

12-2015

Molecular modeling and mutational mapping of the GPR119 binding site

Shane M. Askar
The University of Texas Rio Grande Valley

Follow this and additional works at: <https://scholarworks.utrgv.edu/etd>



Part of the [Chemistry Commons](#)

Recommended Citation

Askar, Shane M., "Molecular modeling and mutational mapping of the GPR119 binding site" (2015).
Theses and Dissertations. 10.
<https://scholarworks.utrgv.edu/etd/10>

This Thesis is brought to you for free and open access by ScholarWorks @ UTRGV. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact justin.white@utrgv.edu, william.flores01@utrgv.edu.

MOLECULAR MODELING AND MUTATIONAL
MAPPING OF THE GPR119 BINDING SITE

A Thesis

by

SHANE M. ASKAR

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

December 2015

Major Subject: Chemistry

MOLECULAR MODELING AND MUTATIONAL

MAPPING OF THE GPR119 BINDING SITE

A Thesis
by
SHANE M. ASKAR

COMMITTEE MEMBERS

Dr. Evangelia Kotsikorou
Chair of Committee

Dr. Frank Dean
Committee Member

Dr. James Bullard
Committee Member

Dr. Hassan Ahmad
Committee Member

December 2015

Copyright 2015 Shane M. Askar
All Rights Reserved

ABSTRACT

Askar, Shane M., Molecular Modeling and Mutational Mapping of the GPR119 Binding Site.

Master of Science (MS), December, 2015, 64 pp., 6 tables, 24 figures, references, 25 titles.

GPR119 receptor's biological role in regulating glucose homeostasis has been studied extensively. Results in the scientific literature indicate that, when activated, GPR119 releases insulin in a glucose dependent manner. Currently the 3D structure of GPR119 has not been resolved.

The goal of this research is to use a combination of homology modeling and ligand docking studies to predict the binding mode of GPR119 ligands. Amino acids implicated to have direct interactions with docked ligands will further be assessed experimentally for their roll in binding and activation of GPR119. Our results indicate that residues W265^{6,48} and R81^{3,28} are likely to be directly involved in ligand binding and activation of the GPR119 receptor. In addition, the R262^{7,36} mutant did not show any involvement in receptor binding or activation. Understanding how GPR119 interacts with its ligands can lead to the development of more effective and selective drugs that are used to treat T2D.

DEDICATION

I dedicate my thesis work to my husband, Soufian, for his patience, support, and encouragement during the challenges of graduate school and life. My husband has been my one and only source of inspiration for continuing my education. Thank you for motivating me to pursue my dreams.

ACKNOWLEDGMENTS

I will always be grateful to Dr. Evangelia Kotsikorou, my research advisor, for including me in her research group. For the past three years, my research at UTPA/UTRGV has been my heart and soul. I hope to feel as passionate about future research projects, as I did while working in her lab. I would also like to say a very special thank you to Dr. Frank Dean for guiding me through the experimental portion of my research. Thank you Dr. Dean for taking time away from your main research project to help me achieve spectacular results.

I would also like to thank Dr. Megan Keniry for her advice, lab protocols, and encouragement, when I was working on cell culture in her lab. Thank you Dr. Keniry for providing me with an awesome, sterile place to do cell culture. Last but not least, I would like to sincerely thank my lab members, Angel Tamez, Ed Villarreal, John Hamilton, Tevin Luna, Crystal Silvola, and Isaac Rodriguez for providing stimulating journal club meetings and input regarding my seminar and thesis presentations.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER I. INTRODUCTION	1
1.1 Type 2 Diabetes Mellitus	1
1.2 G Protein-Coupled Receptors	2
1.3 G Protein-Coupled Receptor Ligands	6
CHAPTER II. LITERATURE REVIEW	8
2.1 Glucose-Dependent Insulinotropic Receptor (GPR119)	8
2.2 Published GPR119 Mutations and Molecular Modeling Studies	9
2.3 GPR119 Ligands	10
CHAPTER III. EXPERIMENTAL METHODS	13
3.1 Molecular Modeling and Docking of GPR119	13
3.2 Site Directed Mutagenesis	17
3.3 Measurement of Mutant and Wild Type GPR119 Receptor Activation	21

CHAPTER IV. RESULTS AND DISCUSSION.....	26
4.1 GPR119 Models.....	26
4.2 GPR119 Agonist/Antagonist Diastereomer Pair Dock.....	34
4.3 GPR119 AR231453 Agonist Dock.....	45
4.4 Site-Directed Mutagenesis Results.....	50
4.5 Dose Response Curve Results.....	53
CHAPTER V. CONCLUSION.....	59
REFERENCES	61
BIOGRAPHICAL SKETCH	64

LIST OF TABLES

	Page
Table 1: Three Letter and Single Letter Abbreviations for the 20 Protein Amino Acids.....	1
Table 2: List of Primers Created for Use in Site-Directed Mutagenesis Experiment.....	18
Table 3: Interaction Energies Between GPR119 R* Receptor Model Residues and Agonist Ligand.....	37
Table 4: Component Interaction Energies Between GPR119 R Receptor Model Residues and Antagonist Ligand.....	41
Table 5: Interaction Energies Between GPR119 R* Receptor Model Residues and the AR231453 Ligand	48
Table 6: LogEC ₅₀ Values and Potency Fold Changes for WT GPR119 and Mutant GPR119 Receptor in the cAMP Assay	58

LIST OF FIGURES

	Page
Figure 1: Topology of a Class A GPCR	3
Figure 2: Agonist and Antagonist Diastereomer Pair of GPR119 Ligands	11
Figure 3: GPR119 Agonists	12
Figure 4: Western Blot Analysis of Transfection Efficiency Results	22
Figure 5: Amino acid Sequence Alignment for GPR119, A ₂ A, and S1P1	28
Figure 6: Intracellular View of Active (R*) and Inactive (R) GPR119 Receptor Models superimposed on TMHs 1, 2, 4, 5, and 7	29
Figure 7: <i>Intracellular View of GPR119 Receptor Models Showing Intact (Inactive) and Broken (Active) Ionic Lock</i>	30
Figure 8: Tryptophan Residue Indicating Rotation of the Chi1 Dihedral Angle	32
Figure 9: Rotamer Changes in the χ_1 Dihedral Angle for GPR119 Residue W238 ^{6,48}	33
Figure 10: Ligand Plot of the Docked Agonist of the Diastereomer Pair Inside the GPR119 Receptor R* Binding Pocket	34
Figure 11: Side View of the GPR119 R* Receptor Model (purple ribbons) with Docked Agonist from the Diastereomer Ligand Pair	35
Figure 12: Top View of the GPR119 R* Receptor Model (purple ribbons) with Docked Agonist from the Diastereomer Ligand Pair	36
Figure 13: <i>Ligand Plot of the Docked Antagonist of the Diastereomer Pair Inside the GPR119 Receptor R Binding Pocket</i>	38

Figure 14: Side View of the GPR119 R Receptor Model (blue ribbons) with Docked Antagonist from the Diastereomer Ligand Pair.....	39
Figure 15: Top View of the GPR119 R Receptor Model (blue ribbons) with Docked Antagonist from the Diastereomer Ligand Pair.....	40
Figure 16: Ligand Plot of the Docked Agonist AR231453 Inside the GPR119 R* Receptor Binding Pocket.....	45
Figure 17: Side View of the GPR119 R* Receptor Model (purple ribbons) with Docked AR231453 Ligand.....	46
Figure 18: Top View of the GPR119 R* Receptor Model (purple ribbons) with Docked AR231453 Ligand.....	47
Figure 19: Nucleotide & Amino Acid Sequence for GPR119 Receptor	51
Figure 20: DNA Sequencing Chromatograms Confirming Desired Mutations in the GPR119 Receptor Sequence.....	52
Figure 21: Experimental Control Dose Response Curve.....	56
Figure 22: Dose Response Curve Comparing cAMP Accumulation in WT and mutant R81L GPR119 Transfected Cells.....	56
Figure 23: Dose Response Curve Comparing cAMP Accumulation in WT and mutant W265A GPR119 Transfected Cells.....	57
Figure 24: Dose Response Curve Comparing cAMP Accumulation in WT and mutant R262L GPR119 Transfected Cells.....	57

CHAPTER I

INTRODUCTION

This chapter introduces the main problem and purpose for conducting this thesis research. In addition, general information for understanding G Protein-Coupled receptor structure and function will also be discussed. Throughout the following pages, either three or single-letter abbreviations for the 20 amino acids will be used and are listed in table 1.

Table 1: Three Letter and Single Letter Abbreviations for the 20 Protein Amino Acids

Amino Acid Abbreviations			
Arginine Arg (R)	Methionine Met (M)	Cysteine Cys (C)	Glutamic Acid Glu (E)
Aspartic Acid Asp (D)	Phenylalanine Phe (F)	Proline Pro (P)	Lysine Lys (K)
Leucine Leu (L)	Valine Val (V)	Tryptophan Trp (W)	Tyrosine Tyr (Y)
Serine Ser (S)	Threonine Thr (T)	Glutamine Gln (Q)	Asparagine Asn (N)
Alanine Ala (A)	Glycine Gly (G)	Histidine His (H)	Isoleucine Iso (I)

1.1 Type 2 Diabetes Mellitus

According to the Centers for Disease Control and Prevention (CDC) 29.1 million people or 9.3% of the U.S. population have diabetes.^[3] This figure includes all age groups, all diabetes types and includes an estimate of people who are undiagnosed.^[3] The CDC has also

reported that in adults type 2 diabetes (T2D) accounts for about 90% to 95% of all diagnosed cases of diabetes.^[3] People with T2D are unable to effectively regulate blood glucose levels. This disease is further characterized by a defect in insulin secretion or by resistance to the blood glucose lowering effects of secreted insulin.^[25] Along with the elevated risk of hyperglycemia, patients suffering from T2D experience severe complications including: high blood pressure, high cholesterol, heart disease and stroke, blindness, kidney disease, and amputations.^[3] In recent years, G protein coupled receptor 119 or GPR119, has become a target for novel T2D treatments. This is due to GPR119's dynamic role in glucose regulation in the human body. Through molecular modeling and mutational mapping methods, a better understanding of the binding-pocket structure of GPR119 can be elucidated. This information will assist in targeting this protein with drugs that produce desired results with limited side effects.

1.2 G Protein-Coupled Receptors

Over 800 G protein-coupled receptor (GPCR) sequences have been identified in the human genome.^[10] GPCRs are grouped according to sequence homology and suspected or known biological functions.^[10] In humans, GPCRs are partitioned into five families named Rhodopsin-like, Adhesion, Glutamate, Secretin, and Frizzled/Taste2 Receptors.^[10] The Rhodopsin-like family, or class A, is the largest with 701 distinct receptors.^[10] Due to their potential as drug targets, the pharmaceutical industry has undertaken a tremendous effort to deduce the physiological function of each human GPCR. However, the physiological role of a large fraction of these GPCRs still remain unknown; these receptors are referred to as orphan GPCRs.^[5]

GPCRs are dynamic, transmembrane proteins. They are formed from a single

polypeptide, which is folded and embedded within the cell plasma membrane.^[18] Despite the diversity of their functional role, GPCRs all maintain the same overall protein topology. Their secondary structure consists of 7 transmembrane α -helices (TMHs) that are connected by three intracellular (IC) and three extracellular (EC) loops. Their N-terminus is located extracellularly and their C-terminus is located within the cytosol.

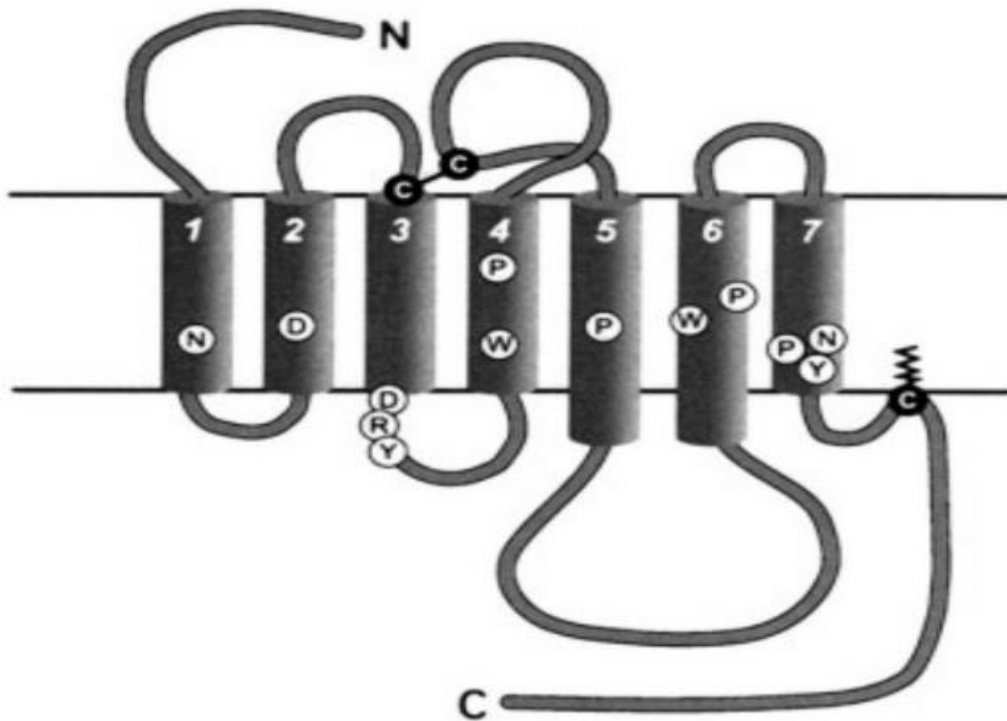


Figure 1: Topology of a Class A GPCR. Circled are conserved amino acids and motifs in each helix.^[18]

The TMHs of GPCRs bundle together within the membrane in a cylindrical shape and the upper portion of the receptor forms the ligand-binding pocket. In different GPCRs, the upper portion of the TMHs, the EC loops, and/or the N-terminus may be involved in the ligand-binding process. For example: According to the Deupi and Kobilka “Small organic agonists often bind within the TMH segments. Peptide hormones and proteins often bind to the N-terminus and extracellular loops joining the TMH domains. However, the size of the ligand alone cannot be used to predict the location of the binding site.”^[5]

The intracellular end of a GPCR includes the three IC loops and the binding site for the heterotrimeric G protein. GPCRs transduce extracellular stimuli to give rise to intracellular signals, through the interaction of their intracellular domains with G proteins.^[10] G proteins regulate concentrations of intracellular secondary messengers based upon the signals they receive from an activated GPCR. G proteins contain three distinct subunits α , β , and γ . The G protein α subunit, and combined $\beta\gamma$ subunits have different signaling functions.^[10] There are at least 18 different human $G\alpha$ subunits that couple to GPCRs, at least 5 $G\beta$ subunits, and at least 11 different $G\gamma$ subunits.^[10] “Heterotrimeric G proteins can be broadly categorized into four major classes based on the identity of the β subunit: G_s , $G_{i/o}$, $G_{q/11}$, and $G_{12/13}$.”^[9]

Each of the seven TMHs of a GPCR contain at least one highly conserved residue. In order to compare amino acid sequences between different GPCRs with limited sequence identity, conserved residues are used to anchor GPCR sequence alignments.^[23] “The conserved residues are: TMH 1 (Asn), TMH 2 (Asp), TMH 3 (Arg), TMH 4 (Trp), TMH 5 (Pro), TMH 6 (Pro), and TMH 7 (Pro).”^[2] The motifs D/ERY from TMH3, FxxCWxP from TMH 6, and NPxxY from TMH 7 can be considered as common “molecular switches” that are found within most GPCRs.^[2] When these “switches” are disturbed from their basal stabilized state, by ligand

binding or receptor conformational flexibility, GPCRs can undergo structural rearrangement to facilitate changes in GPCR signaling activity.^[5]

The DRY motif or “ionic lock” involves the interaction between an Asp or Glu at the intracellular end of TMH 6, and Arg, in the DRY motif on TMH 3.^[5] “This ionic interaction is proposed to hold together the cytoplasmic ends of TMH 3 and TMH 6 in the resting state of different amine receptors.”^[5] The FxxCWxP motif on TMH 6 is often referred to as the “rotamer toggle switch.” It has been suggested that rotameric positions of Phe and Trp from the motif,, modulate the bend angle of TMH 6 around the highly conserved Pro kink, leading to the outward movement of the cytoplasmic end of TMH 6 upon GPCR activation.^[5] In addition, rotameric changes in residues from the NPxxY motif are also important for GPCR activation.^[16] It is important to note that a single GPCR will, more than likely, contain numerous “molecular switches.” However, the motifs discussed previously are the switches that are commonly proposed for the rhodopsin family of GPCRs.^[5]

In general, activation of a GPCR via agonist binding induces the extracellular portions of TMHs 5-7 to move inward, while the cytosolic portions of TMHs 5-7 translate outward.^[23] Therefore, the movements that a GPCR helical bundle undertakes during activation are analogous to the pinching a clothespin. This opening up of the cytosolic side of the receptor creates room for the helical tip of the G Protein’s α subunit to insert inside the receptor.^[23] G Protein dissociation from the GPCR complex occurs upon the exchange of a guanosine diphosphate (GDP) for guanosine triphosphate (GTP).^[18] The G protein α subunit bound to GTP then diffuses away from the receptor and the $\beta\gamma$ dimer.^[18] Both the α subunit and the $\beta\gamma$ dimer move toward target proteins and assist in the stimulation of second messengers that propagate biological signaling cascades.^[18]

Recent studies have shown that GPCR signaling is actually more complex than the basic on/off switch model. “In the simplest model for the conformational dynamics of GPCRs there is an equilibrium between two states, R and R*.”^[12] The inactive receptor state is identified by R and the active state identified by R*. “Where only R* can couple and activate G proteins.”^[12] However, numerous distinct intermediate states between inactive and active state conformations have been discovered in the GPCR Rhodopsin.^[19] “G proteins also exhibit some promiscuity and a single GPCR can couple to and signal via G proteins from multiple classes. This type of activity results in the propagation of signals through multiple biochemical pathways to achieve different cellular responses.^[9] In addition, G protein-independent signaling can occur within a cell via arrestin proteins to increase the diversity of cellular responses.^[9] Arrestins bind to phosphorylated sites on the C-terminus or IC loops of GPCRs and often function to desensitize the receptor to ligand binding and mediate G protein independent signaling pathways.^[9]

1.3 G Protein-Coupled Receptor Ligands

A GPCR’s role in cell signaling and signal transduction is initiated by protein-ligand interactions.^[Cohen] GPCR ligands are highly diverse and include peptides, nucleotides, lipids, amino acids, and glycoproteins.^[4] Since GPCRs and their ligands play a very active role in a multitude of cellular processes, 40-60% of the current drugs on the market target GPCRs.^[4] The diverse nature of GPCRs has led to the development of drugs for the treatment of a number of different maladies including cardiovascular, metabolic, neurodegenerative, psychiatric, and oncologic diseases.^[13]

GPCR regulation includes full/partial agonism, neutral antagonism, inverse agonism, as well as, allosteric regulation. Agonists are ligands that bind to and activate receptors thereby

eliciting a physiological response.^[9] Agonists function to stabilize GPCRs in their active state conformation. Antagonists are classically considered compounds that bind to GPCRs to block binding and activation by agonists, but produce no G-Protein mediated activity of their own.^[6] Inverse agonists bind to receptors as agonist but exert affects opposite to that of an agonist.^[9]

Some GPCRs are known to have a specific level of basal or constitutive activity, which is an activity level in the absence of agonist stimulation.^[6] For example, the amount of second messengers, like cyclic adenosine monophosphate (cAMP), produced in the absence of agonist stimulation is referred to as the GPCR's basal activity. Inverse agonists are agonists that act upon constitutively active GPCRs and their binding reduces signaling mediated activity below the level of basal activity.^[9] Extra binding sites also exist on the surface of GPCRs. These allosteric binding sites are distinct from the main orthosteric binding pocket, located at the central upper portion of most rhodopsin-like receptors.^[6] Allosteric binding sites bind small molecules and function to further regulate GPCR signaling activity.

