5-2020

Elucidating the Binding Pockect of the GPR119 Receptor, a Type 2 Diabetes Target

Matthew Dy Rosales
The University of Texas Rio Grande Valley

ELUCIDATING THE BINDING POCKECT OF THE GPR119 RECEPTOR, A TYPE 2 DIABETES TARGET

A Thesis

by

MATTHEW DY ROSALES

Submitted to the Graduate College of
The University of Texas Rio Grande Valley
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2020

Major Subject: Chemistry

ELUCIDATING THE BINDING POCKECT OF THE GPR119 RECEPTOR, A TYPE 2 DIABETES TARGET

A Thesis
by
MATTHEW DY ROSALES

COMMITTEE MEMBERS

Dr. Evangelia Kotsikorou
Chair of Committee

Dr. Frank Dean
Co-Chair of Committee

Dr. James Bullard
Committee Member

Dr. Shizue Mito
Committee Member

May 2020

ABSTRACT

Rosales, Matthew Dy, Elucidating the Binding Pocket of the GPR119 Receptor, a Type 2 Diabetes Target. Master of Science (MS), May, 2020, 49 pp, 2 tables, 17 figures, references, 25 titles.

An in-house homology model of the GPR119 receptor was used to identify residues which may affect ligand binding and ligand-dependent activity through computational and experimental studies. In a docking study comparing an agonist and its inverse agonist structural analog, Leu5.43^{169}, L6.52^{242} and Ser1.32^{4} appear to be involved in ligand binding. These residues were mutated experimentally to test the predictions of the homology model. The *in vitro* studies indicate that Leu5.43^{169} and Ser1.32^{4} mutations cause ten-fold and six-fold decreases in ligand-induced cAMP formation, indicating their importance in ligand-induced activation. Leu6.52^{242} mutations show minimal effect in cAMP production, indicating a lesser involvement in ligand binding. MD simulations of the homology model bound to an agonist indicate that Leu5.43^{169} has an indirect effect in ligand binding, via interactions with Phe6.51^{241}, whereas Leu6.52^{242} is not facing the binding pocket. Ser1.32^{4} seems to interact occasionally with the ligand headgroup.

DEDICATION

I dedicate this thesis to my family Jocelyn Dy Rosales, Jose Rosales, Anthony Dy Rosales, Jonathan Dy Rosales, and Jolynn Salvador for your continuous support of my efforts and dreams no matter where they lead.

ACKNOWLEDGEMENTS

First, I would like to thank my family for all the support they have given me. Thanks to my parents, Jose Rosales and Jocelyn Dy Rosales, for all the encouragement they've given me. Thanks to my brothers, Anthony Dy Rosales and Jonathan Dy Rosales, for listening to my random ramblings about everything even though you guys are all the way in Washington state. Thanks to my cousin, Jolynn Salvador, for showing interest in my project and my endeavors in general.

Secondly, I would like to thank Dr. Evangelia Kotsikorou for the guidance and the patience to have me in her lab for all these years. I would also like to thank Dr. Frank Dean for lending his experience in tissue culture research and help in trouble shooting assays when things were not working.

Finally, I would like to acknowledge some of my friends Naila Bravo, Jenny Bravo, and Dennis Kim, for the moral support of just being around and being there to hang out and goof off with to get away from the craziness of life.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

The G protein-coupled receptor (GPCR) family is the largest class of transmembrane proteins which controls and mediates various signals and cellular responses throughout the body. The general topology of the protein family consists of seven transmembrane α-helices that are arranged into a bundle. The helices are connected by three intracellular loops (ICL) and three extracellular loops (ECL). As the name implies, the receptor couples to a G protein at its intracellular upon activation. This is a heterotrimeric protein consisting of three subunits, Gα, Gβ, and Gγ. A guanosine diphosphate (GDP) molecule is bound to the heterotrimer. When the GPCR gest activated, a conformational change is transmitted to the intracellular end of the receptor where the G protein is bound. The G protein gets activated and a guanosine nucleotide exchange occurs replacing the associated GDP of the heterotrimer with guanosine triphosphate (GTP) causing the Gα to dissociate from the Gβγ structure. Gα functions as a secondary messenger to propagate the signal to alter the activity of adenylyl cyclase, phospholipase C, or RhoGEF depending on the specific type of G protein. [24]

GPCRs are capable of moving between not only active and inactive conformational states, but potentially also metastable intermediate states. Transitioning of the GPCR into the active conformation causes the cyclic nucleotide exchange of the G protein, otherwise known as activation. This process of activation of the receptor consists of a few conserved movement patterns among the GPCR family. One change seen in previous studies is the rotation of the χ_1

dihedral of Trp6.48[238] toggle switch. [5] This change causes the rotation and outward movement of the intracellular end of transmembrane helix (TMH) 6 between 6 Å and 14 Å from the α-helix bundle. This is followed by a slight inward movement of TMH5 an TMH7. Polar groups on the ligand may form hydrogen bonds with residues such as Ser5.42 and Ser5.46 (residues described using the Ballesteros-Weinstein numbering) inducing a contraction of the ligand binding pocket. Ile3.40 tends to move away from Pro5.50 and Phe6.40 to accommodate the outward motion of TMH6. [18]

The classical view for GPCR activation is thought to follow a sequence of steps. First, a ligand would bind into the receptor within the transmembrane ligand binding pocket or at an extracellular domain. The binding to one of these sites would induce a change from the inactive state or intermediate state to the stable active conformation. The structural change caused by the ligand may follow the "sequential mode" allosteric model or the "concerted action" allosteric model. In the sequential mode model, one change to the conformation of a switch would trigger a cascade of conformational changes down the receptor similar to a line of dominos. A missing domino could hinder further signal propagation down the line. The alternative concerted action allosteric model postulates that, instead of having a line of events or sequential flipping of switches, there are a series of microswitches that collectively stabilize the active form of the receptor. This means that even if one "switch" like the conserved amino acid W6.48 toggle switch were to be missing, the protein could still activate if there were other microswitches situated to accommodate the lack of another. [13] Some wild type receptors and mutated receptors activate independent of an agonist, which is known as constitutive activity or the basal activity of the receptor. Due to the various ways that molecules can interact with the receptor and the

different possible pathways which may be initiated, there is a great need to characterize key aspects of receptors that are vital to individual receptor function.

The best way to obtain structural data for proteins in the atomic level is through X-ray crystallography. The first X-ray crystal structure of a GPCR was that of bovine rhodopsin receptor by Palczewski et. al. in 2000. [22] Since then, dozens of GPCR 3D structures have been resolved either as the wild type receptor or more frequently modified whether by point mutations to the sequence, addition of T4 lysozyme in place of an intracellular loop, or other similar modifications to aid in the crystallization of the protein. [5, 7, 11, 15, 16]

Mutations however are not only useful for the crystallization of the receptors; they can also be used to determine various structure-activity relationships (SAR) of the protein. Through *in vitro* testing of the wild type receptor and mutated receptors, amino acids can be evaluated for their involvement in different functions of the receptor such as ligand binding, surface expression, activation, effects in allosteric modulation, and constitutive activity through a variety of assays. [12, 16, 24] Wet bench experimental data can in turn be used to construct or refine computer homology models of a protein of interest. [8, 20]

The structure and function of a receptor can also be probed via *in silico* studies. These computational studies range from performing molecular docking of ligands to exhaustive molecular dynamics (MD) simulations. The 3-dimensional protein structures used in these studies may come from X-ray crystal structures if available or they may be homology models created based on other proteins of a similar sequence whose structure has been resolved. Docking studies hold the potential to identify which amino acids are important to receptor function and what kind of interactions aid or hinder ligand binding. These predictions aid in drug discovery and development. [8, 20] MD simulations provide the ability to observe small time-scale

events to observe movements like side chain and protein movement to ligand binding and unbinding.[3]

There is a limit to how many compounds can be tested experimentally, but through using simulated models and different computational techniques, researchers can focus on what kind of molecules and substituents will work for the targeted receptor. Reciprocally, experimental results can help to further refine homology models of proteins guide the computational studies leading to more predictive results. Through combination of both types of studies, a better understanding of the protein can be developed as demonstrated in the presented study.

CHAPTER II

LITERATURE REVIEW

2.1 The Significance of the GPR119 Receptor

Originally discovered in 2003 as an orphan receptor, GPR119 belongs to the class A Rhodopsin-like GPCR protein family. [9, 12] GPR119 displays a wide range of possible signaling pathways and is capable of binding to both the Gαs, Gαi, and Gαi protein subunits. One potential signaling pathway of GPR119 is the calcium release channel through Gq protein activity. Another pathway is the cyclic adenosine monophosphate (cAMP) response element (CRE) pathway. [12] In response to receptor activation, the Gαs subunit decouples from the remaining subunits of the G protein heterotrimer, exchanges the bound guanosine diphosphate (GDP) with a guanosine triphosphate (GTP), binds to adenylyl cyclase and induces the conversion of adenosine triphosphate (ATP) into cAMP. Although not as prevalent due to a higher ninding affinity of Gαs, last pathway induced by Gαi activity causes an inhibitory effect on adenylyl cyclase activity. [24] Study performed here focuses on CRE pathway.

GPR119 has been observed to be highly expressed in beta cells of pancreatic islets. For intestinal enteroendocrine cells like those to which GPR119 is endogenous to, a rise in cAMP levels is associated with a rise in hormone secretion. In the specific case of the beta cells of pancreatic islets, insulin secretion increases with the rise in cAMP levels. Due to its potential to increase insulin secretion, GPR119 has been a popular drug target for the pharmaceutical

industry. Targeting this receptor can help develop novel alternative compounds to treat diabetic patients. [8]

2.2 GPR119 Ligands

In 2006 the endogenous ligand for the GPR119 receptor, which is the naturally occurring ligand within the body, was determined to be oleoylethanolamide (OEA) shown below in figure 1. [21] Other lipid metabolites such as 2-monoacylglyerol that are byproducts of triglyceride metabolism via pancreatic lipases have also been found to induce increased GPR119 activity, however, OEA was the most efficacious of those endogenous lipids identified thus far.

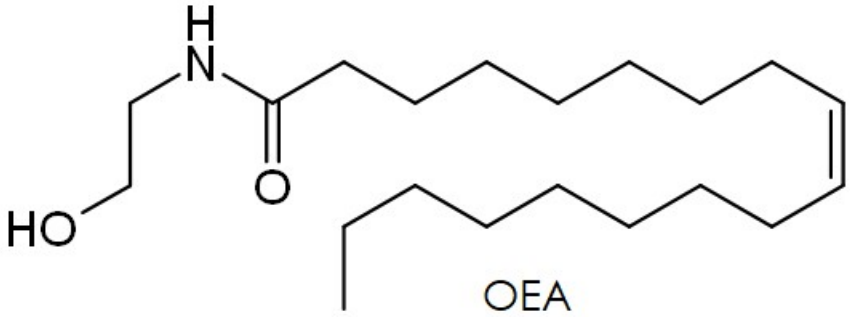

Figure 1: Structure of Oleoylethanolamide (OEA)

In attempts by groups like Arena Pharmaceuticals and Prosidion Limited, many novel synthetic agonists for GPR119 have been developed. Most synthetic agonists which have high agonistic activity contain a heteroaryl moiety serving as the head group with hydrogen bonding capabilities. There is also typically a piperidine ring substituted with another hydrogen bonding group. These two groups are linked by a spacer meant to ensure optimal orientation within the ligand binding pocket of GPR119. Many developed agonists by Arena Pharmaceuticals employs

a pyrimidine spacer while Prosidion Limited uses alkoxy or hetaryloxy spacers. [23] Additionally, many agonists based of the Arena or Prosidion pharmacophore are long flexible structures.

In 2007, AR231453 was developed by Arena Pharmaceuticals. This was the first synthetic nanomolar full agonist of the protein (Figure 2). A full agonist is a compound that induces activation of the GPR119 receptor and initiates the CRE signal pathway that produces insulin in enteroendocrine cells. This ligand has also been observed to function as a positive allosteric modulator for OEA. This means when present, AR231453 increases the potency of OEA. [12, 20]

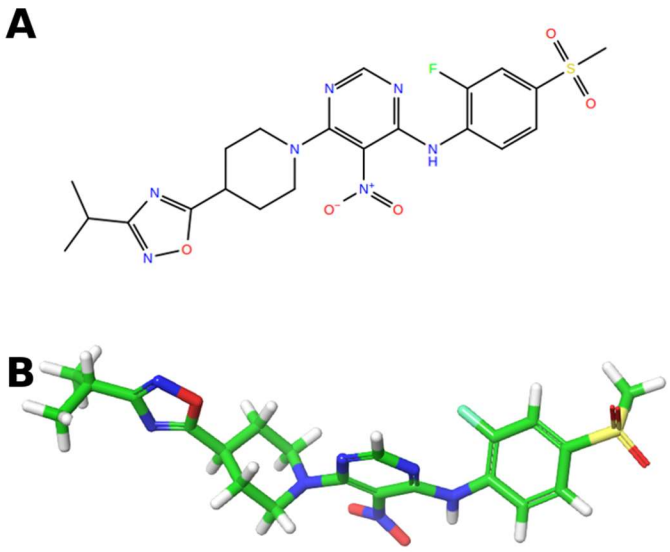

Figure 2: Structure of AR231453. A) 2D structure and B) global minimum structure rendered in tube. Carbon, nitrogen, oxygen and sulfur atoms are displayed in green, blue, red and yellow, respectively.

