SOTOUDEH CHAFI, NOURODDIN, Ph.D. Construction of GPR18 In Silico Model: The Proposed Agonist and Antagonist, The Mechanism of Activation by Itself. (2019) Directed by Dr. Patricia H. Reggio. 130 pp.

GPR18 is an orphan G protein-coupled receptor (GPCR) that has been proposed to be the abnormal cannabidiol (Abn-CBD) receptor, a previously characterized atypical cannabinoid receptor for which the sequence was unknown. Recent studies show that GPR18 exhibits classical cannabinoid receptors (CB1 and CB2) biochemical functions in $CB1^{-/-}$ and $CB2^{-/-}$ knockout mice cells. Prior research has found GPR18 mRNA in a wide range of tissues from spleen, testis, and lymph nodes to peripheral blood leukocytes and phytohaemagglutinin-activated CD4+ T-cells. A variety of diverse therapeutic applications have been reported for the GPR18-ligand signaling system, including cell death via apoptosis, cell migration following central nervous system (CNS) injury or inflammation, regulation of intraocular pressure, sperm physiology, cardiac physiology, immunomodulation, obesity, cancer, pain, and resolution of acute inflammation. Whereas GPR18 is reported to be deorphanized, the lipid N-arachidonoyl glycine (NAGly) has been proposed to be the endogenous ligand of GPR18. However, more recent publications have reported a lack of GPR18 activation by NAGly. These apparent inconsistent findings were claimed to be related to lack of co-localization between NAGly and GPR18 in the body, but our experimental data show that both NAGly and the endogenous lipid mediator, resolvin D2 (RvD2), activate GPR18.

Chapter I details an extensive background about the GPR18 discovery, pharmacological importance, related diseases, and the reported agonists and antagonists.

Chapter II reports all of the biological tests and obtained experimental data. In this chapter, for the first time, two novel agonist and antagonist are proposed, which are more potent than the earlier reported competitors. These two ligands are CBD-derived compounds that researchers could use as a scaffold for future structure-related activity studies.

Chapter III presents a detailed description of the methods employed for the construction of the active and inactive state of GPR18. The significant differences between the GPR18 sequences and the template crystal structure, and the techniques that were used to overcome challenges, are discussed in this chapter.

Chapter IV then answers why the mutation of A3.39N reduces the constitutive activity in GPR18. In this chapter, by presenting the results of two series of Molecular Dynamics simulations, for the first time, a detailed mechanism of activation of GPR18 by itself is clarified.

Chapter V investigates the quality of static model docking jobs presented in Chapter III. To this end, the researcher employed Molecular Dynamics and then analyzed the results. This chapter also explains why the ligand requires parameterization and how this could occur.

The final chapter concerns the conclusions and future works. One crucial part of this project that remains to be done is the mutagenesis study. Chapter VI gives an extensive required mutation to improve the presented model in the future. This chapter also suggests a structure-activity relationship study based upon the discovered scaffold presented in Chapter II to modify the proposed ligands to enhance ligand bindings.

CONSTRUCTION OF GPR18 IN SILICO MODEL: THE PROPOSED AGONIST AND ANTAGONIST, THE MECHANISM OF ACTIVATION BY ITSELF

by

Nouroddin Sotoudeh Chafi

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2019

Approved by

_____
Committee Chair

To my beautiful wife, Faezeh,

the only one who understands me better than I understand myself.

APPROVAL PAGE

This dissertation, written by Nouroddin Sotoudeh Chafi, has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
Patricia H. Reggio

Committee Members _____
Kimberly S. Petersen

_____
Mitchell Croatt

_____
Shabnam Hematian

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Patricia H. Reggio, as my teacher and mentor in computational chemistry. She taught me how to take a critical stance toward scientific research findings.

Besides my advisor, I would like to thank Dr. Dow P. Hurst and Dr. Diane L. Lynch for all they taught me about GPCR modeling and Molecular Dynamics Simulations and their patience in answering my questions.

Special thanks to Dr. Paula Morales for her priceless and constructive collaboration in writing the paper about the mechanism of activation of GPR18.

Also, I am grateful for two research groups that collaborated with us in this project—Dr. Mary Abood and her research group at Temple University for performing the biological tests, and Dr. Nadine Jagerovic and her research team for the synthesis of the chemical compounds at the National Research Council, Madrid (CSIC) in Spain.

Finally, I would like to thank my dissertation committee members, Dr. Mitchell Croatt, Dr. Kimberly Petersen, and Dr. Shabnam Hematian for reviewing and criticism of my dissertation.

TABLE OF CONTENTS

Page

LIST OF TABLES vii

LIST OF FIGURES viii

CHAPTER

I. GPR18, A NOVEL PUTATIVE CANNABINOID RECEPTOR: INTRODUCTION AND PHARMACOLOGY 1

Endocannabinoid System and Classical Cannabinoid Receptor 1
GPR18 Gene Expression and Distribution 3
GPR18 Proposed Agonists 6
GPR18 Proposed Antagonists 14

II. EXPERIMENTAL DATA 17

Introduction 17
Cellular Expression of GPR18 18
Calcium Mobilization of NAGly in CHO/GPR18 19
MAPK Activity of NAGly and RvD2 in CHO/GPR18 Cells 20
Calcium Mobility Activity of Our Novel Proposed Agonist and Antagonist of GPR18 in CHO/GPR18 Cell Lines 22
β-arrestin Activity of Agonists in GPR18 DiscoverX Cells 26
Conclusion 27

III. GPR18 MODELING 29

Introduction 29
Molecular Modeling the Inactive State (R) of GPR18 29
Definition of Torsion Angles 32
Conformational Divergence between Template and GPR18 TMH3 34
Ionic Lock and Toggle Switch 37
Conformational Memory Study of TMH7 38
Extracellular (EC) and Intracellular (IC) Loop Conformations 41
Bundle Minimization 41
Docking of the Putative Antagonist PSB-CB5 in Inactive (R) Model 42
Docking of Our Proposed Antagonist S4 in Inactive (R) Model 43

Molecular Modeling GPR18 Active State (R*) 45
Docking of Putative Agonist in GPR18 Active (R*) Model: Conformational Memory Calculations for NAGly and RvD2 49

IV. TOWARDS A MOLECULAR UNDERSTANDING OF THE CANNABINOID RELATED ORPHAN RECEPTOR GPR18: A FOCUS ON ITS CONSTITUTIVE ACTIVITY 60

Abstract 60
Introduction 61
Results and Discussion 63
Materials and Methods 87
Conclusions 92
Supplementary Materials 92
Author Contributions 92
Funding 92
Acknowledgments 93
Conflicts of Interest 93
Abbreviations 93

V. PARAMETERIZATION OF GPR18 PROPOSED ANTAGONIST S4 94

Introduction 94
Why S4 Needs to be Parameterized 97
Parameter Optimization of Components by ffTK Optimization of Bonds and Angles 101
Optimization of Dihedrals 102
MD Simulation 110

VI. CONCLUSIONS AND FUTURE WORKS 115

REFERENCES 118

LIST OF TABLES

Page

Table 1. Conserved Residues/Motifs Presented in a Sample of Class A GPCRs in TMH7 and 3 31

Table 2. Comparison of Angles and Dihedrals between QM and MM.....99

LIST OF FIGURES

Page

Figure 1. N-arachidonoyl Ethanolamine (AEA), the Potential Precursor of N-arachidonoyl glycine (NAGly) the Proposed Agonist of GPR18.....9

Figure 2. The Only Proposed Biosynthesis Pathways of NAGly from AEA.....9

Figure 3. GPR18 Proposed Agonists and Antagonists.....11

Figure 4. Two Proposed Recent Antagonists for GPR18.....16

Figure 5. Western Blot of CHO/GPR18, Wild-type CHO, and CHO/GPR18/βarr Depicting Expression of GPR18 (38kDA).....18

Figure 6. NAGly Increases $[Ca^{2+}]_I$ in CHO/GPR18 Cells 19

Figure 7. Inhibition of NAGly-induced Increases in MAPK Activity by GPR18 Antagonists in CHO/GPR18 (n=4-19) 21

Figure 8. The Proposed Agonist S5 and Antagonist S4 Based on the Pyrazol-3(5)-ylbenzene-1,3-diol Scaffold 22

Figure 9. S5 Increases $[Ca^{2+}]_I$ in CHO/GPR18 Cells.....23

Figure 10. S4 Alone Has No Effect on $[Ca^{2+}]_I$ in CHO/GPR18 Cells 24

Figure 11. S4 Blocks Nagly Response on $[Ca^{2+}]_I$ in CHO/GPR18 Cells.....25

Figure 12. Antagonism of Intracellular Ca^{2+} Signal by S4 vs. CB-5 26

Figure 13. β-arrestin Activity in GPR18 DiscoverX Cells.....27

Figure 14. Helix Net Representation of hGPR18 Sequence 30

Figure 15. Definition of the Protein Backbone (φ, ψ and ω) and Side Chain (χ1 (chi1), χ2 (chi2), χ3 (chi3), Etc.) Dihedral Angles.....33

Figure 16. Theronine in Gauche Minus Intrahelical Binding 36

Figure 17. Comparison of TMH7 in Straight Model and Bent Model.....39

Figure 18. MD Simulation of GPR18 with an Initial Straight TMH7 (Yellow) Shows That after 50 ns TMH7 Tends to Bend (Green Model) 40

Figure 19. PSB-CB5 (Orange Structure) Docked in R Model of GPR18 43

Figure 20. S4 (Green Structure) Docked in R Model of GPR18 45

Figure 21. Super Position of R Model (Blue) and R* Model (Green) 47

Figure 22. Super Position of R Model (Blue) and R* Model (Green) 48

Figure 23. Superposition of R and R* Models 49

Figure 24. Conformational Memory Results of NAGly.....51

Figure 25. Conformational Memory Results of RvD2 for All of the Csp2-Csp2 Dihedral Angles Which are Populated around 0° and ±180° 53

Figure 26. Conformational Memory Results of RvD2 for All of the Csp3-Csp3 Dihedral Angles Which are Populated Around Gauche Miues, Plus and Trans Conformations 54

Figure 27. Top Row from Left to Right; Superposition of Conformational Memory of RvD2 (Question Mark Shape), NAGly (U-shaped), NAGly (J-shaped).....55

Figure 28. Docking of NAGly (Pink) Inside the GPR18 57

Figure 29. Docking of RvD2 (Pink) Inside the GPR18 58

Figure 30. S5 (Light Blue Structure) Docked in R* Model of GPR18 59

Figure 31. Helix Net of the Human GPR18 Sequence.....64

Figure 32. The Toggle Switch in GPR18 R State 69

Figure 33. The Toggle Switch in the Activated State 70

Figure 34. RMSD vs. Simulation Time Plots for the WT GPR18 R State (Black) and the A3.93N Mutant R State (Red) [72].....71

Figure 35. Stability of the “Ionic Lock” Over the Course of the WT (Black) and A3.93N (Red) Trajectories 73

Figure 36. Heteroatom Distance between R3.50 and D3.49 in the WT GPR18 (Black) and the A3.39N Mutant (Red) 74

Figure 37. Average Distance between Cα Atoms of the Last Five Cytoplasmic Residues in TMH3 (Y3.51 to V(124) and TMH6 (K6.30 to I6.34) in Wild Type GPR18 WT (Black) and the A3.39N Mutant (Red) 75

Figure 38. The Distance between the Sodium Ion and N3.39 in the A3.39N Mutant (Red) [72] and A3.39 in WT (Black) 77

Figure 39. The Distance between the Sodium Ion and N1.50 in the A3.39N Mutant (Red) [72] and WT GPR18 (Black) 77

Figure 40. Comparison of the Sodium Ion (Magenta Sphere) Position in the WT Bundle (Yellow) and A3.39N Bundle (Green) 78

Figure 41. χ1 Rotamer Conformational Change in D7.49 in WT GPR18 (Black) and the A3.39N Mutant (Red) [72] Due to Sodium Ion Movement. 79

Figure 42. WT GPR18 Bundle 80

Figure 43. A3.39N Mutant Bundle 80

Figure 44. Heteroatom to Heteroatom Distance Between Residues Q6.43 and D7.49 in WT GPR18 (Black) and the A3.39N Mutant (Red) [72] 81

Figure 45. Position of Q6.43 in the GPR18 R Model (Highlighted in Purple Thin Tubes) 82

Figure 46. Configuration of the Important Residues (Highlighted in Light Blue Tubes) Involved in Our Proposed Activation Mechanism of GPR18 84

Figure 47. Configuration of the Important Residues (Highlighted in Light Blue Tubes) Involved in the Proposed Activation Mechanism in the A3.39N Mutant 85

Figure 48. Presence of Two Bulky Residues, I3.46 and I6.37 (Magenta Residues in Van Der Waals), Above the Ionic Lock Residues, R3.50/S6.33 (in Magenta Tubes), Makes it More Difficult to Form a Hydrogen Bond in the WT Receptor 86

Figure 49. Superposition of the S4 Global Minima at MP2/6-31G* Level of Theory (Grey Structure) with CHARMM36 Minimized Structure (Orange) Before Parameterization. 98

Figure 50. The Output .str File of Submission of S4 to Paramchem Online Database 101

Figure 51. The Fragment Built from S4 102

Figure 52. Dihedral Scanning Step.....103

Figure 53. Dihedral Parametrization Step 105

Figure 54. Energy-Conformation Plots; the Steps Need to be Taken Before Getting Into the Best Final Fit 107

Figure 55. Superposition of the S4 Global Minima at MP2/6-31G* Level of Theory (Grey Structure) with CHARMM Minimized Structure Before Parameterization (Orange) and After Parameterization (Green) 108

Figure 56. Potential Energy Scan for Targeted Angle N15 C12 C5 and Dihedral N15 C12 C5 C4 109

Figure 57. MD Simulations of Parameterized S4 Inside GPR18 R State Model 112

Figure 58. Comparison of Three Different Conformation of S4 Inside the Binding Site of GPR18.....113

Figure 59. Distance of the Mass Center of the 5 Last Intracellular End Residues of TMH3 and 6 114

CHAPTER I

GPR18, A NOVEL PUTATIVE CANNABINOID RECEPTOR: INTRODUCTION AND PHARMACOLOGY

Endocannabinoid System and Classical Cannabinoid Receptor

The word "endocannabinoid" is formed by combining "endo" and "cannabinoid." "Endo" is the abbreviation of "endogenous," which means having an internal cause or origin. "Cannabinoid" refers to the cannabis plant. Taken together, "endocannabinoids" are cannabis-like chemicals or neurotransmitters that the body naturally produces. The endocannabinoid system has an important role in the diagnosis of many diseases and conditions such as Alzheimer's disease [1], Parkinson's disease [2, 3], Huntington's disease [4, 5], migraine [6], and schizophrenia [7]. The endocannabinoid system is composed of endocannabinoids, cannabinoid receptors, and the enzymes that synthesize and degrade endocannabinoids. Cannabinoid receptors were initially discovered and named after their interaction with Δ^9-tetrahydrocannabinol (Δ^9-THC), a Cannabis plant-derived compound. The endocannabinoid membrane receptor, CB1, was first discovered in rat's brain in 1988 [8]. CB1 receptors' mRNA was found expressed at high level in brain regions, which are involved in mental and physiological processes (such as cognition, memory, metabolic functions, and motor coordination) [9, 10]. Not long afterward, researchers discovered that the CB1 receptor was found not only in rats, but also within other species, including humans. The isolation of a human CB1 receptor cDNA was identified in 1991 [11]. The CB1 receptor is highly expressed in the

mammalian central nervous system (CNS) [12]. It is also found throughout the body in many different tissues such as cardiovascular, reproductive, immune, and gastrointestinal [13]. This wide distribution of the receptor makes it involved in regulation of different disorders like learning and memory [14], reward and addiction [15], pain [16], metabolism [17], and food intake [18]. The second important cannabinoid receptor was identified in 1993. This receptor, which is known CB2, is expressed principally in the immune system and peripheral tissues, and it was first cloned from rat's spleen [19]. However, later, it has been revealed that CB2 is also located in the CNS, liver, and bone [20, 21]. The amino acid sequence of the human CB2 receptor shows 44% homology to the human CB1 receptor. After discovery of the CB1 and CB2 receptors, scientists began researching their roles and the mechanism of the signaling of these receptors. The cannabinoid receptors regulate multiple pathways throughout the coupling to G proteins of the $G_{i/o}$ family, which leads to inhibit adenylyl cyclase in most tissues and cells. This, in turn, regulates ion channels, including G protein-coupled inwardly rectifying potassium channels (GIRK) [22] or N-type voltage-gated calcium channels (Ca v2.2) [23]. In addition to the typical G protein-dependent signaling seen with all GPCRs, cannabinoid receptors regulate, in a G protein-independent manner, by coupling to β-arrestin and desensitization of the receptor [24]. However, CB receptors are responsible for many of the effects of endocannabinoids and cannabinoids, but recent findings suggest that there are other receptors that may serve as a desirable target for certain cannabinoids and activate or inactivate them independently from the CB1 or CB2

receptor. In fact, these newly discovered receptors are orphan GPCRs, namely GPR18, GPR55, and GPR119.

GPR18 Gene Expression and Distribution

G protein coupled receptors (GPCRs) are the membrane proteins that have seven transmembrane helices. GPR18 is an orphan GPCR consisting of 331 amino acids that has been proposed to be the abnormal cannabidiol (Abn-CBD) receptor, a previously characterized atypical cannabinoid receptor for which the sequence was unknown [25]. Recent studies show that GPR18 exhibits classical cannabinoid receptors biochemical functions in $CB1^{-/-}$ and $CB2^{-/-}$ knockout mice cells [26].

The first isolation of GPR18 was done accidentally when Gantz et al. [27] used a relaxed stringency polymerase chain reaction to isolate a peptide receptor (a subclass of G protein-coupled receptors that bind peptide ligands via their seven-transmembrane helices and the extracellular loops). Cloning of the corresponding canine and human genes and genomic library screening designated this novel gene sequence as GPR18 based upon the Human Gene Mapping Workshop nomenclature committee [28]. The nucleotide sequences of the canine and human chromosome region are highly conserved and share 89% nucleotide identity and 92% amino acid similarity between the two species. Comparison of a GPR18 nucleotide sequence with the rat's mu opioid receptor, one of the published seven-transmembrane GPCRs at the time, showed a low similarity (45% nucleotide sequence homology and 34% amino acid similarity). They also sought to find a binding site of several opiate ligands (using the radioligands [^{3}H] bremazocine and [^{3}H]diprenorphine) in stably transfected Hepa and L-cells, but nothing was observed.