CHAPTER II

LITERATURE REVIEW

This chapter examines previously published knowledge and data regarding G Protein-Coupled Receptor 119. GPR119's signaling mechanism, biological role in T2D treatment, similarity to other GPCRs with known structure, and ligands will be discussed in this chapter. The underlying rationale for conducting this research will also be explained.

2.1 Glucose-Dependent Insulinotropic Receptor (GPR119)

GPR119 is a G-protein coupled receptor that is expressed in humans predominantly in the pancreatic β cells and the gastrointestinal tract.^[25] The protein contains the seven TMHs that are characteristic of all GPCRs and is 335 amino acids long. GPR119 is a class A rhodopsin-like receptor, whose endogenous ligands have been identified as oleoylethanolamide (OEA) and 2-monoacylglycerols.^[7] When activated this receptor couples to the $G\alpha_s$ G protein and signals through the G_s pathway, which increases intracellular cAMP levels by stimulating the enzyme adenylate cyclase to cleave adenosine triphosphate (ATP) to cAMP.^[25]

Upon agonist activation of GPR119, intracellular cAMP levels increase, leading to enhanced glucose-dependent insulin secretion from pancreatic β -cells and increased release of the gut peptides GLP-1 (glucagon-like peptide 1) and GIP (glucose-dependent insulinotropic peptide).^[25] GLP-1 and GIP are incretin hormones that are released from gut L and K cells respectively.^[25] Both GLP-1 and GIP act upon related receptors on pancreatic β cells to promote

additional insulin secretion.^[25] Thus, GPR119, GLP-1, and GIP work in conjunction to preserve glucose homeostasis. Pharmaceutical companies like Glaxo-SmithKline, Arena Pharmaceuticals, and Bristol-Myers Squibb have successfully synthesized selective, potent GPR119 agonists to probe GPR119's potential for T2D treatment. Preclinical and clinical trials using potent GPR119 agonists have shown to 1) lower blood glucose without causing hypoglycemia; 2) slow diabetes progression; and 3) reduce food intake.^[25]

Thus far, the 3D structure of GPR119 at the atomic level has not been resolved. Structural knowledge of GPR119 binding pocket, and its mechanism of interaction with endogenous and synthetic ligands, will help in the development of more potent, anti-diabetic drugs that produce desired results, while eliminating unwanted side effects. Of the class A GPCRs that have a published crystal structure in the Protein Data Bank, the most closely related receptor to GPR119 is the A₂A adenosine receptor. In the transmembrane region, GPR119 and the A₂A adenosine receptor share 27.35% sequence identity (exactly matching amino acids) and 73.09% sequence similarity (similarly matching amino acids).^[20] A protein sequence with over 30% sequence identity to a known structure can often be predicted, using homology modeling, with an accuracy equivalent to a low-resolution X-ray structure.^[24] Although the sequence identity of GPR119 and the A₂A adenosine receptor is slightly below 30%, a receptor model can be constructed by homology modeling, and can be refined using additional information obtained from mutational analysis of GPR119 binding pocket residues.

2.2 Published GPR119 Mutations and Molecular Modeling Studies

In August 2014, Engelstoft et al. published the first molecular docking study along with mutational mapping data for GPR119. Their results indicate that GPR119 signals with a high

level of constitutive activity, which is $37 \pm 0.8\%$ of the E_{\max} (maximum response achievable by a drug) obtained from using the synthetic agonist AR231453.^[7] AR231453 is a potent, selective GPR119 agonist synthesized by Arena Pharmaceuticals, Inc. AR231453 was found to significantly increase insulin release in HIT-T15 cells and in rodent pancreatic islet cells.^[25] By contrast, no effect of this compound could be seen in pancreatic islets isolated from GPR119-deficient mice, confirming that its effects were indeed mediated by GPR119.^[25]

The mutational data for GPR119 produced by Engelstoft et al. included 30 mutations at 23 different residue positions. Their results also indicate that EC2 residues are important for constitutive activity, as well as, ligand activation.^[7] However, regarding the docking pose for AR231453 ligand and the residues responsible for receptor activation, our docking and mutational analysis of GPR119 yields different results from the Englestoft et al. study.

2.3 GPR119 Ligands

Currently, the discovery of GPR119 synthetic ligands have focused much attention on the synthesis of receptor agonists. This one-sided view can partly be attributed to the pursuit of novel ligands that activate GPR119 and, therefore; have the potential to become treatments for T2D. In order to accurately elucidate the structural and functional activity of GPR119, agonist and antagonist mechanisms of interaction with the receptor should be examined. In 2011, McClure et al. successfully synthesized and identified an agonist and antagonist diastereomer pair of GPR119 ligands that differed only in the placement of an equatorial or axial ether bridge on a piperidine ring.^[15] The ether bridge served to lock the ligand in either an antagonist or agonist conformation. The agonist [Isopropyl 9-syn-({5-Methyl-6-[(2-methylpyridin-3-yl)oxy]}-

pyrimidin-4-yl}oxy)-3-oxa-7-azabicyclo[3.3.1]nonane-7-carboxylate] is locked into its conformation by an equatorial placed ether bridge. The antagonist [Isopropyl 9-syn-(5-Methyl-6-[(2-methylpyridin-3-yl)oxy]-pyrimidin-4-yl}oxy)-3-oxa-7-azabicyclo[3.3.1]nonane-7-carboxylate] is locked into its conformation by an axial placed ether bridge.

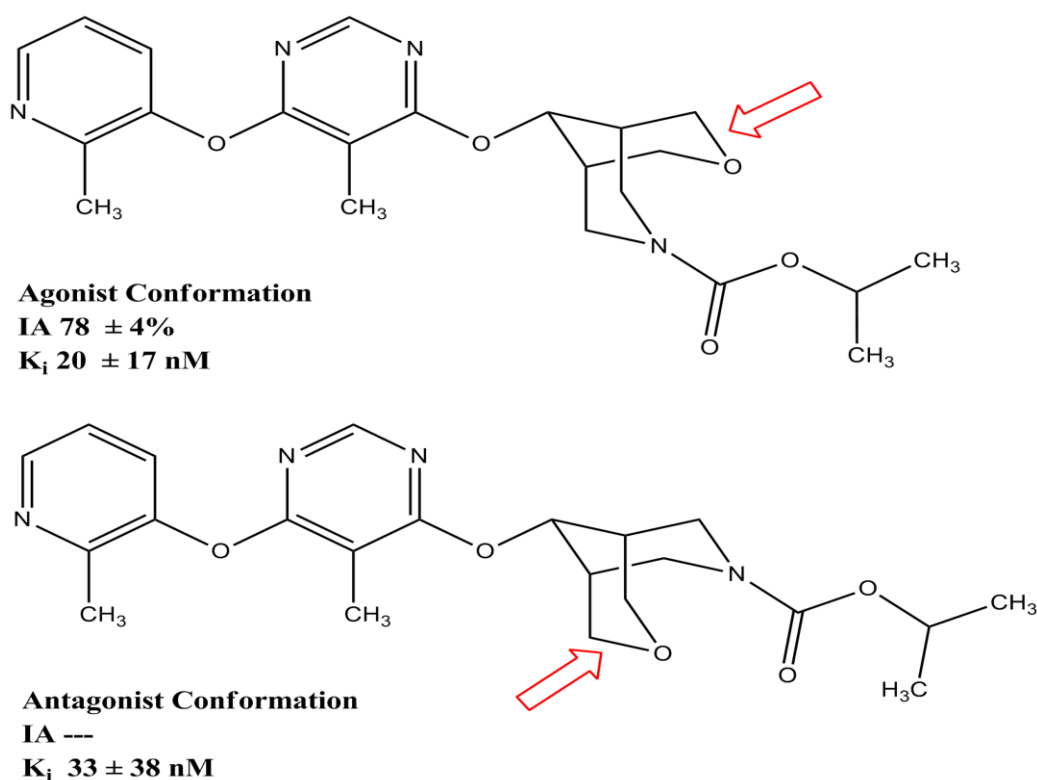


Figure 2: Agonist and Antagonist Diastereomer Pair of GPR119 Ligands. Ligands were synthesized and tested by McClure et al.(2011). Ligands differ only in either an equatorial or axial placement of an ether bridge, which are signified by red arrows. Intrinsic activity (ability of ligand to produce maximum functional response) and K_i (receptor affinity) of each ligand are reported.

Both agonist and antagonist conformations share essentially an equivalent affinity for GPR119 (K_i agonist 20 ± 17nM and antagonist 33 ± 38 nM)^[15] ; however, the ability to produce a

functional response (intrinsic activity) is completely diminished for the antagonist locked conformation and is relatively high for the agonist locked conformation ($78 \pm 4\%$ IA).^[15] An explanation, of the mechanism of the binding mode and effect on activation of the GPR119 receptor in the R* and R state, can be achieved through molecular docking studies of this pair of diastereomers and can be further substantiated by mutational mapping of this receptor using the highly potent agonist, AR231453.

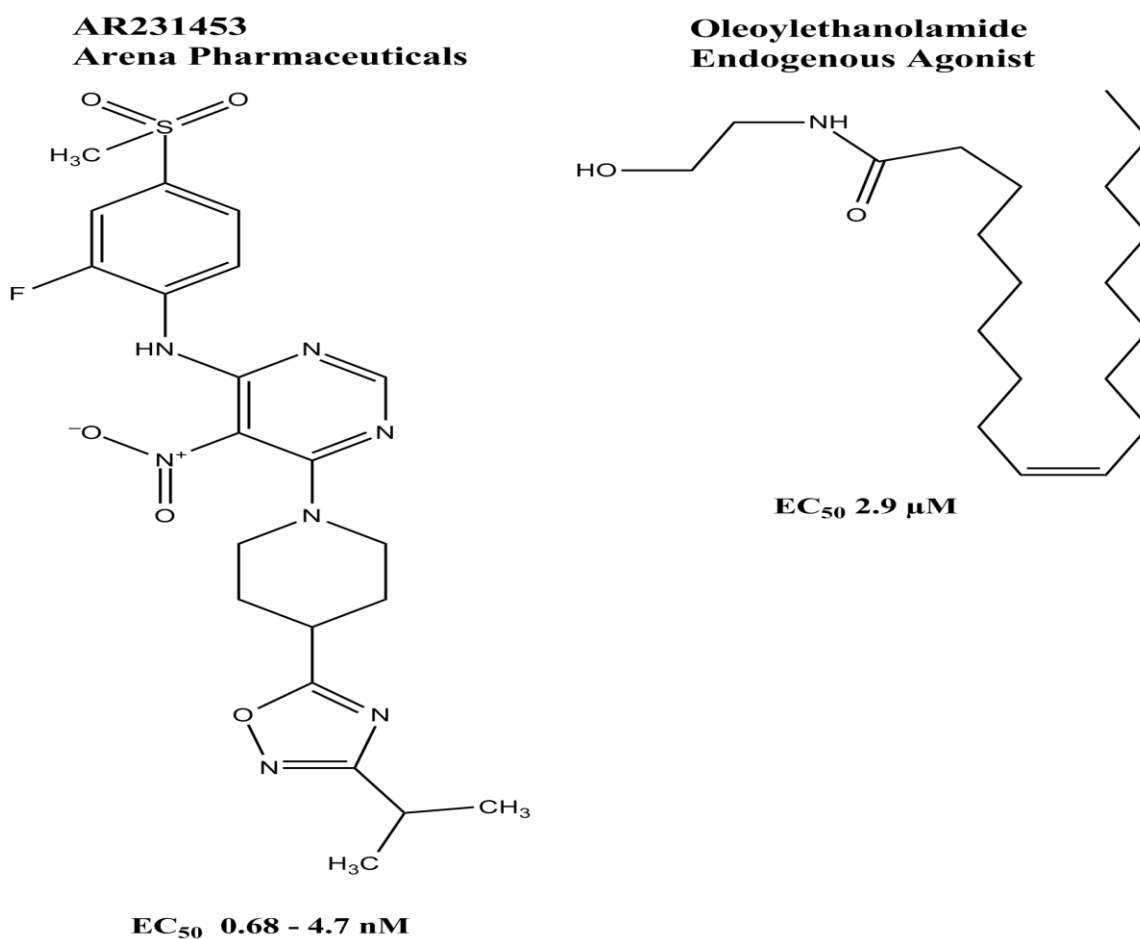


Figure 3: GPR119 Agonists. (Left) AR231453 – most potent, selective agonist for GPR119. AR231453 was utilized in docking studies and to stimulate cAMP production from transiently transfected HEK293 cells. (Right) GPR119's endogenous ligand is shown for comparison.

CHAPTER III

EXPERIMENTAL METHODS

In this section, a description of experimental procedures, as well as, reagents, kits, and software used in modeling or mutational mapping of the GPR119 receptor will be outlined. When discussing specific GPR119 residues absolute sequence numbering followed by Ballesteros and Weinstein numbering^[2] in superscript will be used to identify each residues location within the seven transmembrane helices. For example in R81^{3.28} : 81 is the absolute sequence number for the arginine residue, the 3 in the superscript signifies that this residue is on helix 3, and the 28 in the superscript indicates this residue's position precedes the most conserved residue in that helix, which is arbitrarily assigned to 50. In addition, GPR119 mutations that will be discussed and are listed in the following format: wild type residue, absolute protein sequence residue number, and then the mutant residue. For example in R81L, arginine (R) would represent the wild type residue, 81 would correspond to the residue position in the absolute protein sequence, and leucine (L) would represent the mutant residue.

3.1 Molecular Modeling and Docking of GPR119

Computational methods for modeling of the GPR119 receptor and docking ligands into this receptors orthosteric binding site was employed to predict the binding modes of 3 different ligands (agonist/antagonist diastereomer pair and AR231453) known to bind to GPR119. Ligand interaction energies between GPR119 homology model residues and docked ligands were

analyzed. Residues having the strongest calculated interactions with the ligands were hypothesized to be important for receptor activation.

3.1.1 Ligand Conformational Analysis

Complete conformational analyses were performed on the Arena Pharmaceutical compound, AR231453, and the agonist/antagonist diastereomer ligand pair discussed previously. MacroModel 10.3 (Schrodinger Inc., Portland, OR) was used to explore the conformational space around each rotatable dihedral in each ligand. A coordinate scan was performed on each rotatable dihedral in increments of 60°. Local energy minimum conformers were selected, with respect to potential energy. The selected local energy minimum conformers of each ligand were optimized using Hartree-Fock ab initio method at the 6-31G* level of theory as encoded in Jaguar 7.9 (Schrodinger Inc., Portland, OR). The optimized local energy minimum conformers and the global energy minimum conformer was used in docking studies.

3.1.2 Homology Modeling Using GPCR Crystallized Structures

Homology models for the inactive (R) and active (R*) GPR119 structures were constructed using the Prime module implemented in the Schrodinger Suite 2014. The GPR119 R structure was built by manually aligning on the conserved residues and motifs of the GPR119 amino acid sequence, with Adenosine A₂A receptor (PDB ID 3EML) structure and the sphingosine-1-phosphate receptor (PDB ID 3V2Y) structure. Similarly, the GPR119 R* structure was built by manually aligning the conserved residues and motifs of the GPR119 amino acid sequence with the Adenosine A₂A receptor (PDB ID 2YDO) structure and the Sphingosine 1-phosphate receptor (PDB ID 3V2Y) structure. Energy-based, (algorithm refines residues that do

not come from templates based on energy) composite models (models are produced from specific regions from each PDB template) were produced for both R and R* states. Template regions for building the GPR119 R structure were: PDB 3EML for TMHs 1, 4-7 and the IC1-IC3, EC2, and EC3 and PDB 3V2Y for helices 2 and 3 and EC1. Template regions for building the GPR119 R* structure were: PDB 2YDO for TMHs 1, 4-7 and IC1-IC3, EC2, and EC3 and PDB 3V2Y for TMHs 2 and 3 and EC1.

Preparation and a restrained minimization of the GPR119 R and R* homology models was performed using the Protein Preparation Wizard in Macromodel 10.4 (Schrodinger Inc., Portland, OR). To prepare the protein hydrogen atoms were added using PROPKA at a pH of 7.4 and a restrained minimization was performed using the OPLS 2005 force field until heavy atoms of the receptors converged to a RMSD of 0.3 Å. EC and IC loops were refined using Prime's loop refinement module, employing the variable dielectric surface generalized Born (VSGB) was selected as the solvent model.

3.1.3 Molecular Docking Studies

The automatic docking program Glide 6.2 (Schrodinger Inc., Portland, OR) was used to find the optimal placement of the global minimum conformer, for each GPR119 ligand used in this study, within the GPR119 binding pocket. Glide was used to define the docking area of the GPR119 binding pocket by generating a 30x30x30 Å grid. The center of the grid was defined by the center of mass of following residues: W265^{7.39}, R81^{3.28}, and W238^{6.48}. Glide will only attempt to dock ligands within the area specified by the grid coordinates and will thoroughly explore all potential binding conformations available to each ligand. Flexible docking with standard precision (SP) was applied to all docks and no additional constraints were added to the

docking setup. The maximum number of different docking poses produced by Glide was set to 100.

3.1.4 Ligand/Receptor Minimization

Conjugant gradient energy minimization was performed on accepted ligand/receptor complexes to resolve steric clashes. Minimization was carried out using the OPLS 2005 all-atom force field in Macromodel 10.4 (Schrodinger Inc., Portland, OR). Extended cutoffs, for calculating non-bonded terms for the potential energy function, (nonbonded, 8.0 Å; electrostatic, 20.0 Å; hydrogen bonding, 4.0 Å) were used in each stage of the energy minimization calculation. To represent the different dielectric environments that transmembrane proteins are exposed to, energy minimization of the ligand/receptor complexes were performed in two steps.^[11]

The first step was the minimization of the TMH region. During this step the ϕ , ψ , and ω backbone dihedral angles of the EC and IC loops were highly constrained with 250 kcal/mol of force. Additionally, the EC and IC loops were uncharged during this step. A harmonic constraint was placed on all the TMH backbone torsions (ϕ , ψ , and ω); this was done to preserve the general shape of the helices during minimization. The force was gradually released and the TMH region minimized to convergence. Harmonic constraints of 250 kcal/mol were placed on dihedrals of the ligand to maintain its shape. The conjugant gradient minimization consisted of 1000 steps and in the last 250 steps ligand constraints were released to allow the ligand more flexibility to be able to adapt to the binding pocket. The second step was the minimization of the EC and IC loop regions. Loops were recharged, The TMH region and the ligand were held frozen, but the loops were allowed to relax during this step. The EC and IC loops were

minimized for 5000 steps in a Generalized Born/Surface Area (GB/SA) continuum solvation model for water as implemented in Macromodel 10.4.^[11]

3.2 Site Directed Mutagenesis

From docking experiments using the program Glide, residues R81 and W265 were hypothesized to contribute significantly to receptor binding and activation. Additionally, there is another positively charged residue, R262⁷⁻³⁶, facing the putative binding site, which may or may not play a role in binding and activation of GPR119. To assess the level of residue contribution to receptor activation, these protein residues were mutated to R81L, R262L, and W265A. Mutations were carried out using the QuikChange II Site-Directed Mutagenesis Kit from Agilent Technologies and mutagenic primers were designed following manufacturer's protocol.