In 2015, Engelstoft et. al. discovered the inverse agonist, AR437948, which is nearly structurally identical to the agonist AR437735 as can be seen below in figure 3 and 4. The only

difference between the two compounds is the position of an oxygen and a sulfur atom which results in a thioester bound to an isopropyl group in the agonist. In contrast, the inverse agonist contains a thionoester bound to an isopropyl. Despite the structure of the ligands being nearly identical, they have opposite effects on drug-induced cAMP production. An inverse agonist would decrease receptor activity below that of the basal level. It is still currently unknown as to why this change would have opposite effects. [20]

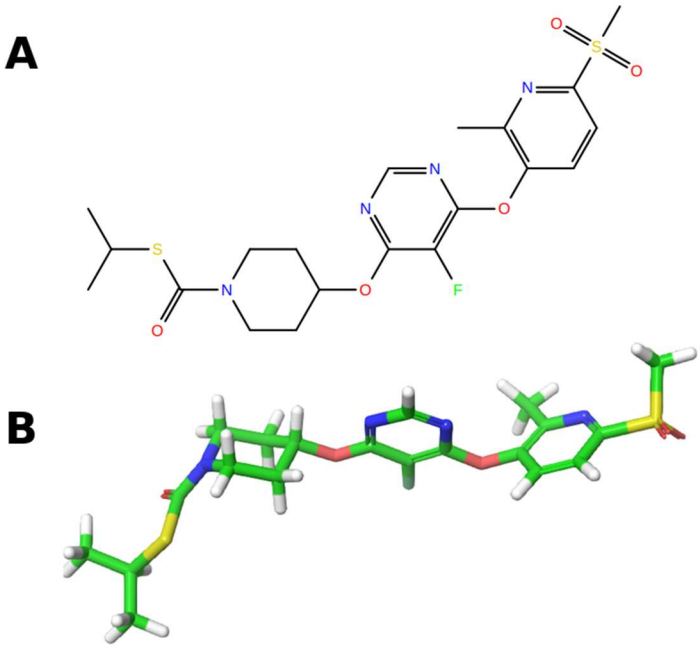

Figure 3: Structure of AR437735. A) 2D structure and B) global minimum structure rendered in tube. Carbon, nitrogen, oxygen and sulfur atoms are displayed in green, blue, red and yellow, respectively.

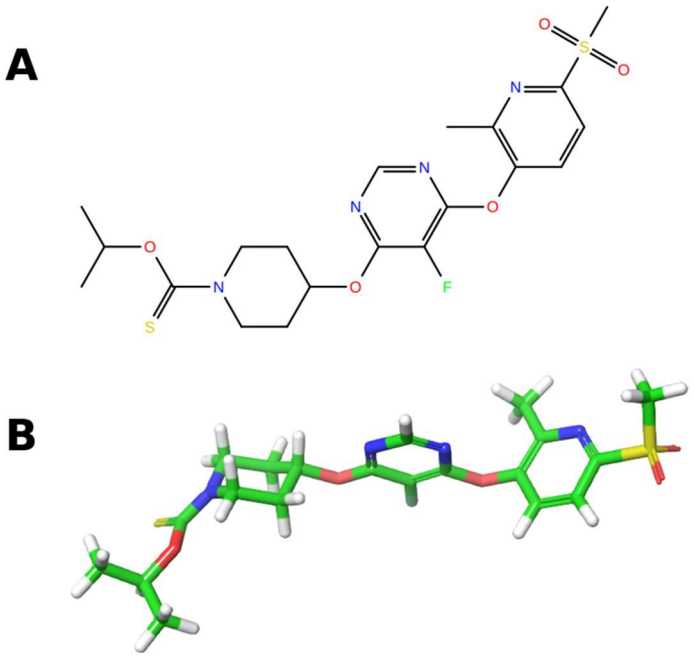

Figure 4: Structure of AR437948. A) 2D structure and B) global minimum structure rendered in tube. Carbon, nitrogen, oxygen and sulfur atoms are displayed in green, blue, red and yellow, respectively.

2.3 Published GPR119 *In Vitro* and *In Silico* Studies

In studies performed by Engelstoft et al in 2014 and 2015, a combination of *in vitro* site directed mutagenesis and *in silico* docking studies were used to aid in determining the role of various amino acids which lined the ligand binding pocket which consists of amino acids in the helical bundle and the extracellular loops.

In 2014, Engelstoft et al tested 30 mutants on 23 residues positions seeking to identify their role in constitutive activity and ligand-induced activity of GPR119. Most of these mutations were located on TMHs 2, 3, 5, 6, 7 but also included some amino acids located on ECL2. The

mutagenesis study reduced the size of the side chains employing alanine mutations or created steric hindrance using a larger amino acid or branched side chain. It was determined that two phenylalanine residues located on ECL2 past Cys 155 which is involved in a disulfide bridge with Cys 2.38[78], have a role in constitutive activity. Ligand binding pocket amino acids found to be involved in basal activity of GRP119 included residues such as Trp6.48[238] believed to function as a microswitch, Arg3.28[81], Met3.29[82], Ser5.44[170], Phe6.51[241], Glu7.35[261] ,Arg7.36[262]. Some of these mentioned amino acids also have an effect on ligand-induced activation such as Phe157 and Phe158 which are located on ECL2 and Trp6.48[238], Trp7.39[265] and Gly7.42[268] located in the binding pocket. It is also determined that for AR231453, the sulfonyl probably binds pointing towards the extracellular aqueous environment based on better docking scores and a larger population of docks converging on this conformation. [8]

The results published in 2015 by the Norn et al were geared toward defining a better way of developing a GPR119 homology model unbiasedly. By using data from the previously mentioned study by Engelstoft et al and more than a handful of point mutations of other residues including Gln252, Glu253, His255, Leu256, and Leu258 from ECL3, a mutationally guided refinement of the homology model was performed. This refined model demonstrated the ability to better identify known GPR119 agonists and sort out decoys compared to a model based solely on other crystal structures. This shows how site directed mutagenesis can be used to enhance the accuracy of homology models in predicting experimental results. [20]

CHAPTER III

METHODOLOGY

3.1 Maintaining HEK293 Cell Line

The human embryonic kidney (HEK293) cell line was being used as a surrogate eukaryotic cell to test the activity of the GPR119 wild type and mutant receptors.

HEK293 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) with the following additives: 10% fetal bovine serum (FBS) (for growth of cells), 1% penicillin/streptomycin (pen/strep)(mitigates unwanted growth of bacteria), 1% antibiotic-antimycotic (anti-anti) (mitigates unwanted growth of fungus). The cells were grown inside of an incubator set to 5% CO_2, humidified at 37°C. Once the plates reach approximately 90% confluency, the cells were passaged onto a new plate. To passage the cells, DMEM with the previously stated additives were aspirated from the 10-cm plate. 5 ml of trypsin EDTA was added to the plate to pick the cells up off of the plate. After removing all of the cells from the plate and pipetting the cells up and down 15 times to break up clumps of cells, 1 ml of the trypsin EDTA solution containing suspended cells was transferred to a new 10-cm plate with 10 ml of DMEM (10% FBS, 1% pen/strep, 1% anti-anti). The plate was then gently rocked back and forth to disperse the cells homogenously throughout the plate.

3.2 Preparing the HEK293 Cells for Transfection

Media from a 10-cm plate of cells was aspirated and 5 mL of trypsin EDTA was added to the 10-cm plate. Once the cells were suspended in the trypsin solution, the solution was then transferred to a 15 mL centrifuge tube. 7.5 ml of DMEM (10% FBS) was added to the 15 ml tube to prevent further activity of trypsin on HEK293 cells. The cells were then centrifuged at 1000 RPM at room temperature for 5 minutes to form a pellet at the bottom of the tube. The supernatant was aspirated and 10 ml of DMEM (10% FBS) was added to the 15-ml centrifuge tube to resuspend the cells. A hemacytometer was used to determine the concentration of the resuspended cells. The cells were then plated into three sets of 6-well plates with an approximate density of 500,000 cells/well. The 6-well plates were then gently rocked to disperse the cells uniformly and placed back into the incubator. The cells were allowed to grow for 2 days before transfection.

3.3 Transfection of HEK293 Cells with Wild Type and Mutant GPR119 DNA

As per the protocol of the Invitrogen Lipofectamine^{TM} transfection kit, 3.75 µL of lipofectamine 3000 reagent, 5 µL of lipofectamine 3000, and 2.5 ng of DNA of the wild type or mutant GPR119 receptor were used to transfect the cells. Seven wells were transfected for the wild type GPR119 receptor DNA and seven other wells were transfected with the mutant GPR119 receptor DNA. Two non-transfected wells were used as controls. 4 hours post-transfection, the media in each of the wells were aspirated and replaced with DMEM (10% charcoal stripped fetal bovine serum, CSFBS). The change to CSFBS at this point is performed because the GPR119 receptor's endogenous ligand is oleoylethanolamide (OEA), which is an

endogenous lipid. As such, FBS contains lipids and therefore contains the potential ligands for binding to the GPR119 receptor and may produce artificially increased levels of cyclic adenosine monophosphate (cAMP). In contrast, CSFBS has been passed through a column containing charcoal which removes hydrophobic compounds such as lipids.

3.4 Application of AR231453 Agonist to GPR119 Transfected HEK293 Cells

The GPR119 agonist AR231453 (Molecular Weight: 505.522 g/mol) was dissolved in 100% dimethyl sulfoxide (DMSO) to make a 10 mM stock solution. The AR231453 was serially diluted into the following concentrations: 10000.0 nM, 1000.0 nM, 100.0 nM, 10.0 nM, 1.0 nM, 0.1 nM, and 0.01 nM. Each dilution was prepared in DMEM, 2.5 % CSFBS and 0.2% DMSO. Prior to addition of the AR231453 agonist, the medium was aspirated and each well was washed with 1.0 mL of Hank's balanced salt solution (HBSS). The HBSS was then aspirated, 1.0 mL of DMEM (no additives) was added to each well, and 1 mL of the diluted drug was added to each respective well. The cells were allowed to incubate with the agonist at 37 °C in the incubator for 30 minutes.

3.5 cAMP ELISA Colorimetric Assay

The cells were gently scraped from the wells at room temperature and transferred to 15-mL centrifuge tubes. The samples were then centrifuged at 1000 RPM and 21°C for 5 minutes. The supernatant was aspirated from each centrifuge tube and the pellet was resuspended in 2.0 mL of phosphate buffered saline without calcium or magnesium ions (PBS). The tubes were then

centrifuged again at 1000 RPM and 21°C for 5 minutes and the PBS was then aspirated from each tube. After adding 286 µL of cold lysis buffer (thimerosal 0.01%, Triton X-100 2.0%) to each sample tube, the samples were placed on ice. Each sample was then transferred to 1.5-mL microfuge tubes and placed back on ice. The pellets were resuspended in the lysis buffer solution and subjected to two cycles of freeze and thaw using a slurry of dry ice ethanol and a 37°C water bath. Each step of the freeze and thaw lasted 3 minutes. The cells were then centrifuged at 13,000 RPM at 4°C for 10 minutes. 200 µL of the supernatant was then transferred over to a new 1.5-mL microfuge tube. From this point forward, all work was done with the cell lysates on ice to prevent degradation of cAMP within the lysates. The cell lysate sample was then mixed with a pipet and 100 µL was transferred to a third 1.5-mL microfuge tube. 50 µL of peroxidase cAMP tracer conjugate, diluted to 1:100 using assay diluent (Thimerosal 0.01%), was added to each of the samples. cAMP standards as described in the manufacturer's protocols were prepared and added in duplicate to the goat anti-rabbit antibody coated 96-well test plate. 25 µL of peroxidase cAMP tracer conjugate was added to each well containing cAMP standards. Each of the samples containing the cell lysates and peroxidase cAMP tracer conjugate was added in duplicate to the goat anti-rabbit antibody coated plate. 50 µL of rabbit anti-cAMP polyclonal antibody, prepared as directed in the manufacturer's manual, was added to each of the tested wells. The goat anti-rabbit antibody coated plate was allowed to sit on an orbital shaker for 2 hours. The contents of the wells were aspirated and washed six times with 200 µL of wash buffer (Thimerosal 0.02%). 100 µL of the substrate solution was added to each tested well and allowed to sit on the orbital shaker for 15 minutes. 100 µL of the stop solution (0.5 N sulfuric acid) was added to each of the wells. The goat anti-rabbit antibody coated plate was then read using the BioRad 480 micro-plate reader at a primary wavelength of 450 nm.

