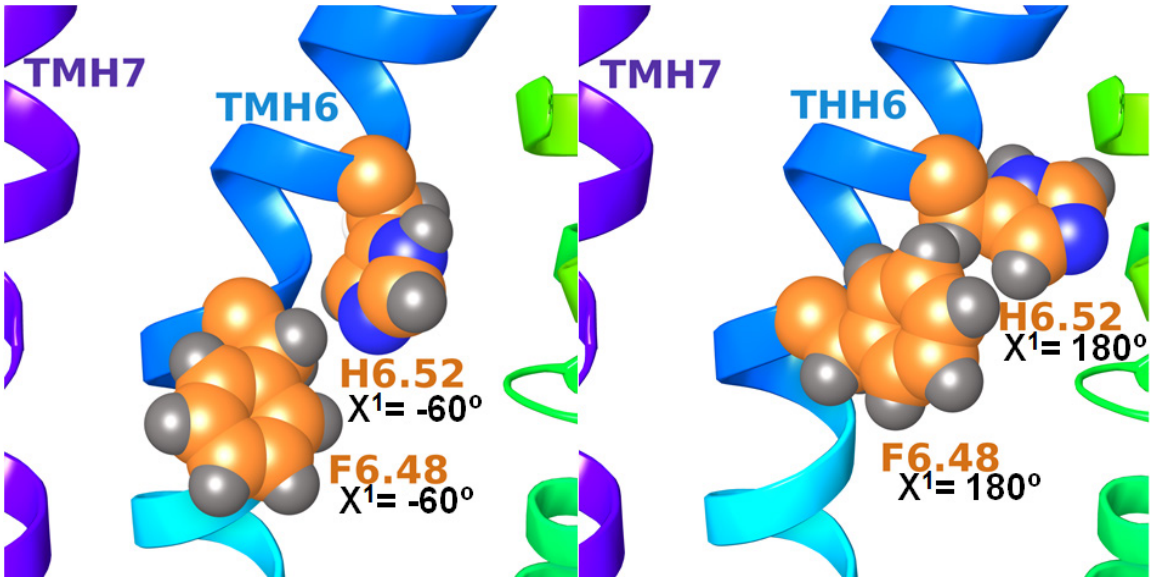



Figure 2. Toggle Switch between (Left) Inactive and (Right) Active States.

When this occurs, TMH6 is able to straighten in the CWXP hinge region, causing the "ionic lock" between R3.50 and E/D6.30 on the intracellular end of the receptor to

break. The break allows the receptor to open up and expose residues that can interact with the C-terminus of the Gα sub-unit of the G protein and causes a cell signal.[3] However, when an antagonist binds, the ligand will block this toggle switch conformational change, making the receptor unable to accept the G-protein and cell signaling will not occur. Many GPCRs have aromatic residues that modulate 6.48 conformation. Another unique characteristic specific to GPR18 is that it lacks the E/D6.30 necessary to form an ionic lock to keep the receptor inactive. Instead of a negatively charged residue, there is a positively charged Lys which would be unable to interact to R3.50. Although there is no GPR18 crystal structure to verify, we hypothesize that the TMH3/TMH6 IC end may be closed due to a serine at position 6.31 that could be involved in a hydrogen bond with R3.50, thereby serving as the "ionic lock."

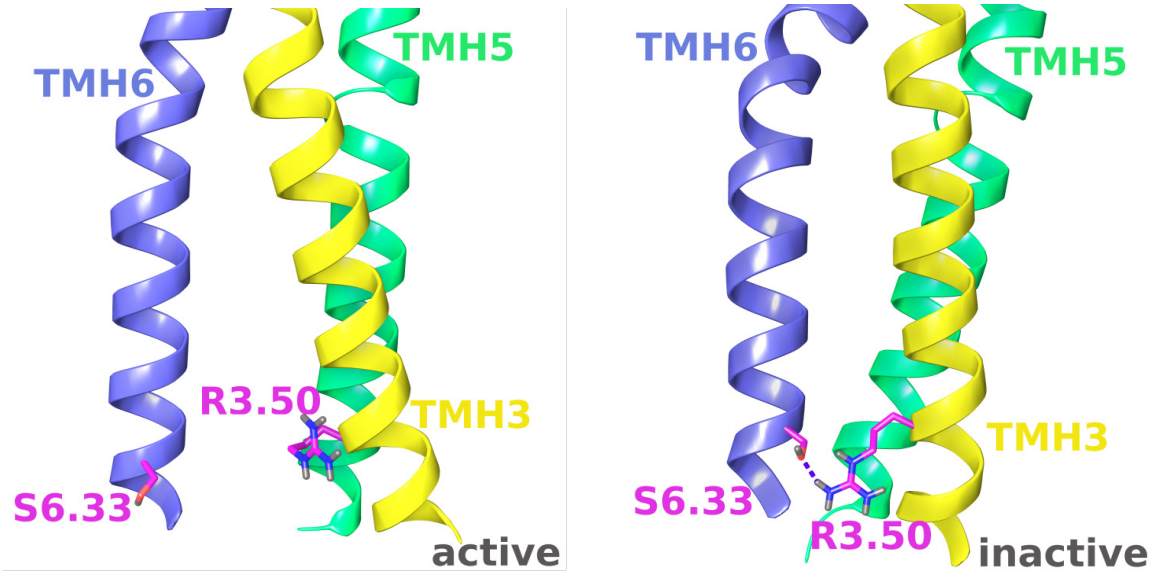



Figure 3. Ionic Lock in Inactive Bundle (Left), and Ionic Lock Broken in Active Bundle (Right).

## GPR18 Ligands

Like most cannabinoid receptors, GPR18 is known to bind small lipophilic molecules and lipophilic endogenous ligands.[1,17] Due to this similarity, GPR18 ligands have been hypothesized to enter the receptor via the lipid bilayer. This type of entry has also been shown for the $CB_2$ receptor.[15,18] NAGly has been reported to be the GPR18 endogenous ligand.[1] NAGly is a derivative of the $CB_1$ endogenous ligand, N-arachidonoyl ethanolamine (AEA).[1] GPR18 is considered a member of the cannabinoid receptor family because of its ability to be activated in cell migration assays by the atypical cannabinoids Abn-CBD, O-1602, as well as by $\Delta^9$-THC.[8] This induced cell migration is blocked by O-1918, or cannabidiol (CBD).[8] This has led many researchers to believe that GPR18 is the Abn-CBD receptor that George Kunos found in 2003 but was never able to cloned.[11]

# CHAPTER II

# HYPOTHESIS AND METHODS

## Goals

The focus of this thesis was to create working computational models of GPR18 that could consequently be studied via mutation and molecular dynamics. This thesis project is a continuation of a previous project in which a homology model of the GPR18 R (Inactive) state was created [16] using the TMH region of the μ-opioid receptor x-ray crystal structure.[19] Using this model, a homology model of the GPR18 R* (Active) state was created here. The most dramatic change in the TMH bundle as the receptor moves from the R to the R* state is the straightening of TMH6. Using Conformational Memories (CM), the Monte-Carlo/simulated annealing technique, possible GPR18 TMH6 conformations were calculated in order to select a straightened TMH6 to represent the R* state. The previously discovered GPR18 inactive state binding pocket [16] was explored in ligand docking studies of five known antagonists. In addition, the endogenous agonist (NAGly) was docked in the newly created R* bundle. Interactions formed through docking were studied to highlight key binding pocket residues necessary for activation or inactivation of the receptor.

## Methods

### Rotameric States in Amino Acid Side Chains

Throughout this thesis, the χ1 dihedral of various side chains in the protein will use rotameric nomenclature used by Shi et al in 2002 to describe their angle.[20] A *trans* χ1 angle has a heavy atom at the γ position in which view from β-carbon to α-carbon is opposite of the backbone nitrogen (180°), making the side chain extend out. Viewed along the same axis, when the γ heavy atom is opposite the backbone carbon, a *gauche*$^+$ ($g^+$) is defined corresponding to an angle of -60°. In the opposite direction, a *gauche*$^-$ ($g^-$) has the γ-carbon heavy atom opposite the α-hydrogen at an angle of +60° (see Figure 4). Both $g^-$ and $g^+$ have the side chain bent inwards towards the helix backbone.

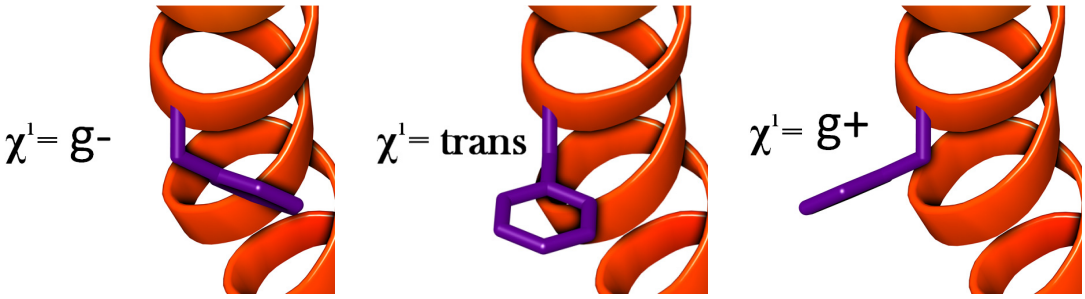



Figure 4. Example of g+, Trans, and $g^-$ Conformation in the χ1 Dihedral of a Phe Side Chain.