3.2.1 Human GPR119 cDNA Clone

The human GPR119 cDNA expression vector consisted of pCMV6 containing the 1008-bp open reading frame of GPR119. The overall size of the vector plus insert was 5.9kb and it was purchased from OriGene (RC216685). This clone was used as a template to create point mutations and/or replace amino acids during site directed mutagenesis. The pCMV6-Entry plasmid vector containing GPR119 cDNA, Myc-DDK tags, and kanamycin/neomycin selectable markers, served as a vehicle for all transformation and transfection experiments.

3.2.2 Mutant Primer Design and Synthesis

Primers used for site-directed mutagenesis were synthesized and HPLC purified by Integrated DNA Technologies and ranged between 29-31 base pairs (bp) length, 58-73% GC

content, a 78.8-80.3 °C melting temperature, and a less than 10% bp mismatch. Sequences of the forward and reverse primers for each mutation are listed in Table 2. Mutagenesis primers were resuspended in 1X TE buffer (10mM Tris, pH 7.5, 1mM EDTA) and final primer concentrations ranged between 177-191 ng/μl.

Table 2: List of Primers Created for Use in Site-Directed Mutagenesis Experiment. Primers listed introduce single or triple amino acid mutations into the GPR119 protein sequence.

R81L Mutation	
Wild Type R (CGG)	
Forward Primer 5'	CC CTG TGC AGC CTG CTG ATC GCA TTT GTC AC 3'
	L C S L L M A F V
Reverse Primer 5'	GT GAC AAA TGC CAT CAG CAG GCT GCA CAG GG 3'

R262L Mutation	
Wild Type R (CGG)	
Forward Primer 5'	C CTA GTG CTG GAA CTG TAC CTG TGG CTG C 3'
	L V L E L Y L W L
Reverse Primer 5'	G CAG CCA CAG GTA CAG TTC CAG CAC TAG G 3'

W265A Mutation	
Wild Type W (TGG)	
Forward Primer 5'	G GAA CGG TAC CTG GCC CTG CTC GGC GTG GG 3'
	E R Y L A L L G V
Reverse Primer 5'	CC CAC GCC GAG CAG GGC CAG GTA CCG TTC C 3'

3.2.3 Mutant Strand Synthesis using Polymerase Chain Reaction (PCR)

In site-directed mutagenesis, a PCR reaction is used to introduce desired mutations into template DNA. The reaction was performed using a double stranded, plasmid DNA template containing the human GPR119 gene sequence.

3.2.3a PCR Mutagenesis Reaction. A PCR reaction was performed for each mutant using a BIO RAD T100 Thermal Cycler and mutagenesis primers specific for each desired mutation.

Each PCR reaction was prepared using the following reagents :

- 1) 5µl 10x reaction buffer
- 2) 2µl purified template GPR119 DNA (200ng)
- 3) 1µl forward primer (177-191ng)
- 4) 1µl reverse primer (177-191ng)
- 5) 2µl dNTPs (25mM)
- 6) 6µl of 25% dimethyl sulfoxide (DMSO)
- 7) 32µl double distilled H₂O
- 8) Lastly 1µl of Pfu Ultra DNA polymerase (2.5U/µl) was added to the reaction

Thermal Cycler parameters were fixed and utilized at 94°C for 1 min and 94°C for 50 sec for template strand denaturation, 56°C for 50 sec for primer annealing, 68°C for 28 min for mutant strand elongation. These steps were repeated 9 times to produce an abundant amount of mutated plasmid DNA. A final elongation step was then set at 72°C for 30 min and after cycle completion reactions were held at 4°C.

3.2.3b Digestion of PCR Reaction. Cleavage of PCR reaction products is essential to remove parental methylated and hemimethylated DNA from the reaction mixture.¹ To facilitate the degradation of parental DNA, 1µl of the restriction enzyme Dpn I was added to each PCR reaction and mixed gently. The cleavage reactions were incubated in a 37°C water bath for 2½ hours.

3.2.4 Bacterial Transformation and Sequencing of GPR119 Plasmid Mutant DNA

Transformation of the Dpn I-digested mutant GPR119 DNA into *E. Coli* cells was performed using the following protocol:

- 1) XL1-Blue cells were thawed on ice for 20 minutes.
- 2) For each reaction 50µl of XL1-Blue cells was pipetted into a pre-chilled 5ml falcon tube.
- 3) 5µl of Dpn I-digested mutant DNA was mixed into each reaction.
- 4) Cell/DNA mixture was placed on ice for 20 minutes.
- 5) Reactions were then heat-shocked in a 42°C water bath for 45 sec.
- 6) 500µl of SOC media preheated to 42°C was added to each reaction.
- 7) Reaction tubes were placed in a 37°C shaker at 200 rpm for 45 minutes.
- 8) 200µl of each reaction was plated on 2xYT Agar plates supplemented with kanamycin at a working concentration of 25 µg/ml.
- 10) Plated reactions were incubated overnight at 37°C and colonies of transformed cells were picked the following day.
- 11) Picked colonies were placed in 3ml of 2xYT-Broth supplemented with kanamycin and were incubated overnight at 37°C with shaking.
- 12) Bacterial cells were harvested the next day by centrifugation at 10,000 rpm for 3 minutes at room temperature.
- 13) Plasmid DNA was purified using the Qiagen QIAprep Spin Miniprep Kit following the manufacturer's protocol.

Purified plasmid DNA was sequenced by Functional Biosciences, Inc. using sequencing primers

Fwd. 5' GGACTTTCCAAAATGTTCG 3'

Rev. 5' ATTAGGACAAGGCTGGTGGG 3'

The resulting sequences were checked for desired mutations. Samples displaying the desired mutations were then midi-prepped to bring mutant DNA concentrations to $\sim 1\mu\text{g}/\mu\text{l}$.

3.3 Measurement of Mutant and Wild Type GPR119 Receptor Activation

GPR119 signals through the $G\alpha_s$ pathway and is known to produce an increase of cAMP upon activation. The importance of specific GPR119 residues for agonist activation can be assessed through comparing levels of cAMP accumulation in the presence of wild type and mutant receptors. The mutation of residues important for agonist activation into residues that either no longer bind the agonist or have weaker interactions with the agonist results in a significant decrease in receptor activation activity.

3.3.1 Mammalian Cell Culture

HEK293 cells were obtained from Dr. Megan Keniry (Biology Dept. UTRGV). Cells were grown in Dulbecco's modified Eagles Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin ($10,000\text{ U ml}/10\text{mg ml}^{-1}$). Cells were cultivated at 37°C in a humidified 5% CO_2 incubator in 100mm plates until ready for experiments (80-90% confluent).

3.3.2 Transient Transfection of Mutant and Wild Type GPR119 DNA

For the transient transfection of the GPR119 wild type gene, Western Blot Analysis was used to assess optimal transfection reagent and DNA concentration volumes. Anti-Myc and Anti-Flag antibodies were used to probe a Western Blot membrane for GPR119 protein. The optimization of transfection reagents and DNA concentrations was performed by Liza Morales Smith (Dr. Dae Joon Kim's Lab; Edinburg Regional Academic Health Center). Results indicated

that 2.5ug of DNA and 3.75µl of transfection reagent was optimal for successful transfection of mutant and wild-type GPR119 DNA into HEK293 cells (Figure 4).

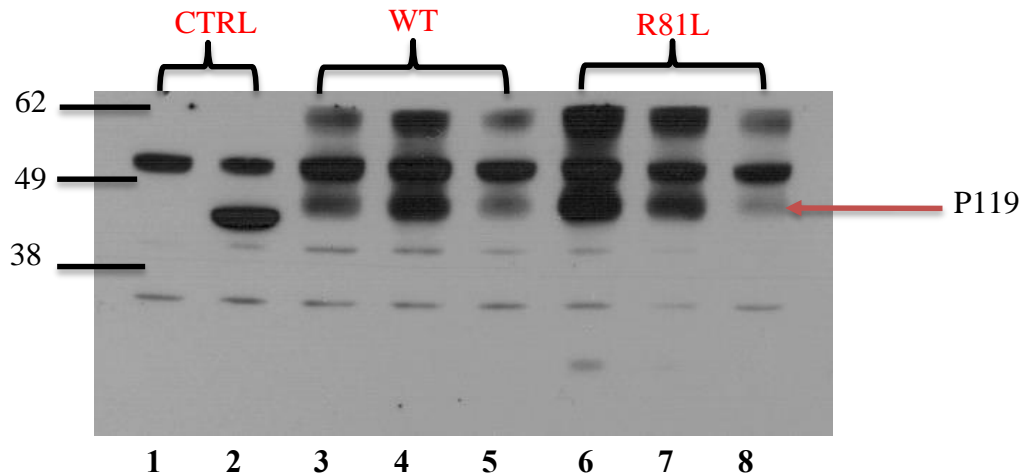


Figure 4: Western Blot Analysis of Transfection Efficiency Results. Blot used Anti-Myc Antibody and lanes are as follows:

1. Cells only
2. TC-PTP (45kD) plasmid [2.5ug] + 3.75ul lipofectamine
3. P119 WT plasmid [1ug] + 3.75ul lipofectamine
4. P119 WT plasmid [2.5ug] + 3.75ul lipofectamine
5. P119 WT plasmid [5ug] + 3.75ul lipofectamine
6. P119 R81L plasmid [1ug] + 3.75ul lipofectamine
7. P119 R81L plasmid [2.5ug] + 3.75ul lipofectamine
8. P119 R81L plasmid [5ug] + 3.75ul lipofectamine

HEK293 cells were plated one day prior to transfection in DMEM supplemented with 10% FBS without antibiotics. Cells were plated in 6-well plates at a density of 500,000 cells/well. The following day, cells were between 80-90% confluent and were transfected with mutant and wild-type GPR119 DNA using the Lipofectamine[®] 3000 Reagent kit (ThermoFisher Scientific, Inc.). In accordance with the optimization results, each transfection was performed using 3.75µl of Lipofectamine[®] Reagent, 2.5µg of mutant or wild type DNA, and 5µl P3000[™] Reagent. Transfections were performed following the manufacturer's protocol. Transfected cells were incubated for 4 hours at 37°C in a humidified 5% CO₂ incubator. Following incubation,

media was completely removed and replaced with 2ml of DMEM supplemented with 10% charcoal stripped FBS per well. Cells were incubated overnight, and experiments were performed the following day.

3.3.3 Stimulation of GPR119 Transfected Cells With Agonist Compound AR231453

The potent GPR119 selective agonist AR231453 (molecular weight: 505.522g/mol) was purchased from Enzo Life Sciences, Inc. and was used to stimulate the transfected GPR119 receptors. A 10mM stock solution of the AR231453 compound was prepared using 10.3mg of compound in 2.03 ml of 100% DMSO. The AR231453 stock solution was subsequently diluted into concentrations: 10,000nM, 1,000nM, 100nM, 10nM, 1nM, 0.1nM and 0.01nM in DMEM with 2.5% csFBS and 0.2% DMSO. Before the application of each drug concentration, cell media was again completely aspirated away and cells were washed with 1ml/well of Hank's Balanced Salt Solution (HBSS). 1ml of DMEM was then reapplied to each well and experiments were performed in DMEM without added supplements. Each compound dilution was added drop-wise to the GPR119 transfected cells. Cells with added compound were then incubated for 30min. at 37°C with 5% CO₂.

3.3.4 Preparation of Cell Samples for cAMP Assay

After AR231453 application and incubation, the cells were gently scraped off from each well. Cells and media from each well were transferred separately into 15ml conical tubes and centrifuged at 4°C, 180 x g for 5 minutes. Media was then removed from the pelleted cells by aspiration. Cells were then lysed with 285µl of lysis buffer per sample (50 mM Tris, 0.1% BSA, 2% Triton X-100, 0.01% Thimerosal, pH 6.0) and incubated on ice for 20 minutes. Samples were

transferred into 1.7ml microcentrifuge tubes and centrifuged at 13,000rpm for 10 minutes. 200 μ l of each sample was placed into a fresh 1.7 microcentrifuge tube and placed on ice.

3.3.5 Preparation of cAMP Standards for Assay Application

cAMP standards were prepared according to the manufacturer's protocol. Standards followed a 10-fold dilution pattern and ranged from 100 μ M to 100pM. Additionally, a cAMP standard series of 3-fold dilutions was prepared, ranging from 32nM to 10pM. New cAMP standards were prepared and tested for each experiment.

3.3.6 cAMP ELISA Assay

A colorimetric cAMP ELISA kit, purchased from Cell Biolabs Inc., was used to assess the level of cAMP accumulation in GPR119 wild-type and mutant transfected HEK293cells after activation by AR231453. The assay is a competitive enzyme immunoassay whose reagents and antibodies compete for cAMP binding. The colored product that is formed upon assay completion is inversely proportional to the amount of cAMP in the sample. Colorimetric cAMP ELISA microtiter plates were prepared and tested following the manufacturer's protocol and absorbance was read at 450nm on a BioRad 480 micro-plate reader.

3.3.7 Data Analysis

Assay data was analyzed using the SigmaPlot 11 software (Systat Software Inc., San Jose, CA). Standard curve graphs were generated by plotting known concentrations, on a logarithmic scale, vs. absorbance. Unknown sample concentrations were determined from the standard curve as picomolar (pM) concentrations of cAMP. After the unknown concentrations

were determined, the sigmoidal does-response w/ hillslope equations in SigmaPlot was used for graphing and determining EC_{50} values for WT and mutant GPR119 constructs.

CHAPTER IV

RESULTS AND DISCUSSION

This chapter discusses the molecular modeling and docking results used as a guide for selecting amino acid mutations important for receptor activation. In addition, experimental results confirming desired mutations and dose response curves showing how selected mutations effect GPR119 activation will be presented.

4.1 GPR119 Models

Inactive (R) and active (R*) GPR119 models were constructed using the computational software Prime, as described in the experimental methods section. The selection of templates used to construct each GPR119 model with Prime was based on sequence identity to known crystalized structures available in the Protein Data Bank (PDB). When comparing amino acid similarity within the TMH regions of various class A GPCRs in the PDB, it was found that the adenosine A₂A receptor (A₂A) and the sphingosine-1-phosphate receptor (S1P1) share a 27.35% and 25.11% sequence identity to GPR119 respectively.^[20]

4.1.1 Template Selection for GPR119 Models

The choice to use two different models (A₂A and S1P1) for GPR119 R and R* construction was based on the lack or presence of key structural features in TMH2 and TMH3. In TMH2, the A₂A structure contains a proline (P61^{2.59}) towards its extracellular end. This proline

creates a kink in TMH2 of the A₂A crystal structures found in the PDB. The GPR119 sequence doesn't contain any prolines within its TMH2 region. Therefore, TMH2 should be a relatively straight helix for the GPR119 structure. S1P1 was chosen to model the TMH2 region of GPR119 because the S1P1 TMH2 sequence does not contain a proline throughout the middle portion of its TMH2 helical segment, as seen in figure 5. S1P1 does contain a proline (P79^{2,38}) toward its intracellular end; however, this proline does not cause any kinks within TMH2 for S1P1.

The extracellular end of TMH3 in GPR119 contains a positively charged arginine residue (R81^{3,28}), see figure 5. Upon inspection of the ligands known to bind to GPR119, the majority of these ligands contain highly electronegative functional groups, such as, sulfone, carboxylate, and carbamate groups. GPR119's R81^{3,28} has the potential to be an important residue for ligand binding and activation of this receptor. The correct placement of this residue in the GPR119 models is crucial for capturing the correct orientation of each ligand when bound to the GPR119 receptor. The S1P1 receptor (PDB ID 3V2Y) structure places GPR119's R81^{3,28} within the binding pocket for both R and R* models and was therefore used as a template for the TMH3 region of GPR119. It is also important to note that GPR119 contains a second positively charged arginine R262^{7,36} that could potentially be facing into the ligand-binding pocket. However, our experimental results indicate that R262^{7,36} does not play a significant role in receptor activation.

Figure 5 shows an example of the alignment created that was used to construct homology models for R and R* states of GPR119. Sequences were aligned on the conserved residues and motifs shown in yellow and blue respectively. Although, the amino acid sequence between Adenosine A₂A receptor structures published in the Protein Data Bank are exactly the same, some are captured by x-ray crystallography in the R* and R conformation, different Adenosine A₂A receptor structures were used to capture structural differences between the R* and R states.

GPR119	(1)	M ESSFSFGVILAVLASLIATNTLVAVAVLLLIHKNDG	
A2A	(2)	I MGSSVYITVELAIAVLAAILGNVLCWAVWLNSNLQN-	
S1P1	(42)	E NSIKLTSVVFILICCFIILENIFVLLTIWKTKKFHR-	
		TMH1	IC1
GPR119	(39)	V SLCFTLNLAVADTLIGVAISGLLTDQLSSPSRPTQ	
A2A	(40)	V TNYFVVS ^Y LAAADIAVGVLAIPFAITISTGFCAA--	
S1P1	(79)	P MYFIGNLALS ^Y DLLAGVAYTANLLLSGATTYKLT-	
		TMH2	EC1
GPR119	(75)	K TLC ^Y SLRMAFVTSSAAASVLTVM ^L ITFD ^R RYLAIKQ ^P FFRYLKIMS	
A2A	(74)	C HGCLFIACFVLVLTQSSIFSLLAIAIDRYIAIRI ^P PLRYNGLVT	
S1P1	(114)	P AQWFLREGSMFVALSASVFSLLAIAIERYITMLKMKLHNGS-N	
		TMH3	IC2
GPR119	(119)	G FVAGACIAGLWLVSYLIGFLPLGIPMFQQTAYKG-----	
A2A	(118)	G TRAKGIIAICWVLSFAIGLTPMLGWN ^N CGQPKEGK ^N HNSQGCGE	
S1P1	(157)	N ERLFL ^L ISACWVISLILGGLPIMGWNCISALSSCSTVLP ^L YH-	
		TMH4	EC2
GPR119	(154)	---QCS ^H FAV ^F --HPHFVLT ^L SCV ^G FFPAM ^L LLFVFFYCD ^M LKIA	
A2A	(162)	GQVAC ^L LEDV ^V PMN ^Y MVYFNFFACV ^L V ^P LL ^L MLGVYLRI ^F LAAR	
S1P1	(200)	-----KHYI ^L PCIT ^V FTLL ^L LSIVILY ^C RIYSLVR	
		EC2	TMH5
GPR119	(193)	SMHS ^Q QIRKMEHAGAMA ^G GGYRS ^P RTPS	
A2A	(206)	RQLK ^Q MESQPLP-GERA---RS ^T TLQK-	
S1P1	(230)	TRSRRLTFRKNISK-----AS ^R RSSE	
		IC3	
GPR119	(214)	D FKALRTVSVLIGS ^F FALS ^T WTPFLITGIVQ ^V AC ^Q EC-HLY	
A2A	(228)	E VHAAKSLAIVGLFALCWLPLHIINCFT ^F FC ^P DC ^S HAP	
S1P1	(250)	N VALLKTVIIVLSVFIACWAPLFILL ^L LDV ^G CKVKTCDI	
		TMH6	EC3
GPR119	(258)	I VLERYLWLLGVGNSLLNPLIYAYWQ ^K EVRLQLYH ^M AL	
A2A	(267)	L WLMYLAIVLSHTNSV ^V NPF ^I YAYRIREFRQ ^T FRKIIR	
S1P1	(289)	L FRAEYFLVLAVLNSGTNPIIY ^T LTN ^K EMRRAFIRIMS	
		TMH7	H8

Figure 5: Amino acid Sequence Alignment for GPR119, Adenosine A₂A, and S1P1 receptors. Conserved residues are highlighted in yellow, conserved motifs are highlighted in blue, identically matching residues are highlighted in red, and helical regions are in bolded font.