3.6 Analysis of cAMP ELISA Colorimetric Assay

Using Microsoft excel, an absorbance vs [cAMP] for the cAMP standards. The absorbance of each sample was then compared to the plot of the cAMP standards to determine the concentration of cAMP present in each lysate. SigmaPlot 11 (Systat Software Inc., San Jose, CA) was used to graph the concentration of cAMP in the lysate samples in nM vs log concentration of the AR231453 agonist. Based on the cAMP levels found in the controls due to an internal cAMP reservoir, the levels of cAMP of the samples have been adjusted to remove cAMP activity which was not induced by the compound. The sigmoidal dose dependent curve was set to produce a hillslope of 1.0. The log EC_{50} reported by SigmaPlot 11 are then converted to values for the concentration of cAMP in nM.

3.7 Conformational Search of AR437735 and AR437948

A conformational search was run for all ligands involved in the computational studies to identify the global energy minimum conformation. AR437948 and AR437735 were built using Maestro (Schrodinger modeling software). The rotatable bonds of each molecule were identified. Each bond was rotated based on the hybridization of the two atoms involved. SP^3-SP^3 bonds were rotated 3-fold, SP^3-SP^2 were rotated 6-fold, and SP^2-SP^2 bonds were rotated 4-fold. Following each rotation, the resulting structures were minimized using molecular mechanics (Macromodel, Schrodinger modeling software) and compared to all the other conformers to remove duplicate conformations. The unique structures were used as input for the next rotational bond search. Once all 8 rotatable bonds each molecule had, had been explored, 163 unique

conformations had been observed and 14 conformations with the lowest energy were determined to be representative of possible conformations.

3.8 Optimization of AR437735 and AR437948

All 14 representative conformers of AR437948 were optimized using the Jaguar module of Schrodinger modeling software. Hartree-Fock theory was used employing 6-31G* basis set. Of the 14 representative conformers, 12 of the conformers resulted in a shape which remained within 10 kcal/mol of the global minimum conformer and considered for further study. The same was done for the agonist (AR437735) conformations.

3.9 Docking of AR437948 into the GPR119 Inactive Homology Model

An in-house developed GPR119 inactive homology model, which was based on the A2A and CB1 receptors was used to dock the inverse agonist. A docking grid was created using Glide Grid Generation function where the center of the grid was selected to be in the center of the binding pocket. Specifically, the center was based on the center of mass of the amino acids Arg3.28[81], Cys255 (ECL3), Phe157 (ECL2), Trp6.48[238], Arg7.36[262], and Trp7.39[265]. All 12 conformations of AR437948 were docked both flexibly and rigidly into the ligand binding domain of GPR119 using Glide (Schrodinger modeling software).

3.10 Minimization of GPR119-AR437948 Complex

Since GPR119 is a transmembrane protein, it must be minimized in a way that accounts for both the aqueous environment of the cytosolic region and the hydrophobic environment of

the lipid bilayer. All charged residues were mutated to the uncharged form with exception of residues which were already involved in a salt bridge or hydrogen bond to another amino acid. The ligand-receptor complex was then minimized using the OPLS 2005 force field, distance dependent dielectric constant of 2.0, and extended interaction cutoff distances to emulate the hydrophobic lipid surroundings. The backbone of the protein was constrained with 1000.0 kcal/mol force, which was gradually reduced to no constraint force over the course of the minimization. Following the minimization of the hydrophobic region of the receptor, the uncharged residues were mutated back to the original residue and the transmembrane region of the protein was frozen. The protein was then minimized using an implicit water solvent to allow the loops to relax to a lower energy state.

3.11 Docking Analysis

Minimized ligand-receptor complexes which were considered for further analysis met the following criteria. First, the position AR437948 must be able to block both the rotation of the chi 1 (χ_1) and the chi 2 (χ_2) torsional angles of the tryptophan toggle switch, Trp 6.48^{238}. Secondly the SO_2 moiety of the ligand must point towards the extracellular end of the receptor.

Using Schrodinger2014, the interaction energies between the ligand and the surrounding amino acids were calculated using a distance dielectric value of 3 and extended cutoff interaction distances. The ligand was extracted from the inactive GPR119 binding pocket, its dihedrals were frozen and its atoms were allowed to optimize using the same parameters as discussed in section 3.8. This was done to calculate the energy of the ligand in the docked conformation.

The agonist, AR437735, was then superimposed with the same dihedral conformation as the docked AR437948. Steric clashes made by the thioester group as opposed to the thionoester group of the inverse agonist were noted.

3.12 Docking of AR231453 into the Active GPR119 Receptor

The active state homology model of the GPR119 receptor was prepared using the Schrodinger modeling suite. The agonist AR231453 was flexibly docked into the active GPR119 receptor using the Induced Fit docking protocol to allow flexibility for both the ligand and the protein. This method is better for obtaining possible bound conformation of the ligand. The van der Waals scaling for both the ligand and the protein atoms was the default value of 0.50. The preliminary round of docking with Glide is followed by a Prime Refinement step which optimizes the side chains within 5 Å of the ligand to better fit the docked molecule. The backbone of the protein was excluded from the prime refinement step to prevent alterations to the backbone of the helices. The molecule is then re-docked into the optimized binding pocket of the receptor. This procedure resulted in eight possible docked conformations. The ligand/receptor complex with the ligand in proximity to Trp6.48[238] was chosen to be used for further analysis.

3.13 Preparation of Active GPR119 Simulations in VMD193

Using the visualization software VMD193, the active homology model of the GPR119 receptor, the AR231453 ligand, and conserved waters of the hydrophobic core of the receptor were merged and the related PSF and PDB files were created. The disulfide bridges between

Cys3.25[78] and Cys155 (ECL2) and between Cys6.61[251] and Cys254 (ECL3) that disappear when the protein file is prepared were reformed using the DISU patch. The combined GPR119 structure was then imbedded into an 85 Å x 85 Å 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) hydrated lipid bilayer. Overlapping lipids were removed and TIP3 waters were added to the top and bottom of the simulation cell so that there was an additional 15 Å above and below the protein. Overlapping waters in the system were removed. The system was then ionized to a 0.15 M NaCl concentration to closer match cytosolic NaCl concentrations.

3.14 Protonation of 2 Amino Acids

Previous studies indicate that for the cannabinoid type 2 (CB2) receptor, the Asp3.49 and Asp6.30 are believed to be protonated during activation. The equivalent Asp3.49[102] and Asp6.49[220] were mutated to their protonated form for the active GPR119 receptor using the ASPP patch through VMD193. Topology parameters for the charged forms are contained within the charmm36 "par_all36_prot.prm" from the MacKerell group from the University of Maryland (http://mackerell.umaryland.edu/charmm_ff.shtml).

3.15 Equilibration of Molecular Dynamics Simulation

For all simulations periodic boundary conditions were used to simulate an infinitely large membrane without the presence of a vacuum at the edge of the simulation cell. Topology and parameters for proteins, lipids, carbohydrates, sodium, water, and ions were obtained from the charmm36 topology and parameter files found in the MacKerrel group website (http://mackerell.umaryland.edu/charmm_ff.shtml). The general "cgenff" parameters were also used in conjunction with an in-house developed topology and parameter file to describe the

ligand AR231453. The cutoff distance was set to 12.0 Å with a switch distance of 10.0 Å. The particle mesh Ewald grid was set to a dimension of 90 Å x 90 Å x 108 Å. The SHAKE algorithm to allow for 2.0 fs timesteps.

The simulation cell was minimized for 2000 steps with the receptor/ligand complex and lipid head groups frozen. The minimization was followed by a 0.5 ns equilibration at 310 K under constant moles, volume, and temperature conditions (NVT) using a 1 fs timestep. The lipid headgroups and protein were frozen in place to allow for the waters and the hydrophobic lipid tails to adjust. The Langevin damping coefficient was set to 1 to maintain a constant temperature.

For the remainder of the simulations, a 2.0 fs timestep was used. For a gradual equilibration of the lipids and protein, the constraint force imposed upon them was slowly released over a period of 0.5 ns under constant moles, pressure, area, and temperature (NPAT) conditions. Pressure of the system was set to 1.0 atm and the constant area condition allows for maintaining a constant area of the lipid membrane to prevent unrealistic packing or dispersing. The Langevin piston was used to maintain pressure.

Final equilibration was continued using the collective variable hydrogen bond on Ser6.43^{233} and Ser6.47^{237} to maintain a hydrogen bond to back bone of Val6.39^{229} and Ser6.43^{233} respectively to prevent backbone hydrogen bonding to TMH7 which distorts the helix. The simulation is continued using NPAT conditions for 4 ns until the RMSD was constant and system considered to be equilibrated for production runs.

3.16 Voronoi Tessellation Monte Carlo Integration to Calculate Area per Lipid

By using a Voronoi Tessellation Monte Carlo (VTMC) Integration program given by Mori et al, the average area per lipid can be observed. A configuration file was made where the input file was set to the pdb file of the simulation cell and parameters such as the cell size, cell center, radius of the pseudo-probe were set. The system is then divided into a 2D representation of the top half of the protein and the bottom half. The program then uses pseudo-probes of the designated size and randomly probes sections of both halves to approximate the average area occupied per lipid head group. To monitor that the area per lipid maintains the experimental determined value, VTMC is employed on the first step of the minimization equilibration process and at the end of the MD simulations. [19]

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Results of Docking Study

Rigid docks were performed using all 12 inverse agonist conformers which yielded 59 poses. These docks were not used for further analysis because in many cases, the ligand was positioned too high that it was partially outside the extracellular end of the receptor or buried too deeply, such that it extended below the tryptophan Trp6.48[238] which is considered to act as a toggle switch during the activation of the receptor. In a few cases, the ligand did dock near or at the desired area in the putative receptor binding pocket, however it had in some cases minor and in some cases major steric clashes with residues in the binding pocket.

Docking of a flexible AR437948 molecule into a rigid receptor produced 48 poses. Out of those 48, only 20 poses had the electron dense sulfonyl group pointed toward the extracellular end of the receptor and the ligand was close enough to Trp6.48[238] that it could block its $\chi 1$ and/or $\chi 2$ dihedral angles from rotating since a change in the conformation (potentially rotation of the $\chi 1$ dihedral angle from g+ to trans or rotation of the $\chi 2$ dihedral angle) has been postulated as one of the steps transitioning the receptor from the inactive to the active conformation. From the 20, six poses were positioned to block both $\chi 1$ and $\chi 2$ rotations of the tryptophan toggle switch. The remaining poses blocked the rotations of either $\chi 1$ or $\chi 2$ but not both. The rotation of Trp6.48[238] was either blocked by the isopropyl group of AR437948 or the thionoester sulfur

atom. After the ligand-receptor complex was minimized, only one of the docks out of the six did the ligand maintain its position near the toggle switch in a way that it still blocked both $\chi 1$ and $\chi 2$ rotations as seen in figure 4. The ligand in that dock blocked the movement of Trp6.48[238] with its sulfur atom. Additionally, this dock had two edge-to-face aromatic stacking interactions with Trp7.39[265] and Phe1.35[7]. The pyridine ring of AR437948 and Trp7.39[265] stacking interaction has a centroid to centroid distance of 5.2 Å at an angle of 68.2°. The π stacking interaction between the pyridine ring of AR437948 and Phe1.35[7] occurs at an angle of 85.7° and a centroid to centroid distance of 4.9 Å. The ligand also forms a hydrogen bond between one of its sulfonyl oxygen atoms and Arg7.36[262]. The sulfonyl oxygen acts as a hydrogen bond acceptor and has a hydrogen bond length of 1.73 Å with a bond angle of 118.85°.

Because sulfur atoms have a larger ionic radius than oxygen, the carbon-sulfur bond of the thioester in the agonist extends the isopropyl group 0.31 Å farther than the isopropyl group of the inverse agonist. In AR437735, the carbon-oxygen bond length of the carbonyl group is 1.2 Å. In contrast, the AR437948 ligand has a bond length of 1.68 Å for the corresponding carbon-sulfur double bond. Additionally, the carbon-sulfur-carbon angle in the agonist is 100.6° while the carbon-oxygen-carbon angle of the inverse agonist is 125.1°. As a result of the different bond lengths and angles, the distance from the leucine residues, Leu5.43[169] and Leu6.52[242], to the nearest carbon of the isopropyl is 4.20 Å and 3.65 Å respectively for the inverse agonist. On the other hand, the closest carbon of the isopropyl on the agonist to Leu5.43[169] is 2.86 Å and 2.25 Å to Leu6.52[242].