### Conformational Searches of Antagonists

Complete conformational analyses of the GPR18 antagonists: 1,3-dimethoxy-5-methyl-2-[(1R,6R)-3-methyl-6-prop-1-en-2-ylcyclohex-2-en-1-yl]benzene; **1**, (2R,3aS,3bR,5aS,11bS,11cS,13aS)-3a,11b,11c-trimethyl-2-(2-methylprop-1-en-1-yl)-2,3,3a,3b,4,5,5a,6,11,11b,11c,12,13,13a-tetradecahydrobenzofuro[4',5':6,7]indeno[1,2-

b]indole; **2**, (3S,4S,4aR,6aS,12bS,12cS)-4-((E)-4-methoxy-4-methylpent-2-en-1-yl)-4,12b,12c-trimethyl-1,2,3,4,4a,5,6,6a,7,12,12b,12c-dodecahydrobenzo[6,7]indeno[1,2-b]indol-3-ol; **3**, (Z)-2-(3-((4-chlorobenzyl)oxy)benzylidene)-6,7-dihydro-5H-imidazo[2,1-b][1,3]thiazin-3(2H)-one; **4**, and 2-[(1*R*,6*R*)-6-isopropenyl-3-methylcyclohex-2-en-1-yl]-5-pentylbenzene-1,3-diol; **5** were performed to identify the global minimum energy conformation of each compound. Each compound was built in Spartan08 and optimized in the builder module. A conformational search facility was used to generate 7 conformers for the first torsion angle explored. Starting conformations represented sequential increments of 51.4° in the first torsion angle. Each generated conformation was optimized using the Hartree-Fock method at the 6-31G* level of theory. Any high level energy duplicates were eliminated and only unique conformers with lower energies were studied further. This process was repeated for each major rotatable bond within each compound.

**Conformational Memories (CM)**

To create a GPR18 activated state model, all possible conformations of TMH6 were explored using the conformational memories (CM) method. Due to the highly flexible nature of the arachidonic acid portion of the endogenous ligand, NAGly, CM was also used to generate a set of low free energy conformers for NAGly. The CM method employs multiple Monte Carlo/simulated annealing random walks and the CHARMM (**C**hemistry at **HAR**vard **M**olecular **M**echanics) force field.[21] The CM method allows highly flexible molecules to converge in a fewer number of steps and efficiently overcome large energy barriers. The CM method can give a full picture of the range of

conformations allowed for each molecule based on their free energies. These free energies are a sum of both the intrinsic energy of each allowed conformational state and the thermodynamically applied energy conformations. Each conformation chosen are selected because they are considered more favorable to another conformation. These CM calculations are performed in two phases, the Exploratory and Biased Annealing Phases.

**Exploratory phase.** The first phase of CM is to identify the region of conformational space that is most probable for each torsion angle and bond angle. In order to do this, a random walk is used. Each run starts with an initial temperature of 3000 K and decreases in temperature for 18 steps to a final temperature of 310 K. Within each of these steps, 50,000 Monte Carlo steps are applied to each torsion or bond angle being varied. For each step, two dihedral angles and one bond angle from the entire set of variable angles are chosen at random at each temperature and each movement of the two torsion angles and the one bond angle are accepted or rejected using the Metropolis criterion.[22] If a conformation is accepted in this phase, then it is used to create "memories" of each of the accepted torsion angles and bond angles creating a map of the accessible conformational space that the molecule can exist in as a function of temperature.

**Biased annealing phase.** Using the map of populated conformation space identified in the exploratory phase, the second phase of CM explores torsion angles and bond angles that only occur in that "populated conformational space." Biased annealing starts with an initial temperature of 749.4 K and cools to 310 K in seven steps. The final output contained 108 structures at 310 K for TMH6 and 105 structures for NAGly.

**CM study of GPR18 TMH6.**

***TMH6.*** In TMH6 of Class A CPCRs, there is a conserved flexible hinge motif, CWXP. In GPR18 the motif is CFXP, where the usual "toggle switch" Trp (6.48) is replaced with a Phe (6.48). This residue has been found to be important for activation in many GPCRs. Previous studies have shown that a conformation change in the 6.48 $\chi$1 from $g^+ \rightarrow$ trans induces conformational changes in the receptor that create an opening on the intracellular (IC) side of the receptor that allows the G protein to dock with the receptor.[20] In GPR18, H6.52 appears to form the other part of the toggle switch. This residue must undergo a conformational change in the $\chi$1 from g+ $\rightarrow$ trans in order to allow F6.48 to undergo its $\chi$1 g+ $\rightarrow$ trans transition. CM was used to identify a helix that would allow the IC end of the bundle to open. This opening should allow interaction with the G protein. The phi and psi dihedrals of the backbone from i (P6.50) to i-4 (V6.46) region was allowed to vary ±50° while the rest of the phi and psi dihedrals of the backbone were varied ±10°. Amides were varied ±20° and side chains were varied ±180°. Default bond angles were varied ±8° except for special cases like Met C-S-C bonds, Tyr C-O-H bonds, and Lys C-N-H which were all varied ±15°.

**Superimposition and Determination of Helices**

To achieve a packed R* bundle, the GPR18 TMHs of the inactive R bundle were pulled apart 2Å away from a central point. All 108 output conformations from CM were superimposed onto the inactive TMH6 in the inactive GPR18 model. The Cα's from His23 to Leu31 were used as the basis for the superimposition.

Once all outputs were superimposed, a helix that was the most straightened and had the least amount of steric clashes in the bundle was chosen for the new R* model. This new TMH6 also had an F6.48 $\chi$1 of *trans*. The result should be a bundle in which no possible hydrogen bond between the R3.50 and S6.33 "ionic lock" is possible. However, problems arose from the CM outputs not exploring the region of space needed to create an opening big enough for the G protein to interact. Recent studies have proposed a tilting of helices may occur during activation to allow the intracellular region to open up to accommodate the G protein.[23] With this new information, the EC end of the chosen TMH6 was tilted inward 1.5Å. The movement allowed an appropriate intracellular opening.

**Ligand Docking Protocol**

A total of 5 antagonists and 1 agonist were docked into the R and R* bundles respectively. All low energy conformations identified for each ligand were then docked used using the automatic docking program, Glide (Schrodinger Inc.). From the centroid of select residues inside the binding site, Glide generates a grid.[24] The primary interaction site chosen within the binding pocket for Glide docking studies was R5.42 because of its positive charge and its location near the F6.48/H6.52 toggle switch residues. With the grid generated, Glide is able to explore the binding pocket and find the lowest energy ligand/receptor complex. Glide denotes energy favorable docks with low Glide scores. The ligand/receptor complex with the lowest Glide score was then energy minimized using the OPLS 2005 force field, by the same process as described below.

**Minimization Protocol**

Each ligand/receptor complex was energy minimized using Macromodel 9.9 (Schrödinger LLC) to a convergence of 0.1 kcal/mol·Å. In the active bundle, a distance constraint was applied to the α carbon of R3.50 and the α carbon of S6.33 was constrained to a distance of 13.4 Å ± 2.5 Å with a force of 1000 kcal/mol·Å. All charged residues pointing into the bundle where mutated to their neutral forms except for D2.50 and D7.49 which were solvated with 5 waters. The χ1 dihedrals for both Y5.58 and L6.44 were adjusted to g+ to allow TMH5 and TMH6 to pack. The minimization protocol used the OPLS_2005 force field and a distance dependent dielectric with a dielectric constant of 1. All backbone phi / psi dihedrals were constrained with a force of 1000 kcal/mol·Å, while the χ1 dihedrals of the amino acid side chains of each residue were allowed to vary. The energy was allowed to converge to 0.1 kcal/mol, then the torsional constraint force was reduced by half and the bundle was minimized again to the same gradient. This sequence was repeated 6 times. In the 7th stage the torsional constraints were eliminated. An 8.0-Å extended nonbonded cutoff (updated every 10 steps), a 20.0-Å electrostatic cutoff, and a 4.0-Å hydrogen bond cutoff were used in each stage of the calculation. For each minimization, the ligand was present within the binding pocket, to preserve the ligand binding space.

**Loop and Termini Methodology**

To complete the R* bundle, loop segments and termini were built and added to the R* model using Modeller 9.8. Modeller uses a template library of optimal side chain conformations derived from the Protein Data Bank for all amino acids. Each loop

structure is given an objective function ranking value. This value is assigned based on steric interactions and hydrogen bonding of each possible conformation of both the loops and the termini. This Monte Carlo technique developed by Fiser and co-workers [25] orders the output to give energy realistic IC and EC loop conformations. Using the CHARMM force field, three sets of loops/termini were run separately and assigned an objective function ranking value.[21] The first set contained the 3 IC loops and the C-terminus, the second contained the three EC loops, and the last was the N-terminus. By simultaneously running the IC loops in a single calculation and EC loops in a single calculation, more plausible conformations could be determined in the presence of each other. The N-terminus was calculated separately from the EC loops to find outputs appropriate for both with the lowest energies and the desired interactions for the docked agonist. The 300 lowest energy loop outputs from each group were used for further analysis.