4.1.2 Inactive and Active GPR119 Models

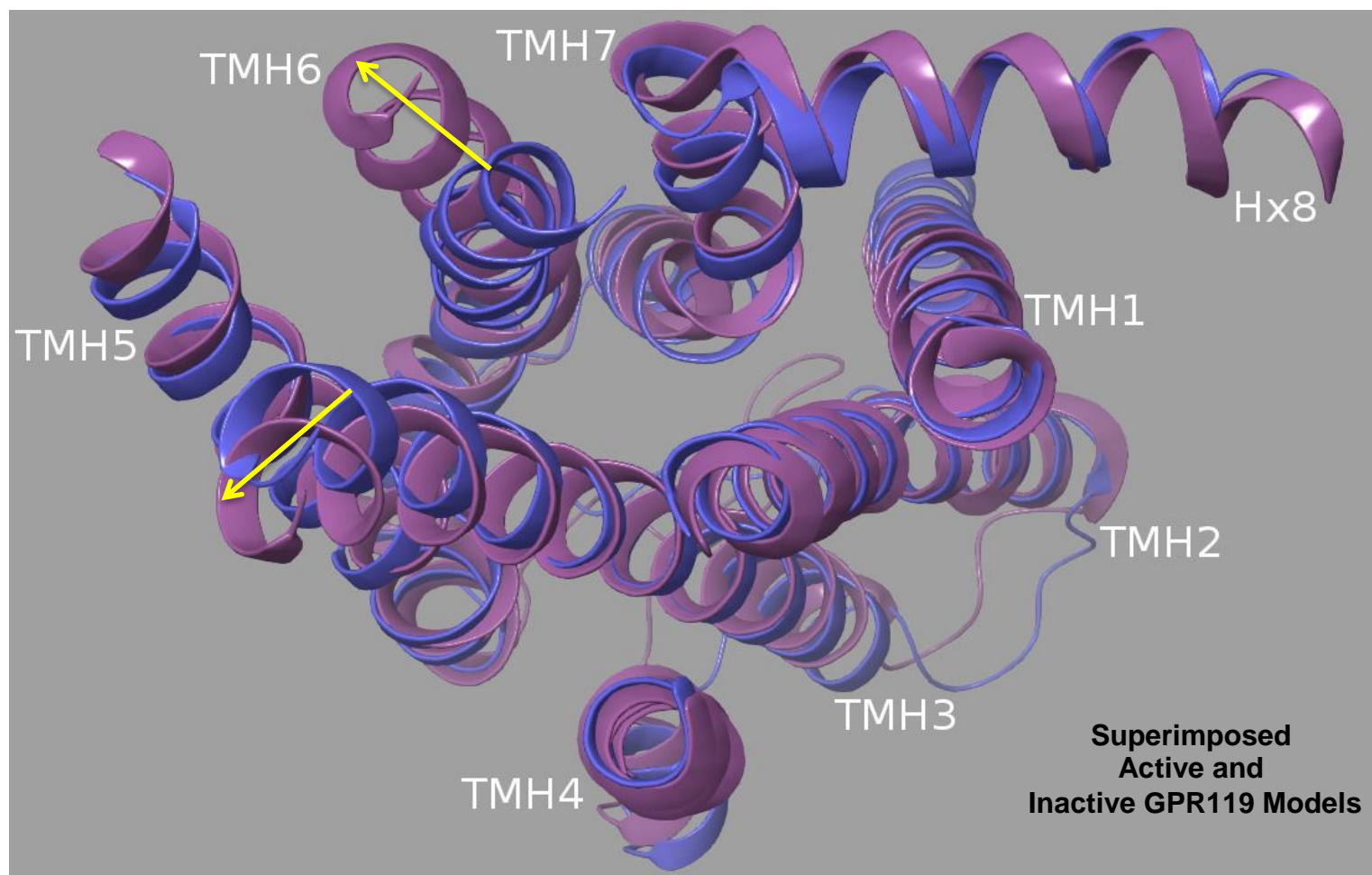


Figure 6: Intracellular View of Active (R^*) and Inactive (R) GPR119 Receptor Models superimposed on TMHs 1, 2, 4, 5, and 7. The R^* model is in purple and the R model is in blue. Yellow arrows indicate the structural changes in TMH3 and TMH6. Intracellular loops are removed for clarity.

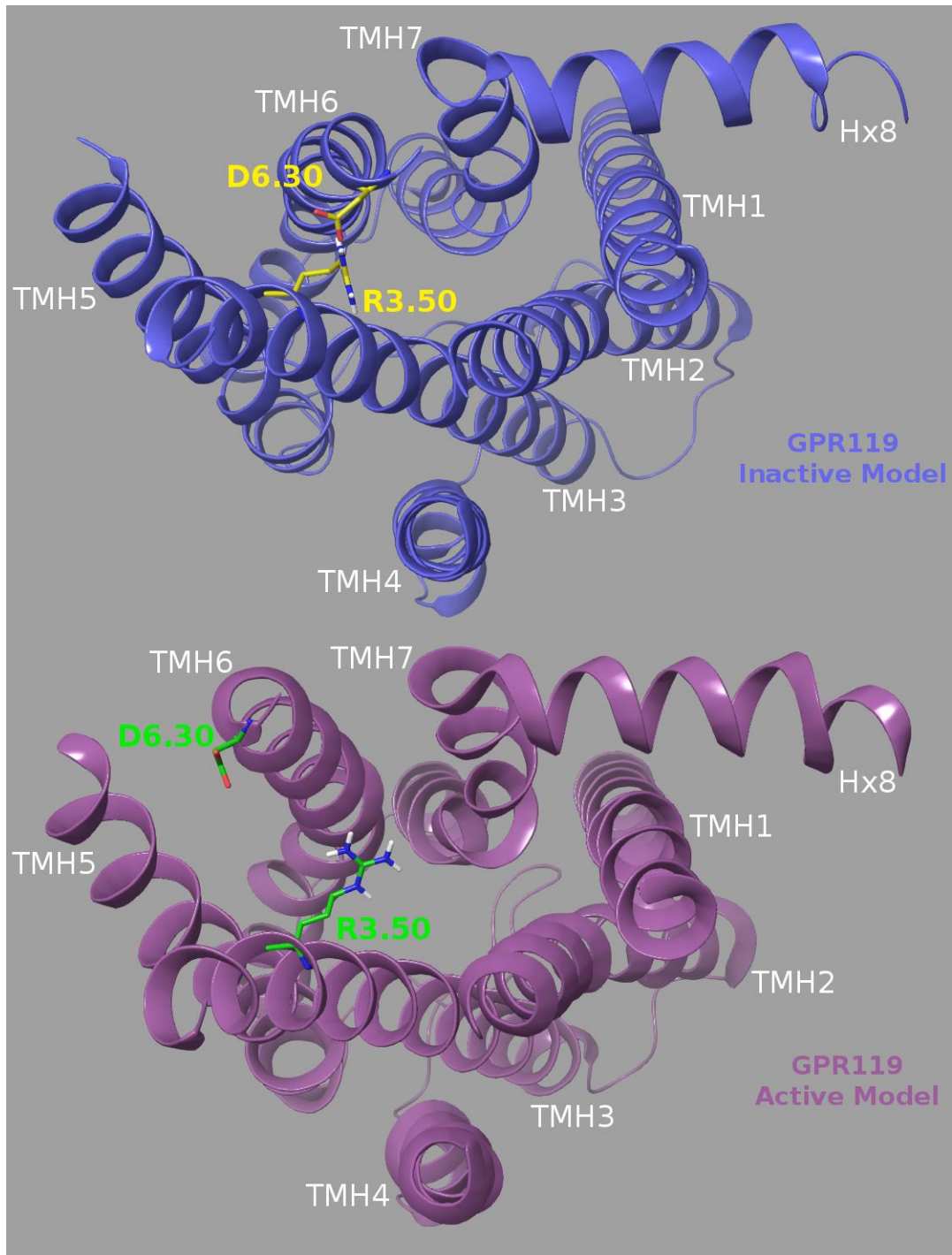


Figure 7: The “ionic lock” in GPR119 is the interaction between the positively charged R3.50 and the negatively charged D6.30 residues that keep the intercellular end of the receptor closed. Intracellular loops have been removed for clarity.

Figure 6 illustrates the main structural differences of the overall shape of the TMHs between the inactive and active GPR119 models. The main difference between the models is the outward movements of the cytoplasmic ends of TMHs 6 and 3. TMHs 1, 2, 4, 5, and 7 remain in similar positions and have the same overall backbone shape inactive and inactive GPR119 receptor models.

Figure 7 illustrates the intact (inactive model) and the broken (active model) ionic lock interaction. The outward movement of the cytoplasmic ends of TMHs 3 and 6 during GPR119 receptor activation results in the disruption of this interaction. The R3.50 residue in this interaction is part of the highly conserved DRY motif of TMH3.

4.1.2 Discussion of Active and Inactive GPR119 Modeling Results.

Experimental evidence to support the outward cytoplasmic rearrangement of both TMH3 and 6 upon GPCR activation include fluorescence spectroscopic studies of the β_2 adrenergic receptor labeled with fluorescent probes, zinc cross-linking studies, and chemical reactivity measurements in constitutively active β_2 adrenergic receptor mutants.^[5] The rearrangement of TMH 3 and 6 between GPR119 R and R* model conformation is shown in Figure 6. The use of multiple templates to construct GPCR homology models is relatively common in the field. Engelstoft and colleagues have reported that the energy of their GPR119 receptor models based on multiple PDB templates were lower compared with the best models developed from single-template PDB structures.^[7]

The use of two Adenosine A_{2A} receptor (A_{2A}) templates, PDB ID 3EML and 3REY, which are inactive and active A_{2A} crystal structures respectively, to construct both R and R* GPR119 receptor models was necessary to capture the structural differences in TMH6 for the GPR119 receptor's R and R* states. However, in the case of TMH3 the sphingosine-1-

phosphate (S1P1) receptor (PDB ID 3V2Y), was used as a template for both GPR119 R and R* models. The TMH3 slight outward movement in the R* GPR119 model can be explained by its loss of the ionic lock interaction with the intracellular end of TMH6. The ionic lock interaction has been proposed by many research groups to hold the ends of TMH3 and TMH6 in GPCRs in a resting state, inactive conformation^[5]. The ionic lock interaction is also found in the inactive state crystal structure of Rhodopsin.^[5] Upon minimization of the initial GPR119 R* receptor model, the cytoplasmic end of TMH3 was able to move outward due to the lack of the ionic lock interaction.

4.1.2a Discussion of Rotameric Differences in TMH6 Residues in the R & R* GPR119 Receptor Models

Four atoms, N-C α -C β -C γ , define the χ_1 (chi1) dihedral angles for amino acid side chains. The rotation of the χ_1 angle takes place around the central C α and C β atoms of the amino acid (figure 8).

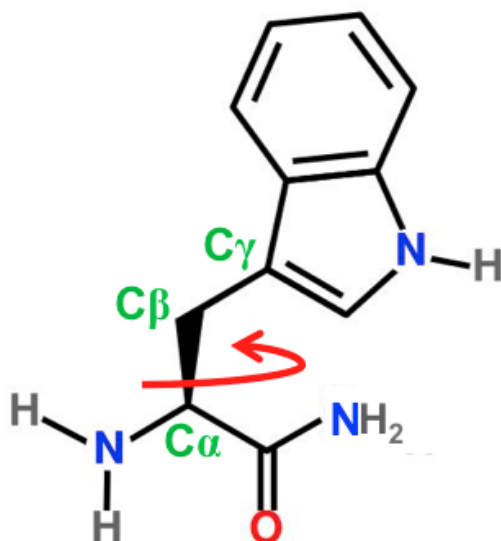


Figure 8: Tryptophan Residue Indicating Rotation of the Chi1 Dihedral Angle

Prior research has indicated that rotamer changes in the χ_1 angles for cysteine and tryptophan from the CWxP motif modulate the bend angle around the highly conserved proline residue in this motif.^[5, 21] These χ_1 angle changes lead to movement of the cytoplasmic end of TMH6, which is correlated with shifting GPCRs between their active and inactive states.^[5] The position of the χ_1 angle for W238^{6,48} is in a trans conformation (roughly $\pm 180^\circ$) for active state GPCRs and in a g+ (roughly -60°) for inactive state GPCRs.^[14, 21] Figure 9 illustrates the different χ_1 conformations for W238^{6,48} in the GPR119 receptor models.

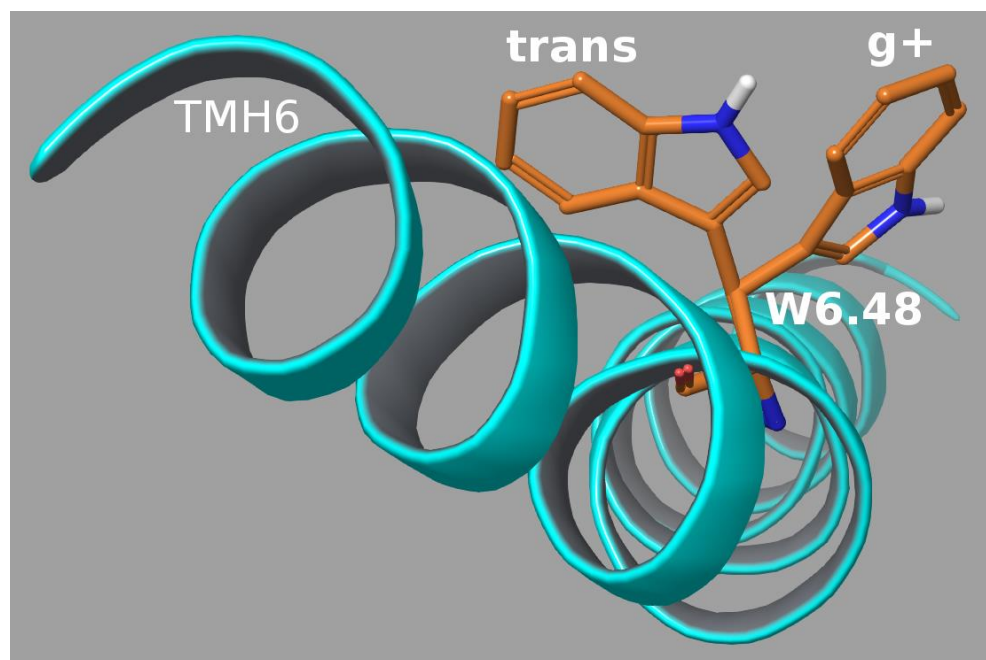


Figure 9: Rotamer Changes in the χ_1 Dihedral Angle for GPR119 Residue W238^{6,48}

4.2 GPR119 Agonist/Antagonist Diastereomer Pair Dock

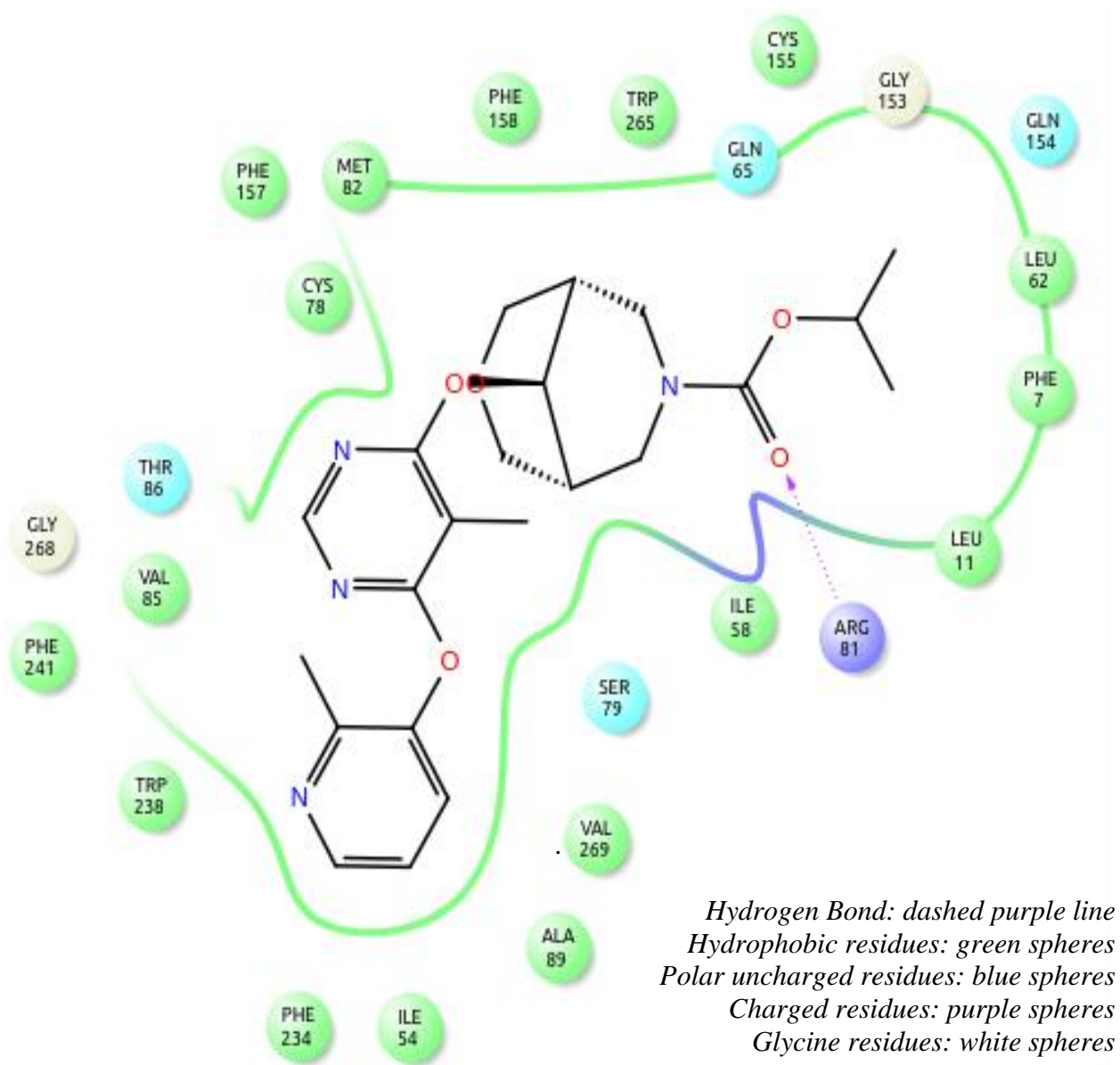


Figure 10: Ligand Plot of the Docked Agonist of the Diastereomer Pair Inside the GPR119 Receptor R* Binding Pocket.