The experiments proved that GPR18 is not an opioid receptor family. The results of the Northern blotting analysis in several different cell types performed by Gantz et al. [28] clarified that GPR18 mRNA is present at the highest level in human testis and spleen, and lesser degree of expression has been found in several other tissues including thymus, peripheral blood leukocytes, appendiceal, lymph node, and small intestine. This pattern of tissue expression suggests a possible role relevant to fertilization and regulation of the immune system for this receptor. No detectable GPR18 mRNA has been observed in other adult tissues, including brain, heart, lung, liver, kidney, pancreas, colon, skeletal muscle, ovary, placenta, prostate, adrenal medulla, and adrenal cortex. Vassilatis et al. [29], in a study about diverse members of the GPCR repertoires of human and mice via RT-PCR/RNA tissue profiling, revealed that GPR18 genes presented in four different expression levels: no expression in amygdala, frontal cortex, hippocampus, liver, and muscle; low expression in cortex, thalamus, adrenal tissue, colon, intestine, kidney, prostate, skin, spleen, stomach, and uterus; moderate expression in lung, ovary, testis, thymus, and striatum; and strong expression in hypothalamus, thyroid, peripheral blood leukocytes, cerebellum, and brain stem. Later analysis of chromosomal localization revealed that GPR18 genes and another orphan GPCR with constitutive activity (constitutive activity or basal activity means that the receptor can be activated by itself without the requirement for binding of an agonist ligand), known as the Epstein-Barr virus (EBV)-induced receptor 2 (EBI2), were both located at chromosome 13 in the region 13q32.3 [30]. Moreover, expression analysis of GPR18 and EBI2 uncovers important similarities, especially in peripheral blood mononuclear cells and lung tissues.

The similar receptor expression pattern and the close genomic proximity suggest a debatable argument that GPR18 and EBI2 have similar biologic functions in the immune system. Kohno et al. [31] was the second group found that GPR18 is highly expressed in HUT102 cells while they were cloning chemokine receptors and GPCRs expressed in ATL cells. They are also the first research group who introduced N-arachidonoylglycine (NAGly) as an endogenous agonist for GPR18, which is discussed in next section. Their findings show that GPR18 mRNA to be highly expressed in several human peripheral lymphocyte subsets and human T-Cell lymphotropic virus-transformed cell lines.

Moreover, the GPR18 gene was significantly expressed in phytohaemagglutinin-activated CD4+ T-cells. Altogether, their results also suggest that GPR18 may be involved in immune system regulation. Qin et al. [32] were the first who suggested that GPR18 plays a critical anticancer target role. Using a quantitative large-scale PCR-based array, they found that GPR18 is highly expressed in melanoma metastases suggesting that it may be highly relevant for the malignant behavior of melanoma itself. Their further studies using siRNA-mediated knockdown revealed that GPR18 is constitutively active and inhibits apoptosis in *in vitro* human melanoma cells, indicating an essential role in tumor cell survival for GPR18. Their research also introduced the first mutation report in the GPR18 sequence, which will be discussed in Chapter IV. The recent studies demonstrate that GPR18 has been localized in the mouse macrophage-derived cell line, RAW264, and plays an essential role in macrophage function [33]. More recent research explored GPR18 expression in the anterior murine eye using immunohistochemical tools. They found that GPR18 expressed in the murine eye (in the corneal epithelium and

interestingly in the trabecular meshwork) may serve as a desirable target for the regulation of elevated intraocular pressure in different forms of glaucoma [34]. Findings also demonstrated GPR18 expression in rostral ventrolateral (RVLM) neurons, which modulated the sympathetic activity and suggested a restraining influence for RVLM GPR18 on blood pressure [35]. Finally, the most recent report confirmed GPR18 expression using RT-PCR and immunocytochemistry experiments and localized the GPR18 protein in the midpiece of human spermatozoa which was unknown before [36]. In this study, the researchers focused on the functional characterization of GPR18 in human spermatozoa because they detected a high level of expression of this receptor for the first time in human spermatozoa. For the detection of GPR18 proteins, the researchers performed RT-PCR and immunocytochemical staining with human spermatozoa; finally, GPR18 was detected at the protein level in all human spermatozoa investigated.

GPR18 Proposed Agonists

The orphan GPCR GPR18 has been hypothesized to be a member of the endogenous cannabinoid system. As previously mentioned, a variety of diverse therapeutic applications have been reported for the GPR18-ligand signaling system, including cell death via apoptosis [32], cell migration following central nervous system (CNS) injury or inflammation [26], regulation of intraocular pressure [34, 37], sperm physiology [36], cardiac physiology [35, 38, 39], immunomodulation [40-42], obesity [43], and pain [44].

Whereas GPR18 is reported to be deorphanized, the lipid N-arachidonoyl glycine, a carboxylic analog of anandamide, has been proposed to be the endogenous ligand of

GPR18 [31]. However, more recent publications have reported a lack of GPR18 activation by NAGly [45-47]. Kohno et al. [31] used two different assays, “intracellular calcium mobilization” and “cyclic adenosine monophosphate or cAMP,” to identify the natural ligand for GPR18. In calcium mobilization assay using a bioactive lipid library in combination with GPR18 expressing L929 cells (stably expressing human GPR18 or mocked transfected) lead to a significant increase in intracellular Ca^{2+} level with the lipid NAGly as a hit in the screen at 10μM concentration. To confirm the findings, they repeated this experiment with K562 and Chinese Hamster Ovary cells (CHO cells), in which similar results were obtained. The cAMP assay has been performed in human GPR18-transfected CHO cells, which resulted in inhibition of forskolin-induced cAMP levels with an IC_{50} value of 20 nM. Pretreatment of the GPR18-transfected CHO cells with pertussis toxin (PTX is a useful tool or protein to study the molecular mechanisms by which GPCRs mediate the biological actions of neurotransmitters; $G_{\alpha i}$ protein is more sensitive to PTX actions than $G_{\alpha o}$ and $G_{\alpha s}$ proteins) completely abolished the inhibition of forskolin-stimulated cAMP production by NAGly and increased intracellular calcium. These data support that GPR18 binding to NAGly leads to coupling to $G_{\alpha i}$ protein. Yin et al. [45] evaluated the ability of the DiscoveRx PathHunter™ β-arrestin assay platform that measures β-arrestin binding to GPCRs by utilizing ligands from bioactive lipid, endocannabinoid, and orphan ligand libraries, as well as NAGly to identify a hit ligand to interact with GPR18. However, the report found that NAGly is an ineffective agonist for GPR18.

Bradshaw et al. [26, 48-50] was the first group that verified GPR18 activation by NAGly. They first hypothesized that NAGly biosynthesis might involve anandamide (AEA) as a precursor (see Figures 1 and 2) with oxidation at the terminal alcohol. As summarized in Figure 2, one pathway is a fatty acid amide (FAAH)-dependent conjugation of glycine to AEA-released arachidonic acid (AA) and the second one is by oxidation of the ethanolamine moiety in anandamide, likely by alcohol dehydrogenase (ADH). As Figure 1 shows, NAGly differs from AEA by the oxidation state of the carbon beta to the amido nitrogen. Despite the structural similarity between NAGly and AEA, and also NAGly biosynthesis pathway dependence on AEA, recent studies support that NAGly has no affinity for CB1, CB2, and GPR55 receptors [51]. Therefore, antinociceptive and anti-inflammatory effects of NAGly[52-55] might occur through an alternative novel cannabinoid receptor, which is GPR18. The second reason is that the anti-inflammatory effects of NAGly are consistent with GPR18 expression in peripheral blood leukocytes and several hematopoietic cell lines [31], as well as being highly expressed in the spleen and testis. McHugh et al.'s [26, 50, 56-58] results revealed that the mouse microglial cell line BV-2 has an endogenous expression of the GPR18 mRNA.

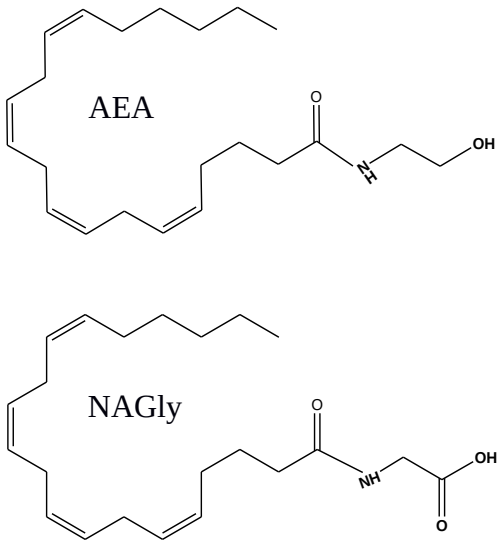


Figure 1. N-arachidonoyl Ethanolamine (AEA), the Potential Precursor of N-arachidonoyl glycine (NAGly) the Proposed Agonist of GPR18.

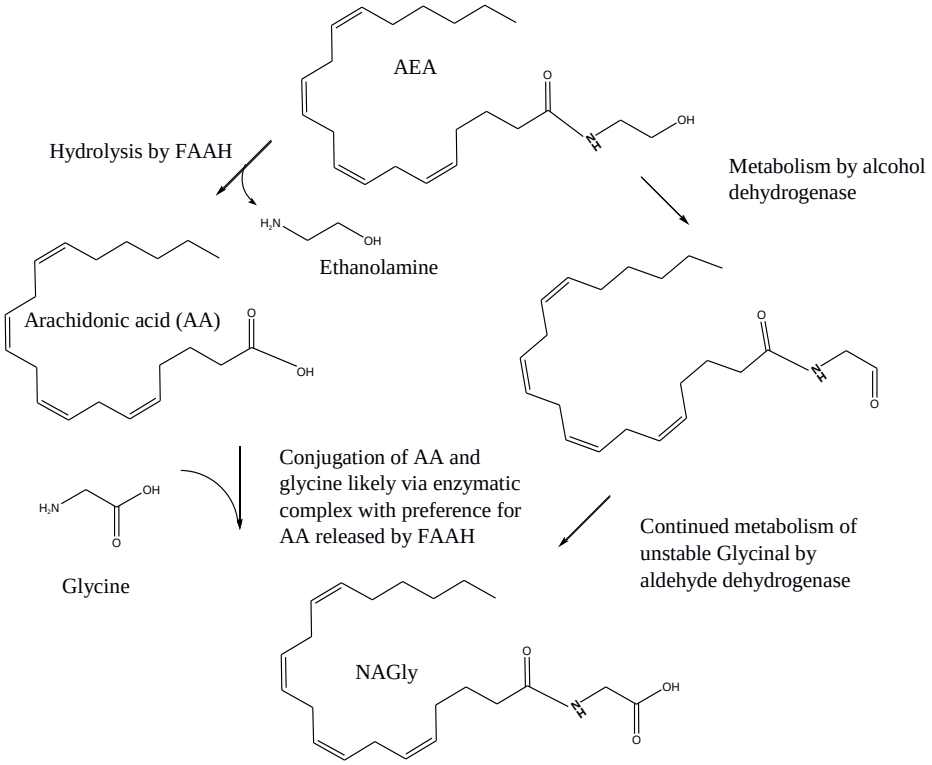
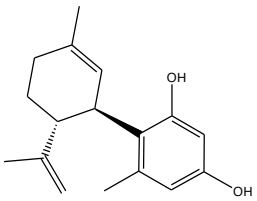
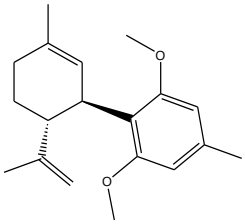


Figure 2.The Only Proposed Biosynthesis Pathways of NAGly from AEA. Adapted from Bradshaw et al. [48].

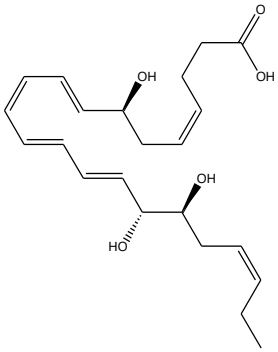
Δ^9-THC and cannabidiol (CBD), shown in Figure 3, known as two essential phytocannabinoids, play a role with GPR18 signaling in both BV-2 cell endogenously expressing mouse GPR18 (mGPR18) and in hGPR18-transfected human embryonic kidney (HEK) cells when evaluated for Extracellular Signal-Regulated Kinase (pERK) activation. The results indicate that AEA, Δ^9-THC, O-1602, and abnormal cannabidiol (AbnCBD), along with NAGly, act as full agonists at GPR18, while a further Cannabis-derived compound CBD is a low efficacy agonist. In the other experiments, AEA, Δ^9-THC, and N-arachidonoyl glycine (NAGly) induced migration of the HEC-1B cells through CB1 receptor-independent mechanisms. Δ^9-THC also activated mitogen-activated protein kinase (MAPK) signaling at nanomolar concentrations in stably transfected HEK-293-GPR18 cells, whereas WIN55212-2 and CP55940 had no effect. Besides these results, Caldwell et al. [34] also offered evidence that the presence of GPR18-NAGly and Abn-CBD interactions in the anterior mouse eye (in the corneal epithelium, and interestingly in the trabecular meshwork) may serve as a desirable target for the regulation of elevated intraocular pressure in different forms of glaucoma. At the time, all of these data were sufficient to suggest GPR18 as a deorphanized receptor, but later studies retained GPR18 as a challenging receptor.



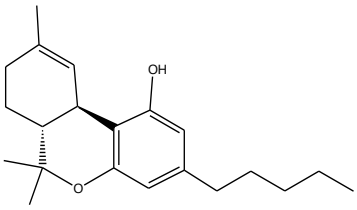
O1602



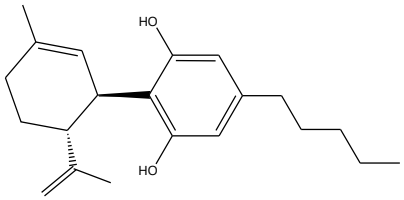
O1918



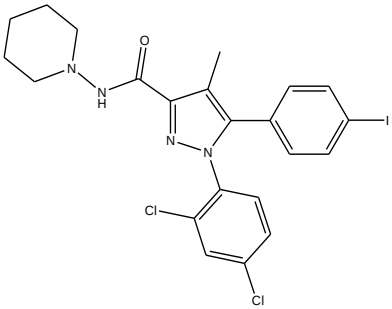
RvD2



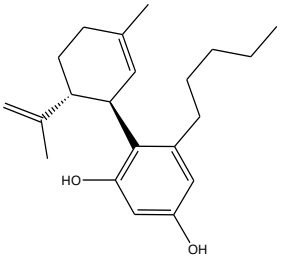
Δ^{9}-THC



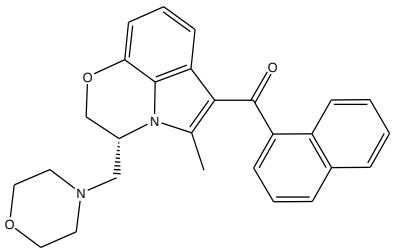
CBD



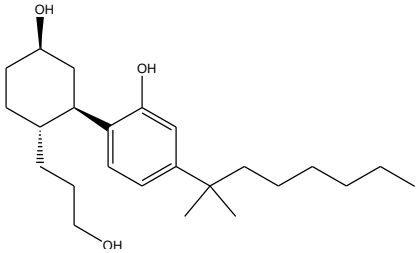
AM251



Abn-CBD



WIN55212-2



CP55940

Figure 3. GPR18 Proposed Agonists and Antagonists.

The second publication with a failed attempt toward finding a relation between NAGly and GPR18 was done by Ikeda et al. [46], which included entirely negative results. GPR18 heterologously expressed in rat sympathetic neurons, but the application of NAGly did not inhibit calcium currents. Not only NAGly but also an investigation of other proposed agonists of GPR18, including anandamide and abnormal cannabidiol, was useless. This failure also produced by the other downstream effectors of GPCRs. The level of cAMP was not changed by NAGly after the application of forskolin to stimulate cAMP production in GPR18-expressing cells. G protein-coupled inwardly rectifying potassium channels were not modulated by GPR18 signaling. Five different GPR18 receptor mutants were generated to test the basal activity of the receptor by monitoring calcium current in superior cervical ganglion neurons during agonist application, but it could not inhibit the induced calcium current. To explore that those results were not specific to the $G_{\alpha i}$ type proteins, they also tested other types of G_{α} proteins not endogenously expressed in superior cervical ganglion neurons like $G_{\alpha s}$, $G_{\alpha z}$, and $G_{\alpha 15}$, but these also failed to induce detectable calcium current changes when the cells were stimulated with the reported GPR18 ligands.

These apparent inconsistent findings sometimes were claimed to be related to a lack of co-localization between NAGly and GPR18 in the body [47]. In contrast to GPR18, in rats, the highest levels of NAGly were measured in the spinal cord, small intestine, kidneys, and skin. NAGly is much less expressed in high-level GPR18 distributed organs like testes and spleen [52]. However, it should be noted that some of the endogenous ligands such as AEA and 2-arachidonylglycerol (2-AG) are known to be

rapidly synthesized on demand. Therefore, NAGly might follow the same pattern and cannot be produced in an in vivo experiment. Finlay et al. [47] sought to find GPR18-mediated cellular signaling through the assays earlier studies reported. They utilized all of the earlier successful experienced cell lines and assays in response to both NAGly and AEA but their results detected neither inhibition of cAMP, induction of pERK and nor regulation of cellular migration and proliferation/apoptosis.

The other novel competitor of NAGly to deorphanize GPR18 is a potent immunoresolvent compound, Resolvin D2 (RvD2, see Figure 3). The complete stereochemistry of RvD2 is 7*S*,16*R*,17*S*-trihydroxy-4*Z*,8*E*,10*Z*,12*E*,14*E*,19*Z*-DHA [59]. A recently published unbiased screening for functional RvD2 GPCRs identified that GPR18 is a target for RvD2. Here both are localized in the immune system, including bone marrow, lymph nodes, tonsil, and specifically in leukocyte cells. Studies suggested that GPR18-RvD2 signaling stimulates the resolution of acute inflammation with a high binding affinity value (K_d ~ 10 nM) [42, 60, 61]. Although studies mainly focused on different levels of pro-resolving actions of resolvins, other new reports suggested that RvD2 coupling to GPR18 can also treat obesity that takes place either in adipose tissue or the obesity-associated hypothalamic inflammation in the central nervous system [41, 62].

The Abood laboratory (our collaborator in this project) has previously studied the effects of NAGly and other cannabinoids on stably expressing HEK293/GPR18 and CHO-K1 GPR18 β-arrestin cell lines through different G protein-coupled pathways (calcium mobilization and ERK1/2 phosphorylation assays), in addition to β-arrestin signaling (using the DisvoveRx PathHunter assay) [63]. In this study, NAGly, AbnCBD,

O-1602, O1918, and Δ^9-THC induce an increase in both calcium mobilization and MAPK activity (also this signaling inhibited by Pertussis toxin), whereas between all of the ligands tested, only Δ^9-THC induces a concentration-dependent response in β-arrestin recruitment assay. The most recent evidence of the effect of the reported GPR18 agonist NAGly was investigated in human spermatozoa [36]. Since the activation of protein kinases is involved in central physiological processes in sperm, this research investigated the phosphorylation status of 40 different protein kinases upon NAGly stimulation in human spermatozoa using a phospho-kinase array. The results show that the highest phosphorylation level was observed for the T308 phosphorylation site of the Akt kinase which has been confirmed by western blot analysis after exposure of an agonist. Chapter II reveals that our experimental data confirm activation of GPR18 with putative proposed agonists NAGly and RvD2 and also a novel agonist S5 is suggested.