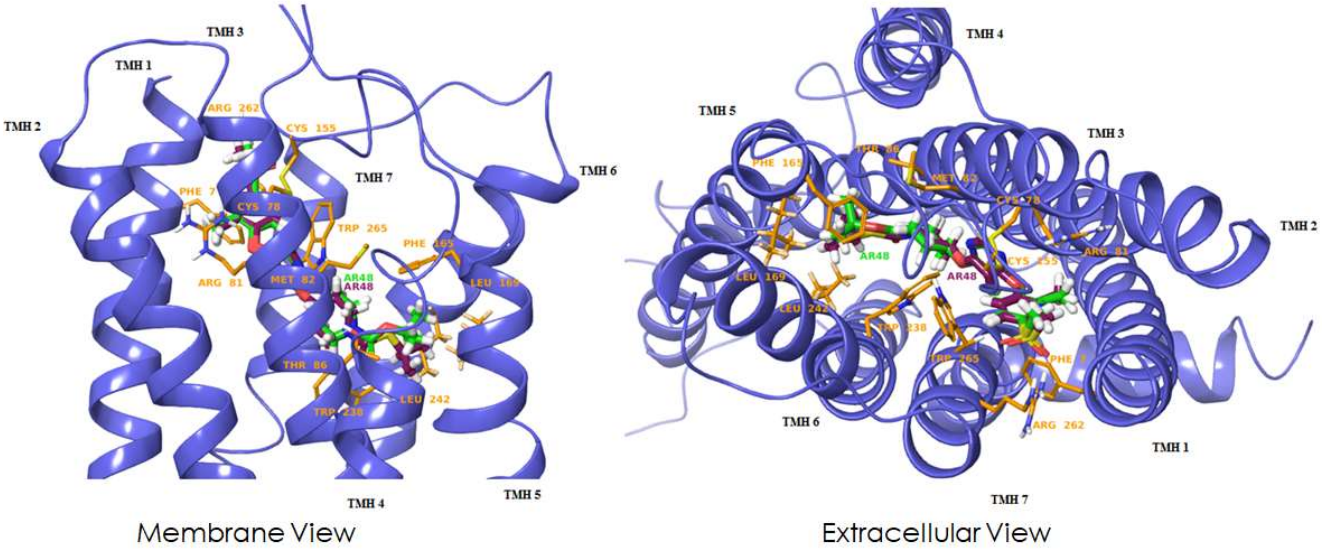

Figure 5: Superimposition of AR437735 on AR437948 within Inactive GPR119 Model. AR437735 (agonist) is colored burgundy and AR437948 (inverse agonist) is colored green.

The inverse agonist docked into the inactive receptor in position where the sulfur of the thionoester blocks the $\chi 1$ and the $\chi 2$ rotations of the Trp6.48^{238}. When the agonist was superimposed upon AR437948 and though they may fit in a similar fashion in the binding pocket, AR437735 failed to block the $\chi 2$ rotations of the toggle switch and had minor to major steric overlaps with Leu5.43^{169} and Leu6.52^{242} which can be seen in figure 6 indicated by the dashed red and orange lines. The conformational cost of the inverse agonist ligand was calculated to be 4.003 kcal/mol compared to the global minimum energy conformer. The corresponding agonist structure, which was constrained to have the same dihedral values as the inverse agonist, had a 4.024 kcal/mol conformational cost compared to its global minimum energy structure. Intuitively, besides the moieties connected to the nitrogen of the piperidine ring, it would seem that the agonist and inverse agonist would bind identically inside the GPR119 receptor.

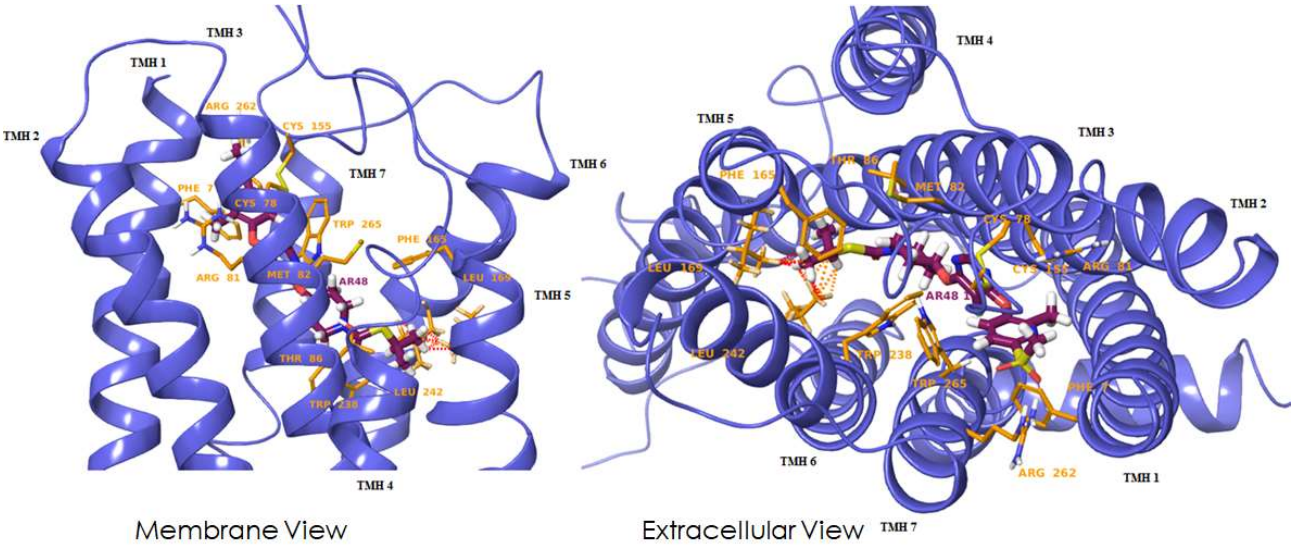

Figure 6: Steric Clashes of AR437735 With L5.43[169] and L6.52[242]. Steric overlap indicated by the orange and red dashed lines.

As can be seen below in table 1, Trp7.39[265], Arg7.36[262], Arg3.28[81], and Met3.29[82] are the amino acids that have the highest energy of interaction with the ligand. These strong interactions suggest that these amino acids may be important for ligand binding. This hypothesis could be tested by mutating these amino acids to see if these mutations would affect the way the ligand binds in the pocket and its ability to induce activation of the receptor.

Table 1: Complete Energy Per Residue Breakdown of Amino Acids Involved in AR437948-GPR119 (Inactive) Complex. Amino acids with only the name and the absolute sequence number followed by a single asterisk belongs to ECL1 or ECL2 if followed by a double asterisk

Amino Acids	VDW(KCal/mol)	Elec(Kcal/mol)	Total(Kcal/mol)
Trp7.39[265]	-7.321	-0.553	-7.874
Arg7.36[262]	-1.606	-4.949	-6.555
Arg3.28[81]	-6.339	0.060	-6.279
Met3.29[82]	-4.990	-0.215	-5.206
Thr3.33[86]	-4.368	-0.149	-4.517
Phe1.35[7]	-3.737	-0.353	-4.091
Cys155**	-2.925	-0.203	-3.128
Phe5.39[165]	-2.921	-0.128	-3.049
Cys3.25[78]	-2.697	-0.315	-3.012
Val3.32[85]	-2.981	0.201	-2.780
Phe6.51[241]	-2.590	-0.012	-2.602
Ser1.32[4]	-1.983	-0.514	-2.497
Trp6.48[238]	-2.327	0.098	-2.229
Leu5.43[169]	-2.249	0.061	-2.189
Leu6.52[242]	-2.139	-0.008	-2.147
Gln154**	-1.884	-0.119	-2.003
Leu1.39[11]	-1.721	0.019	-1.702
Val5.46[172]	-1.541	0.041	-1.500
Ile2.57[58]	-1.369	-0.051	-1.420
Thr5.42[168]	-1.397	0.112	-1.285
Ala3.36[89]	-1.052	-0.108	-1.159
Leu7.32[258]	-1.106	-0.003	-1.109
Phe158**	-0.880	-0.088	-0.968
Val7.43[269]	-0.791	-0.062	-0.853
Leu2.61[62]	-0.758	0.037	-0.721
Gln2.64[65]	-0.410	0.027	-0.382
Gly1.36[8]	-0.342	0.001	-0.341
Ala3.30[83]	-0.311	-0.004	-0.315
Ser1.31[3]	-0.282	-0.030	-0.311
Ser3.26[79]	-0.165	-0.018	-0.183
Ile2.53[54]	-0.110	-0.013	-0.123
Gln74*	-0.197	0.082	-0.115
Gly6.55[245]	-0.100	0.028	-0.073
Total	-65.587	-7.127	-72.715

4.2 Results of ELISA cAMP Assays

In eukaryotic cells, cyclic nucleotide monophosphates may become a substrate for the cyclic nucleotide phosphodiesterase. For assays such as the ones performed in this study, this poses a problem because the cAMP which is being measured is being converted over time into AMP. Cyclic nucleotide phosphodiesterase activity is positively corelated with the temperature of its environment. [25] To mitigate the loss of cAMP while enhancing the yield of cAMP from the lysates, two precautions were taken. First, samples were placed on ice as detailed in sections 3.5. Secondly, the temperature of the water bath used to thaw lysates was optimized. This involved observing the benefit of lowering the thawing temperature during the freeze and thaw step at the cost of time. A 37°C water bath was found to yield the most cAMP for the assays performed in this study.

By employing site-directed mutagenesis to residues of interest, we investigated the potential role Leu5.43[169], Leu6.52[242], and Ser1.32[4] may play in ligand activation, in this case AR231453-induced cAMP production. When compared to the wild type, shifts of the dose response curve for the mutant protein compared to the wild-type indicate the potential involvement the amino acid has in receptor activation, where the larger the shift of the EC_{50}, the larger the role the amino acid may be playing. To help understand what kind of interactions the residues of interest may participate in, they were mutated according to the potential interactions they may have with neighboring residues. The two leucine residues lack the ability to hydrogen bond, participate in polar interactions, and do not have aromatic rings for pi cloud interactions. As such, these residues were mutated to the larger nonpolar methionine or to the smaller nonpolar alanine residue to alter potential steric interactions. Since both Leu5.43 and Leu6.52 are located on α-helices deep within the transmembrane region of GRP119, evidence of involvement

would indicate that such residues would be oriented towards the interior of the GPR119 ligand binding pocket, facing away from the lipid bilayer. For both leucine residues, if there is a lack of a shift in the EC_{50}, this would imply these residues do not interact with the ligand, participate in activation, and instead would potentially be positioned on the outside of the transmembrane region making contacts with neighboring lipids.

The L5.43^{169}A mutation shifts the response of the GPR119 receptor to the agonist, AR231453, to a higher EC_{50} concentration, indicating a decreased response to the ligand (Figure 7). This mutation causes an approximate 9-fold decrease in ligand-induced activity. This change of the EC_{50} suggests that the decreased size of the non-polar side chain has moderate deleterious effects on the ability of the receptor to interact with this ligand.

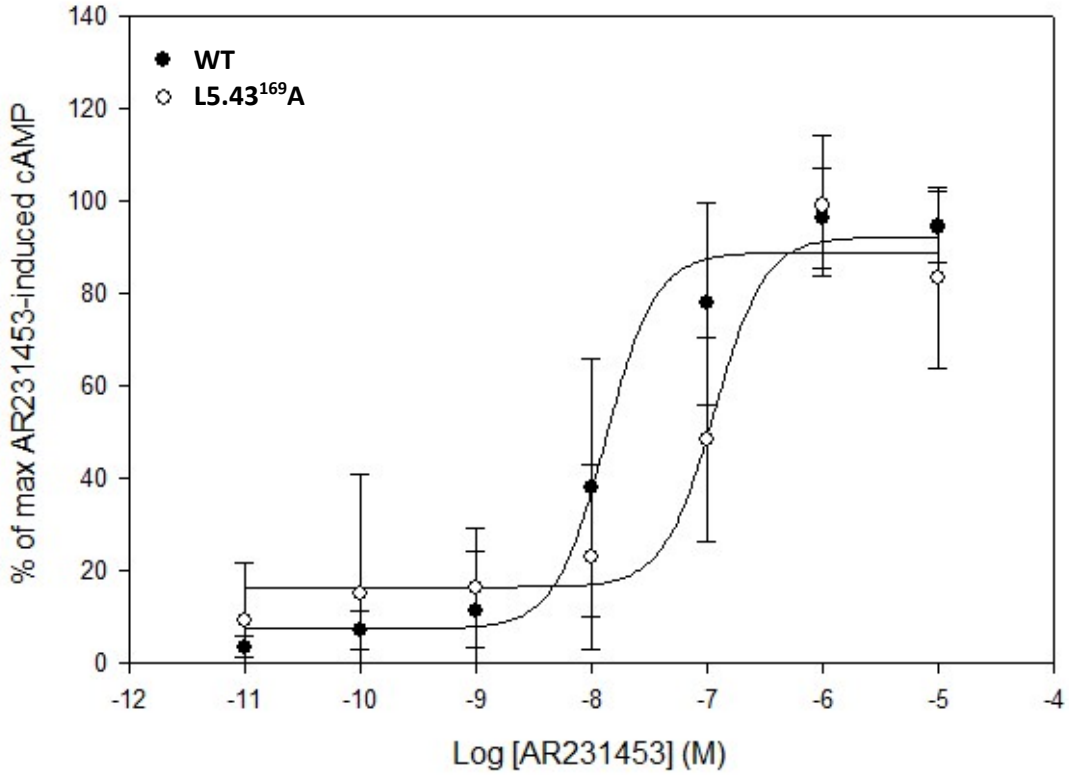

Figure 7: Dose Response Curves of the Wild Type L5.43[169] and the Mutant L5.43[169]A, Displaying Levels of AR231453-induced cAMP. Background cAMP levels removed and scaled to maximum AR231453-induced cAMP production.