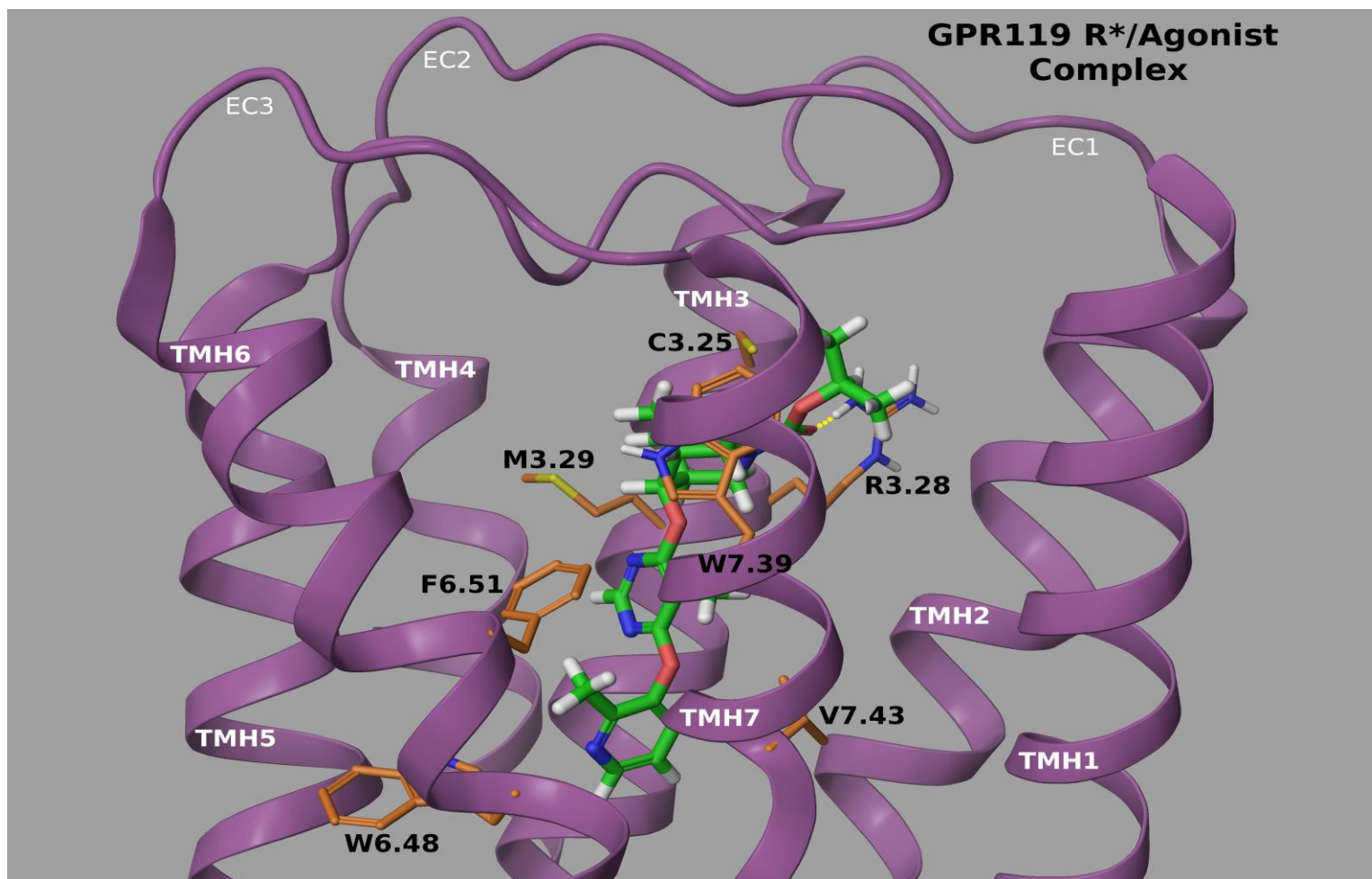


Figure 11: Side View of the GPR119 R* Receptor Model (purple ribbons) with Docked Agonist from the Diastereomer Ligand Pair. Protein residues having significant interactions with the docked agonist ligand are shown in orange. The yellow dotted line represents a hydrogen bond between the residue R81^{3.28} and the ligand

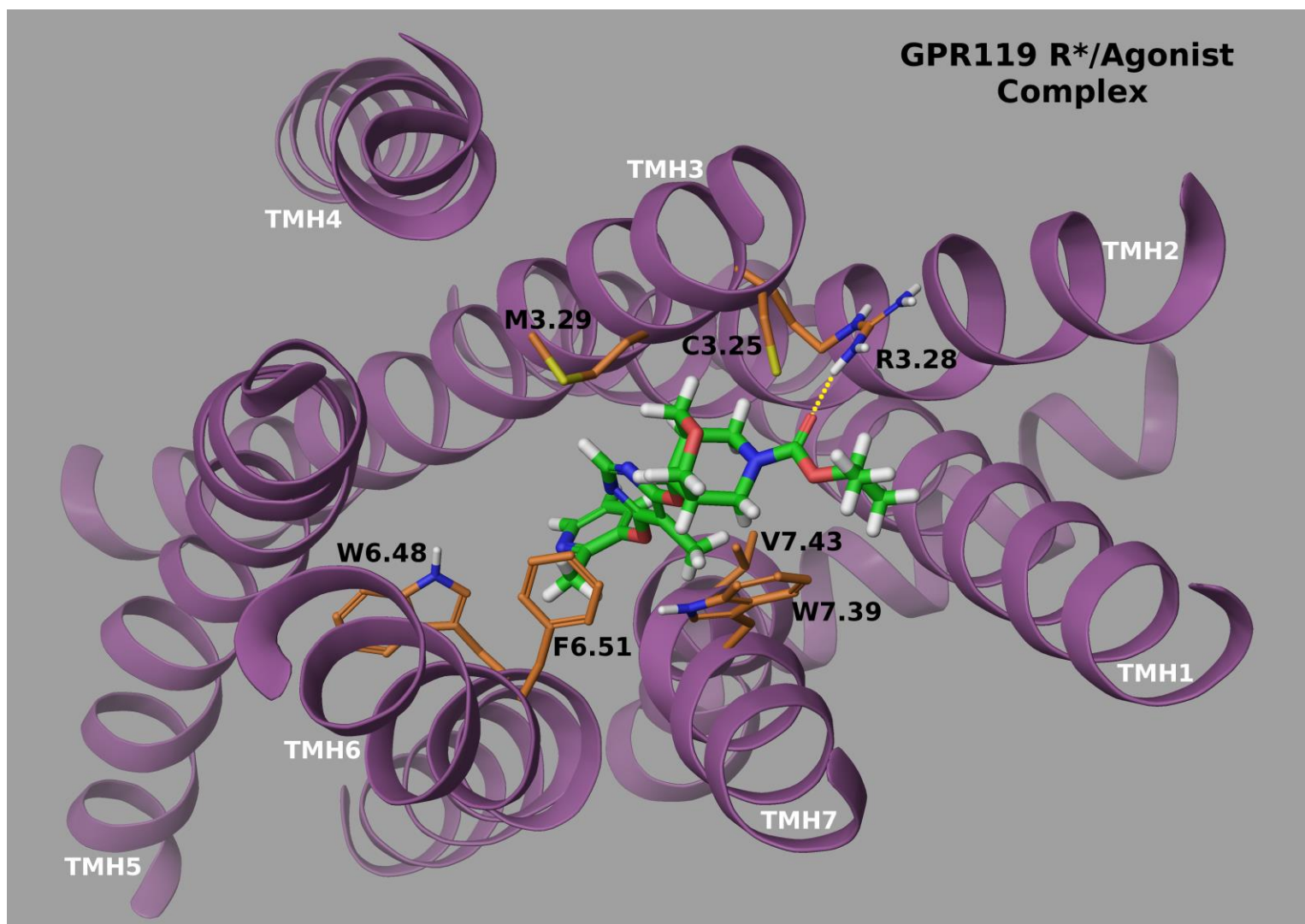


Figure 12: Top View of the GPR119 R* Receptor Model (purple ribbons) with Docked Agonist from the Diastereomer Ligand Pair. Protein residues having significant interactions with the docked agonist ligand are shown in orange. The yellow dotted line represents a hydrogen bond between R81^{3.28} residue and the ligand. EC and IC loops removed for clarity.

Table 3: Interaction Energies Between GPR119 R* Receptor Model Residues and Agonist Ligand. Residues with absolute sequence numbers not followed by Ballesteros and Weinstein numbering are located in the EC loops of the GPR119 Receptor.

GPR119 R* Receptor Agonist Interaction Energies			
Amino Acid	Total Energy (kcal/mol)	Electrostatic Energy	Van der Waals Forces
TRP265 (7.39)	-8.506	-0.0859	-8.4201
ARG81 (3.28)	-7.0426	-6.4126	-0.63
PHE241 (6.51)	-5.8954	-0.23	-5.6654
MET82 (3.29)	-4.1463	-0.1606	-3.9857
CYS78 (3.25)	-3.632	0.0699	-3.7018
VAL269 (7.43)	-3.3539	0.1186	-3.4725
VAL85 (3.32)	-2.6574	-0.0365	-2.621
GLY268 (7.42)	-2.636	0.4122	-3.0483
CYS155	-2.2924	0.3587	-2.6511
PHE157	-2.2656	-0.0338	-2.2317
TRP238 (6.48)	-2.2345	-0.4451	-1.7894
PHE158	-1.4842	0.1184	-1.6026
LEU62 (2.61)	-1.3586	-0.022	-1.3365
GLN154	-1.2076	0.1624	-1.37
ALA89 (3.25)	-1.0153	0.0406	-1.056
ILE54 (2.53)	-0.7961	0.0836	-0.8797
THR86 (3.33)	-0.7115	-0.002	-0.7095
GLY153	-0.7064	-0.1603	-0.5461
SER237 (6.47)	-0.6864	-0.1932	-0.4933
PHE7 (1.35)	-0.5251	0.0589	-0.584
ILE58 (2.57)	-0.4372	0.0205	-0.4577
SER79 (3.26)	-0.4257	0.0368	-0.4625
LEU11 1.39)	-0.4142	0.0172	-0.4314
PHE165 (5.39)	-0.355	0.0096	-0.3646
SER272 (7.46)	-0.31	-0.0794	-0.2306
LEU264 (7.38)	-0.2796	-0.036	-0.2437
ASN271 (7.45)	-0.2282	-0.04	-0.1882
GLN65 (2.64)	-0.2206	0.2057	-0.4263
LEU266 7.40)	-0.1861	-0.0198	-0.1663
PHE234 (6.44)	-0.141	0.0571	-0.1981
SER4 (1.32)	-0.1005	-0.0262	-0.0743
LEU169 (5.43)	-0.0784	0.012	-0.0904
GLY8 (1.36)	-0.0596	0.0176	-0.0772
ARG262 (7.36)	-0.0149	0.2338	-0.2487
Totals	-56.4043	-5.9497	-50.4545
Conformational Cost (kcal)	-0.9864		
Total Docking Energy (kcal)	-55.4179		

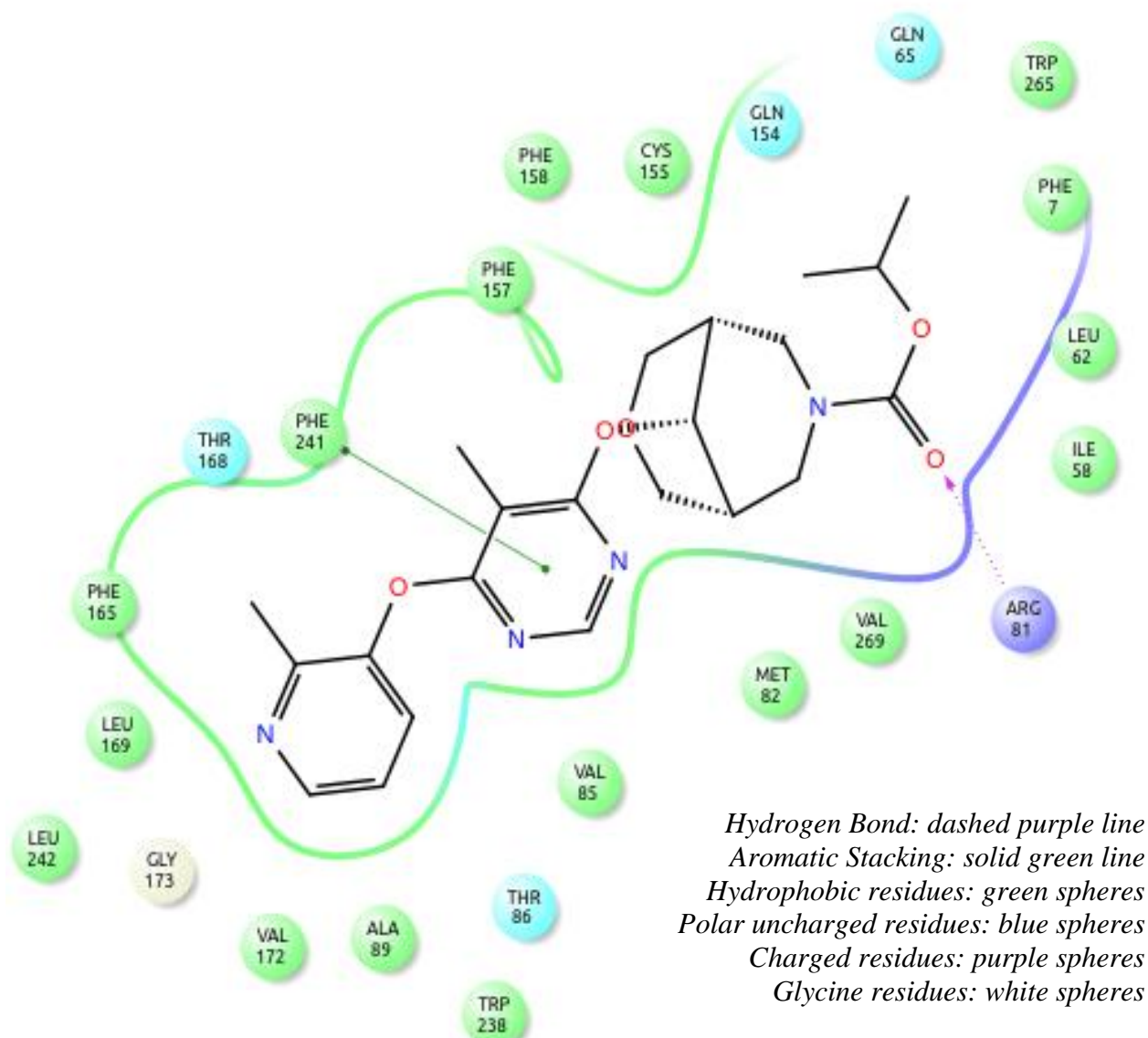


Figure 13: Ligand Plot of the Docked Antagonist of the Diastereomer Pair Inside the GPR119 Receptor R Binding Pocket.

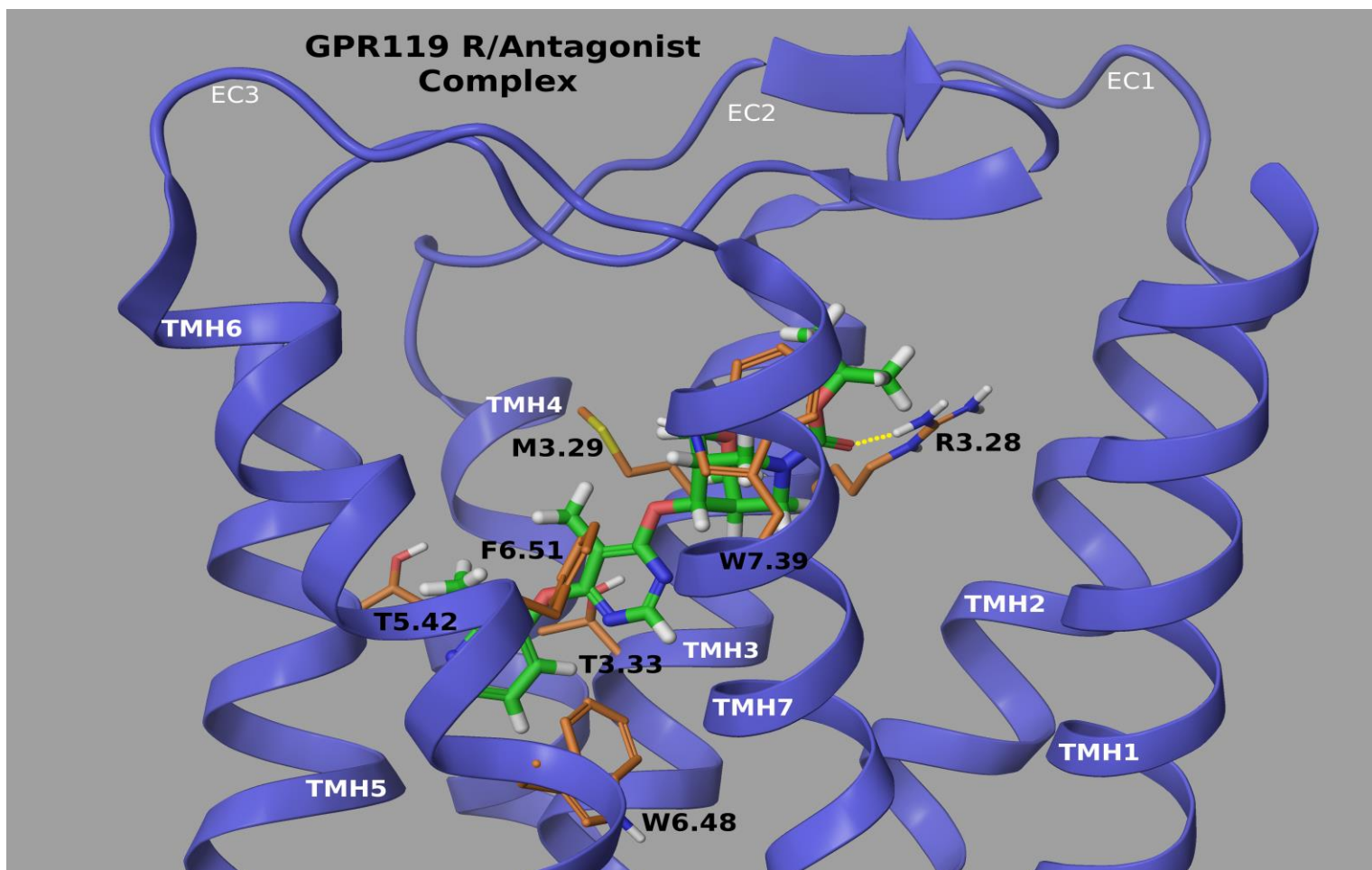


Figure 14: Side View of the GPR119 R Receptor Model (blue ribbons) with Docked Antagonist from the Diastereomer Ligand Pair. Protein residues having significant interactions with the docked antagonist ligand are shown in orange. The dotted line represents a hydrogen bond between R81^{3.28} residue and the ligand.