Once a final output set of loops were selected, the bundle was refined via two sets of minimizations in MacroModel. The first minimization was run at a high dielectric ($\varepsilon$=80; water) with all the transmembrane regions frozen, but loops and termini free to move. After the first minimization, the loop regions, termini and TMH backbones were frozen with the TMH residues free to move. This system was minimized in a low dielectric ($\varepsilon$=1) using extended non-bonded cut offs.

# CHAPTER III

# RESULTS

## Ligand Conformational Searches and Docking

### Antagonists

The binding sites of five antagonists were explored using the previously constructed inactive state GPR18 model. The key interaction site for the antagonists, R5.42, positioned the ligands on top of the F6.48/H6.52 toggle switch residues, preventing the conformational change associated with activation of the receptor. Based on the functional groups within the structures of each ligand, more interactions were able to form in the binding pocket. The final Glide docking score of each antagonist/GPR18 R complex followed the trend of $IC_{50}$ values are displayed in Table 1.

Table 1

The Five Antagonists with Their $IC_{50}$/$EC_{50}$ Values and Their Final Glide Score

| Antagonist | Glide Score | $IC_{50}$ |
|---|---|---|
| **1** | -6.06 | - |
| **2** | -6.49 | 9.91 ± 2.59 μM |
| **3** | -6.46 | 13.4 ± 1.3 μM |
| **4** | -9.97 | 0.28 ± 0.11 μM |
| **5** | -6.92 | - |

*Both **1** and **5** were proven as antagonists through the attenuation of cell migration induced by NAGly.

**Antagonist 1.** 1,3-dimethoxy-5-methyl-2-[(1*R*,6*R*)-3-methyl-6-prop-1-en-2-ylcyclohex-2-en-1yl] benzene (O-1918), **1** has long been known as an antagonist at the two former orphan receptors GPR18 and GPR55. Although there is no $IC_{50}$ value for this compound, Compound **1** was shown to inhibit the NAGly induced migration of BV-2 and HEK293-GPR18 transfected cells.[11,13] Compound **1** has only one crucial rotatable bond and was found to have only 2 conformers (Conformer A, global min and Conformer B, ΔE= 1.03 kcal/mol; see Figure 5). Docking studies identified primary interaction site for the antagonist to be R5.42. Here, the interaction was a cation-pi interaction between the primary interaction site R5.42 and the aromatic portion of Compound **1** (see Figure 6). Another interaction was spotted involving H6.52 which created a pi stacking interaction with the same aromatic portion of antagonist 1 giving the dock a Glide score of -6.06 (see Figure 7).

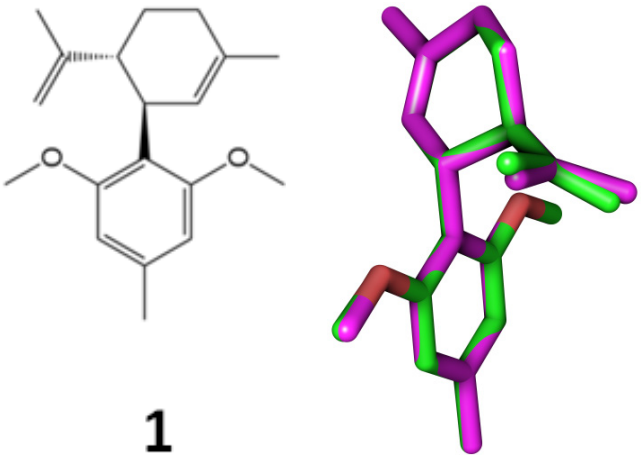



Figure 5. O-1918 (**1)** Conformers. Conformer A: Green; Global Min. Conformer B: Pink; ΔE=1.031 Kcal/Mol.

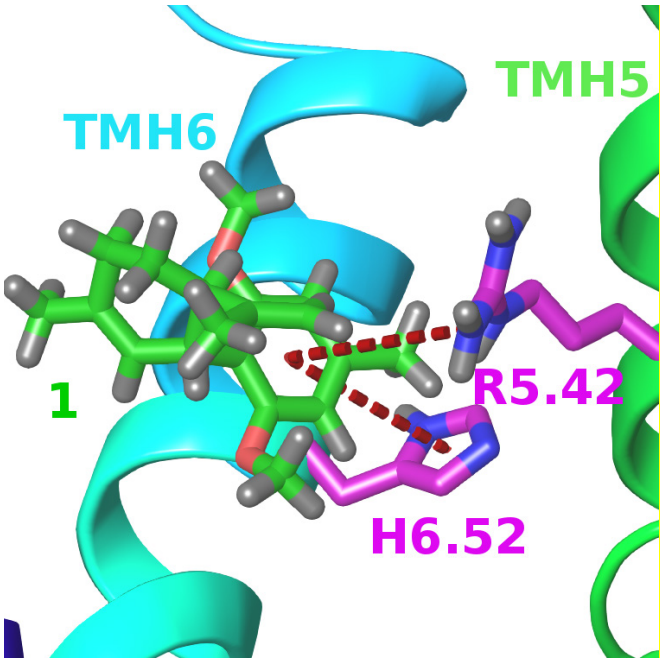



Figure 6. Compound **1** Forming a Cation Pi Interaction with the Primary Interaction Site, R5.42 and Pi Stacking with H6.52.

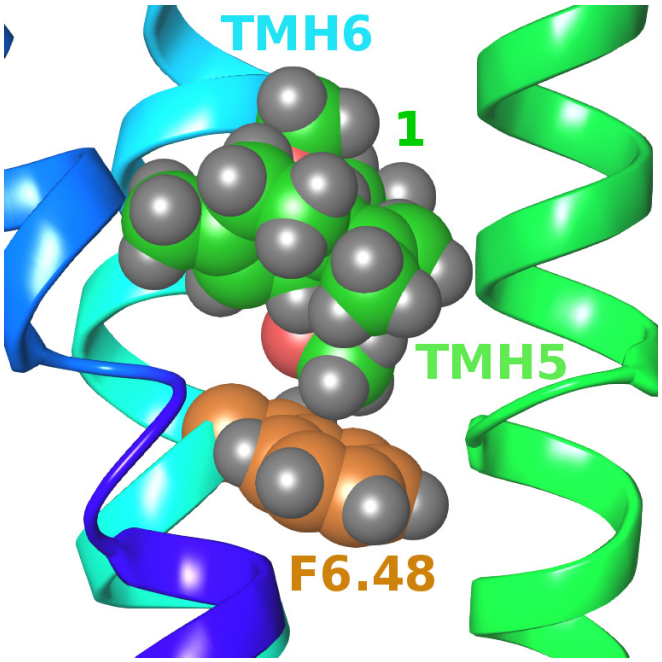



Figure 7. Compound **1** above the Inactive State Toggle Switch, Preventing Activation.

**Antagonists 2 and 3.** Compounds **2** and **3** are marine fungal derived antagonists of GPR18 with an $IC_{50}$ of 9.91 ± 2.59 µM and 13.4 ± 1.3 µM respectively.[26] Compound **2** has only one rotatable bond. Only one conformer was found for **2** due to the limited number of rotatable bonds within the structure (see Figure 8). Inside the binding pocket, **2** formed a cation pi interaction with R5.42 and a hydrogen bond with the backbone carbonyl of F6.51 giving a Glide score of -6.49 (see Figure 9).

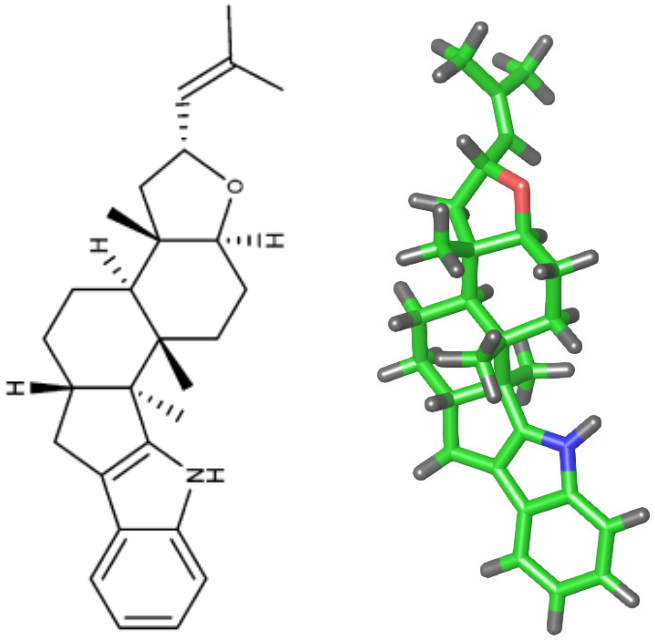


Figure 8. 2D Structure of Compound **2** (Left) and the Global Min of Compound **2** (Right).