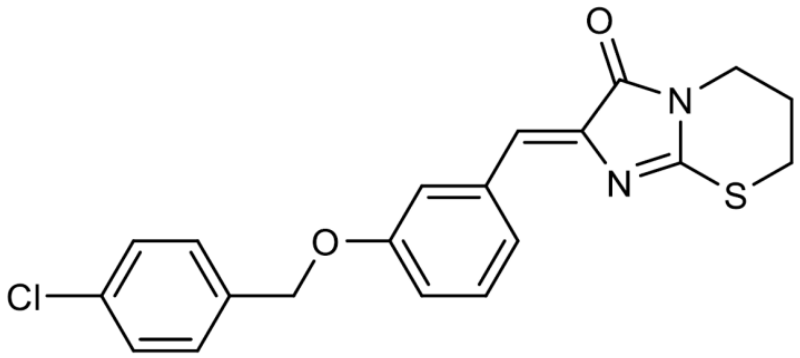
GPR18 Proposed Antagonists

Inactivation of the GPR18 biological effects and pharmacology is lesser-known than its activation. Endometriosis is a chronic painful disorder stimulated by GPR18-NAGly signaling in which endometrial cells grow abnormally outside of the uterus [57]. These results suggest that a GPR18 antagonist might be a treatment for this disorder. The first antagonist for GPR18 suggested by McHugh et al. [26, 50], who showed that CBD is a low-efficacy agonist behaving as partial agonist/antagonist depending on receptor expression level; whereas the CBD analog, O-1918, acts as an antagonist by blocking NAGly-induced migration in BV-2 microglial migration. Although this report was not confirmed by later studies [63], in other experiments, they showed that AM251 and CBD

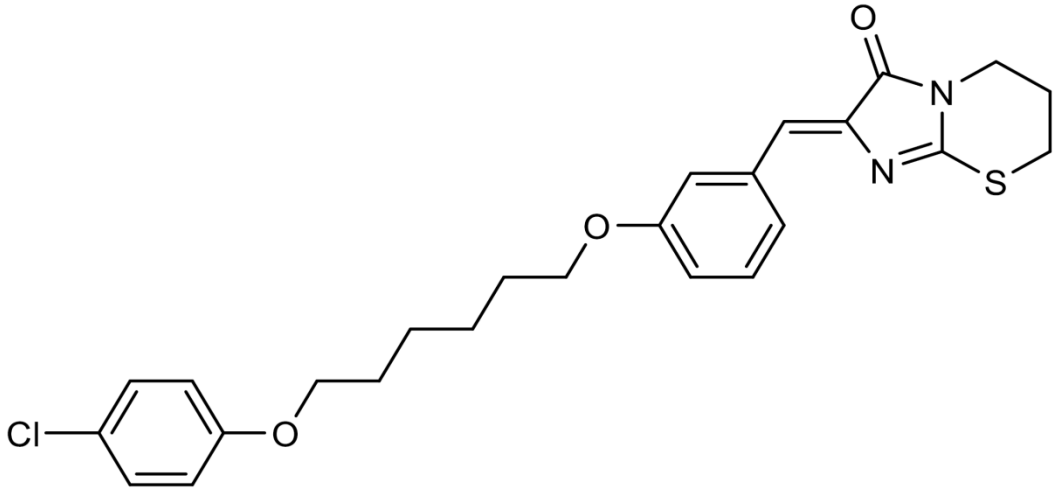
antagonized the HEC-1B migration and MAPK activation by NAGly and Δ^9-THC. It is important to note that AM251 is a widely used CB1-selective antagonist and often is used in combination with Δ^9-THC to suggest a role for CB1 receptors, but here its role deducted to a weak partial agonist/antagonist. Thus far, the first and only published selective GPR18 antagonists are PSB-CB5 (Z-2-(3-(4-Chlorobenzyloxy)benzylidene)-6,7-dihydro-2H-imidazo[2,1-b][1,3]-thiazin-3(5H)-one) and PSB-CB27(Z-2-(3-(6-(4Chlorophenoxy)hexyloxy)benzylidene)-6,7-dihydro-2Himidazo[2,1b][1,3]thiazin-3(5H)-one) (see Figure 4) with IC_{50}~1μM by Δ^9-THC-induced β-arrestin recruitment to GPR18 assay. Using β-arrestin translocation assays based on the β-galactosidase enzyme fragment complementation technology (β-arrestin PathHunter™ assay, DiscoveRx, due to the lack of potent radioligand for GPR18 and GPR55) and radioligand binding assays vs. [^{3}H]CP55,940 for CB1 and CB2, Harms et al. [64] presented an extensive structure-activity relationship study of 49 derivatives investigated at all four GPCRs and revealed that PSB-CB5 and PSB-CB27 are the most potent and selective antagonist for GPR18 in the recombinant cell line (CHO) stably expressing the respective human GPCR. Later studies have confirmed these results [36]. The co-treatment of the selective GPR18 antagonist PSB-CB5 and PSB-CB27 with NAGly (1:1, 1μM) completely blocked the NAGly-induced acrosomal exocytosis in humans.

Furthermore, they also investigated the effects of the suggested antagonists on THC-induced acrosomal exocytosis. Again, these effects could be inhibited by the specific GPR18 antagonists PSB-CB5 and PSB-CB27 (Figure 4). The next chapter would

show that how our experimental data confirm these antagonists and then introduce and compare our novel proposed antagonists with the previous one.



PSB-CB5



PSB-CB27

Figure 4. Two Proposed Recent Antagonists for GPR18.

CHAPTER II

EXPERIMENTAL DATA

Introduction

The experimental data described here were submitted as an NIH R01 grant (Application Number: 1 R01 DA049533-01) High Selectivity Antagonist and Agonist Hits for GPR18, which was a well-established collaboration between Dr. Abood (a well-recognized expert in the pharmacology of cannabinoid and related GPCRs), Dr. Reggio (a well-recognized expert in molecular modeling of GPCRs), and Dr. Jagerovic (from the synthetic lab that carried out the medicinal chemistry component of the research).

As discussed in Chapter I, many earlier investigations, including ours, support activation of GPR18 by the endogenous lipoamino acid N-arachidonoyl glycine (NAGly), the endogenous immunomodulating lipid resolvin (RvD2), and Δ9-tetrahydrocannabinol (Δ9-THC). In contrast, there have been reports indicating a lack of activation by NAGly. Our experimental data generated by stable cell lines expressing the human GPR18 in CHO cells demonstrate that both NAGly and RvD2 activate GPR18. We also employed the recent introduced GPR18 antagonists PSBCB-5 and PSBCB-27 to antagonize NAGly and RvD2 effects.

However, no highly potent GPR18 antagonist has yet been discovered. Therefore, the first goal of this project was to design a nanomolar potency antagonist that could guide *the design* of future high-affinity antagonists and lead to a better understanding of

the physiological role of GPR18. In this way, we recently synthesized a series of GPR18 agonists and antagonists based upon a pyrazol-3(5)-ylbenzene-1,3-diol scaffold and compared them with NAGly, RvD2, and Δ^9-THC, which triggers activation of GPR18. Candidate compounds were tested for their ability to inhibit calcium mobilization, MAPK signaling, and recruitment of β-arrestin in stably transfected CHO cells expressing GPR18.

Cellular Expression of GPR18

We have synthesized both proposed antagonists PSB-CB5 and PSB-CB27 (see Figure 4), and have generated a stable cell line of Chinese Hamster Ovary cells (CHO) expressing human GPR18 (cDNA resource, GP0180000; CHO/GPR18). As Figure 5 shows, the levels of GPR18 expression were qualitatively illustrated in this cell line, wild-type CHO cells (negative control), and a commercially available CHO cell line co-expressing GPR18 and β-arrestin-2 via Western hybridization protocols according to the methodology employed in Dr. Abood's laboratory and as previously described [63].

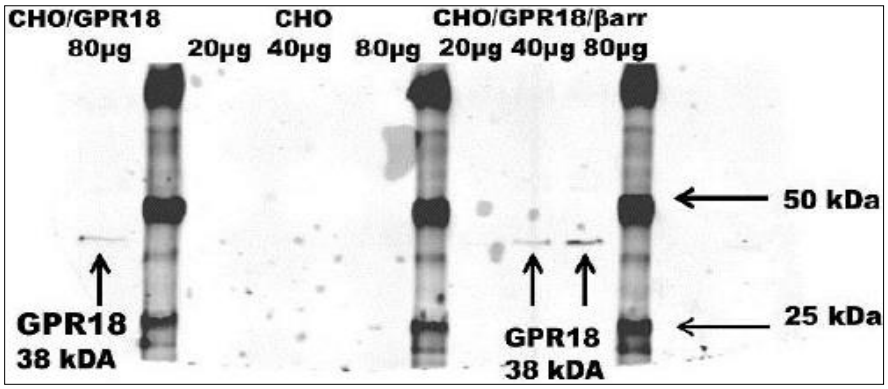


Figure 5. Western Blot of CHO/GPR18, Wild-type CHO, and CHO/GPR18/βarr Depicting Expression of GPR18 (38kDA). The Membrane was Probed with Rabbit, Anti-GPR18 (2ug/ml; Sigma). Note the Lack of Signal at All Concentration in Wild-type CHO Cells.

Calcium Mobilization of NAGly in CHO/GPR18

Intracellular calcium measurements were initially used because it is the most robust assay found for activity at GPR18. Distribution of GPR18 in the human N-terminal GFP-tagged-GPR18 (CHO/GPR18) cells is illustrated using immunocytochemistry activity. Figure 6A (left) depicts the stable expression of the GPR18, which is tagged with the green fluorescent protein (GFP). The color change from green to blue (using Fura-2/AM fluorescence dye to detect intracellular calcium) in

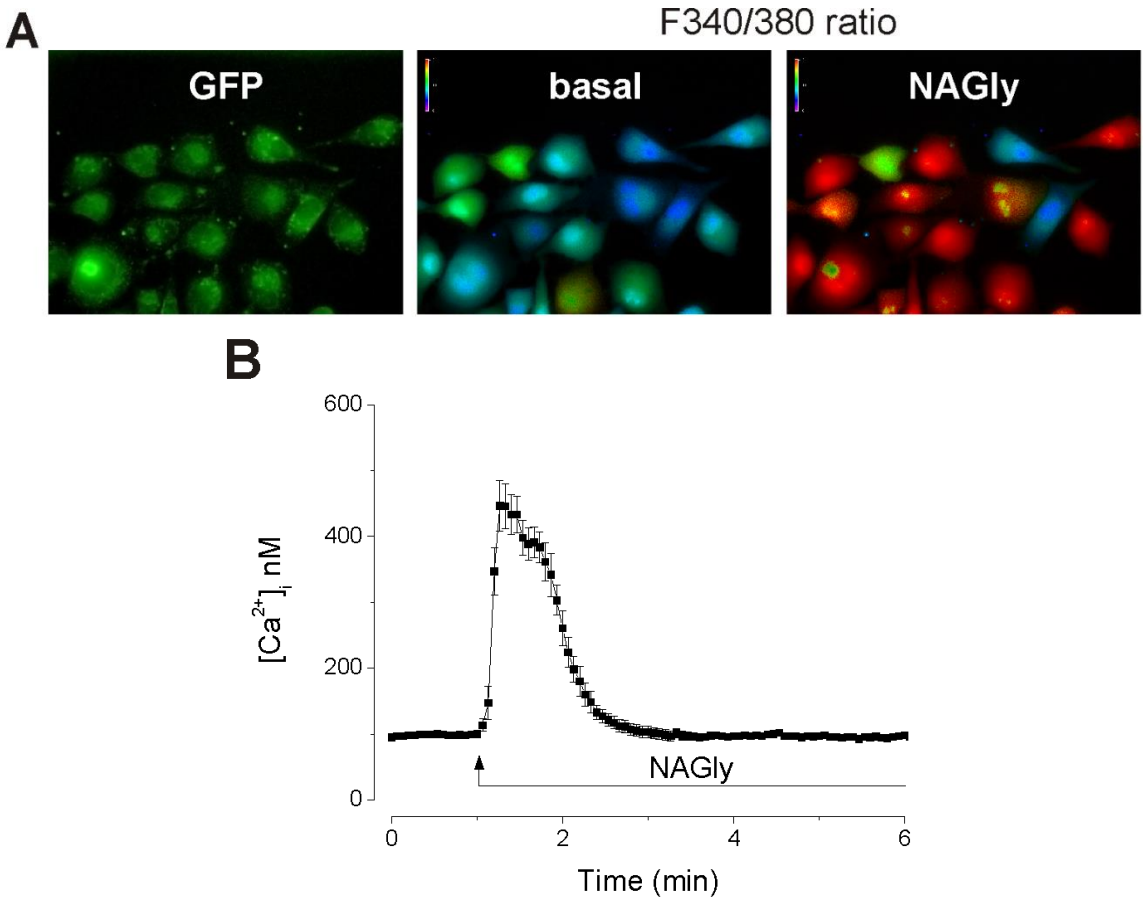


Figure 6. NAGly Increases $[Ca^{2+}]_I$ in CHO/GPR18 Cells. A. GFP Signal, Representative Examples of F340/F380 Ratio before (Basal) and after Treatment with 10 µM NAGly. B. Representative Examples of NAGly-induced Increase in $[Ca^{2+}]_i$. The arrow indicated the addition of NAGly.

Figure 6A (middle) indicates calcium mobilization in the basal activity, which means activation of the receptor is happening by itself. The wavelengths used to excite and detect the fluorophores for GFP is at 488 and 595 nm respectively and 340/380 nm and 510 nm for Fura-2 AM.

Activation of GPR18 after treatment with 10 μM NAGly is also visualized compared with the basal activity color in Figure 6.

MAPK Activity of NAGly and RvD2 in CHO/GPR18 Cells

Our experimental data also demonstrate that both putative antagonists of GPR18 are efficacious as inhibitors of NAGly and RvD2-induced MAPK activation (see Figure 7). As the bar charts in Figure 7 depict, we first tested ERK 1/2 phosphorylation by NAGly and RvD2 (both in 1μM concentration), the two most important proposed agonists for activation of GPR18. Then, these responses were inhibited after the two proposed selective antagonists PSB-CB5 (10μM) and PSB-CB27 (3μM) were treated.

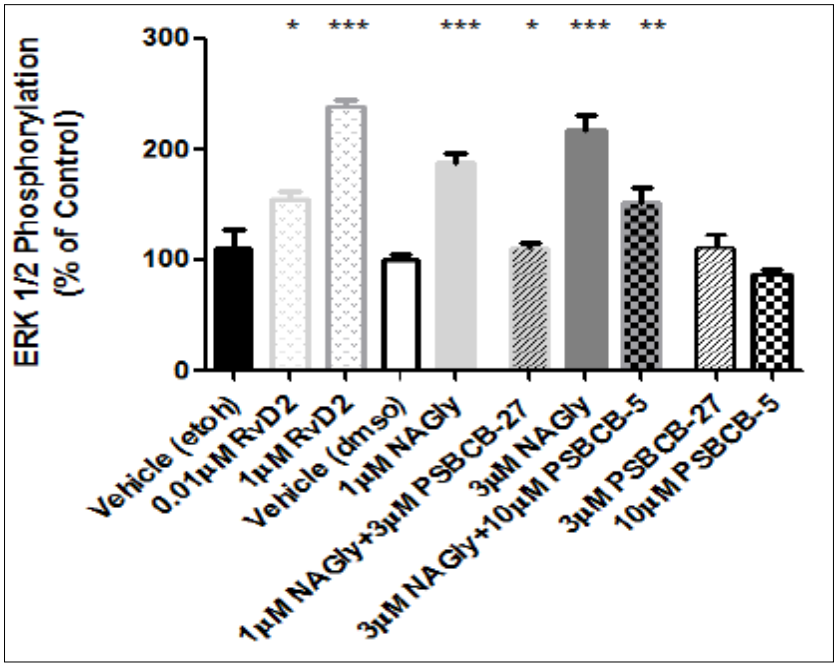


Figure 7. Inhibition of NAGly-induced Increases in MAPK Activity by GPR18 Antagonists in CHO/GPR18 (n = 4-19). Significant Increases in MAPK Activity Upon Exposure to NAGly (5 min) and RvD2 (30 Min). Significance Was Determined by One-way ANOVA with Bonferroni Correction, $p < 0.001$ as Compared to Respective Vehicle Control (DMSO/NAGly;etoh/RvD2). Inhibition of NAGly-induced Increases Were Blocked (30-min pre-incubation) with PSBCB-27 and PSBCB-5 (One-way ANOVA with Bonferroni correction, $p < 0.05$ and $p < 0.01$, Respectively).

Calcium Mobility Activity of Our Novel Proposed Agonist and Antagonist of GPR18 in CHO/GPR18 Cell Lines

Despite controversial discussions, especially about NAGly's ability to bind with GPR18, our experimental data prove that NAGly and RvD2 are two putative endogenous agonists for activation of GPR18. We recently synthesized a series of GPR18 agonists and antagonists based upon a pyrazol-3(5)-ylbenzene-1,3-diol scaffold for their ability to regulate intracellular calcium (Ca^{2+}) in our cell line. The best antagonist from this initial series **S4** (see Figure 8) at 10 μM can reduce NAGly calcium signaling completely, while 10 μM PSBCB-5 can only reduce a NAGly-generated calcium signal by 40 to 50%. The best agonist in these series, **S5** (see Figure 8), has an efficacy comparable to NAGly.

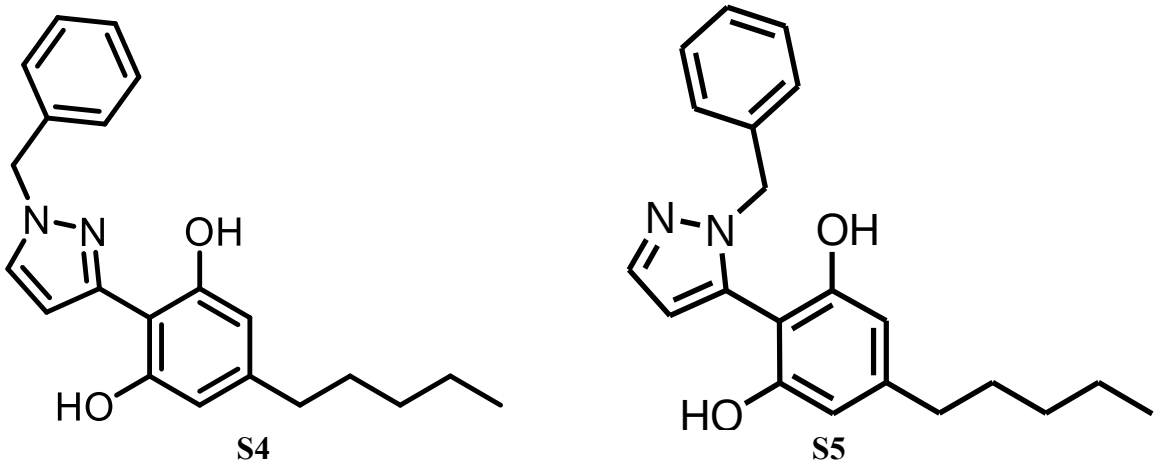


Figure 8. The Proposed Agonist S5 and Antagonist S4 Based on the Pyrazol-3(5)-ylbenzene-1,3-diol Scaffold.

As Figure 9 illustrates, the real-time photographs of calcium mobilization in S5 treated states are very similar to NAGly-treated states in Figure 6, which means that S5 is as efficacious as NAGly. Figure 9A (left) depicts the stable expression of the GPR18,

which is tagged with the green fluorescent protein (GFP), with the color change from green to blue indicating calcium mobilization in basal activity (middle) and activation of GPR18 after treatment with 10 μM S5 (right). Figure 9B reveals the change of intracellular calcium in CHO/GPR18 cells after exposure to 10 μM S5. Figure 9C illustrates the bar graph of the mean increase of calcium following exposure to 10 μM S5.