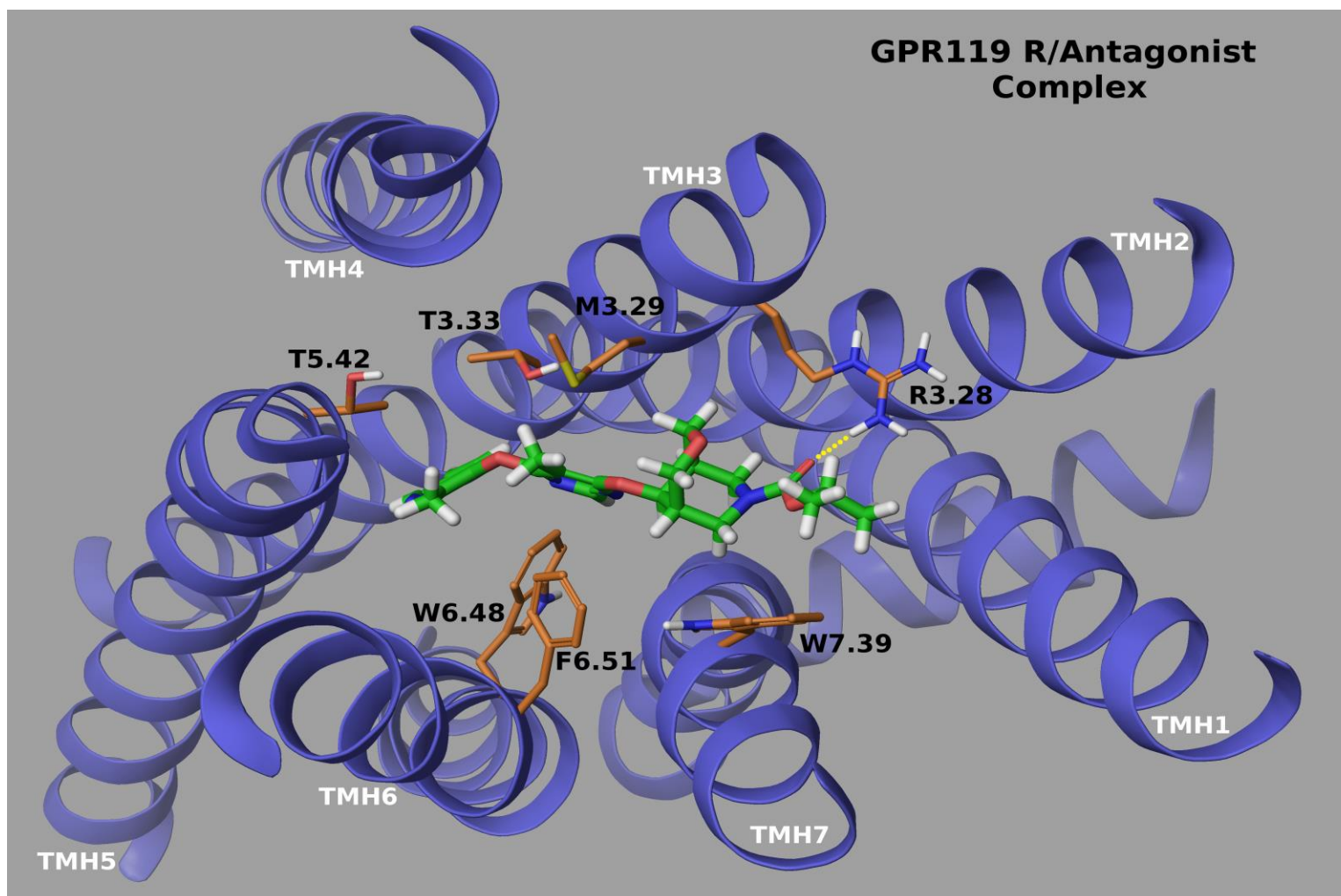


Figure 15: Top View of the GPR119 R Receptor Model (blue ribbons) with Docked Antagonist from the Diastereomer Ligand Pair. Protein residues having significant interactions with the docked antagonist ligand are shown in orange. The yellow dotted line represents a hydrogen bond between R81^{3.28} residue and the ligand. EC and IC loops removed for clarity.

Table 4: Interaction Energies Between GPR119 R Receptor Model Residues and Antagonist Ligand. Residues with absolute sequence numbers not followed by Ballesteros and Weinstein numbering are located in the EC loops of the GPR119 Receptor.

GPR119 R Receptor Antagonist Interaction Energies			
Amino Acid	Total Energy (kcal/mol)	Electrostatic Energy	Van der Waals Forces
ARG81 (3.28)	-8.4576	-6.1297	-2.3279
TRP265 (7.39)	-6.0289	-0.0421	-5.9868
MET82 (3.29)	-5.1049	-0.0628	-5.0421
PHE241 (6.51)	-4.3858	0.021	-4.4068
THR86 (3.33)	-3.7012	0.0793	-3.7805
TRP238 (6.48)	-3.3102	-0.2521	-3.0581
THR168 (5.42)	-3.1743	-0.1309	-3.0434
VAL85 (3.32)	-2.3858	0.0282	-2.414
PHE165 (5.39)	-2.2043	0.1786	-2.3829
VAL172 (5.46)	-2.1425	-0.0489	-2.0935
LEU242 (6.38)	-1.8624	-0.0945	-1.7678
CYS155	-1.7229	-0.0303	-1.6926
PHE157	-1.4928	0.085	-1.5778
LEU169 (5.43)	-1.4533	0.0461	-1.4994
GLN154	-1.0812	0.2726	-1.3538
ALA89 (3.25)	-0.8784	0.0803	-0.9587
CYS78 (2.62)	-0.7794	0.1545	-0.9339
GLY173 (5.47)	-0.6057	-0.1136	-0.492
ILE58 (2.38)	-0.5838	0.0089	-0.5928
PHE158	-0.4582	0.0196	-0.4779
VAL269 (7.43)	-0.4079	-0.0033	-0.4045
LEU62 (2.61)	-0.2832	-0.0121	-0.2711
ALA90 (3.26)	-0.2338	-0.0077	-0.2261
SER156	-0.2141	-0.0062	-0.2079
LEU11 (1.39)	-0.1981	-0.006	-0.1921
GLN65 (2.64)	-0.1737	0.1518	-0.3255
PHE7 (1.35)	-0.1699	0.0288	-0.1988
GLY268 (7.42)	-0.157	0.0101	-0.1671
ILE54 (2.53)	-0.0578	0.0213	-0.0791
VAL93 (3.40)	-0.057	0.0255	-0.0824
ARG262 (7.36)	0.0242	0.1728	-0.1486
Totals	-53.7418	-5.556	-48.1858
Conformational Cost (kcal)	-2.15798		
Total Docking Energy (kcal)	-51.5838		

Figures 10 and 13 are ligand interaction diagrams for the agonist ligand dock and antagonist ligand dock pair. Amino acids interacting with each ligand in the diagrams are located within 4Å of the ligands. Residues are numbered according to the absolute sequence for GPR119 receptor.

Figures 11 and 12 show the docking pose of the agonist and Figures 14 and 15 show the docking pose of the antagonist within the GPR119 receptor-binding pocket. Active and inactive GPR119 receptor models were used when docking the agonist and antagonist of the diastereomer ligand pair respectively. For clarity, only residues having high interactions energies (tables 3 and 4) are displayed even though more residues show interactions with the ligands. In the images, residues are numbered using the Ballesteros and Weinstein GPCR numbering scheme.

Tables 3 and 4 contain the calculated interaction energies between amino acid residues and each docked diastereomer of the ligand pair. Interaction strength was calculated using the OPLS_2005 force field, as implemented in Schrodinger's Macromodel module. The total energy for each residue is a sum of the energy from both electrostatic interactions and van der Waals forces. Negative values indicate attractive forces between ligand and amino acid residues and positive values indicate repulsive forces between ligand and amino acid residues. All of the interaction energies from each residue are summed towards the bottom of Tables 3 and 4 in the totals section. The conformational cost, calculated using Hartree-Fock quantum mechanical theory, for each ligand to adopt the docking pose is subtracted from the total energy to yield the total docking energy for each ligand in the diastereomer pair.

4.2.1 Discussion of Agonist/Antagonist Pair Docking Results

The agonist/antagonist pair docking results contained six residues that interact with both agonist and antagonist ligands. The following residues have a greater than -2.000 kcal/mol

interaction energy and can be found in both agonist and antagonist docks: R81^{3.28}, M82^{3.29}, V85^{3.32}, W238^{6.48}, F241^{6.51}, and W265^{7.39}. Ligand-residue interactions are mostly attractive, Van der Waals interactions for both ligands, with the exception of R81^{3.28}, which has favorable electrostatic interaction with the ligands. The docking cost for each ligand is minimal, which indicates the amount of energy required for each ligand to adopt the docked pose is relatively small. The low docking costs are indicative of probable ligand-binding poses.

The diastereomer pair agonist-binding pocket in the GPR119 receptor encompasses TMHs 2, 3, 6, and 7. The agonist ligand adopts a nearly vertical position in the binding site. The position of the W238^{6.48} residue, from the CWxP motif is in a trans conformation with respect to its χ_1 (chi1) dihedral angle. The positioning of the agonist ligand inside the GPR119 binding pocket prevents the W238^{6.48} from rotating into a g+ (inactive) conformation. Therefore, the receptor would remain in the active state when the diastereomer pair agonist binds. The residue with the largest interaction for the agonist dock is the hydrophobic W265^{7.39} and the only interaction that is predominately an electrostatic interaction is between the agonist carbamate moiety and R81^{3.28}.

There is one more positively charged amino acid that is facing toward the binding pocket, R262^{7.36}, but is partially obstructed by W265^{7.39} and has only weak interactions with the ligand (see last amino acid line of Table 3). The rest of the amino acids with high interaction energies are F241^{6.51}, M82^{3.24}, C78^{3.25}, V269^{7.43}, V85^{3.32}, G268^{7.42}, C155, F157, and the toggle switch residue W238^{6.48}. Most of these interacting amino acids are from TMHs 3, 6, and 7, plus two residues from the EC2 loop. C155 is part of the disulfide-bonded residues that tether the EC2 loop to the top of TMH3. The other, F157, is the aromatic residue second from C155 that has been seen to point in the binding site in most GPCR crystal structures.

The diastereomer pair antagonist-binding pocket in the GPR119 receptor encompasses TMHs 2, 3, 5, 6, and 7. The antagonist ligand adopts a diagonal position inside the GPR119 binding pocket. The W238^{6.48} residue in the antagonist dock is in the g+ conformation. The docking position of the antagonist prevents W238^{6.48} from rotating into the trans (active) conformation. Therefore, the receptor would remain in the inactive state when the diastereomer pair antagonist binds. The residue with the largest interaction for the antagonist dock is R81^{3.28} and this interaction is predominately electrostatic. The antagonist also has a large hydrophobic interaction with W265^{7.39}. Additionally, there is an aromatic stacking interaction between F241^{6.51} and the pyrimidine ring of the antagonist (Figure 13).

In addition, a few different residues, mainly from TMH5 have high interaction energies with the diastereomer pair antagonist. This is due to the more diagonal binding pose of the antagonist. Those residues are T168^{5.42}, F165^{5.39}, and V172^{5.46}.

The other positively charged residue mention earlier, R262^{7.36}, has very weak interactions with the antagonist, just as it has with the agonist ligand of the diastereomer pair. Some of the same amino acids that interact with the agonist interact with the antagonist too. These residues are M82^{3.29}, V85^{3.32}, and the toggle switch residue W265^{6.48}.

4.3 GPR119 AR231453 Agonist Dock

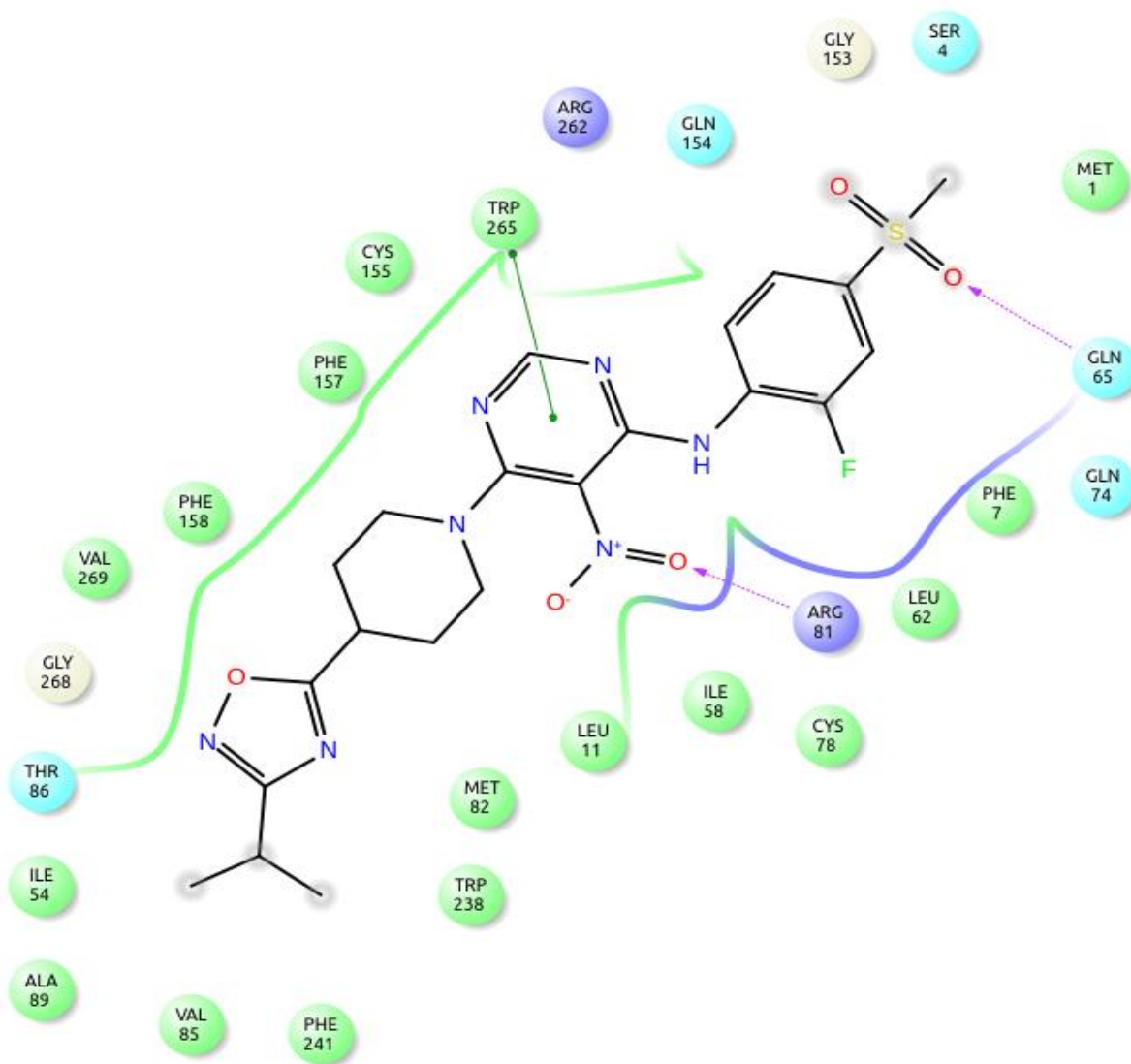


Figure 16: Ligand Plot of the Docked Agonist AR231453 Inside the GPR119 R* Receptor Binding Pocket

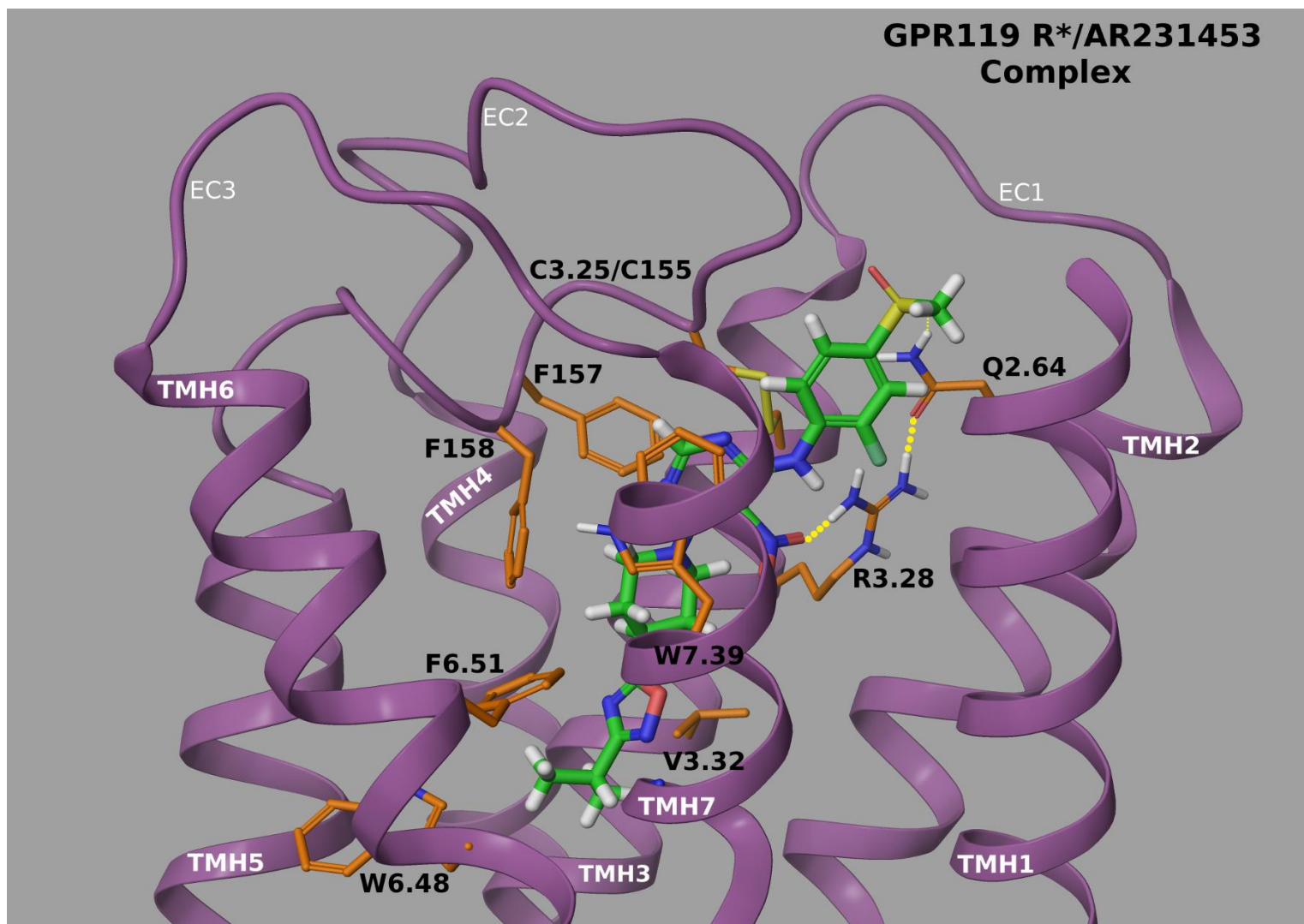


Figure 17: Side View of the GPR119 R* Receptor Model (purple ribbons) with Docked AR231453 Ligand. Protein residues having significant interactions with the docked AR231453 ligand are shown in orange. The yellow dotted line represents hydrogen bonds between R81^{3.28} and Q65^{2.64} with the AR231453 ligand.