The mutation of Leu5.43 to methionine increases the size of the non-polar side chain and increases non-polar interactions (Figure 8). Similar to the results with L5.43[169]A, L5.43[169]M also shifts the AR231453 dose response curve to yield a higher EC_{50} value. This mutation of a leucine to a methionine causes a 10-fold decrease in activity in response to the agonist.

Both modifications of L5.43[169] have an impact whether they resulted in a larger side chain, for methionine, or a smaller R-group, for alanine. This implies that Leu5.43[169] plays a specific role in ligand-induced activation and that the specific side chain interactions are

important. As previously mentioned, for Leu5.43^{169} to have effects within the transmembrane region of GPR119 in reference to AR231453-induced activation, our results indicate that Leu5.43^{169} is positioned towards the interior of the receptor ligand binding cavity, either to assist with ligand binding or to interact with other amino acids involved in propagating conformational changes associated with activation towards the intracellular G protein.

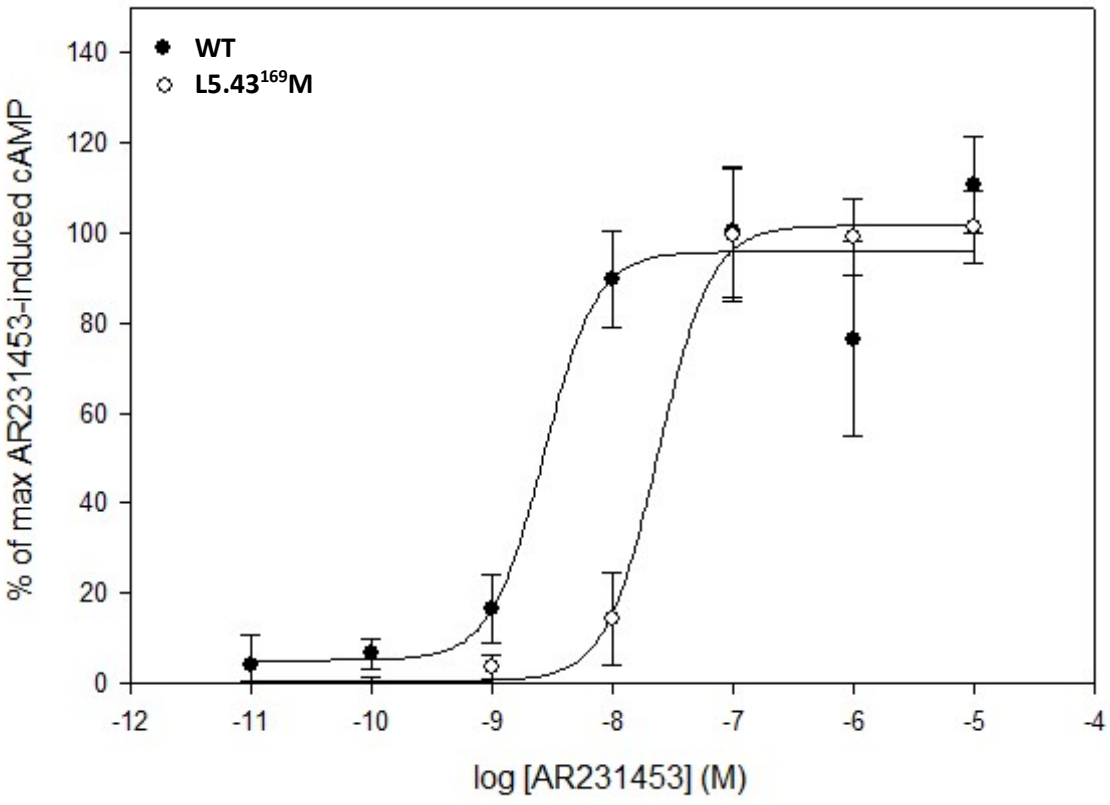

Figure 8: Dose Response Curve of the Wild Type L5.43^{169} and the Mutant L5.43^{169}M, Displaying Levels of AR231453-induced cAMP. Background cAMP levels were removed and scaled to maximum AR231453-induced cAMP production.

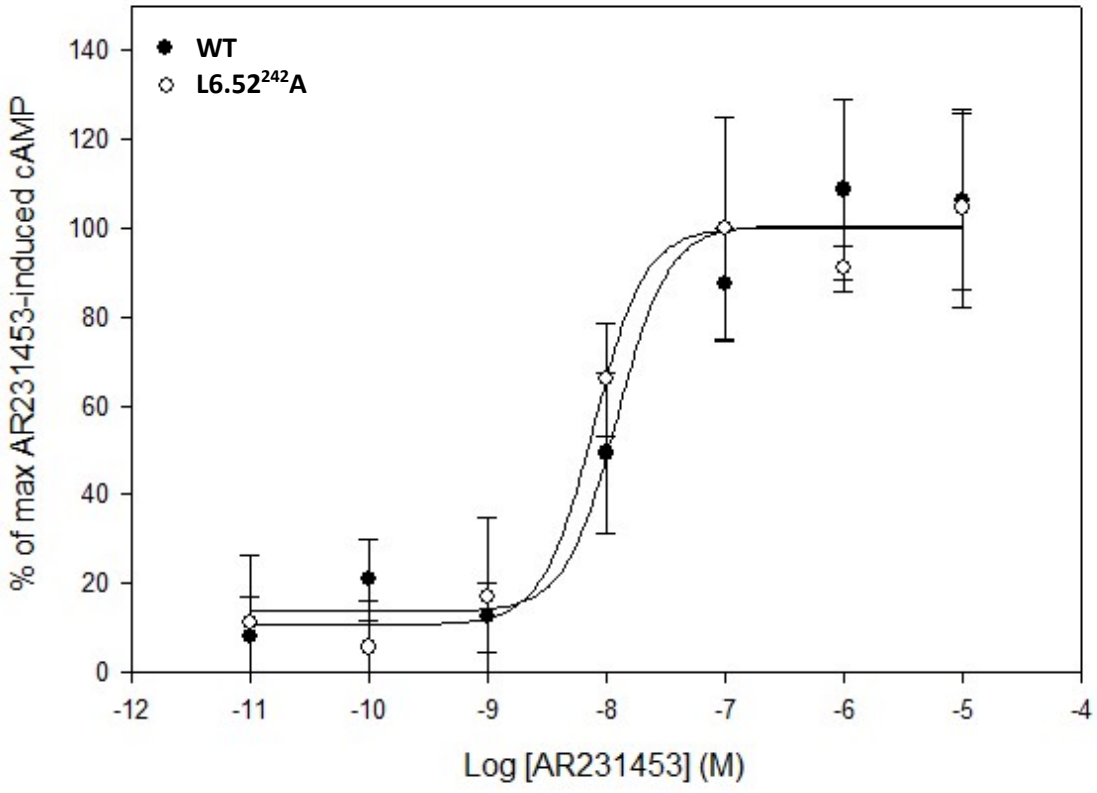

Figure 9: Dose Response Curve of the Wild Type L6.52^{242} and the Mutant L6.52^{242}A, Displaying Levels of AR231453-induced cAMP. Background cAMP levels removed and scaled to maximum AR231453-induced cAMP production.

The alanine mutation of Leu6.52^{242} yielded only a small shift to the left of the wild type dose response curve indicating a slight increase in agonist response (Figure 9). Interestingly, in 2015 Norn et al reported a similar result for the same mutation and observed a 3-fold decrease in receptor activity in response to AR231453. Both results show mild effects, implying that L6.52^{242}A is not likely to be heavily involved in ligand binding nor receptor activation. The opposite direction in the EC_{50} shift observed by Norn et al may possibly be attributed to their use of FBS as opposed to the CSFBS in this study. [20] As mentioned in section 2.2, lipids present in

FBS that was not charcoal stripped may also bind in the ligand binding pocket and activate GPR119. The small opposite shift may be the result of activation by lipids contained within the FBS.

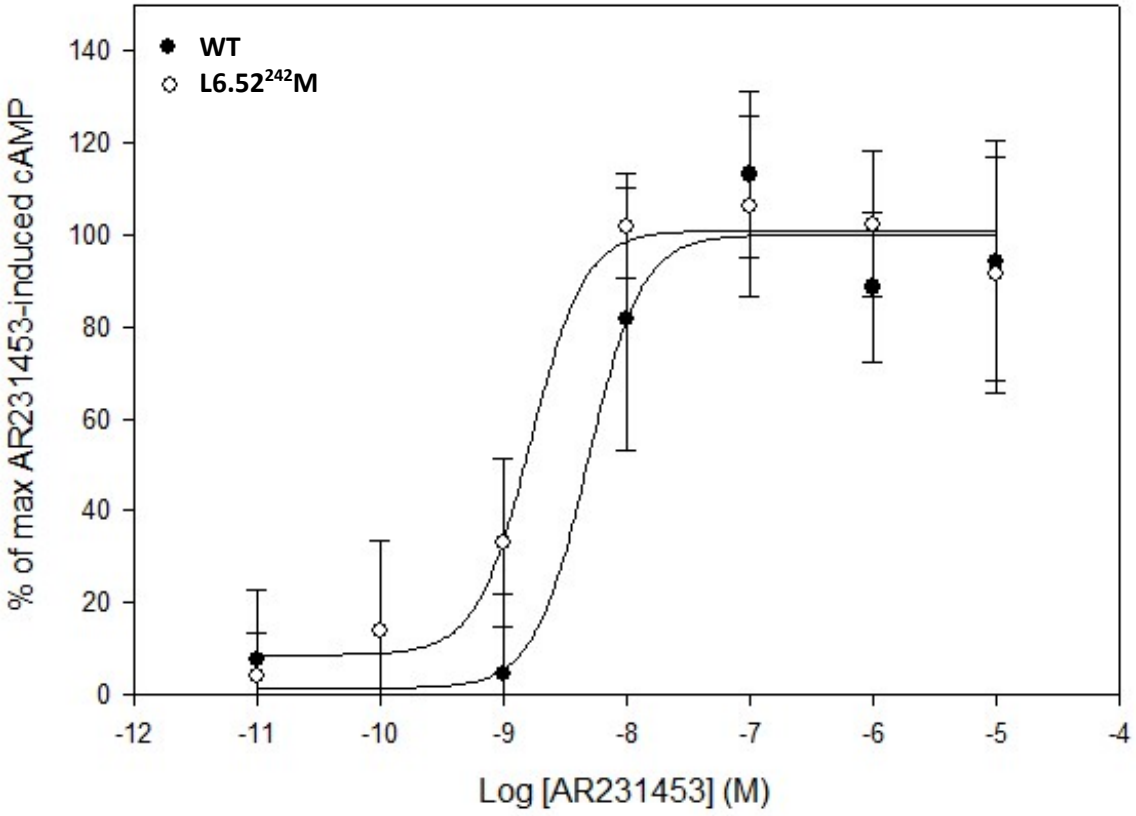

Figure 10: Dose Response Curve of the Wild Type L6.52^{242} and the Mutant L6.52^{242}M, Displaying Levels of AR231453-induced cAMP. Background cAMP levels were removed and scaled to maximum AR231453-induced cAMP production.

The mutant L6.52^{242}A exhibits a small change in dose-dependent activity measured through the AR231453-induced cAMP production (Figure 10). The wild type EC_{50} is 3-fold larger than the mutant EC_{50}, meaning this change does increase the response to AR231453

although small in effect. Like L6.52[242]A, the difference is not large, and this mutation does not largely impact binding or activation.

Since both mutations to change the bulk of the non-polar side chain of L6.52[242] had little effect, the *in vitro* data suggests that this amino acid does not participate in ligand binding or GPR119 activation. This implies that this leucine may be located in a position in which it cannot cause an effect, like being pointed sideways towards another helix, or out towards the surrounding membrane lipids.