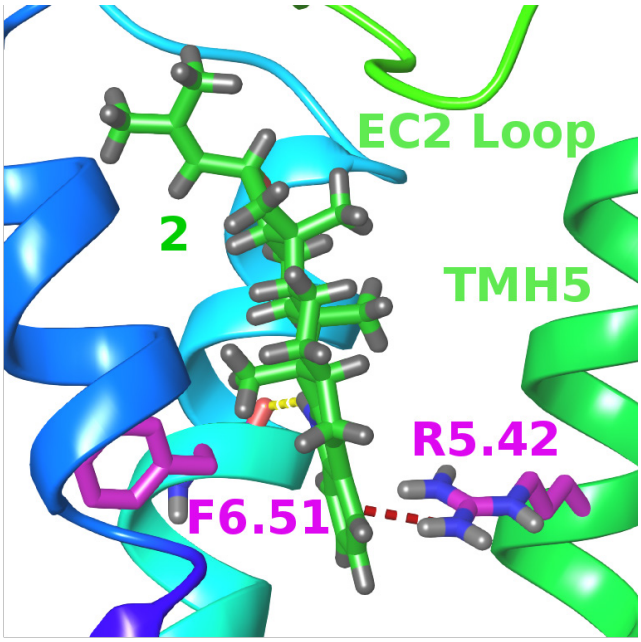



Figure 9. Cation Pi Interaction with R5.42 and Docked Antagonist **2** and Hydrogen Bond with the Backbone of F6.51.

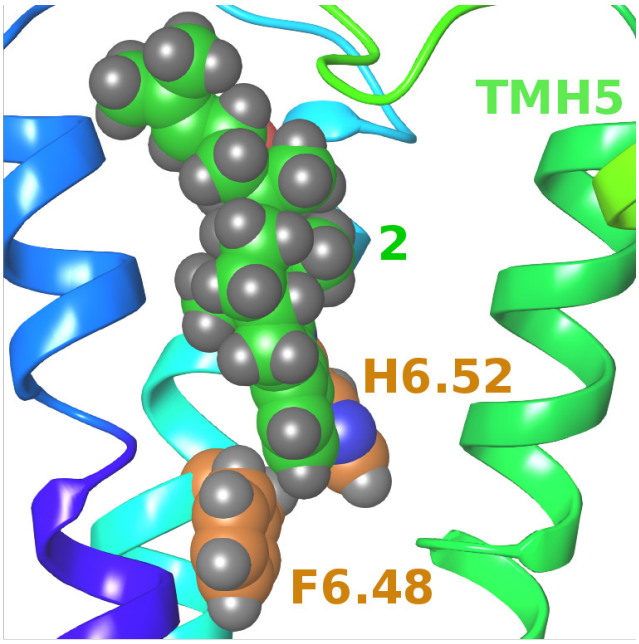



Figure 10. Compound **2** Interacting with the Inactive State Toggle Switch Preventing Conformational Change.

Compound **3** has four rotatable bonds (see Figure 11). Conformational analysis resulted in 83 conformers. When the global min of **3** was docked in the inactive state bundle, its aromatic ring forms a cation pi interaction with R5.42 along with a hydrogen bond with the backbone of F6.51 with a Glide score of -6.46 (see Figure 12).

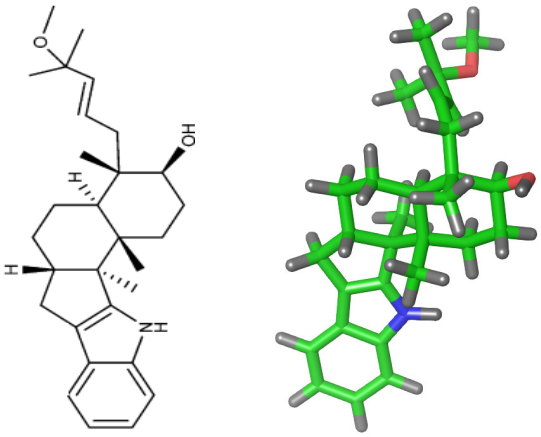



Figure 11. 2D Structure of Compound **3** (Left), and All Conformations of Compound **3** Superimposed on the Fused Ring System (Right).

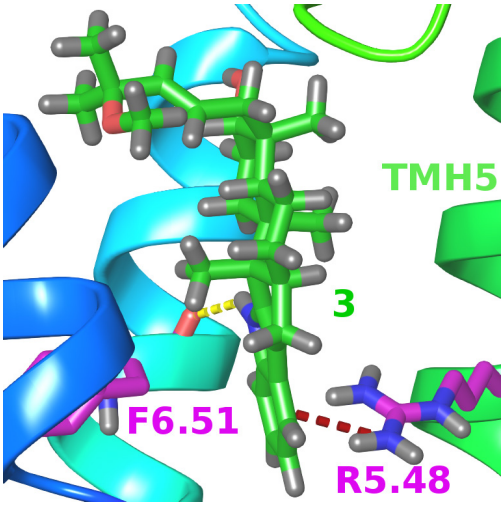



Figure 12. Compound **3** Forming a Cation Pi Interaction with R5.42 and a Hydrogen Bond with the Backbone of F6.51.

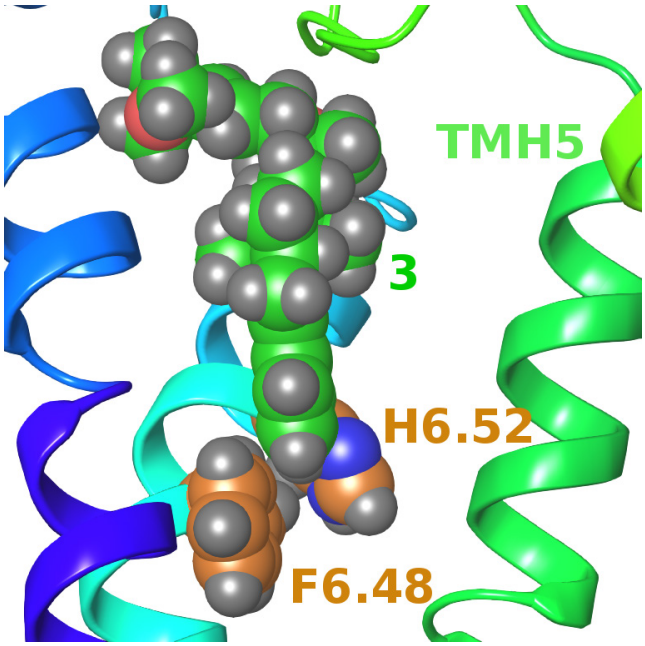
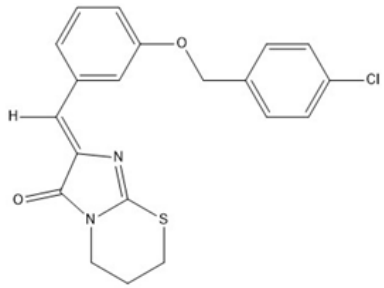



Figure 13. Compound **3** Interacting with the Inactive State Toggle Switch Preventing Conformational Change.

**Antagonist 4.** Compound **4** is a GPR18 antagonist with an $IC_{50}$ of 0.279 +/- 0.111 µM.[27] Compound **4** has 3 rotatable bonds and produced 25 conformers. The conformers had either an extended conformation or a petal conformation. The lowest energy conformers were the extended conformers and the global min was then docked into the inactive bundle. Docking studies indicated that the antagonist forms pi stacking interaction with H6.52, while the primary interaction residue forms a cation pi interaction with the ligand (see Figure 15). A hydrogen bond interaction occurs with a carbonyl on the antagonist and a residues on the EC2 loop, K174, further holding the ligand in place on top of the toggle switch (see Figure 16). A fourth interaction occurs with the chlorobenzene group on antagonist **4** and M7.42 creating a methionine interaction.[28] The

antagonist generated the best Glide score of -9.97 due to its multiple interactions within the binding pocket (see Table 1).




**4**

Figure 14. 2D Structure of Antagonist **4**.

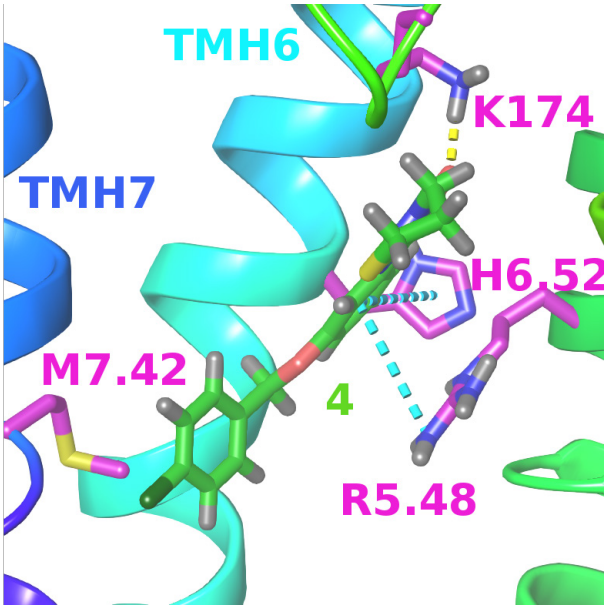



Figure 15. Compound **4** Forming a Hydrogen Bond Interaction with K174 and Two Pi Interactions with R5.48 and H6.52; Antagonist 4 also Forms a Methionine Interaction with M7.42.