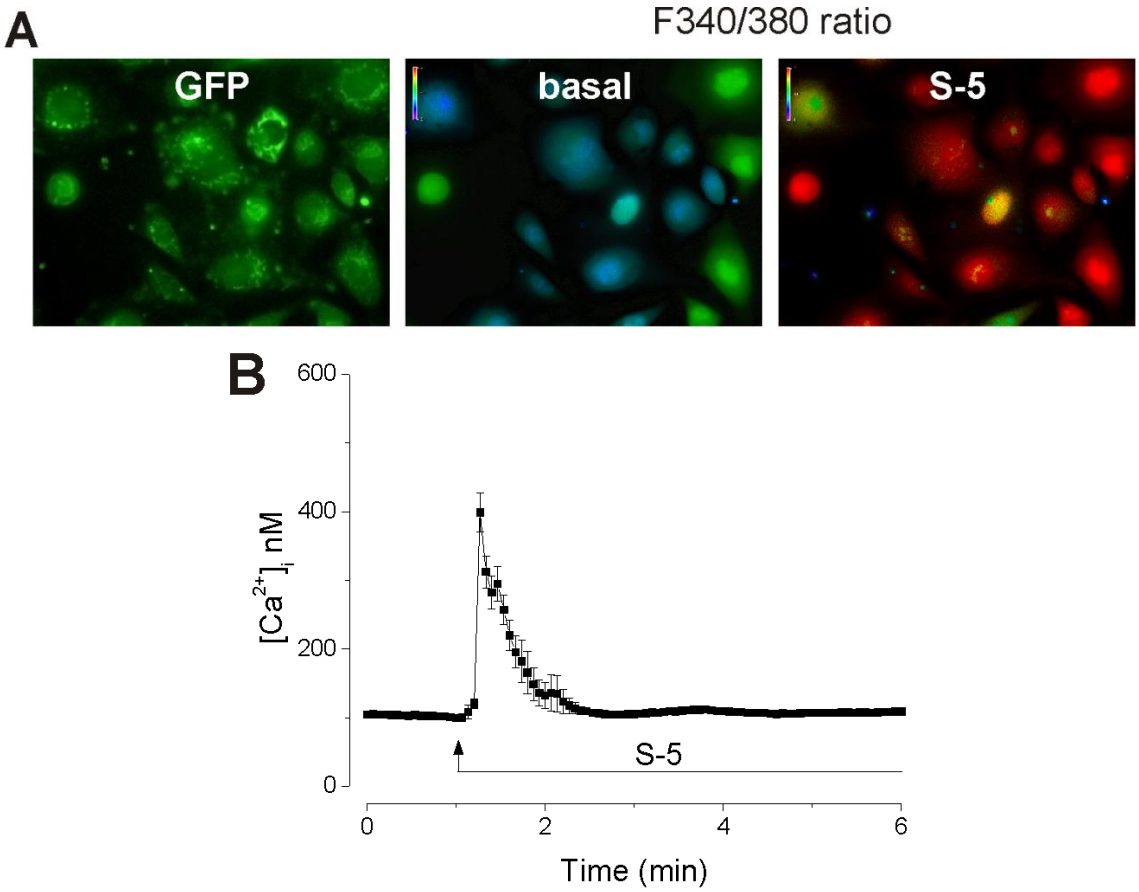


Figure 9. S5 Increases $[Ca^{2+}]_I$ in CHO/GPR18 Cells. A. GFP Signal, Representative Examples of F340/F380 Ratio before (Basal) and after Treatment with 10 μM S5. B. Representative Examples of S5-induced Increase in $[Ca^{2+}]_i$. The arrow indicates the addition of S5.

S4 is an antagonist at GPR18 with a closely related analog of S5. Figure 10 shows that S4 does not affect the basal activity and the intensity and type of colors did not change after S4 treatment, which means that S4 cannot affect the basal activity of GPR18. But as Figure 11 shows, an equimolar concentration of S4 inhibits the NAGly response. The comparison between NAGly treatment states in Figure 6 and NAGly-S4 treatment state in Figure 11 clarifies that S4 is antagonizing NAGly's effects.

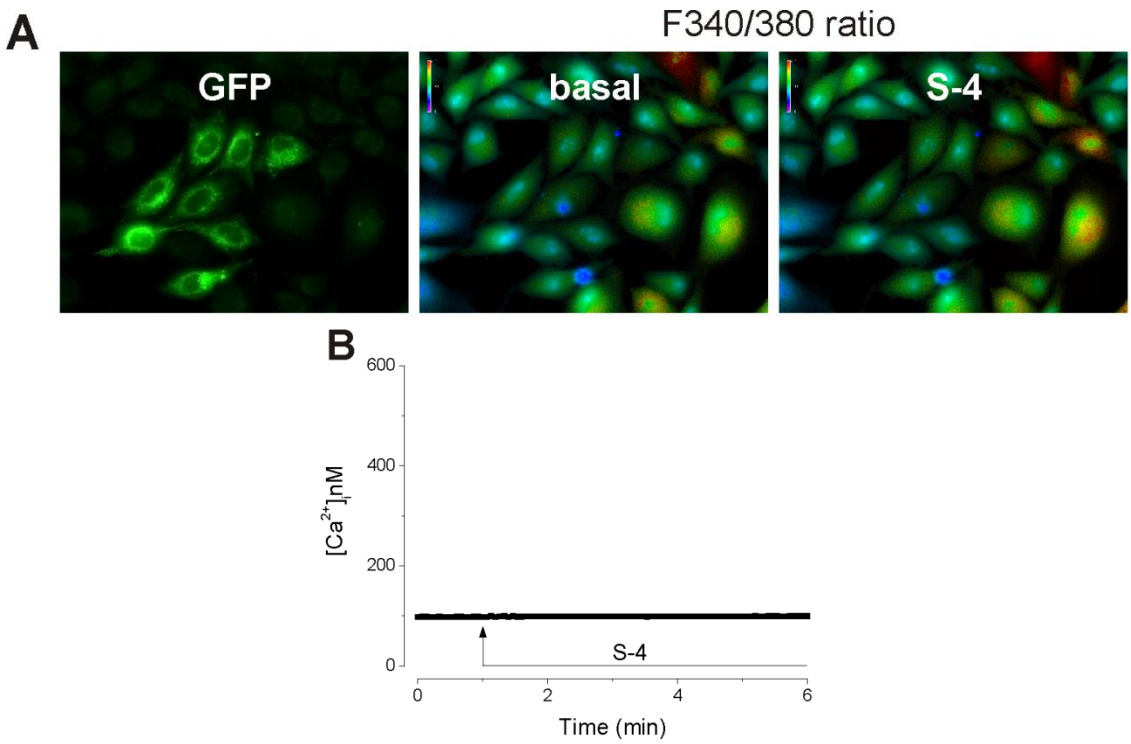


Figure 10. S4 Alone Has No Effect on $[Ca^{2+}]_I$ in CHO/GPR18 Cells. A. GFP Signal, Representative Examples of F340/F380 Ratio before (Basal) and after Treatment with 10 μM S4. B. Representative Examples of S4-induced Increase in $[Ca^{2+}]_i$. The arrow indicates the addition of S4.

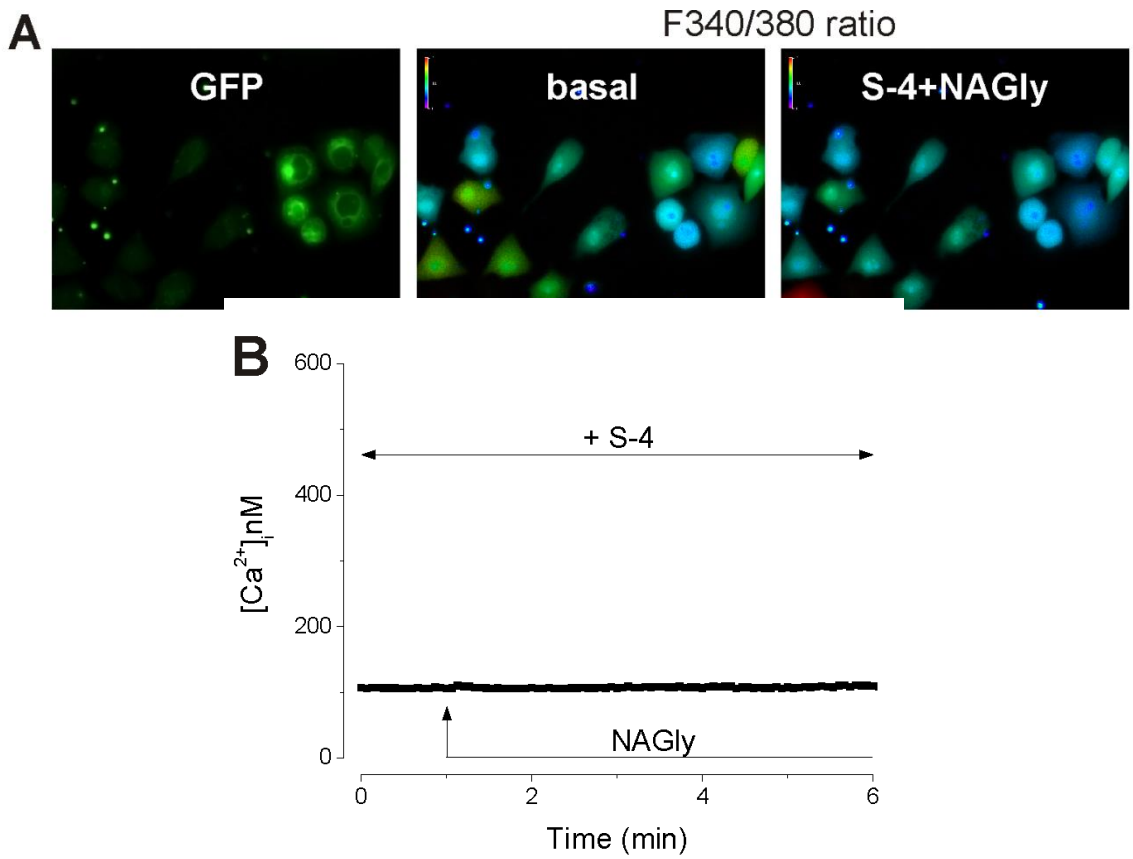


Figure 11. S4 Blocks NAGly Response on $[Ca^{2+}]_I$ in CHO/GPR18 Cells. A. GFP Signal, Representative Examples of F340/F380 Ratio before (Basal) and after Treatment with 10 μM S4+ NAGly in HEK Cells. B. Representative Examples of S4 + NAGly-induced Increase in $[Ca^{2+}]_i$. The top arrow indicates the prsence of S4 in the whole of the experiment and the bottom arrow indicates the addition of NAGly.

Comparison between our proposed antagonist S4 and earlier suggested antagonist PSB-CB5, our experimental data show that S4 is a more efficacious antagonist than PSBCB-5 at inhibiting NAGly stimulation of intracellular calcium (see Figure 12). As the bar charts in Figure 12 depict, S4 antagonism much more significantly affected NAGly-induced calcium mobilizations.

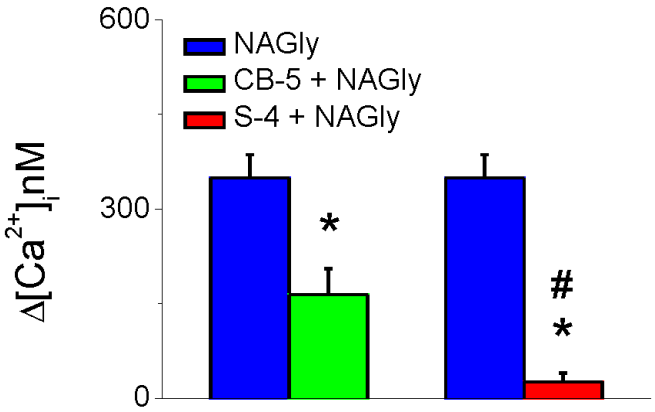


Figure 12. Antagonism of Intracellular Ca^{2+} Signal by S4 vs. CB-5. S-4 (10 μM) Fully Antagonizes the Ca^{2+} Signal Produced by NAGly (10 μM), Whereas CB-5 (10 μM) Only Partially Blocks the Signal. * $p<0.05$ Compared to NAGly, # $p<0.05$ Compared to CB-5. n=45-50 Cells.

β-arrestin Activity of Agonists in GPR18 DiscoverX Cells

Earlier findings suggest that GPR18 activation involves several signal transduction pathways, indicative of biased agonism, as activation of this receptor by a particular ligand, which might provide a possible explanation for the controversial GPR18 activation found in different literature. For instance, previous reports show that while all five ligands Δ^9-THC, NAGly, AbnCBD, O-1602, and O-1918 increase both intracellular calcium and MAPK cellular signaling pathways, in HEK/GPR18 cells, only Δ^9-THC produces concentration-dependent GPR18-mediated β-arrestin activity [63]. Our experimental data revealed that S5 also recruits β- arrestin (see Figure 13). Quantification of β-arrestin recruitment was determined using DiscoveRx PathHunter CHO-K1 cells stably expressing GPR18 fused with a β-galactosidase enzyme fragment, and β-arrestin fused to an N-terminal deletion mutant of β-galactosidase. Activation of GPR18 induces β-arrestin recruitment, forcing complementation of the two β-galactosidase enzyme

fragments. Levels of this active enzyme are a direct result of GPR18 activation and are quantitated using chemiluminescent PathHunter detection reagents containing the β-galactoside substrate [63].

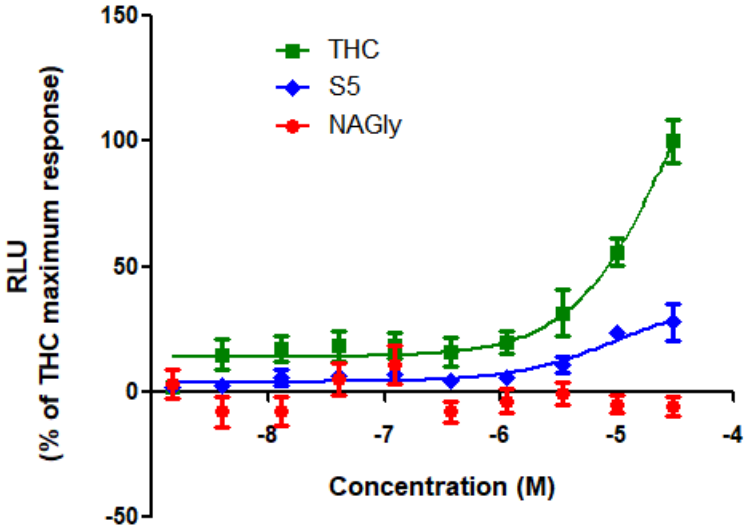


Figure 13. β-arrestin Activity in GPR18 DiscoverX Cells. Data Were Normalized to the Response Produced by Δ^9-THC. n=3 Experiments.

Conclusion

Thus far, potent and specific agonists and antagonists of GPR18 have not been identified, which is the reason for preventing full acceptance of NAGly as an agonist, development of radioligands, and the advancement of GPR18's therapeutic potential. We have found a distinct set of agonists and antagonists based on the pyrazol-3(5)-ylbenzene-1,3-diol scaffold to interact with GPR18, either G protein-dependent or by β-arrestin signaling pathways. Taking into consideration the results presented in this chapter allows us to employ the structure-activity relationship (SAR) technique based on the pyrazol-

3(5)-ylbenzene-1,3-diol scaffold to modify the effect or potency of S4 and S5 by changing their chemical structures.

CHAPTER III

GPR18 MODELING

Introduction

Thus far, there is no structural information available for GPR18 because neither X-ray crystallographic nor nuclear magnetic resonance (NMR) spectroscopic studies have been reported. To bridge this gap, computer models of GPR18 (inactive state, R; activated state, R*) are proposed to be developed here. These models will be used to study ligand binding and activation at GPR18. By employing a GPR18 computer model, the putative binding site of candidate agonists can be identified and then tested via mutational analyses. However, no highly potent GPR18 antagonist has yet been discovered. Therefore, the importance of building a 3D structure of GPR18 is to help design a nanomolar potency antagonist that can guide the design of future high-affinity antagonists and lead to a better understanding of the physiological role of GPR18.

Molecular Modeling the Inactive State (R) of GPR18

A homology model of GPR18 was constructed using the X-ray crystal structure of the delta opioid receptor (DOR) at 1.8 Å resolution as a template downloaded from Protein Data Bank [65]. The DOR was chosen as the template because it has 26% sequence homology (similarity) with the GPR18 and an analogous sodium binding pocket region. Sequence similarities include key Cysteines in extracellular (EC) loop 2 (C(172) in GPR18; C(198) in DOR) that form a disulfide loop with C3.25 on TMH3

(yellow circles in Figure 14). This helps limit the conformational flexibility of the EC2 loop and stabilizes its structure. The GPR18 sequence was first aligned with the sequences of the Rho, CB1, CB2, DOR, β_2-AR, and S1P1 receptors using the highly conserved Class A GPCR residues/motifs as alignment guides (N1.50, D2.50, R3.50, W4.50 (green circles in Figure 14), P5.50, P6.50 (red circles in TMH5 and 6 in Figure 14), and NPXXY in TMH7). Based on this alignment, the DOR model was then mutated to the GPR18 sequence. Two important GPR18 motif variations are CFMP vs. CWXP in TMH6 and DVILY vs. NPXXY in TMH7. Also, GPR18 has a P3.36 that should bend TMH3 (see Table 1).

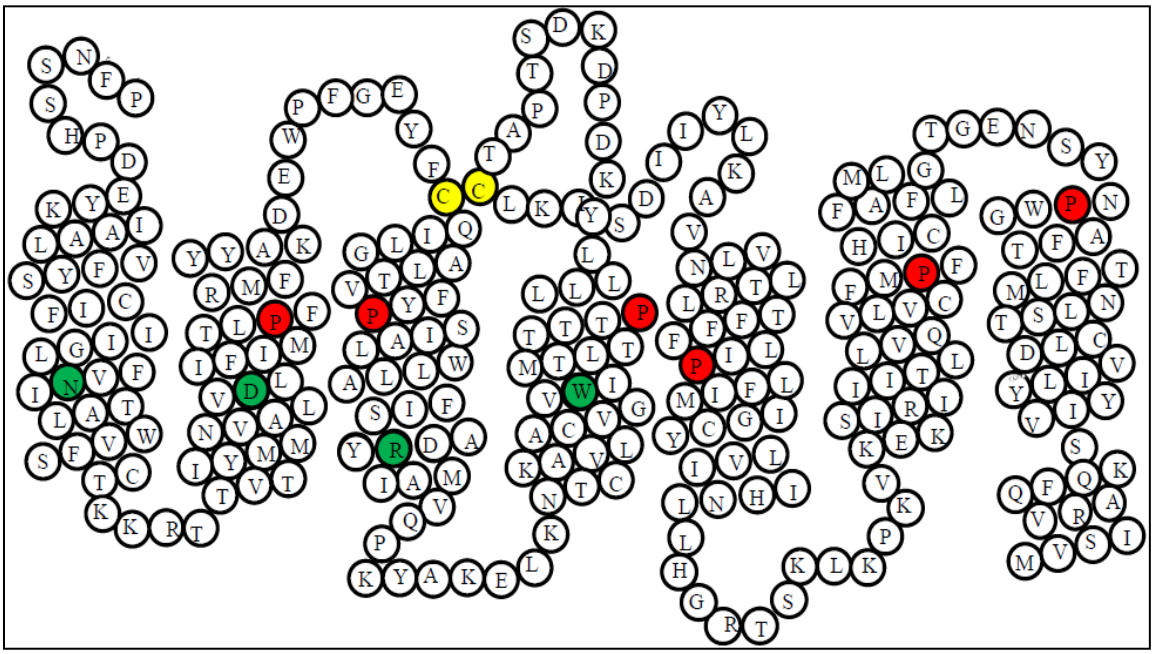


Figure 14. Helix Net Representation of hGPR18 Sequence. Residues in Red are Prolines; All Except the Proline in Helix III are Highly or Reasonably Conserved. Residues in Green are Highly Conserved Residues in GPCRs Class A and Residues in Yellow are Highly Conserved Cysteins Which Form Key Disulfide Bridge in Extracellular Loop 2.