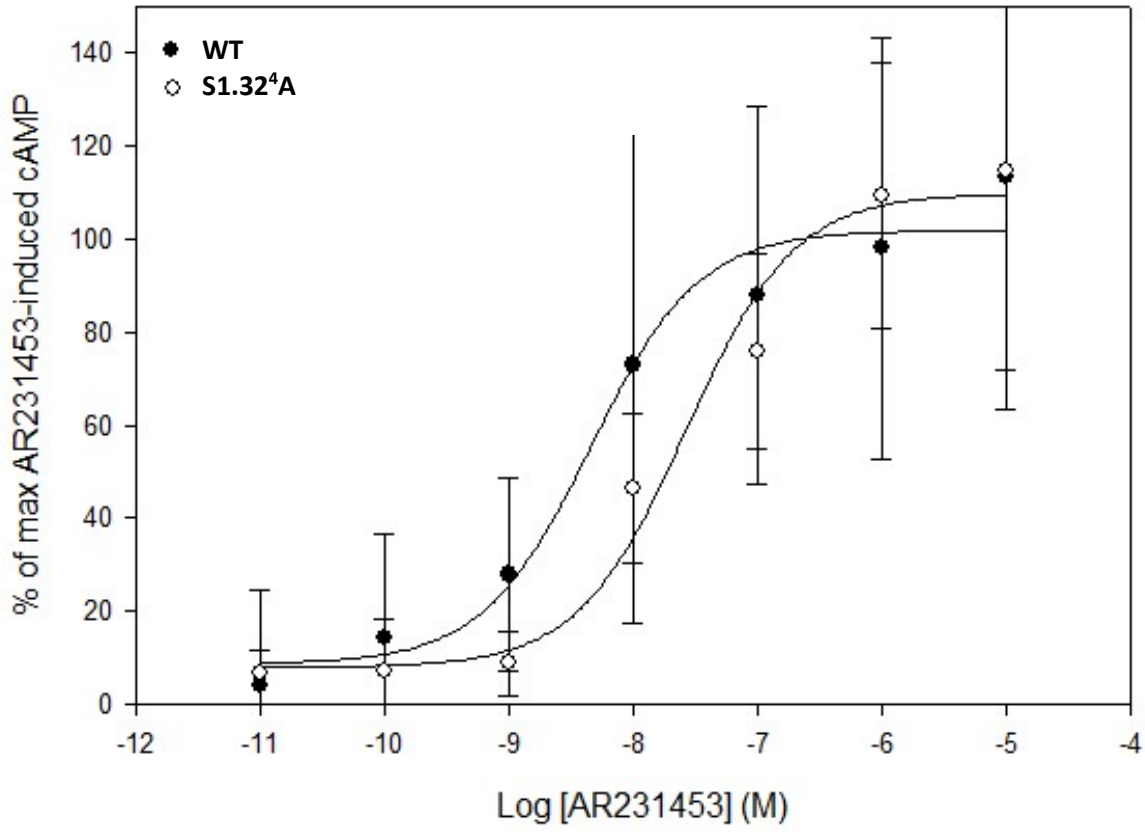

Figure 11: Dose Response Curve of the Wild Type S1.32[4] and the Mutant S1.32[4]A, Displaying Levels of AR231453-induced cAMP. Background cAMP levels were removed and scaled to maximum AR231453-induced cAMP production.

Unlike leucine, S1.32[4] possesses the ability to hydrogen bond. As was shown in previous studies, AR231453 most likely binds in the "SO2-up" position. [8, 20] To study whether this serine residue can have polar or hydrogen bond interactions with the AR231453 ligand head group, an alanine mutation was performed. There was a 6-fold decrease in AR231453 potency as a result of its change (Figure 11). This demonstrates that losing this hydroxyl group at the extracellular end of GPR119 antagonizes the response compared to the wild type. A table summarizing the effects of the mutations performed in this study is shown below in table 2.

Table 2: Summary of Experimentally Determined Mutant EC_{50} Values Compared to the Wild Type. n indicates the number of trails performed for each experiment.

Mutation	EC_{50} for mutant (nM)	EC_{50} for WT (nM)	Fold Change Relative to Wild Type	Effect to Activation relative to Wild type	n
L5.43[169]A	115.2	13.1	9	Decrease	6
L5.43[169]M	23.9	2.7	10	Decrease	4
L6.52[242]A	7.8	12.2	2	Increase	4
L6.52[242]M	1.6	4.8	3	Increase	6
S1.32[4]A	26.6	4.5	6	Decrease	4

4.3 Results of the MD Simulations

After the minimization and equilibration of the active GPR119 homology model in complex with the agonist AR231453 was completed as described in section 3.15, a more relaxed and equilibrated form of the active protein structure has been achieved as can be seen by

observing the RMSD over time (Figure 14). The RMSD of the backbone in the transmembrane region levels out around 1.5 Å.

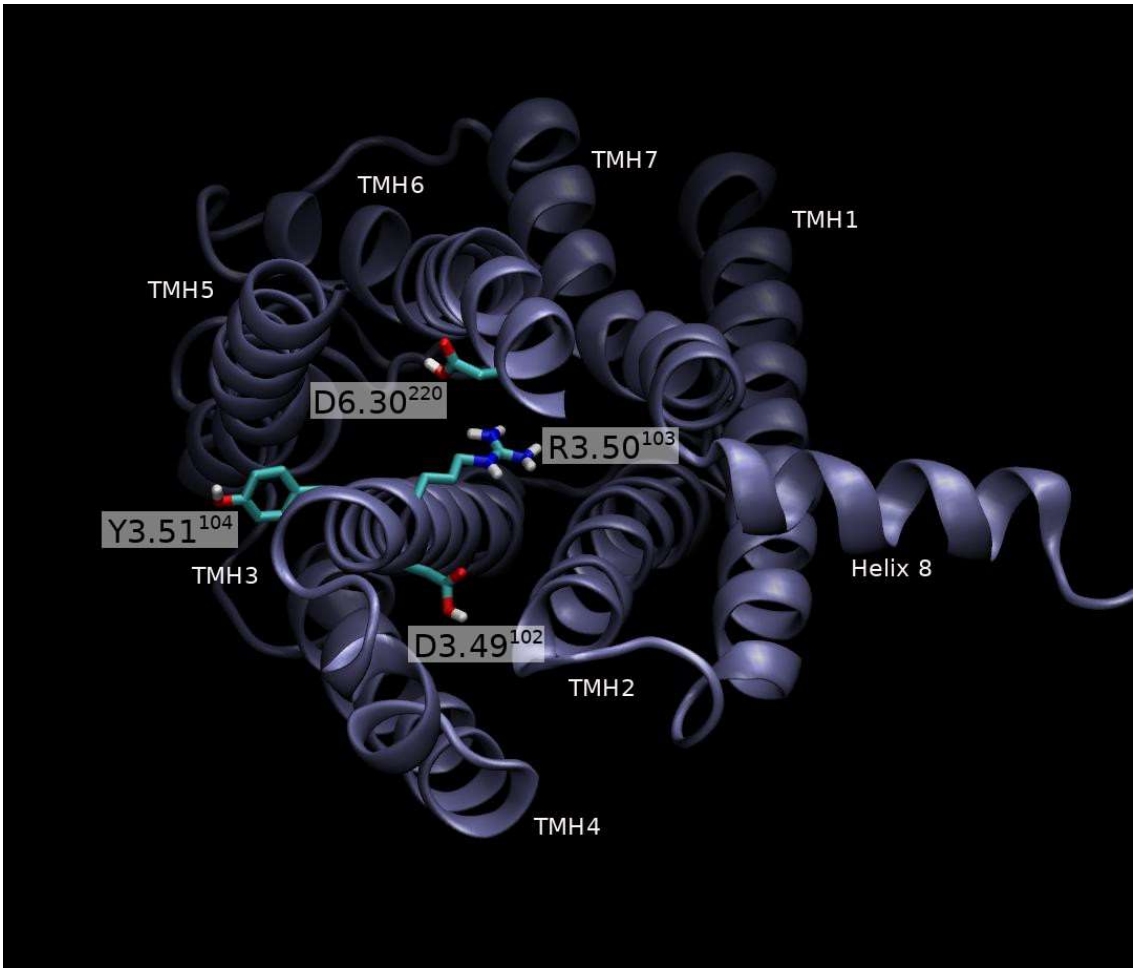

Figure 12: Ionic Lock of the GPR119Receptor Active Homology Model. The ICL3 has been undisplayed for clarity.

Arnis et al discovered that two protons were taken up by rhodopsin upon activation during the transition of metarhodopsin II_a to metarhodopsin II_b. Through flash photolysis and site-directed mutagenesis experiments, it was found that when the conserved Glu3.49[134] was mutated into a glutamine, light-induced proton uptake was inhibited. This implies that proton uptake seen by light-induced activation slows down when there is no Glu3.49[134] to be

protonated. [1] Following this study, it was found that charge-neutralizing mutations of the Glu6.30^{268} and Asp3.49^{130} together and separately cause an elevation in basal activity of the β_2-Adrenergic receptor. This demonstrates that by protonating these residues and disrupting the ionic lock, the protein becomes more active. [2,10] Finally, MD simulations demonstrating binding of the endogenous ligand, 2-arachidonoylglycerol, to the Cannabinoid 2 (CB2) receptor show that after the endogenous ligand had bound to the receptor, the fully active structure had not been achieved due to the reclosing of the ionic lock between Asp6.30^{240} and Arg3.50^{136}. A frame before the closing of the ionic lock was then taken and the Asp3.49^{130} and Asp6.30^{240} residues were protonated which allowed the ionic lock to remain open for the duration of the CB2 receptor MD simulations. [14] To prevent the closing of the ionic lock during MD simulations with the active GPR119 receptor homology model and since the GPR119 receptor has high sequence homology with the CB2 receptor, Asp3.49^{102} and Asp6.30^{220} were mutated to the protonated form using the ASPP patch in VMD193. The ionic lock of the GPR119 receptor remained open throughout the equilibration process (Figure 12).

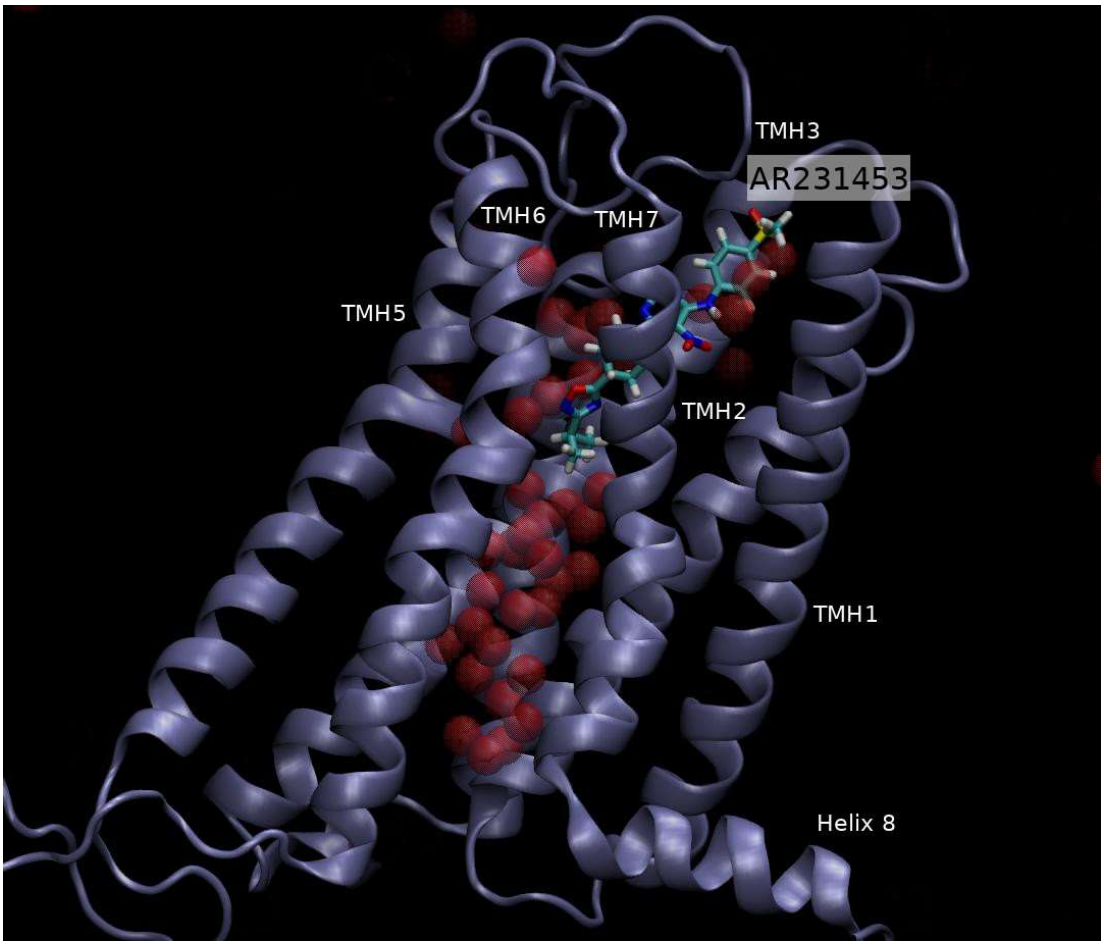

Figure 13: Water Channel of Active the GPR119 Receptor Homology Model in Complex with the Agonist AR231453.

Similar to the CB2 receptor study by Hurst et al in 2010, our models show a continuous water channel flowing between the extracellular and intracellular regions of the system due to the opening of the intracellular end of the GPR119 receptor (Figure 13). [14] Additionally, a water seen to stabilize the proline kink caused by Pro6.50^{240} in the inactive homology model no longer does so due to the straightening of TMH6.