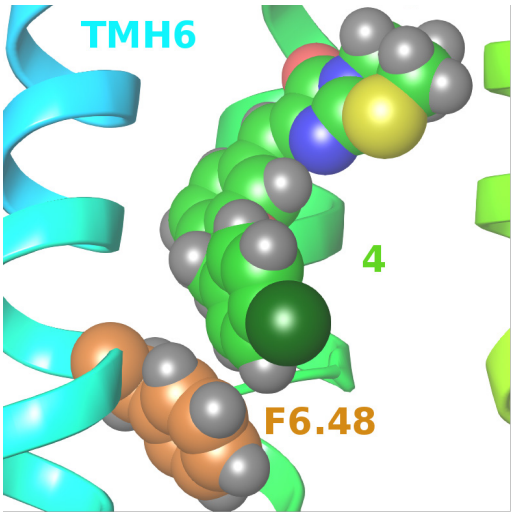



Figure 16. Antagonist **4** Interacting with the Inactive State Toggle Switch Preventing Activation.

**Antagonist 5.** Compound **5** is a small molecule that was found to be a GPR18 antagonist by McHugh et al.[(11)] and was used as the final antagonist to dock in the inactive state receptor model. Compound **5** has three main rotatable bonds that would that governed the conformations of the ligand. **5** produced 8 unique conformers (see Figure 17). When docked, compound **5** formed two hydrogen bond interactions with R5.42 and Y174, preventing the toggle switch from changing conformation, giving it a Glide score of -6.92 (see Figure 18 and Figure 19).

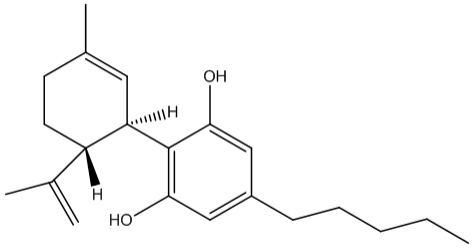



Figure 17. 2D Structure of Antagonist **5**.

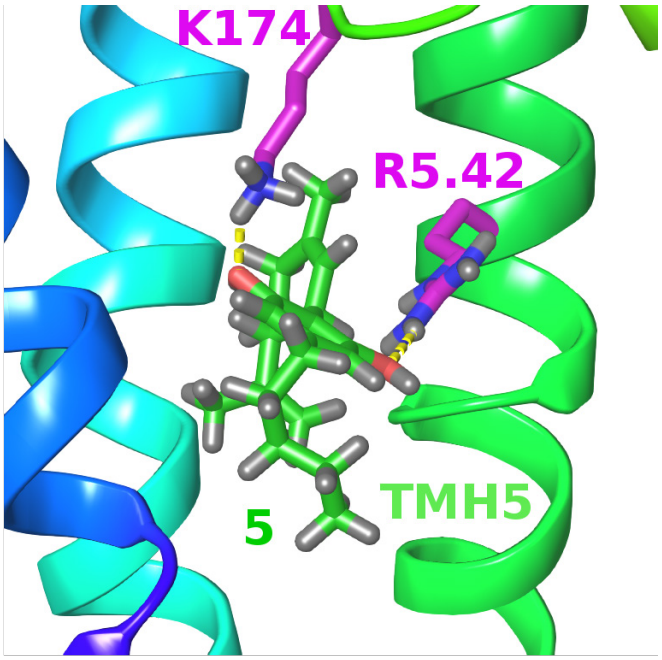



Figure 18. Docked Compound **5** Forming Hydrogen Bonds with Primary Interaction Residue R5.42 and K174.

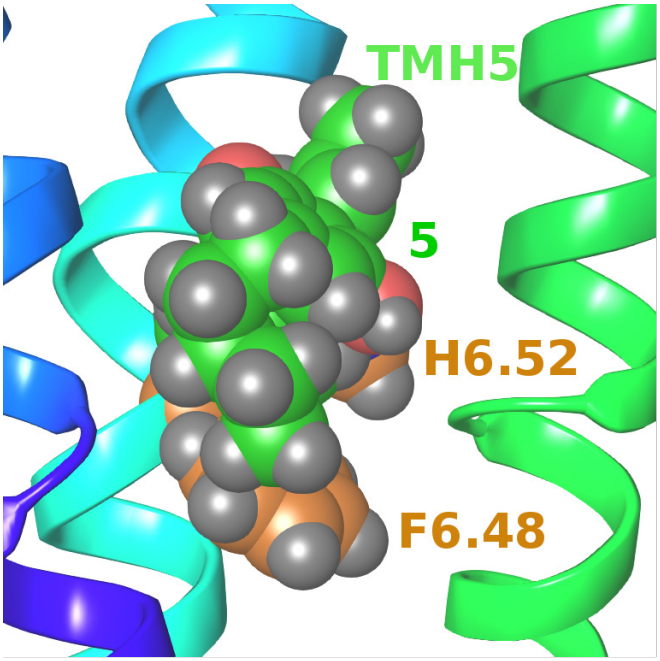



Figure 19. Compound **5** above the Toggle Switch Residues Preventing Activation.

**NAGly**

The GPR18 endogenous agonist, NAGly, was chosen as the sixth ligand for the docking studies (see Figure 20). NAGly was identified as the endogenous ligand of GPR18 via screening of 198 lipids of the Bioactive Lipid Alignment Library against a stable polyclonal population of GPR18-expressing L929 cells and GPR18 stably transfected K562 and CHO cells. This screening used intracellular $Ca^{2+}$ mobilization as the readout. At a concentration of 10 μM, NAGly induced a significant increase in intracellular $Ca^{2+}$ concentration in GPR18-expressing cells as compared to mock-transfected cells.(1) NAGly also inhibited cAMP production in GPR18-transfected CHO cells compared to mock-transfected cells with an $EC_{50}$ of 44.5 nM.(1) Pretreatment of the GPR18-transfected CHO cells with pertussis toxin (PTX) abolished the inhibition of forskolin-stimulated cAMP production by NAGly.(1) These experiments by Kohno and co-workers were the first to give solid evidence that NAGly is a natural ligand for GPR18 and that GPR18 couples to Gi/o proteins.(1)

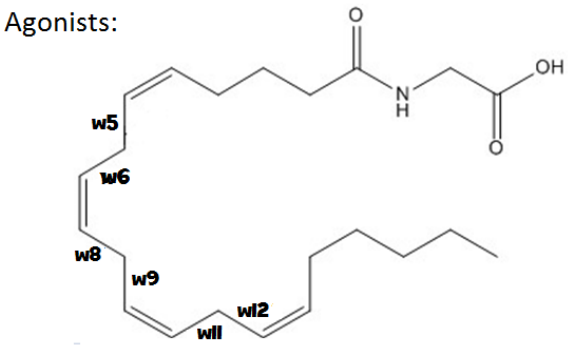



Figure 20. 2D Structure of Compound **6**; NAGly.

NAGly has enough rotatable bonds to make a traditional conformational analysis intractable. Instead, we used the Conformational Memories (CM) method, to generate a set of low free energy conformers of this ligand. This method employs multiple Monte Carlo/simulated annealing random walks and the CHARMM (Chemistry at HARvard Molecular Mechanics) force field.[21] Similar to previously reported CM results for arachidonic acid,[10] NAGly's conformers could be divided into four groups: linear, U-shape, J-shape, and helical.

When measuring the dihedrals in each cluster, NAGly had the same shape-clusters identified previously for the CB1 endogenous agonist, anandamide.[7] In general, when each angle in a pair has the same sign, that region of the acyl chain will be extended. A pair is considered as any two dihedrals adjacent to another for example ω5, consisting of carbons 5-8, and ϖ6, consisting of carbons 6-9. When the two angles in a pair have opposite signs, curvature is introduced into the acyl backbone. In the J-shape (Figure 21), only the (ω5, ϖ6) pair has opposite signs while the first 3 dihedrals and the following homoallylic dihedrals have the same sign. The J-shapes include a subset of structures noted as reverse Js where the head group is near the top and the tail curves into the J. In this sub group it is the (ϖ11, ϖ12) pair that has the opposite signs. In the anandamide paper, for the helical shape (Figure 21), (ϖ5, ϖ6), (ϖ8, ϖ9), and (ϖ11, ϖ12) each had opposite signs. However, our findings indicate that the helical conformers only needed to have at least two dihedrals adjacent to each other with opposite signs to form a helical-like acyl backbone. In the U-shape (Figure 21), (ϖ5, ϖ6) and (ϖ11, ϖ12) have opposite signs which induced curvature near the ends of the molecule rendering it U-

shaped. Most of the conformers fit this pattern however there were some structures that did not. The outliers did have dihedrals in the (ϖ1-ϖ3) that had the same signs then a pair or two in between with opposite signs that induced a curve followed by another pair with the same sign that created the last U-arm. For the last group, the extended conformers we also see that all three pairs of dihedrals have the same sign as their counter within the pair, although there as a few outliers that only contain two pairs with the same sign. For these outliers the pairs are adjacent to each other which still allowed the conformer to extend.