Table 1

Conserved Residues/Motifs Presented in a Sample of Class A GPCRs in TMH7 and 3

	7.34	7.35	7.36	7.37	7.38	7.39	7.40	7.41	7.42	7.43	7.44	7.45	7.46	7.47	7.48	7.49	7.50	7.51	7.52	7.53	7.54	7.55	7.56	7.57	7.58	7.59
Rho	F	M	T	I	P	A	F	F	A	K	S	A	A	I	Y	N	P	V	I	Y	I	M	M	N	K	Q
D3	L	Y	S	A	T	T	**W**	L	G	**Y**	V	N	S	A	L	N	P	V	I	Y	T	T	F	N	I	E
D2L3	L	Y	S	A	F	T	**W**	L	G	**Y**	V	N	S	A	V	N	P	I	I	Y	T	T	F	N	I	E
D2S	L	Y	S	A	F	T	**W**	L	G	**Y**	V	N	S	A	V	N	P	I	I	Y	T	T	F	N	I	E
D2L	L	Y	S	A	F	T	**W**	L	G	**Y**	V	N	S	A	V	N	P	I	I	Y	T	T	F	N	I	E
D1	T	F	D	V	F	V	**W**	F	G	W	A	N	S	S	L	N	P	I	I	Y	A	F		N	A	D
GPR119	L	E	R	Y	L	W	L	L	G	V	G	N	S	L	L	N	P	L	I	Y	A	Y	W	L	K	E
M2	V	W	T	I	G	Y	**W**	L	C	**Y**	I	N	S	T	I	N	P	A	C	Y	A	L	C	N	A	T
LPAR1	Y	E	K	F	F	L	L	L	A	E	F	N	S	A	M	N	P	I	I	Y	S	Y	R	D	K	E
S1P1	R	A	E	Y	F	L	V	L	A	V	L	N	S	G	T	N	P	I	I	Y	T	L	T	N	K	E
CB1	V	F	A	F	C	S	M	L	C	L	L	N	S	T	V	N	P	I	I	Y	A	L	R	S	K	D
CB2	A	F	A	F	C	S	M	L	C	L	I	N	S	M	V	N	P	V	I	Y	A	L	R	S	G	E
Beta-2	V	Y	I	L	L	N	**W**	I	G	**Y**	V	N	S	G	F	N	P	L	I	Y	C	R		S	P	D
A2A	L	M	Y	L	A	I	V	L	S	H	T	N	S	V	V	N	P	F	I	Y	A	Y	R	I	R	E
Beta-1	L	F	V	F	F	N	**W**	L	G	**Y**	A	N	S	A	F	N	P	I	I	Y	C	R		S	P	D
MOR	S	W	H	F	C	I	A	L	G	**Y**	T	N	S	C	L	N	P	V	L	Y	A	F	L	D	E	N
DOR	A	L	H	L	C	I	A	L	G	**Y**	A	N	S	S	L	N	P	V	L	Y	A	F	L	D	E	N
KOR	S	Y	Y	F	C	I	A	L	G	**Y**	T	N	S	S	L	N	P	I	L	Y	A	F	L	D	E	N
CXCR4	W	I	S	I	T	E	A	L	A	F	F	H	C	C	L	N	P	I	L	Y	A	F	L	G	A	K
P2Y12	V	K	E	S	T	L	**W**	L	T	S	L	N	A	C	L	D	P	F	I	Y	F	F	L	C	K	S
5HT2A	L	L	N	V	F	V	**W**	I	G	**Y**	L	S	S	A	V	N	P	L	V	Y	T	L	F	N	K	T
GPR18	W	G	A	F	T	T	F	L	M	N	L	S	T	C	L	D	V	I	L	Y	Y	I	V	S	K	Q

	3.22	3.23	3.24	3.25	3.26	3.27	3.28	3.29	3.30	3.31	3.32	3.33	3.34	3.35	3.36	3.37	3.38	3.39	3.40	3.41	3.42	3.43	3.44	3.45	3.46	3.47	3.48	3.49	3.50	3.51	3.52	3.53
Rho	T	G	C	N	L	E	G	F	F	A	T	L	G	G	E	I	A	L	W	S	L	V	V	L	A	I	E	R	Y	V	V	V
D3	I	C	C	D	V	F	V	T	L	D	V	M	M	C	T	A	S	I	L	N	L	C	A	I	S	I	D	R	Y	T	A	V
D2L3	I	H	C	D	I	F	V	T	L	D	V	M	M	C	T	A	S	I	L	N	L	C	A	I	S	I	D	R	Y	T	A	V
D2S	I	H	C	D	I	F	V	T	L	D	V	M	M	C	T	A	S	I	L	N	L	C	A	I	S	I	D	R	Y	T	A	V
D2L	I	H	C	D	I	F	V	T	L	D	V	M	M	C	T	A	S	I	L	N	L	C	A	I	S	I	D	R	Y	T	A	V
D1	S	F	C	N	I	W	V	A	F	D	I	M	C	S	T	A	S	I	L	N	L	C	V	I	S	V	D	R	Y	W	A	I
GPR119	T	L	C	S	L	R	M	A	F	V	T	S	S	A	A	A	S	V	L	T	V	M	L	I	T	F	D	R	Y	L	A	I
M2	V	V	C	D	L	W	L	A	L	D	Y	V	V	S	N	A	S	V	M	N	L	L	I	I	S	F	D	R	Y	F	C	V
LPAR1	S	T	W	L	L	R	Q	G	L	I	D	T	S	L	T	A	S	V	A	N	L	L	A	I	A	I	E	R	H	I	T	V
S1P1	A	Q	W	F	L	R	E	G	S	M	F	V	A	L	S	A	S	V	F	S	L	L	A	I	A	I	E	R	Y	I	T	M
CB1	N	V	F	L	F	K	L	G	G	V	T	A	S	F	T	A	S	V	G	S	L	F	L	T	A	I	D	R	Y	I	S	I
CB2	A	V	F	L	L	K	I	G	S	V	T	M	T	F	T	A	S	V	G	S	L	L	L	T	A	I	D	R	Y	L	C	L
Beta-2	F	W	C	E	F	W	T	S	I	D	V	L	C	V	T	A	S	I	E	T	L	C	V	I	A	V	D	R	Y	F	A	I
A2A	H	G	C	L	F	I	A	C	F	V	L	V	L	T	Q	S	S	I	F	S	L	L	A	I	A	I	D	R	Y	I	A	I
Beta-1	F	F	C	E	L	W	T	S	V	D	V	L	C	V	T	A	S	I	E	T	L	C	V	I	A	L	D	R	Y	L	A	I
MOR	I	L	C	K	I	V	I	S	I	D	Y	Y	N	M	F	T	S	I	F	T	L	C	T	M	S	V	D	R	Y	I	A	V
DOR	L	L	C	K	A	V	L	S	I	D	Y	Y	N	M	F	T	S	I	F	T	L	T	M	M	S	V	D	R	Y	I	A	V
KOR	V	L	C	K	I	V	I	S	I	D	Y	Y	N	M	F	T	S	I	F	T	L	T	M	M	S	V	D	R	Y	I	A	V
CXCR4	F	L	C	K	A	V	H	V	I	Y	T	V	N	L	Y	S	S	V	L	I	L	A	F	I	S	L	D	R	Y	L	A	I
P2Y12	F	V	C	Q	V	T	S	V	I	F	Y	F	T	M	Y	I	S	I	S	F	L	G	L	I	T	I	D	R	Y	Q	K	T
5HT2A	K	L	C	A	V	W	I	Y	L	D	V	L	F	S	T	A	S	I	M	H	L	C	A	I	S	L	D	R	Y	V	A	I
GPR18	Y	F	C	Q	I	L	G	A	L	T	V	F	Y	P	S	I	A	L	W	L	L	A	F	I	S	A	D	R	Y	M	A	I

Note. The far-left column is the abbreviation of the GPCRs. The last one is GPR18. The first row shows the Ballesteros–Weinstein residue numbering. In this numbering system in GPCRs that uses the X.YY format; X denotes the TMH number and (YY) denotes residue position relative to the most conserved residue in the helix (X.50) [67] The residues in yellow are highly conserved, in green are reasonably conserved and in red is Proline. Cysteins in position 3.25 incorporates in disulfide bridge. As the alignment shows P3.36 is unique in GPR18. Also, unlike other GPCRs, there is not a NPXXY motif in TMH7.

To explore the consequences of helix bending residues, such as P3.36 in TMH3, and the lack of P7.50 in TMH7, we employed the combined Monte Carlo (MC) and Simulated Annealing (SA) technique, Conformational Memories (CM). CM allows an exploration of the flexibility introduced by helix bending residues such as Proline and outputs a set of low free energy helix conformations from which a new helix conformer can be chosen.

Definition of Torsion Angles

Figure 15 illustrates [66] the definition of the protein backbone dihedral angles. In a protein chain three dihedral angles are defined as φ (phi), ψ (psi) and ω (omega). Dihedral angle φ (phi) is rotation about the bond between N and Cα bond, involves C(O)-N-Cα-C(O). Dihedral angle ψ (psi) is rotation about the bond between Cα-C(O) bond involves N-Cα-C(O)-N. Dihedral angle ω (omega) is the rotation about the peptide bond C(O)-N bond involves Cα1-C(O)-N-Cα2. Since carbonyl pi electrons and nitrogen lone pair are delocalized and from resonance structure, the Omega angle tend to be planar, but phi and psi are flexible; therefore, the rotation happens around them. The side chain torsion angles are named χ1 (chi1), χ2 (chi2), χ3 (chi3), etc.; between them, χ1 (chi1) is the most studied. Three important states of χ1 (chi1) dihedral angles tend to cluster near 180°, 60°, and −60°, which are called the trans, gauche plus, and gauche minus conformations, respectively. The most abundant conformation is gauche plus in which the gamma side chain atom is opposite to the residue's main chain carbonyl group when viewed along the Cβ-Cα bond. Trans conformation is when the side chain gamma atom is opposite the main chain nitrogen. Finally, one of the rarest, especially in the

transmembrane region of GPCRs, is gauche minus. It occurs when the side chain is opposite the hydrogen substituent on the Cα atom. This conformation is not stable because the gamma atom is in the proximity of the backbone CO and NH atoms. As discussed in the next section, this conformation is adopted by the Serine and Threonine residues, which makes a kink in the alpha helix.

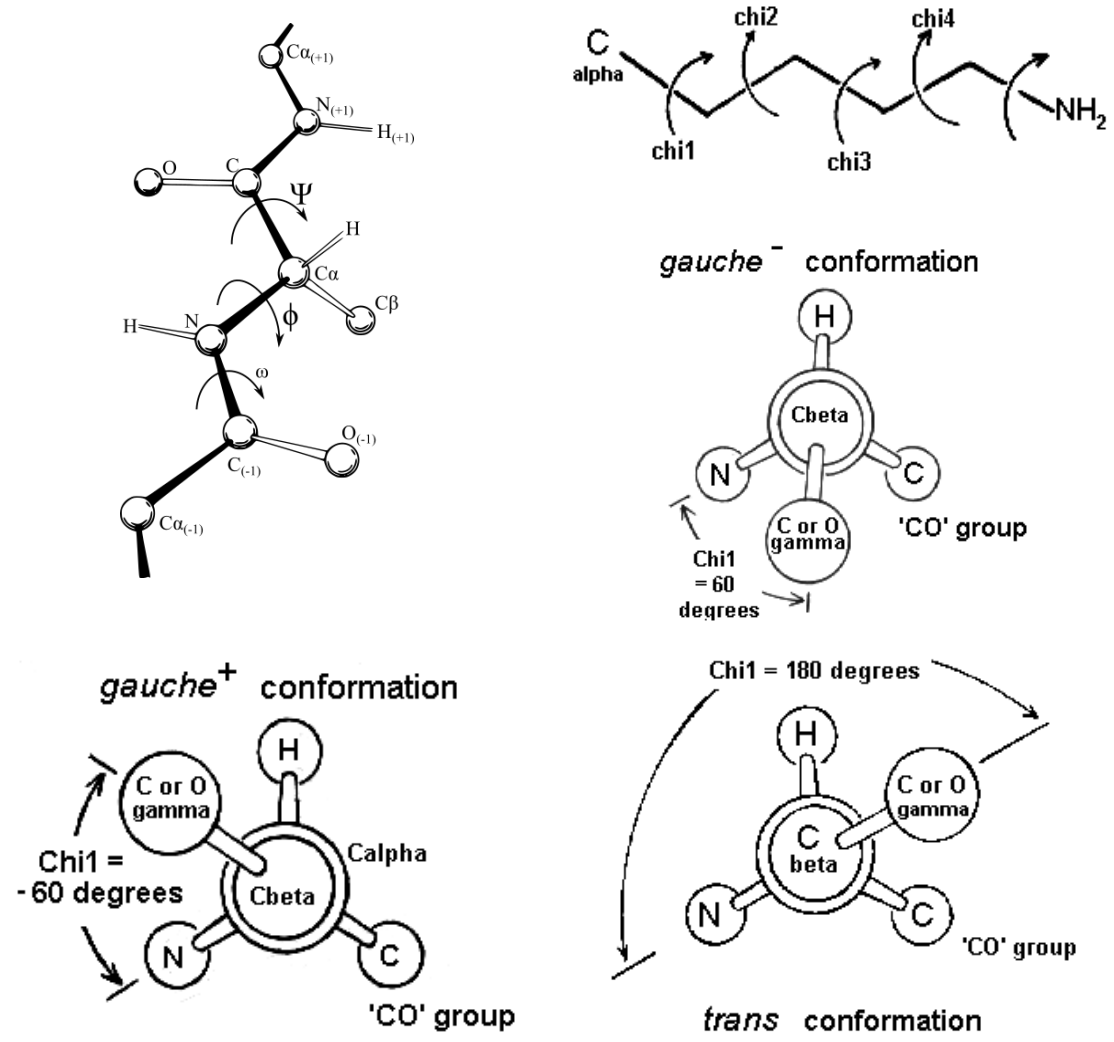


Figure 15. Definition of the Protein Backbone (φ, ψ and ω) and Side Chain (χ1 (chi1), χ2 (chi2), χ3 (chi3), Etc.) Dihedral Angles. Adapted from Richardson et al. [66].

Conformational Divergence between Template and GPR18 TMH3

One of our initial experiments was to explore the most stable conformation of GPR18 helices with unique sequences to find the origin of the differences between the template and GPR18. In the study of GPCRs, the CM method has been used successfully in earlier studies [68-70]. The new dihedral angles will be extensively and effectively explored in two different phases. The first phase is called exploratory. In this phase, the starting temperature for the heating step in simulated annealing (SA) is 3000 K, with cooling to 310 K in 18 temperature steps in the exploratory phase. At each temperature, 50,000 MC steps are applied. With each step, two dihedral angles and one bond angle are varied in the range of ±180° for dihedrals and of ±8° for bond angles. The second phase is called the biased sampling phase because CM explores only regions that were highly populated in the first phase. Because of this new distribution, we only explore nine temperature steps, so that the initial temperature is 749.4 K and the final temperature is 310 K, with 50,000 Monte Carlo steps at each temperature. In this work, all calculations are performed using a distance-dependent dielectric, all-atom calculations with CHARMM22 force field [71-73] and without the CMAP corrections [74]. In the GPR18 sequence, there is a P3.36 in TMH3 that is unique. Since the presence of Proline necessarily introduces a kink in the helix (well documented in previous reports [75]), the CM method was used to explore the possible typical conformational space of TMH3. To achieve this aim, an initial straight helix was built, and then some variations around the Proline kink region were introduced. This region includes the residues from P3.36 to T3.32. In our model the range of variations is: $-120<phi<40$, $-70<psi<0$ and

-160<*omega*<160. These values were extracted from the Protein Data Bank (pdb) [65]. Finally, the 103 structures generated by CM were superimposed from P3.36 to Q3.56 using the Superposition tool in Maestro (**Schrödinger Release 2017-4**: Maestro, Schrödinger, LLC, New York, NY, 2017). Results were compared with the template. Although TMH3 in the DOR crystal structure lacks a P3.36 (M3.36 instead) but still is bent crucially around this region (with an average bend angle of 16.5°, calculated in ProKink program). The only reason that might explain this severe bent is the presence of Thr3.38 in gauche minus dihedral because as Ballesteros et al.'s [76] studies previously proved, Serines and Threonines can act as hinge residues. This bending is produced by an intrahelical hydrogen bond between the backbone carbonyl oxygen (at position i-4 for alpha helix and i-3 for 3/10 helix) and the O^{γ} atom of the Serine or Threoine in gauche minus dihedral ($0° < \chi1 < 120°$) (see Figure 16). However, for two reasons, even the presence of this type of intrahelical interaction is not acceptable for the severe bent observed in DOR TMH3. First, the average Ser/Thr kink, which has been measured in the other alpha helices in the earlier reports, is around 12°. Second, the measurement of the bend angle in DOR TMH3 crystal shows it is 16.5°, which causes a wide and open extracellular region. We believe that the crystal packing contact in the biological units might be the main reason for such a significant deviation, which is a critical problem in X-ray crystallization. Researchers in this field usually underestimate this problem, which causes many later mistakes and understanding in GPCR modeling.