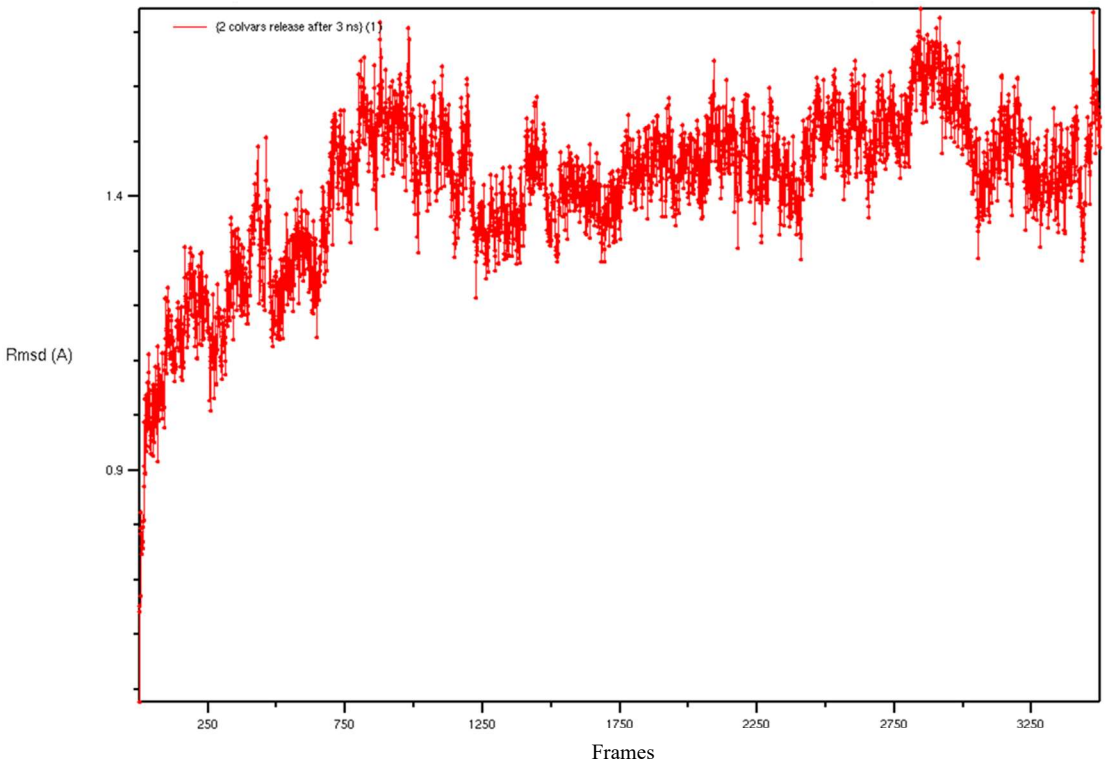

Figure 14: RMSD of GPR119 Active Homology Model Over 7 ns. Every 500 frames = 1.0 ns.

By using our equilibrated active GPR119homology model in complex with the agonist AR231453, some of the previously discussed cAMP assay results can be explained. Starting with the Leu5.43^{169} mutations, according to our homology model of GPR119, Leu5.43^{169} faces the inside of the binding pocket (Figure 15). Though the leucine side chain does not have direct contacts with the isopropyl tail of AR231453, it is in direct contact with Phe6.51^{241}, which in turn is in direct contact with the isopropyl group of the ligand. It appears that Leu5.43^{169} enforces a direct contact between Phe6.51^{241} and the ligand. The importance of maintaining the Phe6.51^{241} contact with the ligand was demonstrated by an alanine mutation of Phe6.51^{241} by Norn et al, which resulted in an 18-fold increase of the EC_{50} of AR231453. [20]

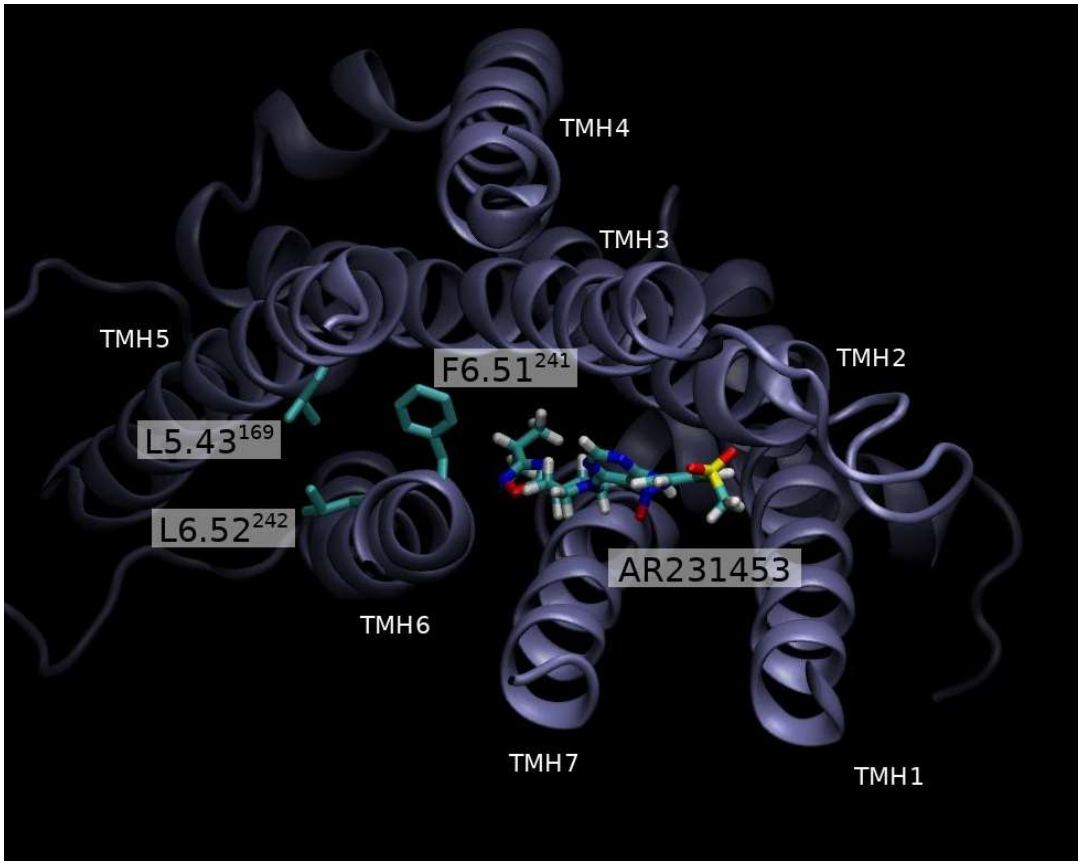

Figure 15: The GPR119 Receptor Active Homology Model in Complex with the Agonist AR231453 Depicting Locations of Leu5.43[169], Phe6.51[241], and Leu6.52[242]. The ECL2 and ECL3 loops and parts of TMH6 have been undisplayed for clarity.

Comparing cAMP assay results with the MD simulations results, our active model does not show Leu6.52[242] to be pointed towards the binding pocket or to have interactions with the ligand. While in our inactive model, this amino acid was pointed into the ligand binding pocket, the TMH6 of GPCR proteins is known to rotate as a part of activation as mentioned in Chapter 1. This then presents the issue, that TMH6 would have to have rotated to an active conformation before the agonist, which is supposed to trigger these conformational changes, is bound. Conventionally, an agonist would induce a change within the ligand binding pocket. This would be followed by conformational changes of the connecting region, sometimes referred to as the

hydrophobic core, which consists of tightly packed hydrophobic residues that communicate changes towards the intracellular interface. The intracellular interface would then change to accommodate the binding of the G-protein. [16] However, Latorraca explains through observing MD simulations of another GPCR, the β_2-adrenergic receptor (β_2AR), that activation of the β_2AR does not occur in sequential order. Instead the ligand binding pocket, connecting region, and the intracellular interface are able to change independently of the other two. Through stabilizing the active conformation in one area, it can influence the states of the other two regions. [18] This means for Leu6.52^{242}, it is possible that TMH6 of the GPR119 receptor could have moved into the active conformation preceding agonist binding and explaining the minimal effects of L6.52^{242} mutations.

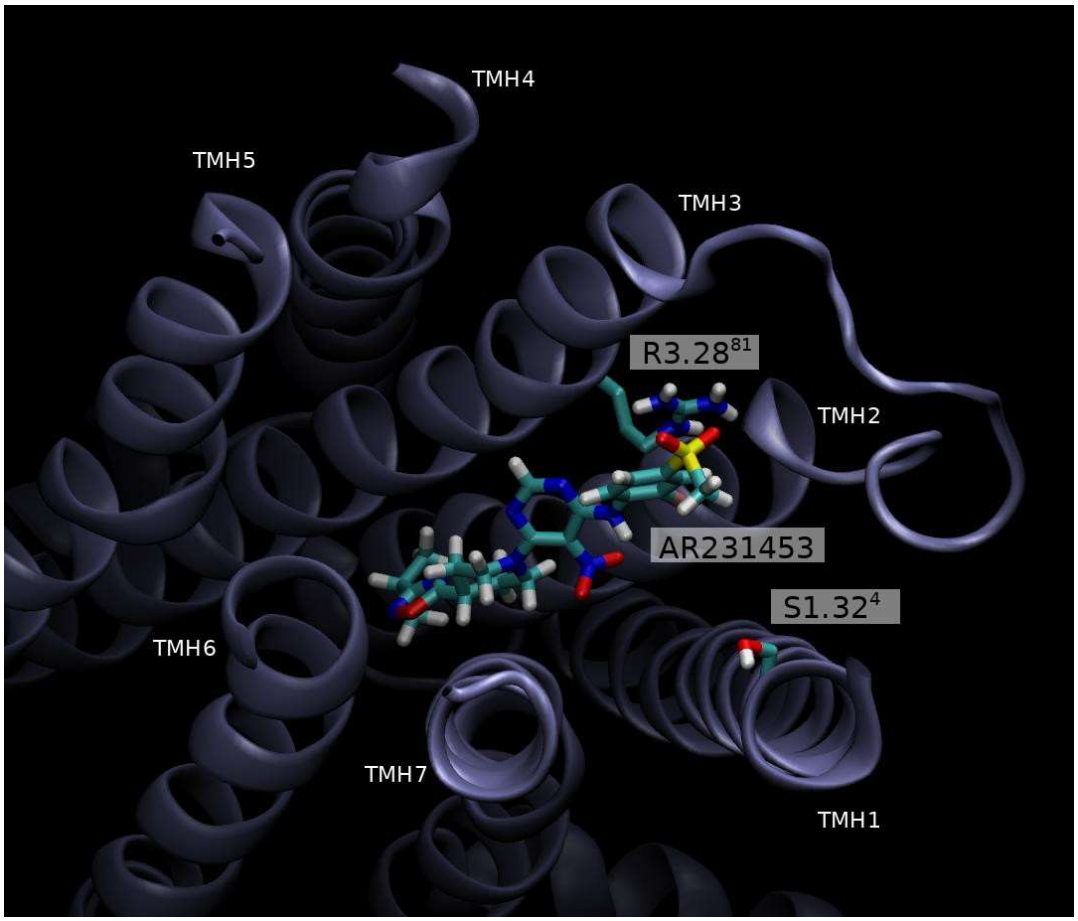

Figure 16: The GPR119 Receptor Active Homology Model in Complex with the Agonist AR231453 Depicting the Locations of Ser1.32[4] and Arg3.28[81]. The ECL2 and ECL3 loops and parts of TMH6 have been undisplayed for clarity.

The mutation results of Ser1.32[4]Ala are not as explicitly explained by the MD simulations as were the results of the two leucine amino acids. As can be seen in figure 16, Ser1.32[4], which is on the first turn of THM1 and exposed to water, lies close to the electron dense sulfone group of AR231453. However, the sulfone oxygen atoms face toward the positively charged Arg2.38[81] occasionally forming hydrogen bonds with it. The methyl group attached to the sulfone faces toward Ser1.32[4] potentially only forming transient hydrophobic interactions with it. It is plausible the methyl sulfone group could rotate and Ser1.32[4] could form hydrogen bonding interactions, but since both the methyl sulfone group and the serine are solvated during the course of the MD simulations study, they would not be long-lasing. It is plausible that even though Ser1.32[4] is heavily solvated in water, it may still assist orient the AR231453 ligand inside the GPR119 receptor pocket.

4.4 Voronoi Tessellation Monte Carlo Integration

Through X-ray scattering experiments, the average area occupied by POPC was determined to be 68.3 ± 1.5 $Å^2$/lipid. [17] To prevent the unrealistic packing of lipids in the system during the MD simulations study, NPAT conditions were used to constrain the area of the membrane. The area per lipid was measured at the start and end of the simulations to check that they occupied a reasonable amount of space. Through use of the VTMC program by Mori et al, the average area per lipid of the initial structure was 67.44 $Å^2$/lipid which is in agreement with

the experimentally determined area per lipid for POPC. After running 7 ns of NPAT MD simulations, the area per lipid was calculated to be 65.09 $Å^2$/lipid. The results show that the lipids are packed somewhat tighter than the experimentally determined value. This appears to be the result of the relaxation of the active GPR119 receptor by the end of the equilibration. This is apparent as the purple region representing the receptor atoms widens in the post simulation representation of the simulation cell (Figure 17 A-D). As a result of the opening of the receptor, the lipids packed a little closer together to maintain the constant area condition. Figure 17 C-D also shows a small opening in the middle indicative of the water channel. Overall, the NPAT simulations have held the area per lipid in close agreement with the experimentally determined findings for POPC.