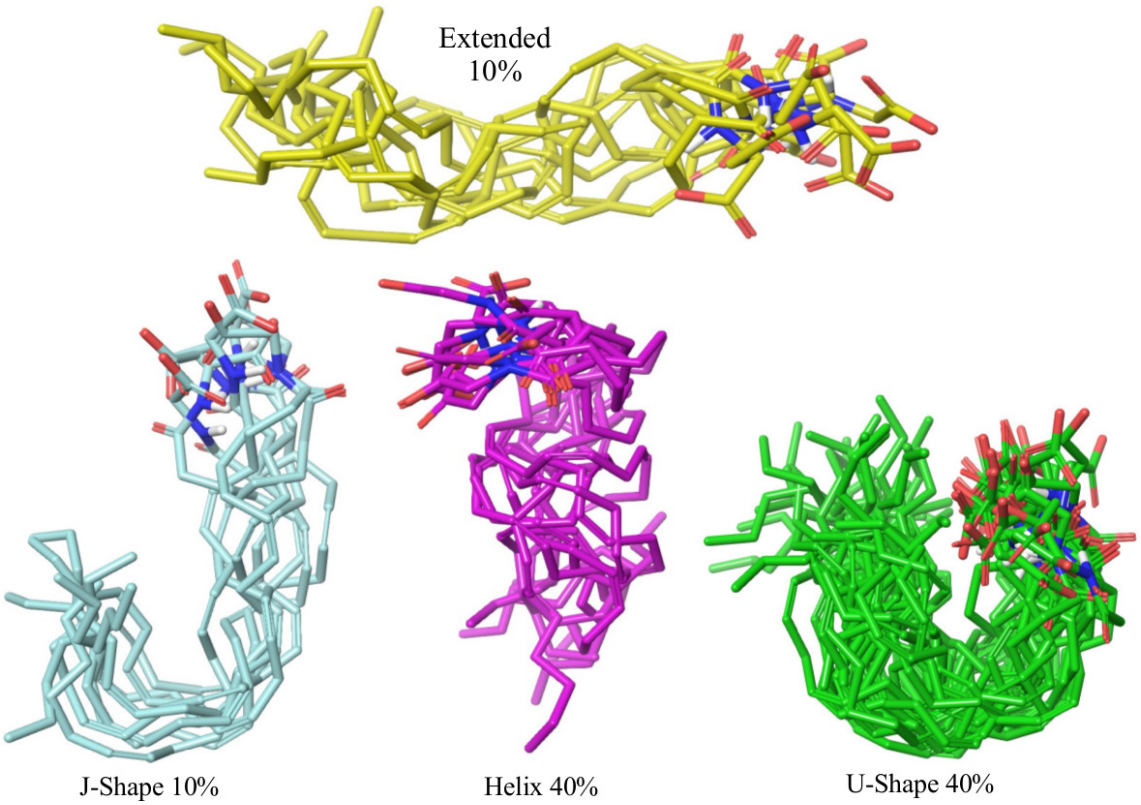



Figure 21. The Four Conformational Shapes NAGly Adopted in its CM Output with Their Percent Occurrence.

Table 2

Torsion Angles of NAGly

| Torsion Angle | Corresponding Carbons |
|---|---|
| $\omega_5$ | C5, C6, C7, C8 |
| $\omega_6$ | C6, C7, C8, C9 |
| $\omega_8$ | C8, C9, C10, C11 |
| $\omega_9$ | C9, C10, C11, C12 |
| $\omega_{11}$ | C11, C12, C13, C14 |
| $\omega_{12}$ | C12, C13, C14, C15 |
| $\omega_{16}$ | C16, C17, C18, C19 |

Before loops could be added to the bundle, NAGly was docked into the newly modified R* bundle. By doing this, the bundle could be minimized around the endogenous ligand in the binding pocket. The primary interaction site for NAGly at GPR18 was found to be another arginine, R2.60, on the opposite side of the bundle from the antagonists' primary interaction site. Here, a hydrogen bond is formed between the NAGly carboxylic oxygen and R2.60 on TMH2. Another hydrogen bond between NAGly's carbonyl group and the EC2 loop residue, K174, is also formed (see Figure 22). This binding site permits the conformational change in the toggle switch associated with activation (see Figure 23).

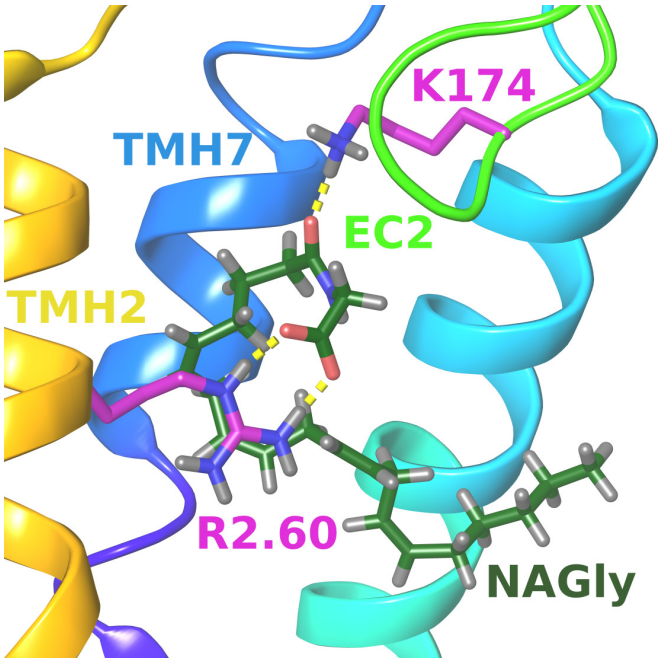



Figure 22. NAGly Docked into the Active State Bundle Showing the Electrostatic Interactions between R2.60 and Kl74 in the EC2 Loop.

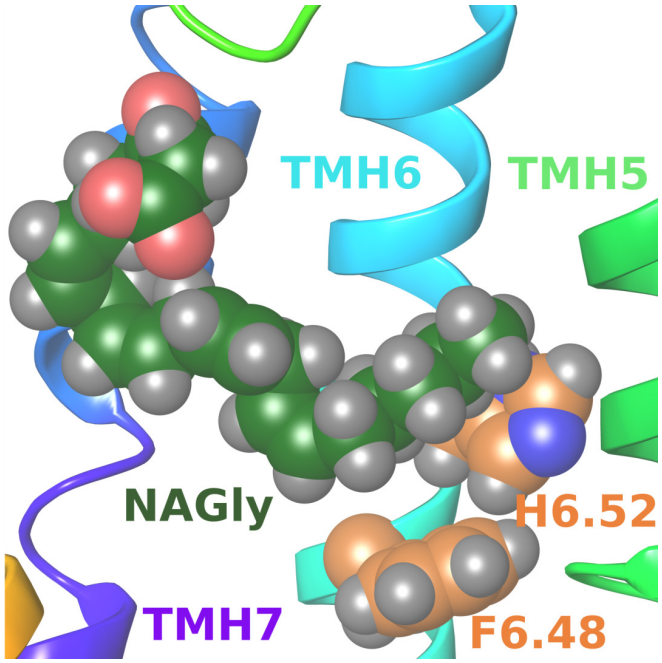



Figure 23. NAGly Positioned above the Already Activated Toggle Switch Residues.

**GPR18 R* Model Based on GPR18 R Model**

Using the previously constructed GPR18 (R) structure as a template, an active (R*) bundle was constructed. The transition from the inactive to active state for Class A GPCRs, involves a conformational change in TMH6 that causes the IC end of TMH6 to move away from the TMH bundle, breaking the ionic lock between the IC ends of TMH3 and TMH6. Figure 24 shows the CM output for GPR18 TMH6 with the EC ends of the helices superimposed. The cyan colored conformer was chosen for incorporation in the R* GPR18 model. Figure 25 illustrates the conformational differences between the final GPR18 R and R* models.