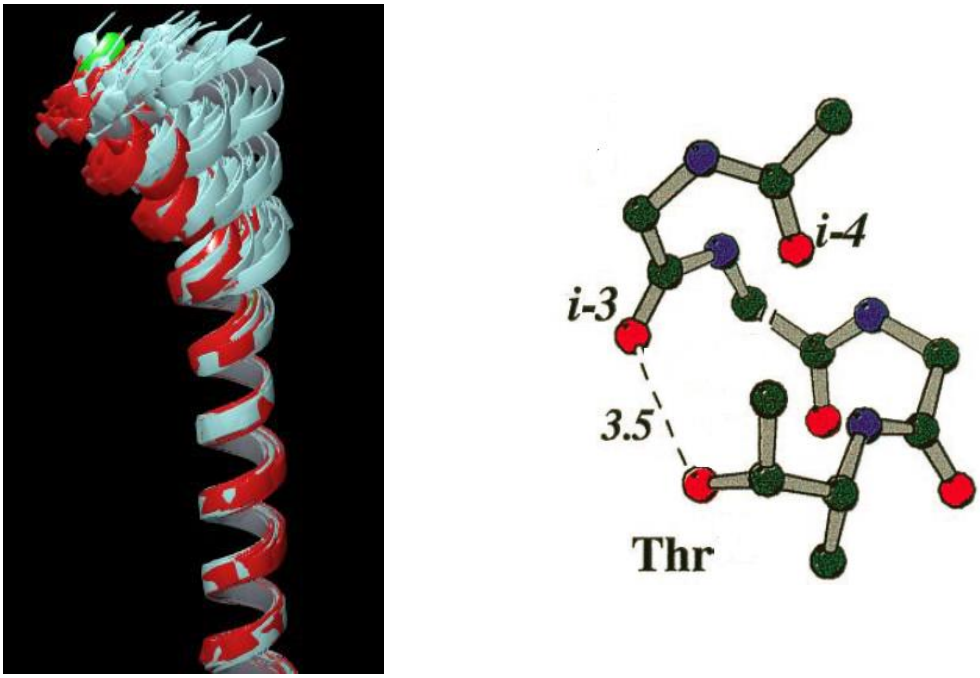


Figure 16. Theronine in Gauche Minus Intrahelical Binding. Dashed Line Shows the Distance of O^{γ}-O^{i-3} (Left) Superposition of TMH3 Conformation Memory Products (Right). The Helix in Green is for DOR TMH3 and Helices in Red and Light Blue are GPR18 TMH3 CM Outputs. Adapted from Ballesteros et al. [76].

After doing a CM job on TMH3 in GPR18, the outputs show that the population rate of the extracted conformers mainly tends to have a bend angle less than TMH3 in DOR (or the conformers of TMH3 in GPR18 tend to be straighter than TMH3 in DOR). Figure 16 depicts an overlay of the one helix in green (which is TMH3 in DOR) and the helices in light blue (which are TMH3 in GPR18 produced by CM job). The comparison of the green helix with the light blues reveals two different populations: the light blue straight helices and the light blue plus red bent helices. The light blue straight helices are highly populated (around 62%), but the light blue bent helices and the red helices are slightly populated in CM job-produced conformers. Therefore, we concluded that TMH3 in GPR18 tends to be straighter than TMH3 in DOR and has to be picked from the high

population. In this way, it has to have a bend angle higher than 16.5°. Thus, we finally picked the conformer with a bend angle of 16.8°.

Ionic Lock and Toggle Switch

In all of class A GPCRs, helix 6 plays a vital role in the change between receptor inactive and active states. When the receptor is activated, one of the significant changes happens at the intracellular (IC) end of the TMH6. The TMH6 intracellular end moves away from the transmembrane helix bundle, making room for coupling with the G-protein. This mechanism has been done by two important switches which are shared between TMH3 and 6. The first one is the "ionic lock salt bridge," which in GPR18 is formed between R3.50 and S6.33. The role of this salt bridge is to limit the relative mobility of the cytoplasmic ends of TMH3/6 in the inactive state and by disruption in the active state lead to an open bottom of the receptor to interact with the G-protein. "Toggle switch" is the second switch, which is the conformational change in bulky residue at position 6.48 from gauche plus ($240° < \chi1 < 360°$) in the inactive state to trans ($120° < \chi1 < 240°$) in the active state. This conformational change straightens TMH6 and breaks the ionic lock. The x-ray crystal structure of the β_2-adrenergic receptor complexed with $G_{\alpha s}$ protein shows that TMH6 has straightened by flexing its proline kink at P6.50, forming an opening at the intracellular end of the bundle into which the $G_{\alpha s}$ protein has inserted its C-terminus [77]. This conformational change also impacts binding pocket residues. W6.48 is a highly conserved residue in Class A GPCRs. It has been shown to change its conformational state within the binding pocket upon receptor activation ($\chi1$ gauche plus $\rightarrow$ trans). In the inactive state, W6.48 is typically held in its $\chi1$ gauche plus conformation

by another binding pocket residue. Together, this pair of residues is known as the "toggle switch." [78] In GPR18 inactive state model reported here, F6.48 and M7.42 form the toggle switch. Since antagonists/inverse agonists stabilize the inactive state of the receptor, a critical factor in the design of antagonist/inverse agonist analogs here is that each ligand either directly or indirectly blocks the F6.48 (χ1 gauche plus $\rightarrow$ trans) transition, thus keeping GPR18 in its inactive state. In contrast, GPR18 agonist binding promotes the F6.48 (χ1 gauche plus $\rightarrow$ trans) transition.

Conformational Memory Study of TMH7

GPR18 lacks $N^{7.49}P^{7.50}XXY^{7.53}$ motif in TMH7. Instead, it has $D^{7.49}V^{7.50}XXY^{7.53}$. At first glance, it would appear that the lack of P7.50 would lead to a straight helix, but there are some problems with a straight TMH7. First, a straight TMH7 creates much space at the top of the receptor that cannot be filled by extracellular loops (specifically EC 2 & 3), allowing water molecules to get into the receptor. Second, a straight TMH7 changes the position of some critical residues. For instance, as Figure 17 illustrates, N7.43 is typically a binding pocket residue, but if TMH7 is straight, its Cα position will be near lipid face. T7.39 can also be a ligand interaction site and is typically more accessible to the binding crevice while in a straight helix seven; T7.39 is usually facing more towards TMH1 than TMH3. Third, Y7.53 is slightly removed from its normal position in the sodium pocket (sodium ion binding pocket and its importance are discussed with details in Chapter IV). Lastly, our initial MD simulation results show that a straight TMH7 is not stable and tends to be bent. Superposition of the extracted

structure of those trajectories with our template protein structure (DOR) in Figure 18 reveals the similarity between the TMH7s.

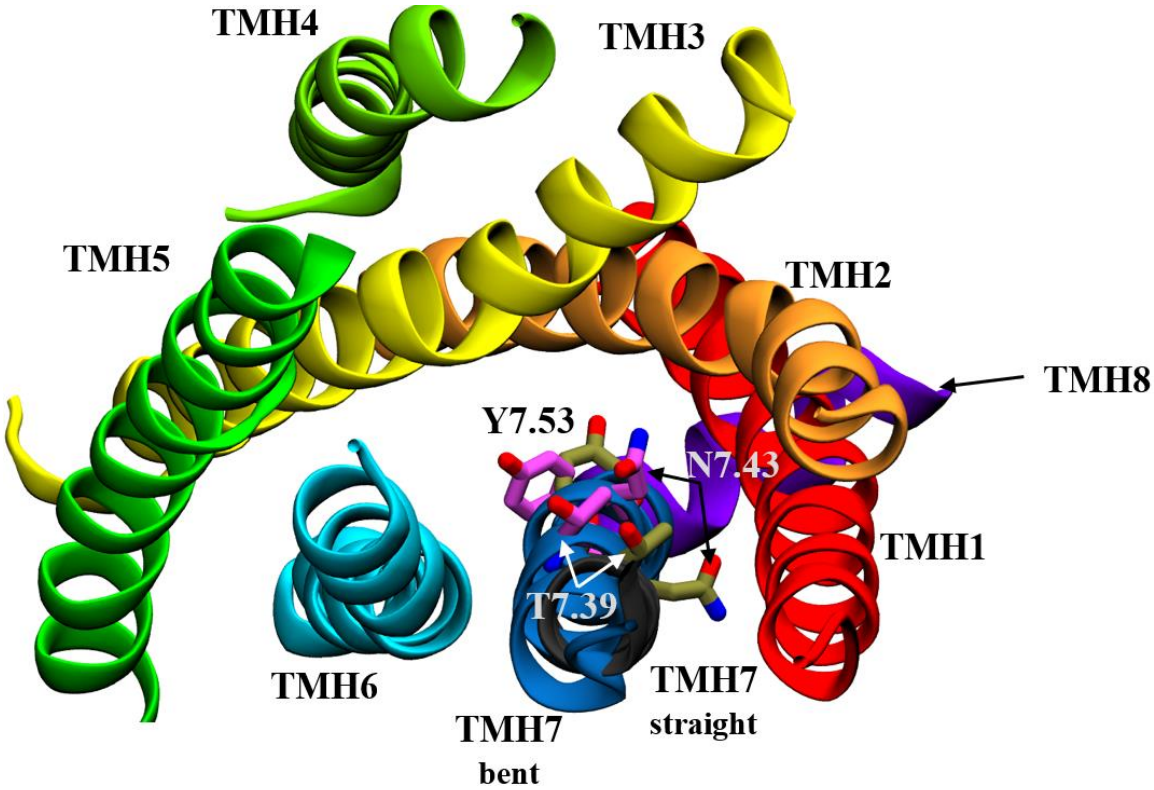


Figure 17. Comparison of TMH7 in Straight Model and Bent Model. Comparison of Position of Important Residues N7.43, T7.39 and Y7.53 in Straight Model (Highlighted in Brown) and Bent Model (Highlighted in Pink). TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (Cyan), TMH7 Bent (Dark Blue), TMH7 Straight (Black), TMH8 (Violet). Superposition of TMH7 in Bent Model with the TMH7 Extracted from Initial MD Simulation.

Taken together, we should, therefore, seek an alternate motif or group of residues on TMH7 that might induce a helix distortion/hinge similar to a Proline. As previously mentioned, Ballesteros et al. [76] showed that Serines and Threonines could act as hinge residues. In GPR18 TMH7, the best residues for this goal are S7.45 and T7.46 because they are very close to the normal Proline position in TMH7. We used the CM method to

explore the effect of S7.45 and T7.46 on TMH7 conformations. The CM study explored gauche minus/gauche plus, gauche plus/gauche minus and gauche minus/gauche minus conformations of S7.45 and T7.46 respectively, by variation of the chi1 dihedral angle of these two residues to the chosen value. The variation range was the same as that used for the Proline kink in TMH3 for phi, psi and omega angles. The range of variations for gauche plus, trans and gauche minus are ($240° < \chi1 < 360°$), ($120° < \chi1 < 240°$) and ($0° < \chi1 < 120°$) respectively. Altogether, the gauche minus/gauche minus conformation gave us the best fit for the position of important TMH7 residues.

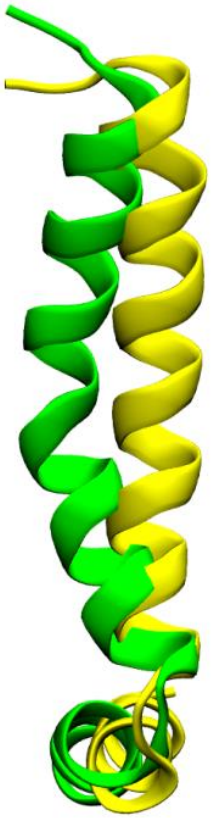


Figure 18. MD Simulation of GPR18 with an Initial Straight TMH7 (Yellow) Shows That after 50 ns TMH7 Tends to Bend (Green Model).

Extracellular (EC) and Intracellular (IC) Loop Conformations

Because GPR18 recognizes lipid-derived ligands, it is possible that like the S1P1 and LPA receptors, the EC loops and N-terminus of GPR18 close the top of the bundle, restricting water access to the binding pocket. The DOR crystal structure has its EC2 loop splayed out, providing water molecules free access to the binding crevice. As a result, we need to model the GPR18 EC loops and N terminus so that they cover the binding site. As Figure 14 shows, the EC-2 is the longest loop (L159-V184) in GPR18, and it includes C172 which forms a disulfide bridge with C3.25 for structural stability. MODELLER [79, 80] was used to generate the N terminus (Met-1 to Pro-18) and the IC and EC loops (IC1: Thr-51 to Thr-55, IC2: Pro-126 to Lys-133, IC3: His-216 to Val-226, EC1: Glu-86 to Gly-90, EC2: Leu-159 to Val-184, EC3: Thr-259 to Tyr-264). A limited distance between Cys-172 and Cys3.25 was used in the EC-2 loop calculation so that the highly conserved disulfide bridge was possible. The final loops were energy minimized to a gradient of 0.05 kJ/mol.A° with 2500 maximum iterations using the Generalized Born/Surface Area (GB/SA) continuum solvation model for water and the OPLS3 [81-83] forcefield in MacroModel 10.7.014 **(Schrödinger Release 2017-4**: MacroModel, Schrödinger, LLC, New York, NY, 2017), while the TMH region was frozen.

Bundle Minimization

The resultant model was minimized before starting MD simulations using our published protocol [20]. In the current study, minimization has been performed by MacroModel for homology modeling. The minimization protocol used the OPLS3[17-19] all-atom force field in MacroModel (**Schrödinger Release 2017-4**: MacroModel,

Schrödinger, LLC, New York, NY, 2017). An extended Van der Waals cutoff (8.0 Å, updated every 10 steps), a 20.0 Å Electrostatic cutoff, and a 4.0 Å hydrogen bond cutoff were used in each stage of the calculation. Subsequently, the inactive state, R model was equilibrated in a fully hydrated POPC bilayer using our published protocol [85].

Docking of the Putative Antagonist PSB-CB5 in Inactive (R) Model

PSB-CB5 has been identified as a selective antagonist for GPR18 [36, 86]. The reports reveal that it produces a concentration-dependent inhibition of Δ^9-THC-induced β-arrestin recruitment to GPR18 (an IC_{50} value of 0.279 μM). Figure 19 illustrates the PSBCB-5/GPR18 R Complex. In the complex, R2.60 forms a hydrogen bond with the carbonyl oxygen of PSBCB-5 (N-O distance =3.1 Å; NH-O angle=170°). F6.51 has an aromatic stacking interaction with the imidazo[2,1-b]thiazine ring of PSBCB-5 (centroid to centroid distance = 4.7 Å; angle = 45°). The central phenyl ring of PSBCB-5 has an aromatic stacking interaction with F6.55 (centroid to centroid distance = 5.4 Å; angle = 69°). The EC2 loop residue, Y(180), has an aromatic stacking interaction with the para-chloro ring of PSBCB-5 (centroid to centroid distance = 5.5 Å; angle = 89°). PSBCB-5 also has strong Van der Waals interactions with T3.32.

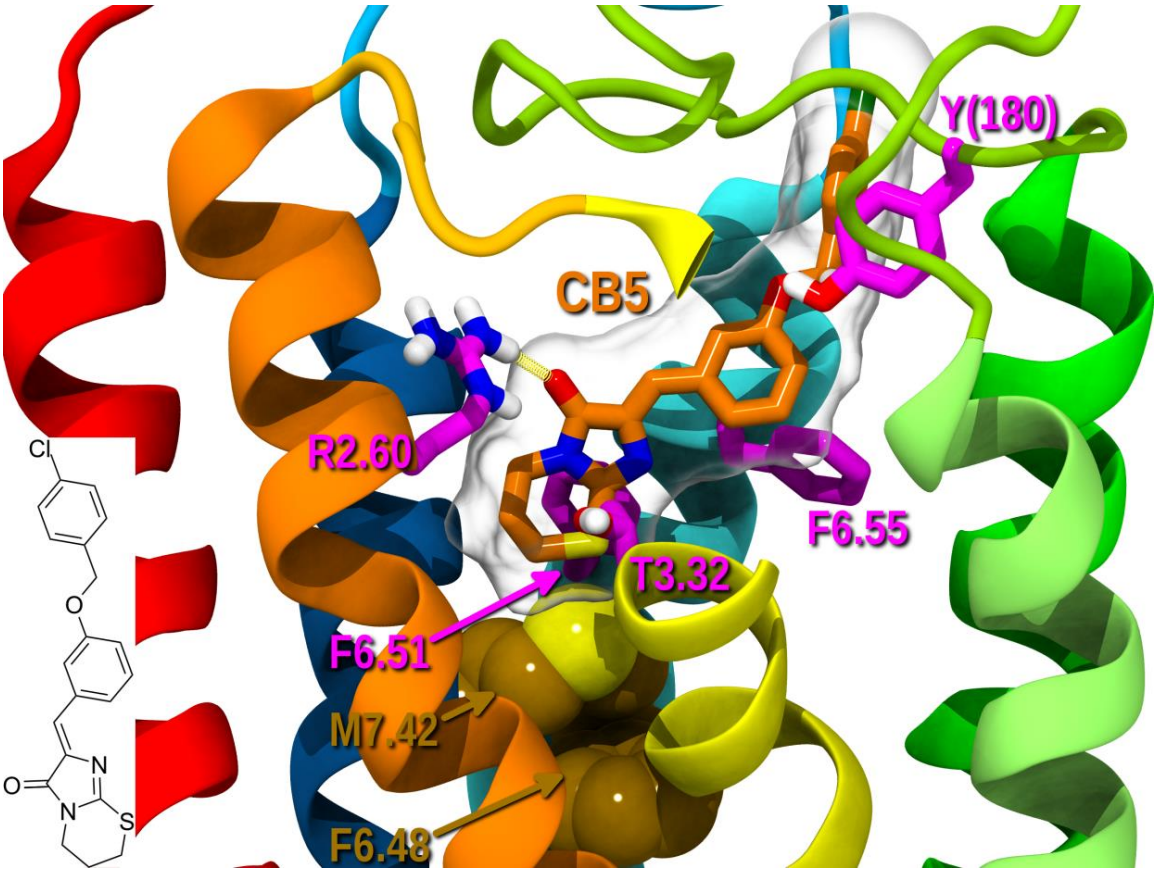


Figure 19. PSB-CB5 (Orange Structure) Docked in R Model of GPR18. Hydrogen Bond between Carbonyl Oxygen of S4 and R2.60 is Shown with Yellow Line. The Putative Interaction Residues Highlighted in Pink. Toggle Switch Residues (M7.42 and F6.48) are Contoured at Their VdW Radius in Brown Color. Color Code for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (in Cyan) and TMH7 (Dark Blue). 2D Structure of PSB-CB5 Inserted in Far Left of the above Picture beside TMH1.

Docking of Our Proposed Antagonist S4 in Inactive (R) Model

Our experimental data on the GPR18 antagonist activity of the pyrazol-3(5)-ylbenzene-1,3-diol derivative S4 allows us to propose the S4 scaffold for further elaboration. The global minimum energy conformer of S4 was determined from a complete conformational analysis at the HF 6-31G* level using Spartan ‘08

(Wavefunction, Inc., Irvine CA). Our dock of the global minimum energy conformer of S4 (purple, transparent VDW surface) in the GPR18 model is illustrated in Figure 20. The S4 global minimum has an internal hydrogen bond between one of the resorcinol ring hydroxyl groups and a pyrazole ring nitrogen. This hydrogen bond causes both the resorcinol and the pyrazole ring to be in the same plane favoring S4 to be elongated. In the S4/ GPR18 R complex, the benzyl ring has a cation-pi interaction with R2.60 (4.0 Å). One hydroxyl oxygen of the resorcinol ring forms a strong hydrogen bond with R5.42 (O-N distance 3.2 Å; NH-O angle = 154°). The resorcinol ring also has a tilted-T aromatic stacking interaction with F6.55 (centroid to centroid distance=5.1 Å, relative ring angle = 77°). The pyrazole ring also has aromatic stacking interactions with F6.51 (centroid to centroid distance=4.9 Å, relative ring angle = 53°) and F6.55 (centroid to centroid distance=5.8 Å, relative ring angle = 87°). Finally, S4 also establishes strong Van der Waals interactions with T3.32, while the S4 aliphatic chain has Van der Waals interactions with L5.38. S4 is an antagonist because its benzyl ring sits directly above M7.42 (gold, contoured at its VdW radius in Figure 21). Consequently, this compound stabilizes the GPR18 toggle switch in its “off” state (F6.48 $\chi 1$ = gauche plus).