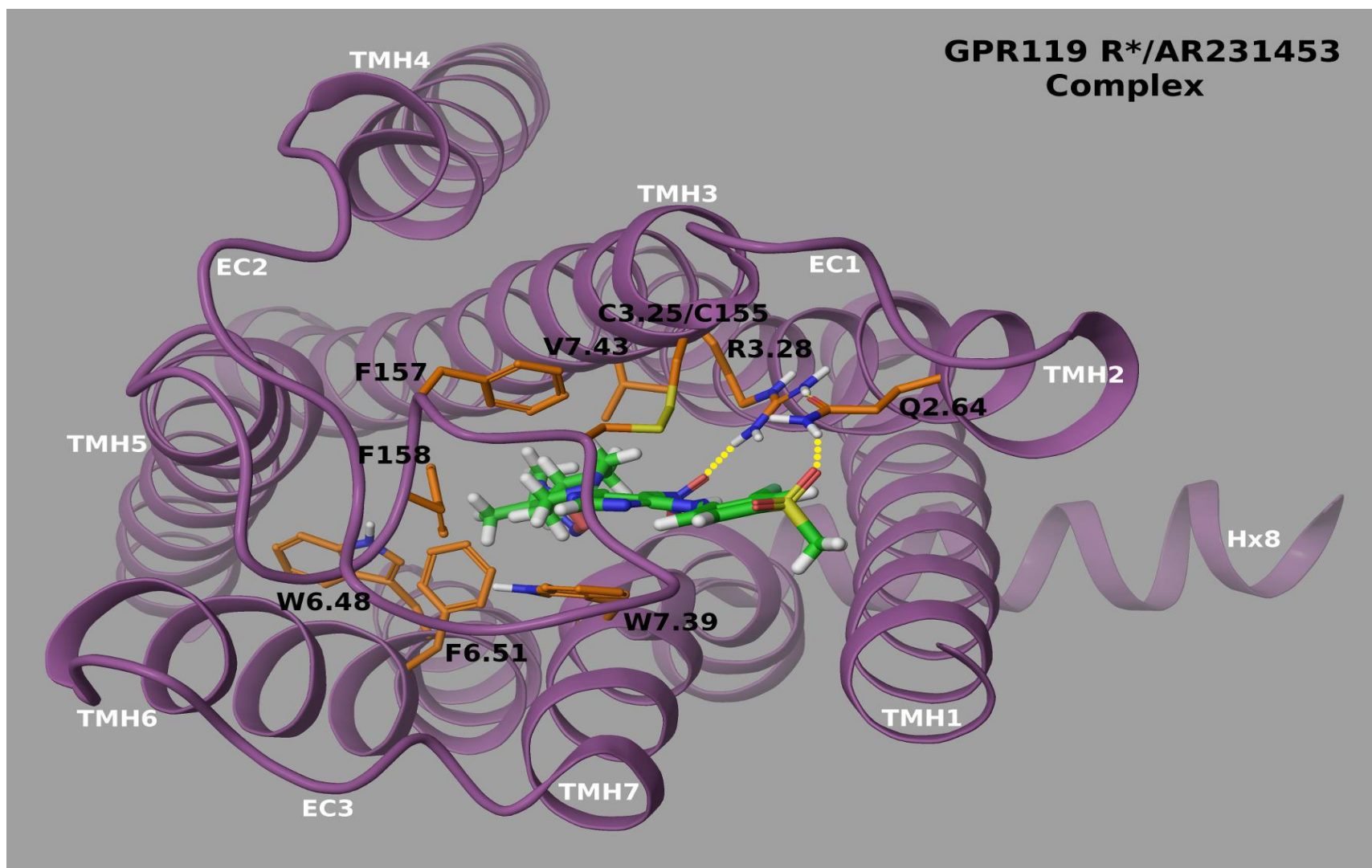


Figure 18: Top View of the GPR119 R* Receptor Model (purple ribbons) with Docked AR231453 Ligand. Protein residues having significant interactions with the docked AR231453 ligand are shown in orange. The yellow dotted line represents hydrogen bonds between R81^{3.28} and Q65^{2.64} with the AR231453 ligand.

Table 5: Interaction Energies Between GPR119 R* Receptor Model Residues and the AR231453 Ligand. Residues with absolute sequence numbers not followed by Ballesteros and Weinstein numbering are located in the EC loops of the GPR119.

GPR119 R* Receptor AR231453 Interaction Energies			
Amino Acid	Total Energy (kcal/mol)	Electrostatic Energy	Van der Waals
TRP265 (7.39)	-10.0973	0.1180	-10.2152
ARG81 (3.28)	-5.1671	-3.9682	-1.1989
CYS 155	-4.8989	-0.1062	-4.7927
GLN65 (2.64)	-3.9091	-4.0555	0.1464
PHE241 (6.51)	-3.9009	-0.1108	-3.7901
PHE 158	-3.6644	-0.2335	-3.4310
CYS78 (3.25)	-3.2001	-0.2824	-2.9177
PHE 157	-2.9436	0.0298	-2.9734
GLY268 (7.42)	-2.7667	-1.0653	-1.7014
VAL85 (3.32)	-2.6951	-0.0212	-2.6738
GLY 153	-2.3817	-0.8257	-1.5560
PHE7 (1.35)	-2.3332	0.0275	-2.3607
VAL269 (7.43)	-2.1014	0.0743	-2.1757
MET82 (3.29)	-1.9865	0.3317	-2.3182
GLN 154	-1.6864	0.3319	-2.0183
MET1 (1.29)	-1.6078	0.1119	-1.7197
SER4 (1.32)	-1.6045	-0.7194	-0.8851
LEU62 (2.61)	-1.0216	-0.0505	-0.9712
ILE58 (2.57)	-0.9606	0.0018	-0.9624
TRP238 (6.48)	-0.9251	0.0486	-0.9737
LEU11 (1.39)	-0.9213	-0.0664	-0.8549
ALA89 (3.25)	-0.6602	0.0525	-0.7127
GLN 74	-0.5552	-0.1784	-0.3767
ILE54 (2.53)	-0.4276	0.0202	-0.4478
THR86 (3.33)	-0.4118	0.0644	-0.4762
PHE234 (6.44)	-0.2175	0.0019	-0.2194
ARG262 (7.36)	-0.1915	0.9041	-1.0956
PHE5 (1.33)	-0.1506	0.0017	-0.1523
SER237 (6.47)	-0.0950	0.0588	-0.1539
GLY270 (7.44)	-0.0463	0.0251	-0.0714
SER70	-0.0076	0.1976	-0.2052
Totals	-63.5364	-9.2818	-54.2546
Conformational Cost (kcal)	3.2636		
Total Docking Cost (kcal)	-60.2728		

Figure 16 shows the ligand interaction diagram for the AR231453 agonist. Amino acids interacting with this ligand in the diagram are located within to 4Å of the ligand. Residues are numbered according to the absolute sequence for the GPR119 receptor.

Figures 17 and 18 are the side and top view of the docking pose of AR231453 in the GPR119 receptor-binding pocket. The active GPR119 receptor model was used when docking the AR231453 compound. For clarity, only residues having high interactions energies (table 5) are displayed, but more amino acids have smaller interactions with the ligand. In the images, residues are numbered using the Ballesteros and Weinstein GPCR numbering scheme.

4.3.1 Discussion of the AR231453 Agonist Ligand Docking Results

The following residues have a greater than -2.500 kcal/mol interaction energy with the AR231453 compound: Q65^{2.64}, C78^{3.25}, R81^{3.28}, V85^{3.32}, F241^{6.51}, W238^{6.48}, W265^{7.39}, C155, F157, and F158. Residues from the terminal end of the EC2 loop (F157 and F158) interact with both the agonist from the diastereomer pair and the AR231453 compound. Prior research has reported that both of these phenylalanine residues are important for receptor activation.^[7] Ligand-residue interactions are mostly attractive Van der Waals interactions, with the exception of the interactions with residues Q65^{2.64} and R81^{3.28}, which are mainly electrostatic. The docking cost is minimal, which indicates the amount of energy required for the AR231453 to adopt the docked pose is relatively small.

The GPR119 binding site encompasses TMHs 1, 2, 3, 6, and 7. This ligand adopts an almost vertical position in the binding site, reminiscent of the diastereomer pair agonist. The position of the AR231453 ligand inside the GPR119 R* receptor-binding pocket prevents the toggle switch residue, W238^{6.48}, from rotating into a g+ (inactive) conformation. Therefore, the

receptor would remain in the active state when this agonist binds. The residue with the largest interaction for the AR231453 dock is the hydrophobic W265^{7.39}. The only interactions that were predominately electrostatic were between the residue Q65^{2.64} and the sulfone group of the ligand, as well as, the residue R81^{3.28} with the nitro group of the ligand.

4.4 Site-Directed Mutagenesis Results

Two GPR119 amino acids (R81^{3.28} and W265^{7.39}) that face inward toward the binding pocket in the GPR119 receptor models and also have high interaction energies with the agonist ligands, were chosen for mutation to determine their role in ligand binding and receptor activation. Additionally, a third amino acid, R262^{7.36}, was chosen for mutation to determine this residue's role, if any, in the activation of GPR119 receptor.

R262^{7.36} is the only charged amino acid, other than R81^{3.28}, in the vicinity of the GPR119 receptor-binding site. In the GPR119 homology models R262^{7.36} has limited access to the binding pocket and exhibited negligible interactions with the ligands. Mutational analysis will confirm if this residue contributes minimally to GPR119 receptor activation.

Both R81^{3.28} and R262^{7.36} residues were mutated to the amino acid leucine. The leucine mutant would be capable of retaining some of the original arginine's hydrophobic bulk, while eliminating the positive charge of the arginine and the hydrogen bonding potential. The W265^{7.39} residue was mutated to an alanine. The W265A mutation was chosen to access the affects, if any, on GPR119 activation by replacing a large aromatic binding-pocket residue with a relatively small amino acid residue.

Each mutation was confirmed by DNA sequencing. GPR119 DNA sequencing results were conducted and prepared by Functional Biosciences, Inc.. Sequencing chromatograms, with confirmed mutants highlighted in blue, are shown in figure 20.

ATG GAA TCA TCT TTC TCA TTT GGA GTG ATC CTT GCT CTC CTG GCC TCC CTC ATC
 M E S S F S F G V I L A V L A S L I
 ATT GCT ACT AAC ACA CTA GTG GCT GTG GCT GTG CTG CTG TTC ATC CAC AAG AAT
 I A T N T L V A V A V L L L I N K N
 GAT GGT GTC AGT CTC TGC TTC ACC TTG AAT CTG GCT GTG GCT GAC ACC TTG ATT
 D G V S L C F T L N L A V A D T L I
 GGT GTG GCC ATC TCT GGC CTA CTC ACA GAC CAG CTC TCC AGC CCT TCT CGG CCC
 G V A I S G L L T D Q L S S P S R P
 ACA CAG AAG ACC CTG TGC AGC CTG **CGG** ATG GCA TTT GTC ACT TCC TCC GCA GCT
 T Q K T L C S L **R** M A F V T S S A A
 GCC TCT GTC CTC ACG GTC ATG CTG ATC ACC TTT GAC AGG TAC CTT GCC ATC AAG
 A S V L T V M L I T F D R Y L A I K
 CAG CCC TTC CGC TAC TTG AAG ATC ATG AGT GGG TTC GTG GCC GGG GCC TGC ATT
 Q P F R Y L K I M S G F V A G A C I
 GCC GGG CTG TGG TTA GTG TCT TAC CTC ATT GGC TTC CTC CCA CTC GGA ATC CCC
 A G L W L V S Y L I G F L P L G I P
 ATG TTC CAG CAG ACT GCC TAC AAA GGG CAG TGC AGC TTC TTT GCT GTA TTT CAC
 M F Q Q T A Y K G Q C S F F A V F H
 CCT CAC TTC GTG CTG ACC GTG TGG TGC GTT GGC TTC TTC CCA GCC ATG CTC CTC
 P H F V L T L S C V G F F P A M L L
 TTT GTC TTC TTC TAC TGC GAC ATG CTC AAG ATT GCC TCC ATG CAC AGC CAG CAG
 F V F F Y C D M L K I A S M H S Q Q
 ATT CGA AAG ATG GAA CAT GCA GGA GCC ATG GCT GGA GGT TAT CGA TCC CCA CGG
 I R K M E H A G A M A G G Y R S P R
 ACT CCC AGC GAC TTC AAA GCT CTC CGT ACT GTG TCT GTT CTC ATT GGG AGC TTT
 T P S D F K A L R T V S V L I G S F
 GCT CTA TCC TGG ACC CCC TTC CTT ATC ACT GGC ATT GTG CAG GTG GCC TGC CAG
 A L S W T P F L I T G I V Q V A C Q
 GAG TGT CAC CTC TAC CTA GTG CTG GAA **CGG** TAC CTG **TGG** CTG CTC GGC GTG GGC
 E C H L Y L V L E **R** Y L **W** L L G V G
 GTG GGC AAC TCC CTG CTC AAC CCA CTC ATC TAT GCC TAT TGG CAG AAG GAG GTG
 V G N S L L N P L I Y A Y W Q K E V
 CGA CTG CAG CTC TAC CAC ATG GCC CTA GGA GTG AAG AAG GTG CTC ACC TCA TTC
 R L Q L Y H M A L G V K K V L T S F
 CTC CTC TTT CTC TCG GCC AGG AAT TGT GGC CCA GAG AGG CCC AGG GAA AGT TCC
 L L F L S A R N C G P E R P R E S S
 TGT CAC ATC GTC ACT ATC TCC AGC TCA GAG TTT GAT GGC
 C H I V T I S S S E F D

Figure 19: Nucleotide & Amino Acid Sequence for GPR119 Receptor. Mutation targets circled.

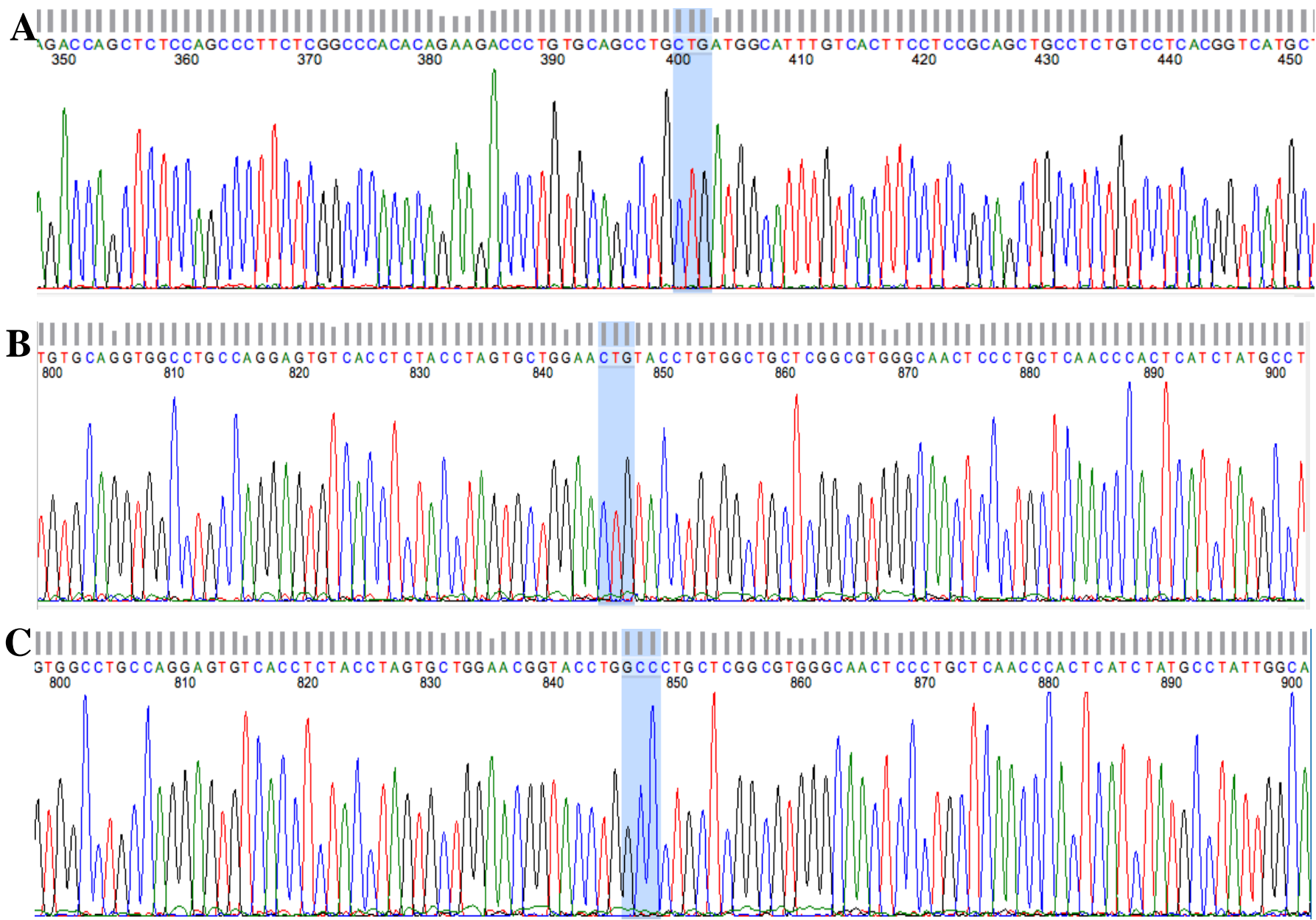


Figure 20: DNA Sequencing Chromatograms Confirming Desired Mutations in the GPR119 Receptor Sequence. Confirmed mutations are highlighted in blue. **A:** R81L was mutated from the CGG codon to CTG. **B:** R262L was mutated from the CGG codon to CTG. **C:** W265A was mutated from the TGG codon to GCC.

4.5 Dose Response Curve Results

Figures 21, 22, 23, and 24 are dose response curves for the wild type (WT) and the mutant transfected cells or the non-transfected cells. cAMP has been induced by the agonist AR231453. The curves have been normalized to the maximum cAMP production that was produced by transfected WT GPR119 cells. To examine drug potency effects between the control, mutant constructs, and WT GPR119 receptor, EC_{50} values for AR231453 obtained from each curve were compared. Table 6 lists a summary of the potency change results between the WT and mutant GPR119 receptor expressing cells

Graphically, a decrease in drug potency is indicated by a rightward shift in a mutant dose response curve when compared to the standard WT dose response curve. This shift can be seen in Figure 22. In Figure 22, R81L has a substantial reduction in AR231453 drug potency, which indicates that a higher concentration of drug is needed for the R81L mutant to produce as much cAMP as the wild type. On the other hand, an increase in drug potency is indicated by a leftward shift in a mutant dose response curve, when compared to the standard WT dose response curve. Figure 24 shows a slight leftward shift (away from the WT) for the R262L mutant. Thus, the R262L mutant is able to produce a slightly higher amount of cAMP, when compared to the wild type.

Flat lines in dose response curves indicate that cells are no longer responsive to the different doses of applied drug. Flat curves can be seen in Figures 21 and 23. The flat line in Figure 21, which is the experimental control, shows that cells not transfected with any type of GPR119 receptor DNA do not respond to the AR231453 drug. The W265A response curve, in Figure 23, also indicates that this mutant is no longer responsive to the different doses of the AR231453 compound. Interestingly, roughly 20% of cAMP is still being produced by the

W265A mutant. However, the cAMP production is independent of the different concentrations of drug applied to the cells.

4.5.1 Discussion Dose Response Curve

The experimental control in Figure 21 shows that no detectable level of cAMP is produced in HEK293 cells, despite stimulation with the AR231453 compound. This compound stimulates an increase in cellular cAMP levels only in the presence of transfected GPR119 receptors, as seen in Figures 22, 23 and 24. According to a study published in 2011 by Atwood et al., GPR119 receptor mRNA was not detectable in HEK293 cells.^[1] Our results in conjunction with Atwood's study indicate that AR231453 is a highly selective agonist for the GPR119 receptor and under standard cell culture conditions cAMP is not produced at a detectable level.

The R81L mutant exhibits a 29 fold reduction in drug potency according to Figure 22 and Table 6. This result indicates that GPR119 receptor activation has been substantially reduced by this mutation. Potentially, the AR231453 compound is able to bind to GPR119 and illicit activation; however, the loss of electrostatic interaction with the arginine residue inhibits full activation of the receptor. Interestingly, R81^{3,28} in GPR119 is in the same location as other positively charged residues that play a role in ligand binding and activation in other GPCRs (One of these cases is residue K192^{3,28} in Cannabinoid 1 receptor).^[22]

The W265A mutant has become non responsive to the AR231453 compound, see Figure 23. The position of this residue, which is in the middle of the GPR119 receptor binding pocket, and the lack of response of the receptor upon this residue's mutation, suggests that this residue interacts directly with the AR231453 compound. Previous studies of GPR119 have identified this receptor as having a high level of constitutive activity (activation in absence of

stimulation).^[7] The approximately 20 % cAMP response seen in Figure 23 from the W265A mutant could likely be attributed to constitutive activity for this mutant receptor.

The R262L mutant increased cAMP production by 2 fold, see Table 6. This slight increase could likely be attributed to indirect effects and it is highly unlikely that this residue has any direct interaction with AR231453.