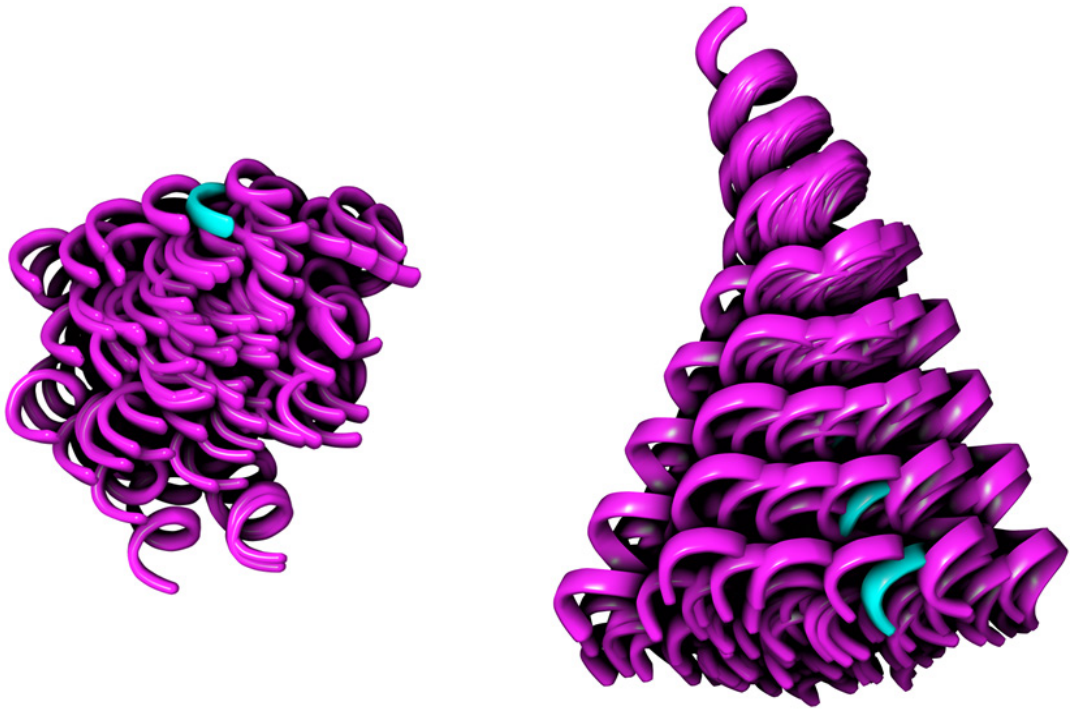


Figure 24. IC View of CM Output of TMH6, Blue Helix was Selected for R* Bundle (Left); View from the Lipid Bilayer of CM Output of TMH6, the Cyan Helix was Selected (Right).

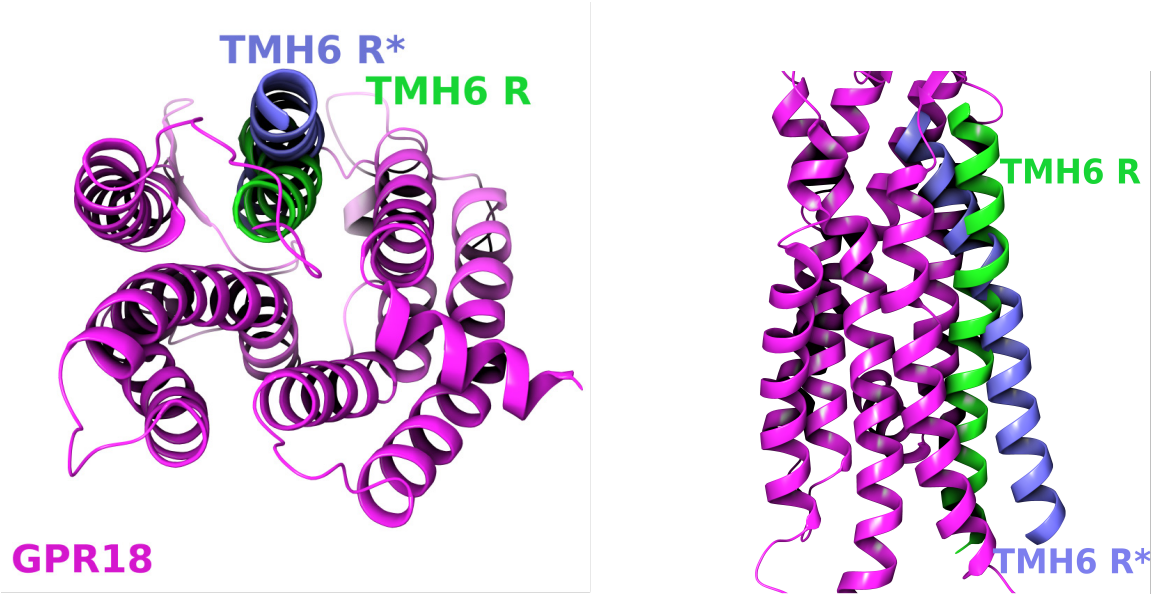



Figure 25. IC (Left) and Lipid Bilayer View of TMH6 in Inactive Conformation (Green) and Active Conformation (Blue) (Right).

## Loop and Termini Construction and Modeling

### Intracellular (IC) Loops and C Terminus

The loops were built in Maestro and connected to the bundle with as minimal steric clash as possible. Using Modeller, two groups of loops were run to calculate all possible conformers, in a high dielectric which simulates an aqueous biological environment. The first group contained the three IC loops and the C terminus, while the second group contained the three EC loops and the N terminus. The IC loops and the C terminus were run first because they had fewer constraints in their conformations. The intracellular loops for GPR18 where relatively small compared to the extracellular loops which limited the number of conformations the loops could take. Also, unlike the EC loops; there were no disulfide bonds that would play a part in the final loop conformation.

Using the same Modeller settings, conformations of all three IC loops and the C terminus were run together so that their outputs would incorporate conformations with respect to each other and minimize steric clash. Due to the lack of any mutation data on the IC loop and C terminus of GPR18, there was no specific loop or terminus conformation in mind. Once the loops were modeled, the output was analyzed for a set of low energy conformations that would not block a G protein from inserting into the IC end of the bundle and displayed a C-terminus conformation consistent with the presence of a lipid bilayer. Once an appropriate output was selected, all three loops and the c terminus was attached to the active bundle.

**Extracellular (EC) Loops**

Once the appropriate IC loops were added to the bundle, the EC loops conformations were calculated using Modeller with the same settings as the IC loops. Both the EC1 and EC3 loops were relatively small with 5 and 6 residues respectively. The EC2 loop, however, is much longer with 26 residues. Cys172 on EC2 loop is also part of an important disulfide bond with Cys3.25 (94) on TMH3, which was the only structural constraint specified. In all published x-ray crystal structures of Class A GPCRs containg a disulfide bridge between a Cys in EC-2 and a Cys at position 3.25 in the receptor, the second EC-2 residue after the Cys faces down into the ligand binding pocket. This was used to screen possible EC-2 loops for the GPR18 model. The EC-2 loop conformation incorporated into the model had Lys174 facing into the binding pocket. The chosen loops and sidechains were then minimized in a high dielectric.

**N Terminus**

The N Terminus of GPR18 contains 19 residues. Modeller was used to calculate a conformation of this terminus in a high 2.8 dielectric to simulate the aqueous environment. No structural restrictions were put on the N terminal segment, but outputs were selected based on conformations with respect to the lipid bilayer. Any conformations that had the terminus dipping into the lipid bilayer were immediately rejected. Overall, an output was chosen that draped over the top of the bundle. This would allow ligands to enter into the binding pocket from the lipid bilayer without interference, but would preclude entrance from the extracellular milieu.

**Summary**

A homology model of the active state of the GPR18 receptor was constructed using a recently built homology model of the receptor in its inactive state. Conformational Memories was used to explore all possible conformers of TMH6. A more straightened conformer of TMH6 was chosen for the activated state model. This conformer precludes the TMH6 from forming the ionic lock between S6.33 and the conserved R3.50. Without the ionic lock, the receptor would not be able to stabilize its inactive state allowing the extracellular end to open up and promote an interaction with the G protein. With the two bundles (both inactive and active) completed, several known GPR18 antagonists and the endogenous ligand were selected to be docked into the bundles to explore their key interactions. All five of the antagonists' conformational searches were done using Spartan while conformations of the endogenous ligand were found using Conformational Memories. All identified low energy conformers of each

ligand were docked using Glide into their respective bundles. The key interaction site for the antagonists, R5.42, positioned the ligands above the binding site toggle switch residues (F6.48 and H6.52) preventing the activation of the receptor. Other residues, like Y160, further stabilized a ligand in this binding site. Glide scores for the 5 antagonists followed the trend in their $IC_{50}$ values (Table 1). The primary interaction site for the agonist, NAGly, was found to be R2.60. Interaction of the head group of NAGly with R2.60, gave the ligand enough docking room such that the toggle switch residues (F6.48 and H6.52) were unconstrained and therefore transition to the active state was possible. Another interaction between NAGly and EC-2 loop residue, K174, also helped anchor the ligand at GPR18. Mutation studies currently in progress in Dr. Mary Abood's lab will test the importance of the interaction site residues identified here and further our understanding of GPR18.