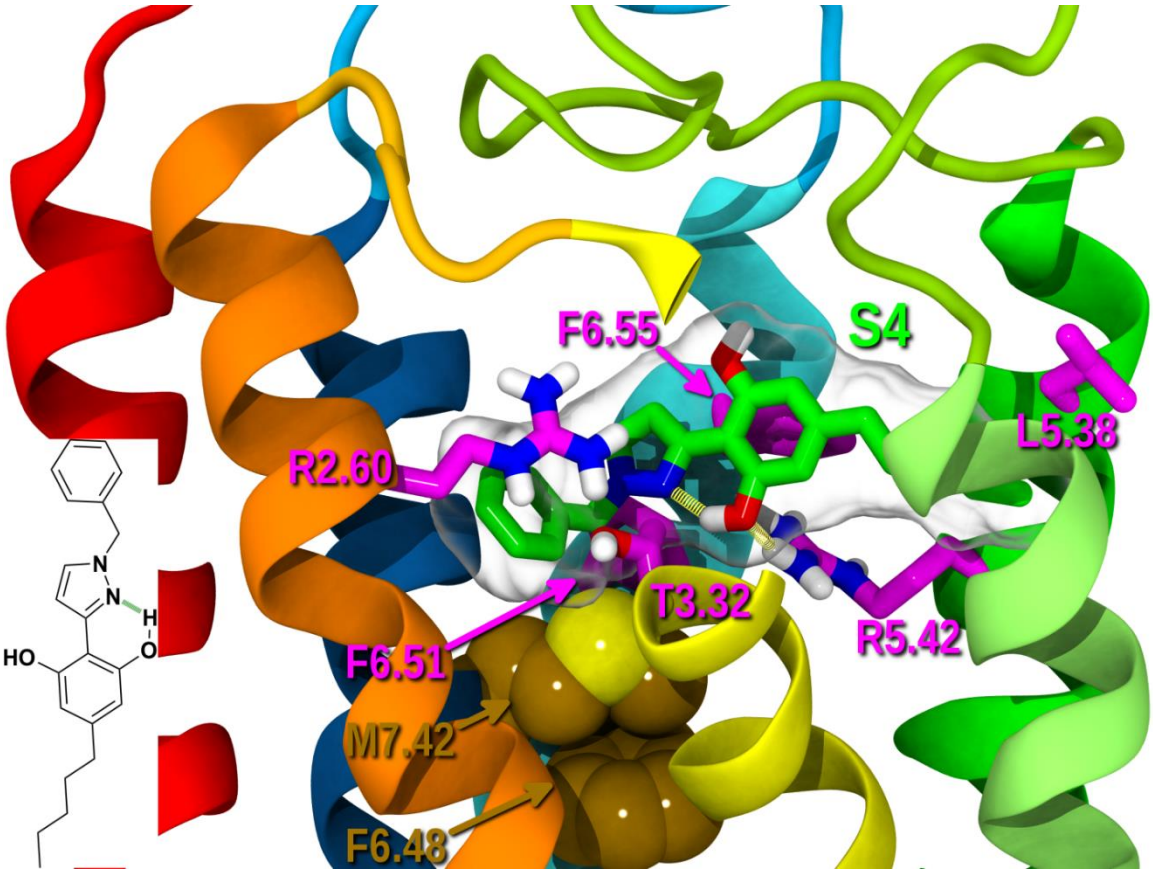


Figure 20. S4 (Green Structure) Docked in R Model of GPR18. The Internal Hydrogen Bond of S4 and Hydrogen Bond between Resorcinol Ring and R5.42 are Shown with the Yellow Lines. The Putative Interaction Residues Highlighted in Pink. Toggle Switch Residues (M7.42 and F6.48) are Contoured at Their VdW Radius in Brown Color. Color Code for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (in Cyan) and TMH7 (Dark Blue). 2D Structure of S4 Inserted in Far Left of the above Picture beside TMH1.

Molecular Modeling GPR18 Active State (R*)

Binding an agonist and subsequent receptor activation are accompanied by significant conformational changes in the receptor inactive state structure. The first change occurs in the cytoplasmic end of the TMH6. The outward movement of TMH6 from the receptor core leads to breaking the ionic lock (R3.50 and S6.33 interaction) and

increasing the distance between TMH6 and TMH3. On the other hand, the movement of TMH6 outward provides additional space for TMH7 to move toward the bundles (see Figure 21 and 22). These movements make adequate room for G-protein interaction. P6.50 plays the role of the flexible hinge, which mediates the transition from R to R* [87]. In the activated state, the toggle switch residue (F6.48 in GPR18) undergoes a gauche plus to trans $\chi 1$ transition change, which creates an open space for the inward movement of TMH7 (see Figure 23).

Taken together, we create our R* model with modification of the explained R model. In the R* bundle, a less kinked TMH6 conformer (bend angle 22°) was substituted for TMH6 in the R bundle. The TMH6 conformer used here was output from CM calculations. The straightening of TMH6 ensures that the TMH3-TMH6 “ionic lock” is broken. The $\chi 1$ of H6.52 was also changed to trans. Due to the changes in TMH6; IC3 loop was remodeled using Modeller. The resultant bundle was then minimized.

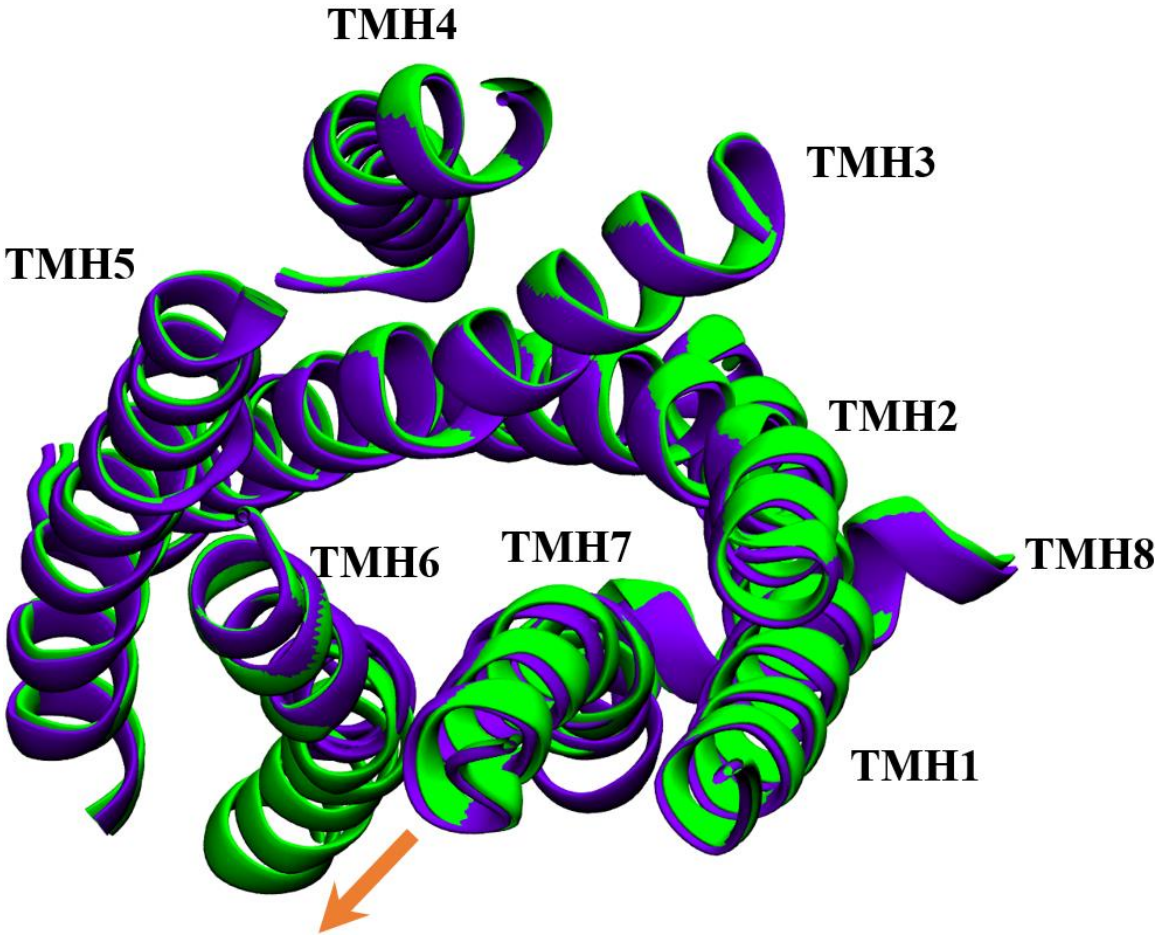


Figure 21. Super Position of R Model (Blue) and R* Model (Green). View from the Extracellular Part of the Receptor. The Orange Arrow Shows That TMH6 in Active Model Clearly Moved away from the Bundle Which Causes to Open Bottom of the Receptor for Interaction with G Protein.

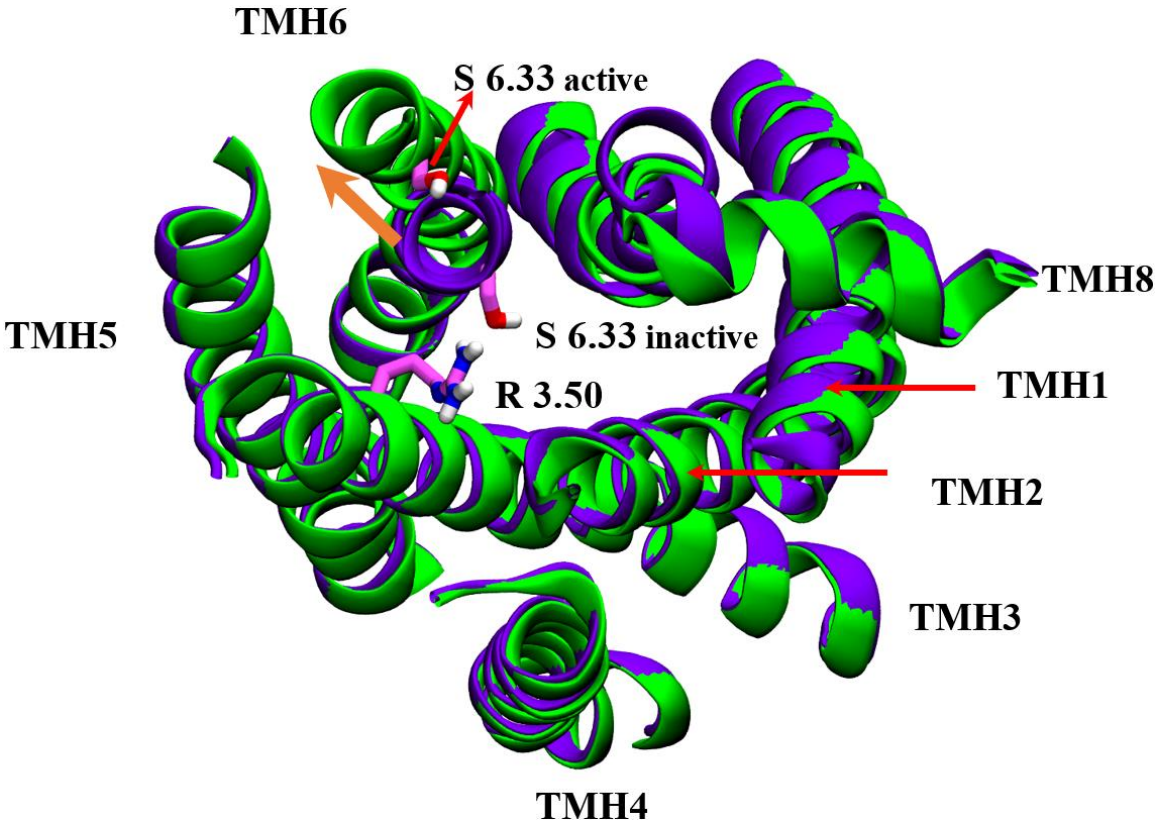


Figure 22. Super Position of R Model (Blue) and R* Model (Green). View from the Cytoplasmic End of the Receptor. The Orange Arrow Shows That TMH6 in Active Model Clearly Moved away from the Bundle. This Movement Causes to Break Ionic Interaction between R3.50 and S6.33 (Highlighted Residues in Pink) and Then Opening Bottom of the Receptor to Interact with G Protein.

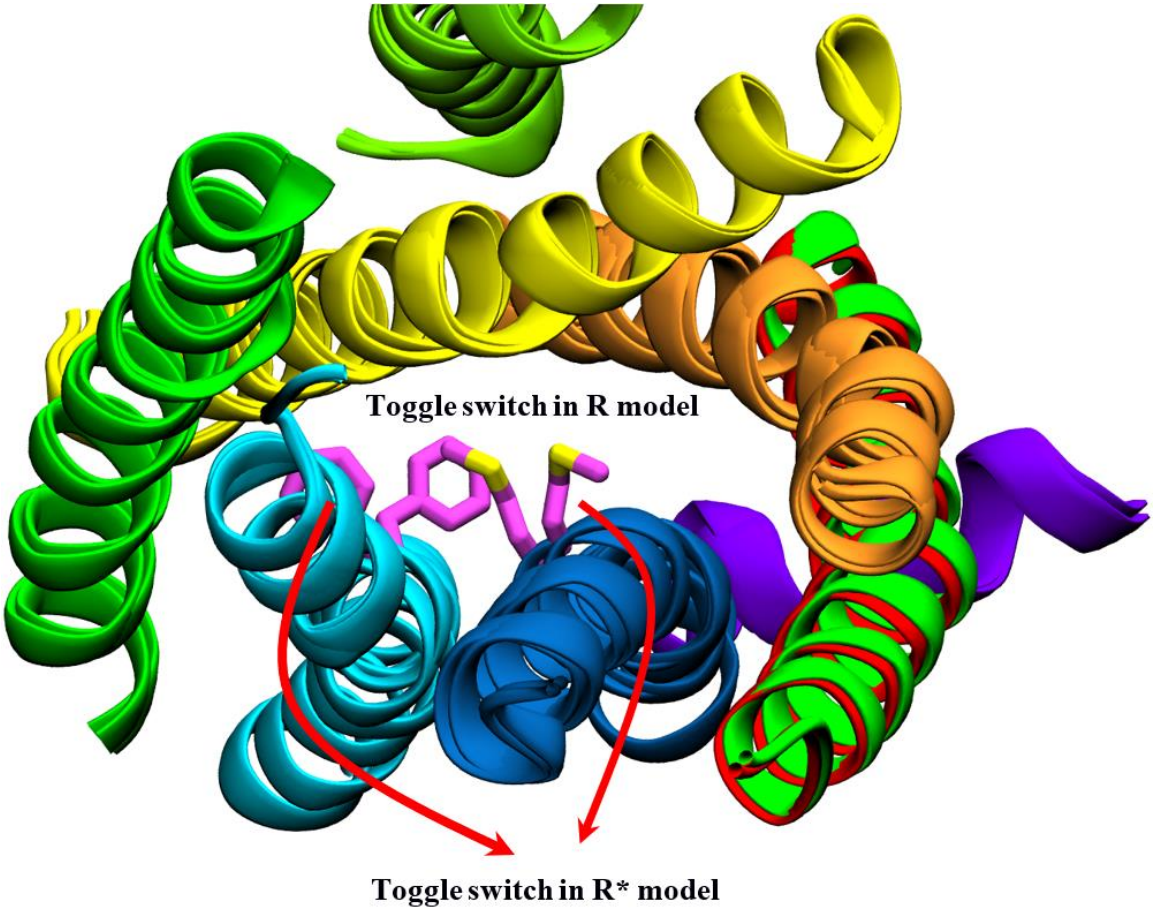


Figure 23. Superposition of R and R* Models. View from the Extracellular Part of the Receptor. Color Codes for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (Cyan), TMH7 Bent (Dark Blue), TMH7 Straight (Black), TMH8 (Violet). Toggle Switch Residues F6.48 and M7.42 Conformations in Both Models are Highlighted in Pink.

Docking of Putative Agonist in GPR18 Active (R*) Model: Conformational Memory Calculations for NAGly and RvD2

In work reported here, we performed a conformational analysis of NAGly and RvD2 to examine their flexibility and the best conformation before doing the docking job. To this end, we used the Monte Carlo/Simulated annealing program, Conformational Memories (CM) [68, 69] to explore the dihedral conformation space ($e^{-\Delta E/kT}$). In this

method, dihedral angles are explored by Monte Carlo (MC) random moves. Since exploring all of the global or local minima in a large search space is time-consuming, this method is combined with simulated annealing (SA). SA can proceed to find all minima without being trapped in a local minimum. Thus, this method can sample all of the conformations that each ligand can produce.

The CM method consists of two phases. In the Exploratory phase, the starting temperature for the heating step in SA is 3000 K, with cooling to 310 K. The cooling formula for lowering the temperature in 18 steps was $T_{n+1} = T_n (3000/310)^{(1/18)}$. At each temperature, 50,000 MC steps are applied. With each step, two dihedral angles and one bond angle are varied in the range of ±180° for dihedrals and of ±8° for bond angles. If this new value is accepted, the new dihedral angle is recorded, but if it is not accepted, then CM uses the value from the previous run. The process of accepting or rejecting random dihedral values is according to the standard Metropolis criterion. The recorded values are shown in an axis that covers all possible 360° angle motions. The method is repeated until the for each temperature step, ending at 310K.

Figure 24a shows conformations for rotation around the amide bond in NAGly. Thus, it is expected that the high population is around trans angles (±180°). The strong double bond leads to a flat conformational search either in high temperatures or low temperatures. Figure 24b shows the memories for the different (Csp3-Csp2-Csp2-Csp3) dihedral angles in NAGly, which has a C=C as the second and third atom. Here, the memories have a high population around 0°, consistent with this being a cis double bond either in high temperatures or low temperatures.