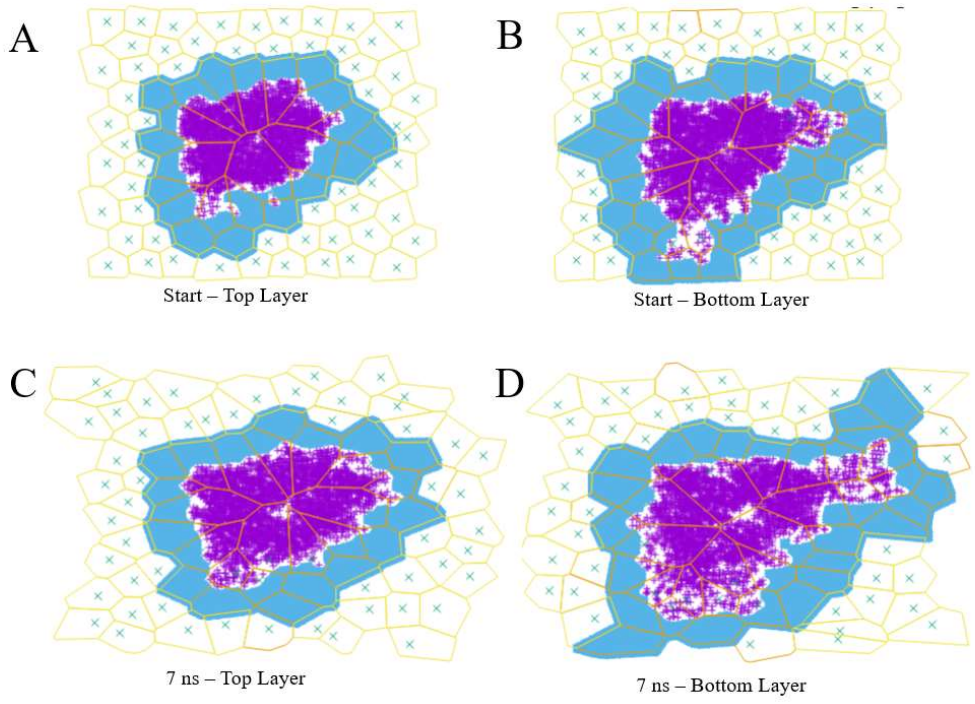

Figure 17: VTMC Diagrams of Start and End of MD. A and B represent the Voronoi tessellation of the top and bottom lipid layers at the beginning of the MD equilibration. C and D represent the Voronoi tessellation of top and bottom lipid layers of the system after 7 ns of NPAT

equilibration. Purple crosses represent the GPR119 receptor atoms. The blue region represents the area taken up by boundary lipids to GPR119. The green “x” represents the headgroup of non-boundary POPC lipids. The yellow lines represent the boundary of the space occupied by each of the lipids.

CHAPTER V

CONCLUSION

Through the docking study of AR437948 and AR437735, it was suggested that Leu5.43[169] and L6.52[242] may interact with the ligands and be involved in ligand binding or activation. Additionally, Ser1.32[4] was also chosen for further investigation based on potential polar interactions it may have with the electron dense ligand head groups. Through the use of MD simulations of a GPR119 receptor homology model and ELISA cAMP assays, our results indicate that Leu5.43[169] and Ser1.32[4] play a moderate role in the GPR119 receptor response to the standard full agonist AR231453. Leu5.43[169]Ala, Leu5.43[169]Met, and Ser1.32[4]Ala increased the AR231453 EC_{50} value by approximately 9-fold, 10-fold, and 6-fold respectively which means that the mutations decrease the overall agonist-induced activation of the GPR119 receptor. Leu5.43[169] is proposed to interact indirectly with the isopropyl tail of the ligand through interactions with the ring of Phe6.51[241]. Previous *in vitro* results by Engelstoft et al supports the importance of Phe6.51[241] and would agree with the GPR119 receptor active homology model using in this study. In agreement with Engelstoft at al, Leu6.52[242] mutations had a minor effect on agonist-induced activity, which according to the GPR119 receptor active homology model used in this study is pointed away from the ligand binding pocket.

Using the software NAMD in conjunction with a set of charmm36 topology/parameters for the protein, membrane and solvent, and a set of in-house developed topology/parameters for the agonist AR231453, an active GPR119 receptor homology model bound with AR231453 was

equilibrated in an explicit physiological transmembrane environment. The MD simulations produced an equilibrated GPR119 receptor structure that was used to interpret the mutational results. This structure may be used for further MD simulations to study the behavior of the receptor and also for virtual high throughput screening studies to discover and develop potent agonists that could potentially lead to new chemotherapeutic treatments for type 2 diabetes.

REFERENCES

1. Arnis, S., & Hofmann, K. P. (1993). Two different forms of metarhodopsin II: Schiff base deprotonation precedes proton uptake and signaling state. *Proceedings of the National Academy of Sciences*, *90*(16), 7849–7853. doi: 10.1073/pnas.90.16.7849

2. Ballesteros, J. A., Jensen, A. D., Liapakis, G., Rasmussen, S. G., Shi, L., Gether, U., & Javitch, J. A. (2001). Activation of the β2-Adrenergic Receptor Involves Disruption of an Ionic Lock between the Cytoplasmic Ends of Transmembrane Segments 3 and 6. *Journal of Biological Chemistry*, *276*(31), 29171–29177. doi: 10.1074/jbc.m103747200

3. Bokoch, M. P., Jo, H., Valcourt, J. R., Srinivasan, Y., Pan, A. C., Capponi, S., … Coughlin, S. R. (2018). Entry from the Lipid Bilayer: A Possible Pathway for Inhibition of a Peptide G Protein-Coupled Receptor by a Lipophilic Small Molecule. *Biochemistry*, *57*(39), 5748–5758. doi: 10.1021/acs.biochem.8b00577

4. Bruzzese, A., Gil, C., Dalton, J. A. R., & Giraldo, J. (2018). Structural insights into positive and negative allosteric regulation of a G protein-coupled receptor through protein-lipid interactions. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-018-22735-6

5. Cong, X., Fiorucci, S., & Golebiowski, J. (2018). Activation Dynamics of the Neurotensin G Protein-Coupled Receptor 1. *Journal of Chemical Theory and Computation*, *14*(8), 4467–4473. doi: 10.1021/acs.jctc.8b00216

6. Deupi, Xavier, and Brian Kobilka. "Activation of G Protein–Coupled Receptors." *Advances in Protein Chemistry Mechanisms and Pathways of Heterotrimeric G Protein Signaling*, 2007, pp. 137–166., doi:10.1016/s0065-3233(07)74004-4.

7. Doré, A. S., Robertson, N., Errey, J. C., Ng, I., Hollenstein, K., Tehan, B., … Marshall, F. H. (2011). Structure of the Adenosine A2A Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. *Structure*, *19*(9), 1283–1293. doi: 10.1016/j.str.2011.06.014

8. Engelstoft, M. S., Norn, C., Hauge, M., Holliday, N. D., Elster, L., Lehmann, J., … Schwartz, T. W. (2014). Structural basis for constitutive activity and agonist-induced activation of the enteroendocrine fat sensor GPR119. *British Journal of Pharmacology*, *171*(24), 5774–5789. doi: 10.1111/bph.12877

9. Fredriksson, R., Höglund, P. J., Gloriam, D. E., Lagerström, M. C., & Schiöth, H. B. (2003). Seven evolutionarily conserved human rhodopsin G protein-coupled receptors lacking close relatives. *FEBS Letters*, *554*(3), 381–388. doi: 10.1016/s0014-5793(03)01196-7

10. Ghanouni, P., Schambye, H., Seifert, R., Lee, T. W., Rasmussen, S. G. F., Gether, U., & Kobilka, B. K. (2000). The Effect of pH on β2Adrenoceptor Function. *Journal of Biological Chemistry*, *275*(5), 3121–3127. doi: 10.1074/jbc.275.5.3121
11. Hanson, M. A., Roth, C. B., Jo, E., Griffith, M. T., Scott, F. L., Reinhart, G., … Stevens, R. C. (2012). Crystal Structure of a Lipid G Protein-Coupled Receptor. *Science*, *335*(6070), 851–855. doi: 10.1126/science.1215904

12. Hassing, H. A., Fares, S., Larsen, O., Pad, H., Hauge, M., Jones, R. M., … Rosenkilde, M. M. (2016). Biased signaling of lipids and allosteric actions of synthetic molecules for GPR119. *Biochemical Pharmacology*, *119*, 66–75. doi: 10.1016/j.bcp.2016.08.018

13. Holst, B., Nygaard, R., Valentin-Hansen, L., Bach, A., Engelstoft, M. S., Petersen, P. S., … Schwartz, T. W. (2009). A Conserved Aromatic Lock for the Tryptophan Rotameric Switch in TM-VI of Seven-transmembrane Receptors. *Journal of Biological Chemistry*, *285*(6), 3973–3985. doi: 10.1074/jbc.m109.064725

14. Hurst, D. P., Grossfield, A., Lynch, D. L., Feller, S., Romo, T. D., Gawrisch, K., … Reggio, P. H. (2010). A Lipid Pathway for Ligand Binding Is Necessary for a Cannabinoid G Protein-coupled Receptor. *Journal of Biological Chemistry*, *285*(23), 17954–17964. doi: 10.1074/jbc.m109.041590

15. Jaakola, V.-P., Griffith, M., Hanson, M., Cherezov, V., Chien, E., Lane, J., … Stevens, R. (2008). The 2.6 A Crystal Structure of a Human A2A Adenosine Receptor bound to ZM241385. doi: 10.2210/pdb3eml/pdb

16. Krumm, B. E., Lee, S., Bhattacharya, S., Botos, I., White, C. F., Du, H., … Grisshammer, R. (2016). Structure and dynamics of a constitutively active neurotensin receptor. *Scientific Reports*, *6*(1). doi: 10.1038/srep38564

17. Kučerka, N., Tristram-Nagle, S., & Nagle, J. F. (2006). Structure of Fully Hydrated Fluid Phase Lipid Bilayers with Monounsaturated Chains. *Journal of Membrane Biology*, *208*(3), 193–202. doi: 10.1007/s00232-005-7006-8

18. Latorraca, N. R., Venkatakrishnan, A. J., & Dror, R. O. (2016). GPCR Dynamics: Structures in Motion. *Chemical Reviews*, *117*(1), 139–155. doi: 10.1021/acs.chemrev.6b00177

19. Mori, T., Ogushi, F., & Sugita, Y. (2011). Analysis of lipid surface area in protein-membrane systems combining voronoi tessellation and monte carlo integration methods. *Journal of Computational Chemistry*, *33*(3), 286–293. doi: 10.1002/jcc.21973

20. Norn, C., Hauge, M., Engelstoft, M. S., Kim, S. H., Lehmann, J., Jones, R. M., … Frimurer, T. M. (2015). Mutation-Guided Unbiased Modeling of the Fat Sensor GPR119 for High-Yield Agonist Screening. *Structure*, *23*(12), 2377–2386. doi: 10.1016/j.str.2015.09.014

21. Overton, H. A., Babbs, A. J., Doel, S. M., Fyfe, M. C., Gardner, L. S., Griffin, G., … Reynet, C. (2006). Deorphanization of a G protein-coupled receptor for oleoylethanolamide and its use in the discovery of small-molecule hypophagic agents. *Cell Metabolism*, *3*(3), 167–175. doi: 10.1016/j.cmet.2006.02.004

22. Palczewski, K. (2000). Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science*, *289*(5480), 739–745. doi: 10.1126/science.289.5480.739

23. Ritter, Kurt, et al. "G Protein-Coupled Receptor 119 (GPR119) Agonists for the Treatment of Diabetes: Recent Progress and Prevailing Challenges." *Journal of Medicinal Chemistry*, vol. 59, no. 8, Sept. 2015, pp. 3579–3592., doi:10.1021/acs.jmedchem.5b01198.

24. Ritter, S. L., & Hall, R. A. (2009). Fine-tuning of GPCR activity by receptor-interacting proteins. *Nature Reviews Molecular Cell Biology*, *10*(12), 819–830. doi: 10.1038/nrm2803

25. Wada, H., Osborne, J. C., & Manganiello, V. C. (1987). Effects of pH on allosteric and catalytic properties of the Guanosine Cyclic 3,5-phosphate cGMP-stimulated cyclic nucleotide phosphodiesterase from calf liver. *Biochemistry*, *26*(20), 6565–6570. doi: 10.1021/bi00394a042

BIOGRAPHICAL SKETCH

Matthew Dy Rosales was born on October 6, 1994. He attended and graduated from The Science Academy of South Texas in 2013. He graduated from University of Texas Rio Grande Valley, formerly University of Texas Pan American, with a BS Chemistry Pre-Med degree graduating as a cum laude in the spring of 2017. In the spring of 2016, he joined Dr. Evangelia Kotsikorou's computational chemistry computer where he did a docking study of AR437735 and AR437948 into the GPR119 inactive homology model. Over the Summer of 2017 he completed a docking study of the binding function 3 (BF-3) surface for the androgen and estrogen receptor using DDT and three similar analogs. The following Fall semester, he continued his research on GPR119 in the MS Chemistry program where he started MD and began working with Dr. Frank Dean to perform ELISA cAMP assays. He earned his MS Chemistry degree in May 2020 from the University of Texas Rio Grande Valley. His email is matthew.rosales106@gmail.com.