## WORKS CITED


1. Kohno, M., Hasegawa, H., Inoue, A., Muraoka, M., Miyazaki, T., Oka, K., and Yasukawa, M. (2006) Identification of N-arachidonylglycine as the endogenous ligand for orphan G-protein-coupled receptor GPR18. *Biochemical and biophysical research communications* **347**, 827-832

2. Gantz, I., Muraoka, A., Yang, Y. K., Samuelson, L. C., Zimmerman, E. M., Cook, H., and Yamada, T. (1997) Cloning and chromosomal localization of a gene (GPR18) encoding a novel seven transmembrane receptor highly expressed in spleen and testis. *Genomics* **42**, 462-466

3. Weis, W. I., and Kobilka, B. K. (2008) Structural insights into G-protein-coupled receptor activation. *Current opinion in structural biology* **18**, 734-740

4. Hamm, H. E., Deretic, D., Arendt, A., Hargrave, P. A., Koenig, B., and Hofmann, K. P. (1988) Site of G protein binding to rhodopsin mapped with synthetic peptides from the alpha subunit. *Science* **241**, 832-835

5. Devane, W. A., Dysarz, F. A., 3rd, Johnson, M. R., Melvin, L. S., and Howlett, A. C. (1988) Determination and characterization of a cannabinoid receptor in rat brain. *Molecular pharmacology* **34**, 605-613

6. Li, S., Huang, S., and Peng, S. B. (2005) Overexpression of G protein-coupled receptors in cancer cells: involvement in tumor progression. *International journal of oncology* **27**, 1329-1339

7. Pridgeon, J. W., and Klesius, P. H. (2013) G-protein coupled receptor 18 (GPR18) in channel catfish: expression analysis and efficacy as immunostimulant against Aeromonas hydrophila infection. *Fish & shellfish immunology* **35**, 1070-1078

8. McHugh, D., Hu, S. S., Rimmerman, N., Juknat, A., Vogel, Z., Walker, J. M., and Bradshaw, H. B. (2010) N-arachidonoyl glycine, an abundant endogenous lipid, potently drives directed cellular migration through GPR18, the putative abnormal cannabidiol receptor. *BMC neuroscience* **11**, 44-58

9. Yin, H., Chu, A., Li, W., Wang, B., Shelton, F., Otero, F., Nguyen, D. G., Caldwell, J. S., and Chen, Y. A. (2009) Lipid G protein-coupled receptor ligand identification using beta-arrestin PathHunter assay. *The Journal of biological chemistry* **284**, 12328-12338

10. Barnett-Norris, J., Guarnieri, F., Hurst, D. P., and Reggio, P. H. (1998) Exploration of biologically relevant conformations of anandamide, 2-arachidonylglycerol, and their analogues using conformational memories. *Journal of medicinal chemistry* **41**, 4861-4872

11. McHugh, D., Page, J., Dunn, E., and Bradshaw, H. B. (2012) Delta(9) -Tetrahydrocannabinol and N-arachidonyl glycine are full agonists at GPR18 receptors and induce migration in human endometrial HEC-1B cells. *British journal of pharmacology* **165**, 2414-2424

12. Qin, Y., Verdegaal, E. M., Siderius, M., Bebelman, J. P., Smit, M. J., Leurs, R., Willemze, R., Tensen, C. P., and Osanto, S. (2011) Quantitative expression profiling of G-protein-coupled receptors (GPCRs) in metastatic melanoma: the constitutively active orphan GPCR GPR18 as novel drug target. *Pigment cell & melanoma research* **24**, 207-218

13. Gatley, S. J., Gifford, A. N., Volkow, N. D., Lan, R., and Makriyannis, A. (1996) 123I-labeled AM251: a radioiodinated ligand which binds in vivo to mouse brain cannabinoid CB1 receptors. *European journal of pharmacology* **307**, 331-338

14. Ballesteros, J. A. Weinstein, H. (1995) Integrated Methods for the Construction of Three-Dimensional Probing of Structure-Function Relations in G Protein-Coupled Receptors. *In Methods in Neurosciences* **25**, 366-429

15. Hurst, D. P., Grossfield, A., Lynch, D. L., Feller, S., Romo, T. D., Gawrisch, K., Pitman, M. C., and Reggio, P. H. (2010) A lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. *The Journal of biological chemistry* **285**, 17954-17964

16. Schmeisser, M. G. (2013) Creation of a GPR18 Homology Model Using Conformational Memories. in *Chemistry*, University of North Carolina at Greensboro

17. Li, J., Edwards, P. C., Burghammer, M., Villa, C., and Schertler, G. F. (2004) Structure of bovine rhodopsin in a trigonal crystal form. *Journal of molecular biology* **343**, 1409-1438

18. Pei, Y., Mercier, R. W., Anday, J. K., Thakur, G. A., Zvonok, A. M., Hurst, D., Reggio, P. H., Janero, D. R., and Makriyannis, A. (2008) Ligand-binding architecture of human CB2 cannabinoid receptor: evidence for receptor subtype-specific binding motif and modeling GPCR activation. *Chemistry & biology* **15**, 1207-1219

19. Manglik, A., Kruse, A. C., Kobilka, T. S., Thian, F. S., Mathiesen, J. M., Sunahara, R. K., Pardo, L., Weis, W. I., Kobilka, B. K., and Granier, S. (2012) Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321-326

20. Shi, L., Liapakis, G., Xu, R., Guarnieri, F., Ballesteros, J. A., and Javitch, J. A. (2002) Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *The Journal of biological chemistry* **277**, 40989-40996

21. Whitnell, R. M., Hurst, D. P., Reggio, P. H., and Guarnieri, F. (2008) Conformational memories with variable bond angles. *Journal of computational chemistry* **29**, 741-752

22. Zhang, R., Hurst, D. P., Barnett-Norris, J., Reggio, P. H., and Song, Z. H. (2005) 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. *Molecular pharmacology* **68**, 69-83

23. Ahuja, S., Hornak, V., Yan, E. C., Syrett, N., Goncalves, J. A., Hirshfeld, A., Ziliox, M., Sakmar, T. P., Sheves, M., Reeves, P. J., Smith, S. O., and Eilers, M. (2009) Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. *Nature structural & molecular biology* **16**, 168-175

24. Kotsikorou, E., Madrigal, K. E., Hurst, D. P., Sharir, H., Lynch, D. L., Heynen-Genel, S., Milan, L. B., Chung, T. D., Seltzman, H. H., Bai, Y., Caron, M. G., Barak, L., Abood, M. E., and Reggio, P. H. (2011) Identification of the GPR55 agonist binding site using a novel set of high-potency GPR55 selective ligands. *Biochemistry* **50**, 5633-5647

25. Fiser, A., Do, R. K., and Sali, A. (2000) Modeling of Loops in Protein Structures. *Protein Science : a publication of the protein society* **9**, 1753-1773

26. Harms, H., Rempel, V., Kehraus, S., Kaiser, M., Hufendiek, P., Muller, C. E., and Konig, G. M. (2014) Indoloditerpenes from a marine-derived fungal strain of Dichotomomyces cejpii with Dntagonistic activity at GPR18 and cannabinoid receptors. *Journal of natural products* **77**, 673-677

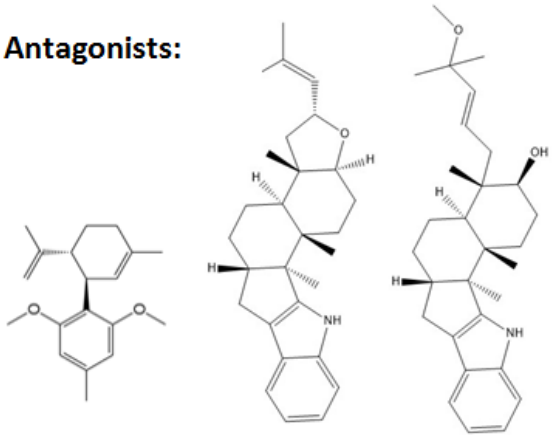
27. Rempel, V., Atzler, K., (2014) Bicyclic imidazole-4-one derivatives: a new class of antagonists for orphan G protein-coupled receptors GPR18 and GPR55. Med. Chem. Commun.

28. Valley, C., Cembran, A., (2012) The Methionine-aromatic Motif Plays a Unique Role in Stabilizing Protein Structure. *Journal of Biological Chemistry* **287**, 34979 –34991

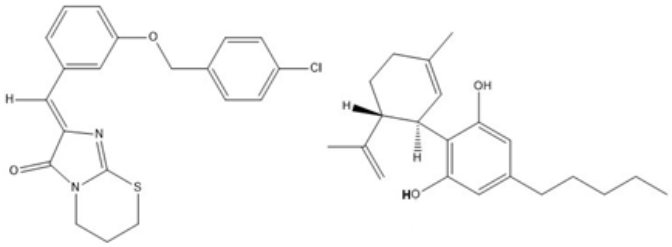
# APPENDIX A

## GPR18 ANTAGONISTS AND AGONIST

**Antagonists:**

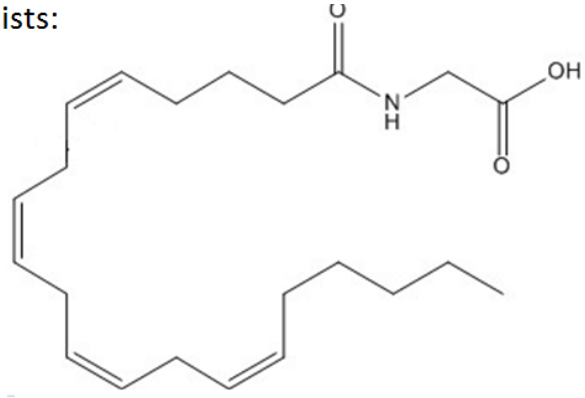



**1** **2** **3**




**4** **5**

Agonists:




**NAGly; 6**

# APPENDIX B

## GPCR SEQUENCE ALIGNMENT