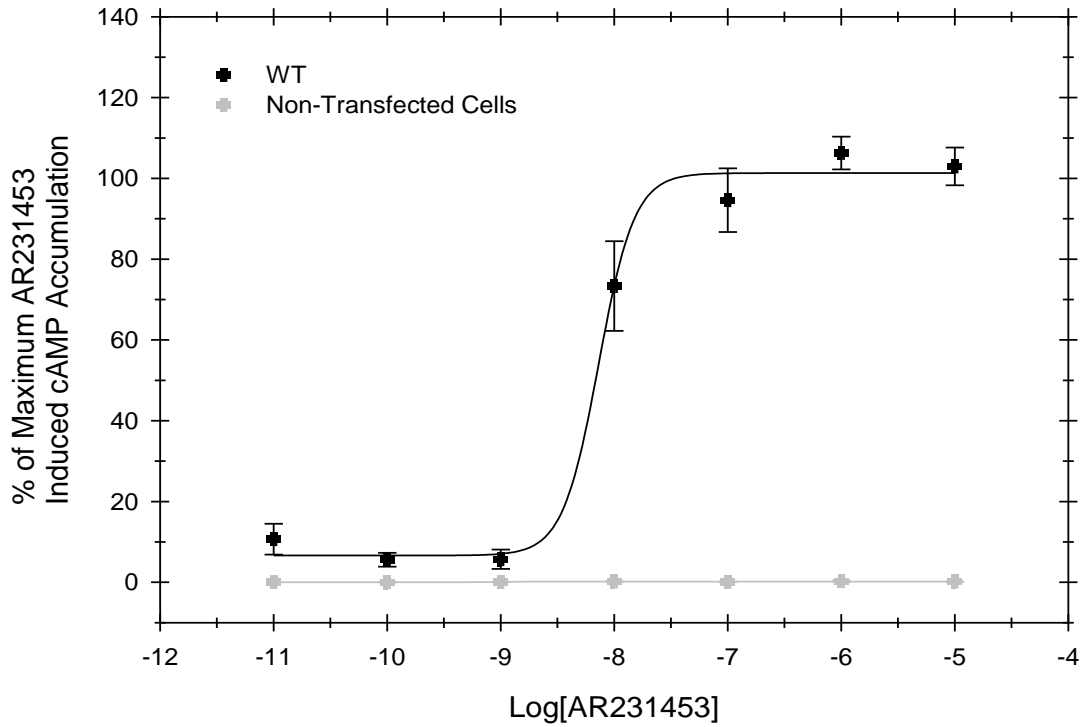


Figure 21: Experimental Control Dose Response Curve. The graph compares cAMP Accumulation in WT and Non-Transfected GPR119 Cells. Cells were induced by the AR231453 compound.

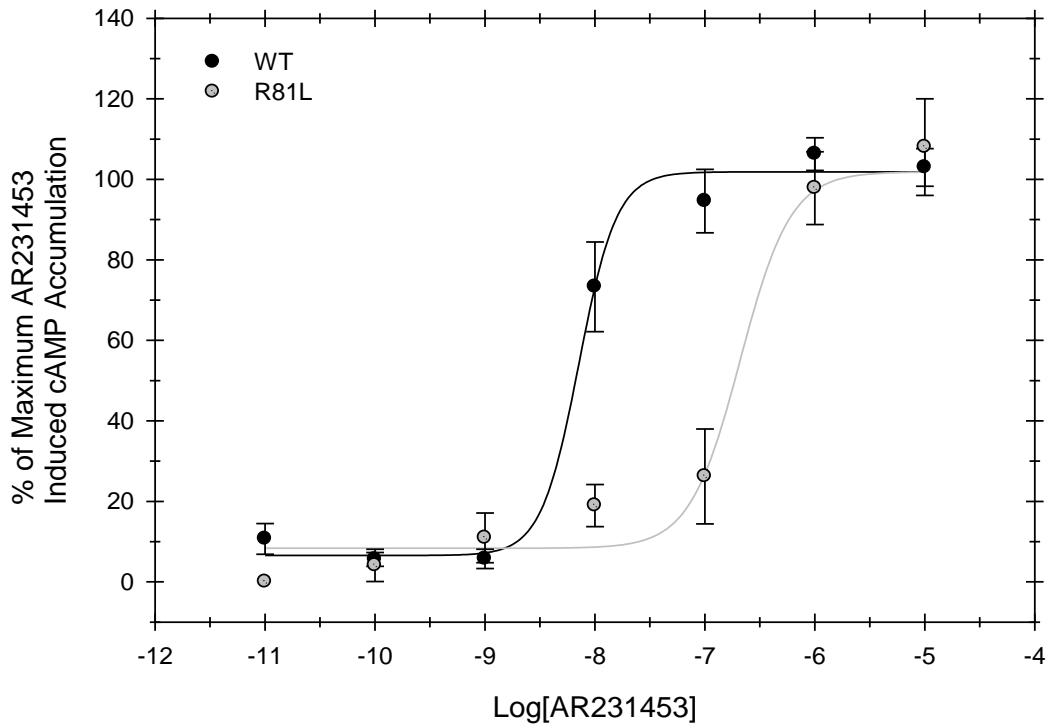


Figure 22: Dose Response Curve Comparing cAMP Accumulation in WT and mutant R81L GPR119 Transfected Cells. Cells were induced by the AR231453 compound.

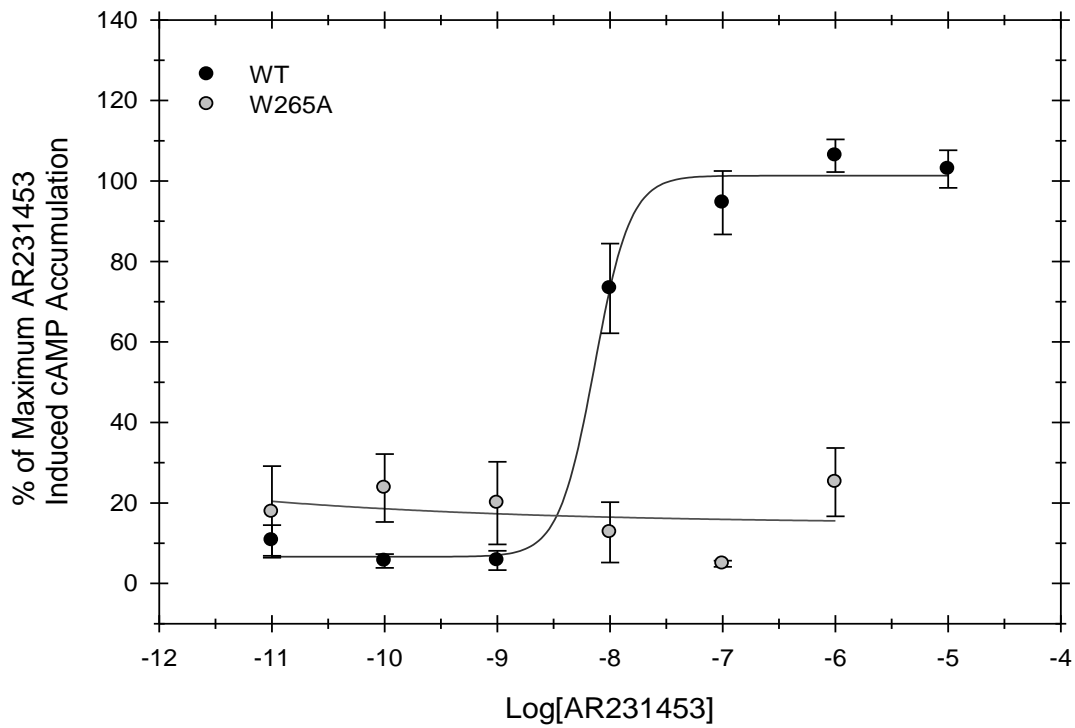


Figure 23: Dose Response Curve Comparing cAMP Accumulation in WT and mutant W265A GPR119 Transfected Cells. Cells were induced by the AR231453 compound.

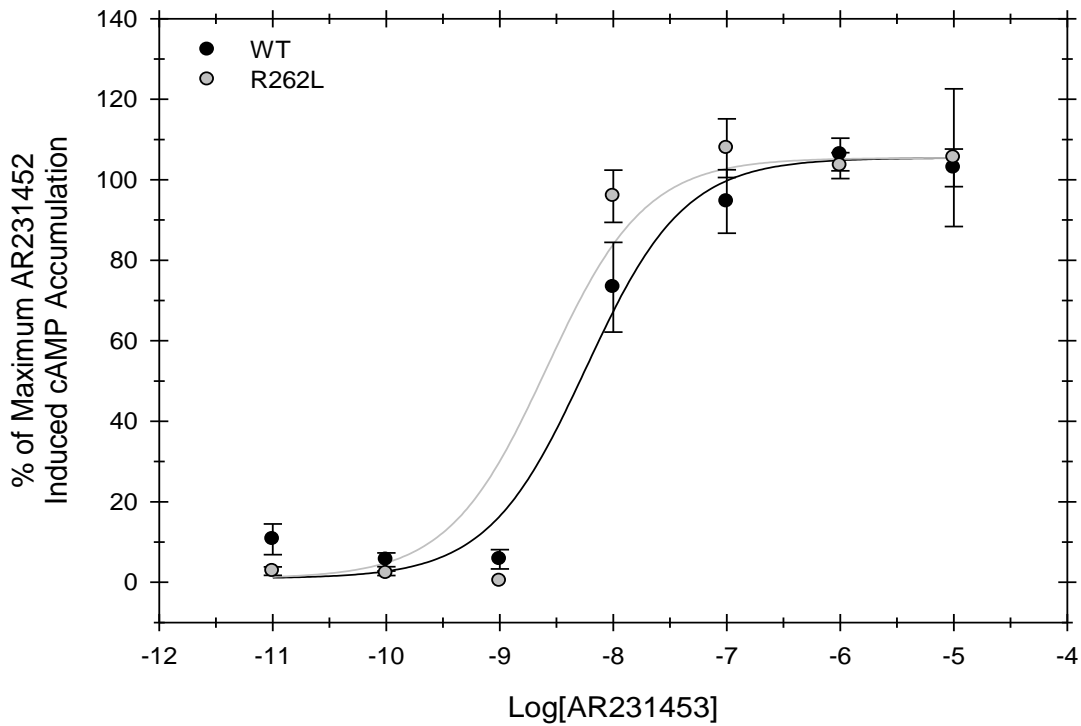


Figure 24: Dose Response Curve Comparing cAMP Accumulation in WT and mutant R262L GPR119 Transfected Cells. Cells were induced by the AR231453 compound.

Table 6: LogEC₅₀ Values and Potency Fold Changes for WT GPR119 and Mutant GPR119 Receptor in the cAMP Assay.

Mutant^a	logEC₅₀^b	Δ EC₅₀^c
WT	-8.14 ± 0.43	
R81L (3.28)	-6.68 ± 0.22	29 fold decrease
W265A (7.39)	NR	NR
R262L (7.36)	-8.59 ± 0.22	~ 2 fold increase

NR means no response

a) Absolute sequence residue number for mutant and with Ballesteros and Weinstein position in parentheses.

b) Logarithm of the EC₅₀. Measurements are followed by ± standard error.

c) Potency change

CHAPTER V

CONCLUSION

The homology models of the GPR119 receptor were built using A_{2A} Adenosine and Sphingosine1-Phosphate receptors as templates. Structural features important for GPCR transitions between the R or R* conformations (rotomer toggle switch, ionic lock) were incorporated into the models. The rotameric conformation and function of the toggle switch residue W238^{6,48} in the R and R* models is supported by the docking of the diastereomer pair of ligands used in this study.

The highly hydrophobic binding site of the GPR119 receptor has two positively charged residues, R81^{3,28} and R262^{7,39}. Docking results showed that: i) residues R81^{3,28} and W265^{7,39} have high interaction energies with the all three of the docked ligands, ii) R81^{3,28} forms a hydrogen bond with all three ligands (carbonyl moiety of the diastereomer pair and nitro-oxygen of AR231453), iii) W265^{7,39} has hydrophobic interactions with all three ligands (aromatic stacking interactions with the AR231452 nitro-pyrimidine ring and high van der Waals interactions with the bridged ring of the diastereomer pair), and iv) R262^{7,36} has only weak interactions with all three ligands. Based on the homology models and the docking results of the ligands, the three amino acids R81^{3,28}, W265^{7,39} and R262^{7,36} were selected for mutation to confirm if they are involved in the ligand binding and receptor activation.

Mutational results showed that: the R81L mutation caused a 29 fold decrease in AR231453 potency, the W265A mutation caused the receptor to become nonresponsive to the

AR231453 drug, and the R262L mutation resulted in a slight increase in AR231453 potency. These results correspond to a diminished ability for the R81L mutant receptor to activate in the presence of the agonist, the W265A mutant receptor can activate at a basal level which does not increase with increasing drug dose, and the R262L receptor's activation level is not affected by the mutation.

Finally, the computational results and the predictions for the GRP119 receptor binding site are in good agreement with the mutational data. Agonist interaction with the W265^{7,39} residue is exceptionally important for GPR119 activation. Agonist interaction with the R81^{3,28} residue is important for achieving the maximal response from GPR119, while using low doses of drug. R262^{7,39} proved to be not important for binding or activation. Overall these results demonstrate that computational studies (homology modeling and ligand docking) can give important insight about: i) the structure of a protein, even if the 3D structure is unknown and ii) how ligands bind in the binding pocket of the protein.

REFERENCES

1. Atwood BK, Lopez J, Wager-Miller J, Mackie K, & Straiker A. Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. *BMC Genomics*. 2011;12:14.
2. Ballesteros, J. A. & Weinstein, H. (1995) *Methods Neurosci.*, ed. C. S. Stuart, Academic Press, vol. 25, pp. 366–428.
3. Centers for Disease Control and Prevention. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014*. Atlanta, GA: U.S. Department of Health and Human Services; 2014
4. Cohen, L. S., Fracchiolla, K.E., Becker, J., & Naider, F. (2014). GPCR structural characterization: Using fragments as building blocks to determine a complete structure. *Peptide Science*, 102(3), 223-237.
5. Deupi, X. & Kobilka, B. (2007). Activation of G Protein–Coupled Receptors. *Advances in Protein Chemistry*, 74, 137-159.
6. Eglen, R. M., & Reisine, T. (2011). GPCRs revisited: New insights lead to novel drugs." *Pharmaceuticals*, 4, 244-272.
7. Englestoff, M. S., Norn, C., Hauge, M., Holliday, N. D., Elster, L., Lehmann, J., Jones, R. M., Frimurer, T. M., & 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), 5574-5789.
8. Gether, U. (2000). Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrinology Review*. 21(1), 90-113.

9. GPCR Activation: What Moves Where? Van der Kant, R. Available online: <http://swift.cmbi.ru.nl/gv/GPCR/> (accessed on 21 May 2014).
10. Kroeze, W. K., Sheffler, D. J., & Roth, B. L. (2003) G-protein-coupled Receptors at a glance. *Journal of Cell Science*, 116, 4867-4869
11. 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., Reggio P. H. (2011) Identification of the GPR55 agonist binding site using a novel set of high-potency GPR55 selective ligands. *Biochemistry*. 50(25), 5633–5647.
12. Lebon, G., Warne, T., Edwards, P. C., Bennett, K., Langmead, C. J., Leslie, A. G., & Tate, C. G. (2011). Agonist-bound Adenosine A2A receptor structures reveal common features of GPCR activation. *Nature*, 474(7352), 521-525.
13. Lundstrom, K. "An Overview on GPCRs and Drug Discovery: Structure-based Drug Design and Structural Biology on GPCRs." *G Protein-coupled Receptors in Drug Discovery*. New York: Humana, 2009. 51-66.
14. Miao, Y., Nichols, S. E., & McCammon, J. A. (2014). Free energy landscape of G-protein coupled receptors explored by accelerated molecular dynamics. *Journal of Physical Chemistry*
15. McClure, K. F., *et al.* (2011). Activation of the G-Protein-Coupled Receptor 119: A conformation-based hypothesis for understanding agonist response. *Journal of Medicinal Chemistry*, 54(6), 1948-1952.
16. Mirzadegan, T., Benko, G., Filpek, S., & Palczewski, K. (2003). Sequence analyses of G-Protein-Coupled Receptors: Similarities to Rhodopsin. *Biochemistry*, 42(10), 2759-2767.
17. Ning, Y., O'Neil, K., Lan, H., Pang, L., Shan, L. X., Hawes, B. E., & Hedrick, J. A. (2008). Endogenous and synthetic agonists of GPR119 differ in signalling pathways and their effects on insulin secretion in MIN6c4 insulinoma cells. *British Journal of Pharmacology*, 155(7), 1056-1065.

18. O'Connor, C. M. & Adams, J. U. "4.2 G-Protein-Coupled Receptors Play Many Different Roles in Eukaryotic Cell Signaling." *Essentials of Cell Biology*. Cambridge, MA: NPG Education, 2010.
19. Paul S.-H., Park, D., Lodowski, T., & Palczewski, K. (2008). Activation of G Protein–Coupled Receptors: Beyond two-state models and tertiary conformational changes. *Annual Review of Pharmacology and Toxicology*, 48, 107-141.
20. Reche, Pedro. "Sequence Identity and Similarity Calculator." Immunomedicine Group Universidad Complutense Madrid, Web. 25 Oct. 2014.
<<http://imed.med.ucm.es/Tools/sias.html>>.
21. Shi, L., Liapakis, G., Xu, R., Guarnieri, F., Ballesteros, J. A. & Javitch, J. A. (2002) β_2 adrenergic receptor activation. *Journal of Biological Chemistry*, 277(43), 40989-40996.
22. Song ZH, Bonner TI. A lysine residue of the cannabinoid receptor is critical for receptor recognition by several agonists but not WIN55212-2. *Mol. Pharmacol.*1996;49:891–896
23. Van der Kant, R., & Vriend, G. (2014) "Alpha-bulges in G Protein-Coupled Receptors. *International Journal of Molecular Sciences*, 15, 7841-7864.
24. Xiang, Z. (2006). Advances in homology protein structure modeling. *Current Protein Peptide Science*, 7(3), 217-227.
25. Zhu, X., Huang, W & Qian, H. (2013). "GPR119 Agonists: A Novel Strategy for Type 2 Diabetes Treatment." *Diabetes Mellitus - Insights and Perspectives*. Prof. Oluwafemi Oguntibeju (Ed.). Available from: <http://www.intechopen.com/books/diabetes-mellitus-insights-and-perspectives/gpr119-agonists-a-novel-strategy-for-type-2-diabetes-treatment>

BIOGRAPHICAL SKETCH

Shane M. Askar. She was born in Susquehanna, Pennsylvania in 1981. She obtained her high school diploma from Susquehanna Community H.S. in 1999. She first attended Ivy Tech Community College in Indianapolis, Indiana, in 2007. Shane then finished her undergraduate studies at Metropolitan State University of Denver (Metro State) in 2012, with a B.S. in Molecular Biology. As an undergrad, she had the opportunity to gain work experience by working in two different labs. She first worked in Dr. Pepper Schedin's lab, in the Division of Oncology at Anschutz Medical Campus, in Aurora, Colorado. She also worked for Dr. Rebecca Ferrell's microbiology laboratory at her home institution, Metro State. Shane was accepted to the University of Texas Pan American (UTPA) in the fall of 2012 to pursue a master's degree in chemistry. While at UTPA/UTRGV, she worked as a teaching assistant for multiple professors and as a research assistant for Dr. Evangelia Kotsikorou and Dr. Frank Dean. Finally she completed her masters in chemistry and received her M.S. degree in the fall of 2015. Her permanent address is:

Shane Askar

1325 E. Daffodil Avenue

McAllen, TX 78501