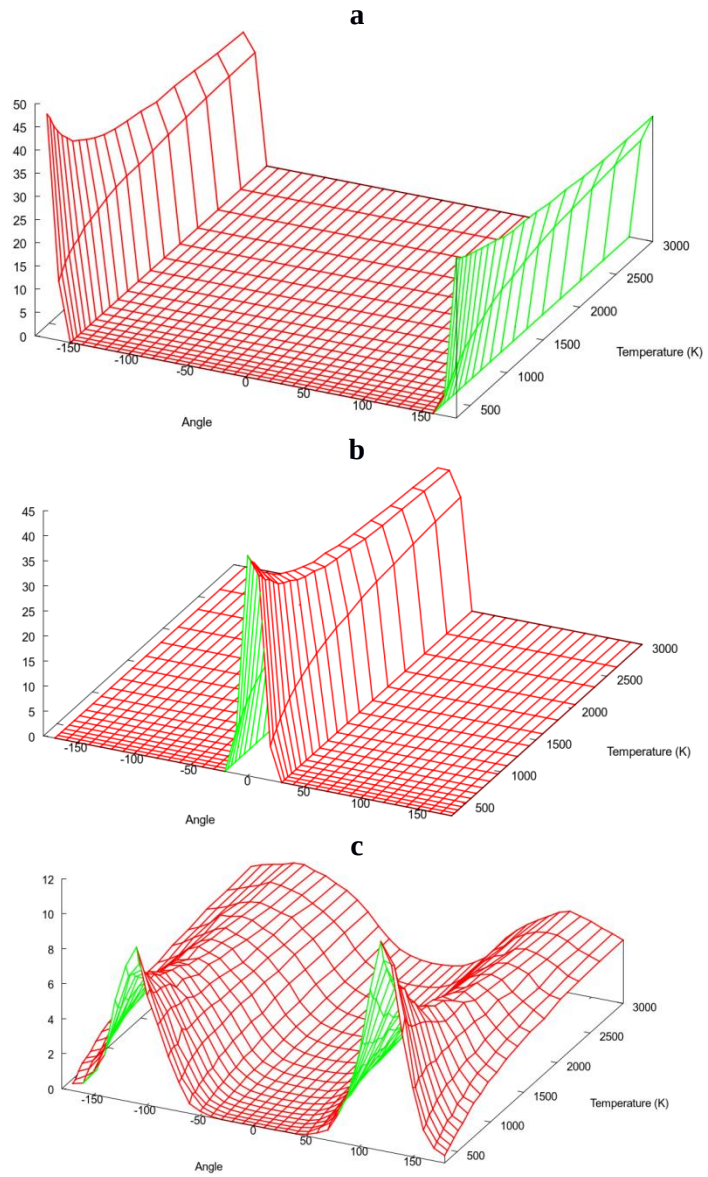


Figure 24. Conformational Memory Results of NAGly. (a) the Amide Bond are Populated around ± 180° (trans); (b) around a Csp3-Csp2-Csp2-Csp3) Dihedral Angles Which are Populated around 0°; (c) the (Csp2-Csp3-Csp2-Csp2) Dihedral Angle Which are Populated around ±120°.

Figure 24c illustrates the memories for the (Csp2-Csp3-Csp2-Csp2) dihedral at high temperatures explored many angles from -180° to +180° (the population for cis (0°) is slightly near 0). As the temperature is lowered, angles particularly around trans (±180°)

disappear and are not accessible. We conclude that at lower temperatures, only conformations around ±120° are most populated and will therefore be used in next phase. Figure 24a and b show that the range of motion at high temperatures is not fully explored by the first step as well as those for Figure 26c because to fully explore the range of angles would require bond breaking. A similar pattern occurs for RvD2 CM job. Figures 25 and 26 depict the memories for Csp2-Csp2 and Csp3-Csp3 dihedral angles, respectively.

The second step is biased sampling. This step is called biased sampling because we explore only in areas that were highly populated in the first phase, rather than exploring all of the dihedral regions. To limit sampling only in the areas in which a dihedral has been accepted and to skip the areas with no population, we use a new procedure. In this procedure, we use only populations at the final temperature (310 K) to search for low-energy conformations. Biased sampling is done by a uniform distribution of random numbers, which are divided into 36 intervals corresponding to 36 dihedral intervals. By this new distribution, we only explore seven temperature steps so that the initial temperature is 749.4 K and the final temperature is 310 K, with 50,000 Monte Carlo steps at each temperature. One hundred eight structures are output at 310 K. All conformers were then superimposed on heavy atoms and clustered. Clusters were analyzed based on their dihedral angles. One hundred eight structures were output at 310 K. The superposition of CM job output structures is shown in Figure 27. We typically dock the lowest enthalpy conformer of a ligand in the receptor. Since all of CM job outputs are produced in the body's temperature (310 K), either is acceptable for the

docking job. In this way, there is no need to sort them based upon their energy values and the best fit inside the binding crevice with the strongest interactions chosen as the final structure.

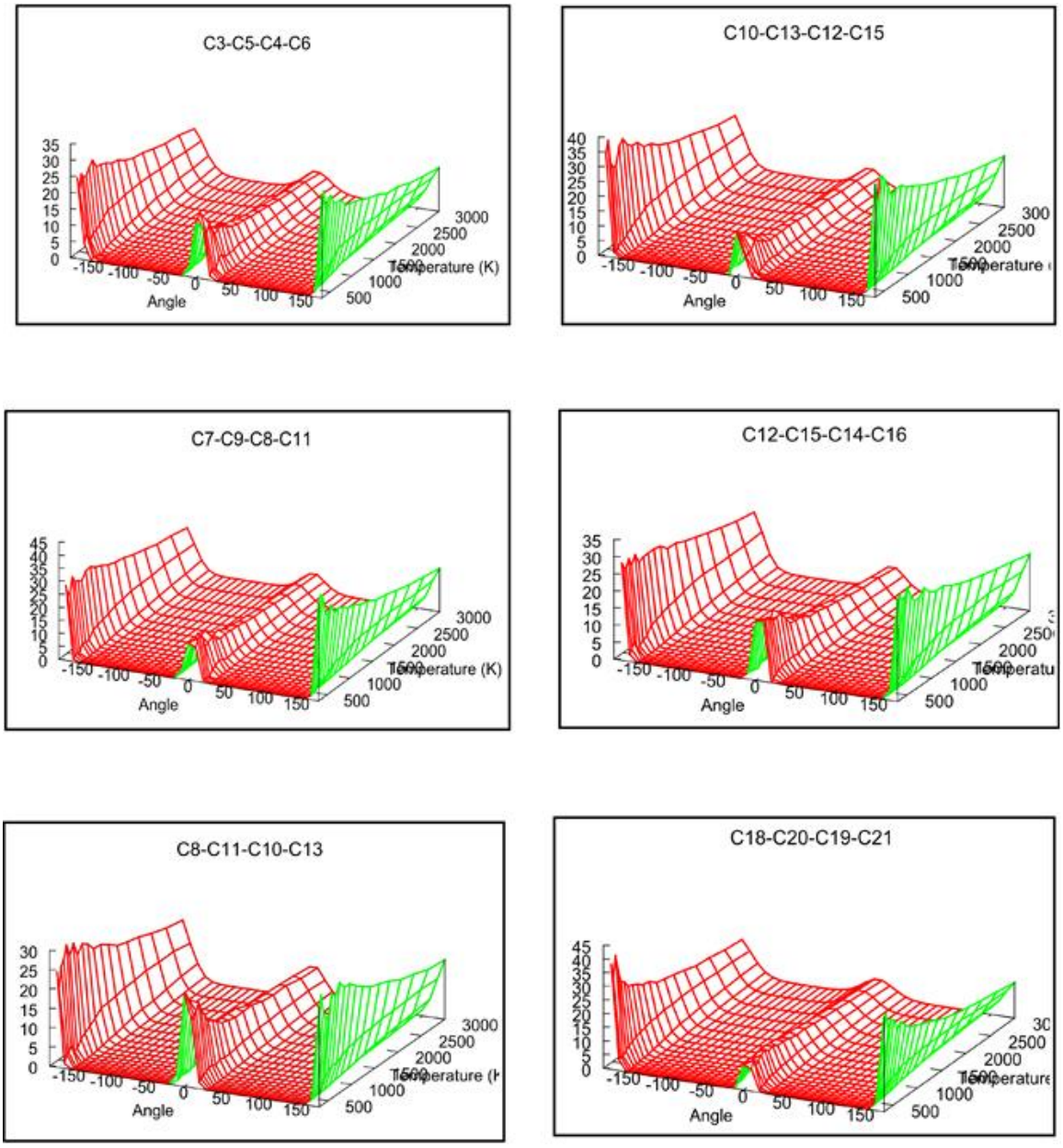


Figure 25. Conformational Memory Results of RvD2 for All of the Csp2-Csp2 Dihedral Angles Which are Populated around 0° and ±180°.

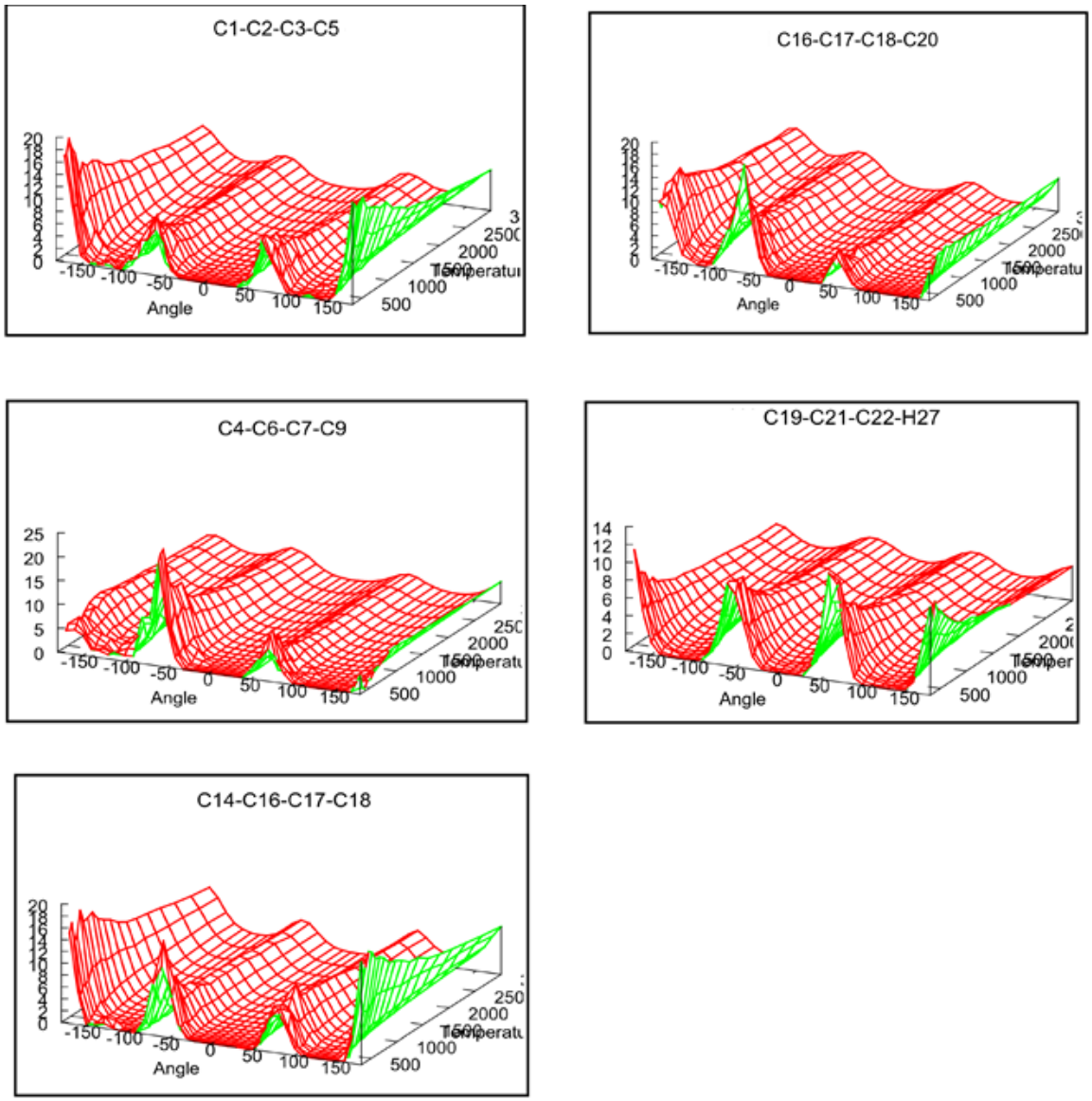


Figure 26. Conformational Memory Results of RvD2 for All of the Csp3-Csp3 Dihedral Angles Which are Populated Around Gauche Miues, Plus and Trans Conformations.

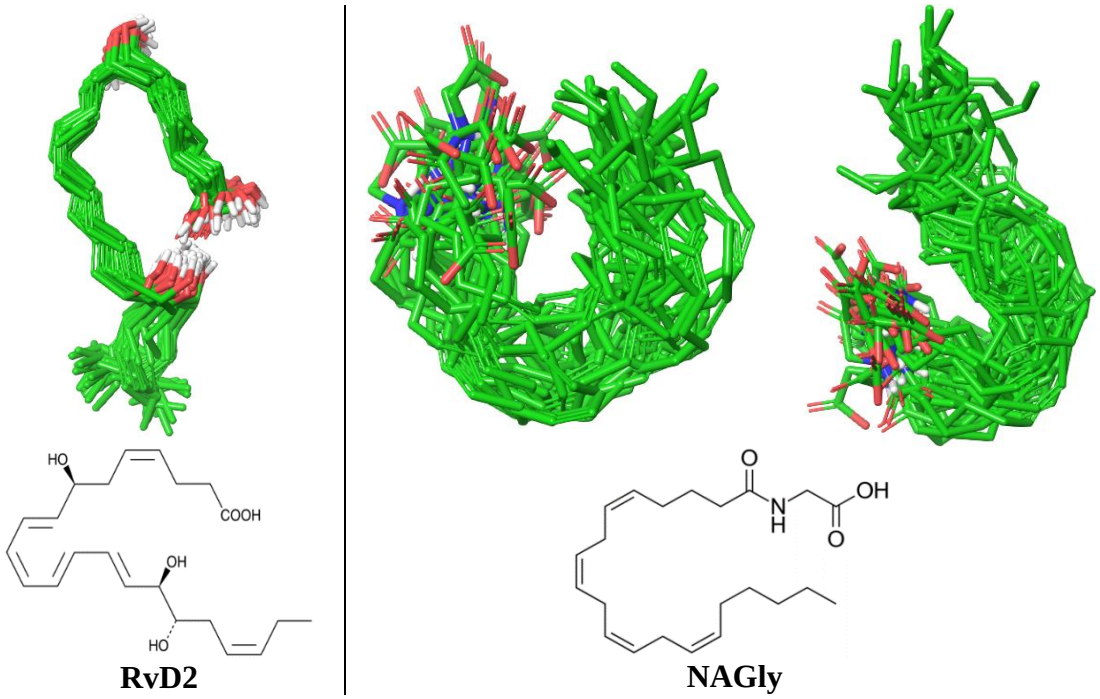


Figure 27. Top Row from Left to Right; Superposition of Conformational Memory of RvD2 (Question Mark Shape), NAGly (U-shaped), NAGly (J-shaped). 2D Structure of NAGly and RvD2 Inserted in the Bottom Row.

NAGly evoked a calculated IC_{50} value of 20 nM in a concentration-dependent, pertussis toxin-sensitive inhibition of cAMP assay [31]. RvD2 has been reported to have a higher affinity (K_d =10nM) [60]. RvD2 has three hydroxyl groups in addition to its polar head group. These hydroxyls are distributed over the length of the molecule, increasing the potential for hydrogen bonding for this ligand. NAGly and RvD2 both have four double bonds in the middle of their structures, but the four cis homoallylic double bonds in NAGly are separated by methylene carbons (like most endocannabinoids), while the four double bonds in RvD2 are conjugated (see Figure 22).

Moreover, the last five carbons at the end of NAGly are connected by single bonds; RvD2 again is constrained by a cis double bond in this region. Earlier studies [88]

have shown that the rotational energy barrier for rotation about sp3-sp2 single bond is low (i.e., for E ≤ 1.0 kcal/mol). The result is a molecular region that can adopt many different conformations. Conversely, saturated carbon acyl chains (like the NAGly tail) have an energy barrier for rotation about sp3-sp3 single bonds that is higher (i.e., for E ≥ 3.0 kcal/mol), which results in confining these carbons to trans (±180°) or gauche (±60°) configurations. As a result, saturated fatty acids are much less flexible and favor more extended conformations. Alternatively, regions containing homoallylic double bonds have a high degree of flexibility and end to curve in this region and form folded structures. Reports also show that the conjugated parts of an acyl chain introduce stiffness into the structure so that it is challenging to adopt an extended conformation. Using the same protocol as described earlier, our conformational studies on NAGly and RvD2 revealed that RvD2 (either protonated or unprotonated head group) adopts predominantly only "question mark" shapes, while NAGly adopts mainly two J-shaped (head group close to separated double bonds) and U-shaped (tail comes close to head group).

Our dock of the global minimum energy conformer of NAGly in the GPR18 active model is illustrated in Figure 28. The carboxylic head group in NAGly forms a strong hydrogen bond with positively charged residue R2.60. The flexibility of the homoallylic double bonds region of NAGly allows it to easily break the interaction of M7.42 and F6.48 in the toggle switch area. As previously mentioned, it is necessary for an agonist to intervene in the toggle switch residue interaction and help to F6.48 (χ1 gauche plus → trans) transition.

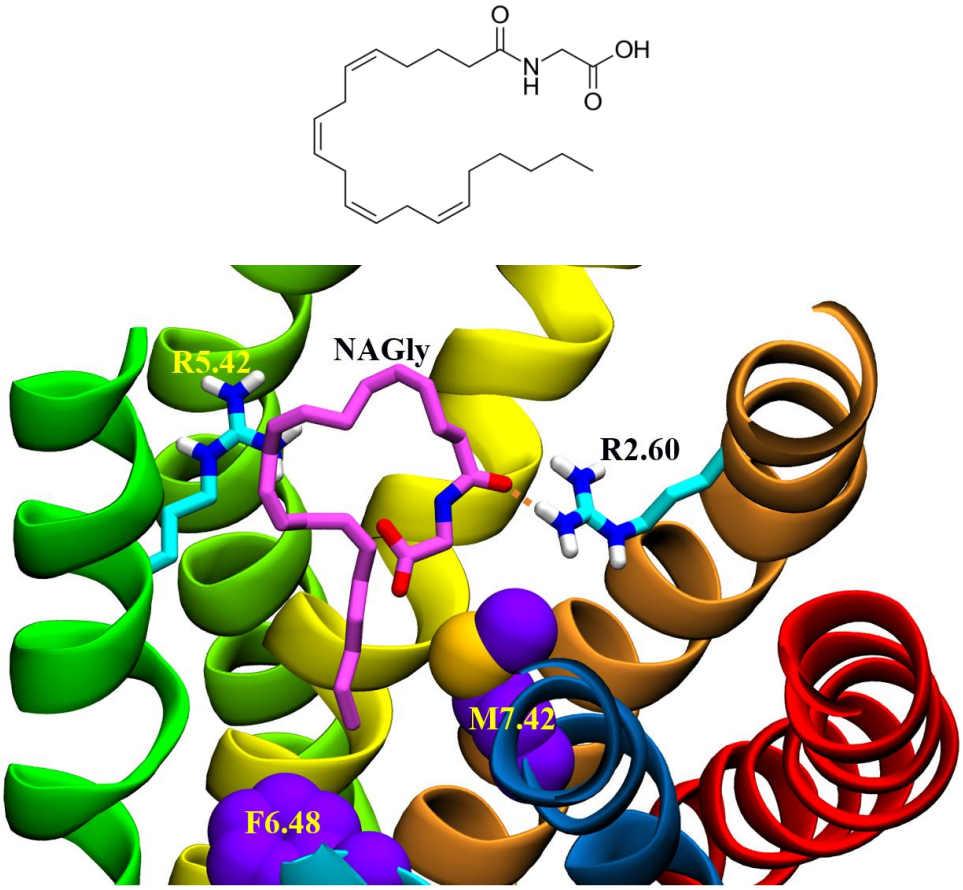


Figure 28. Docking of NAGly (Pink) Inside the GPR18. Orange Dashed Line is Showing Hydrogen Bonding with R2.60 (Cyan) and No Interaction with R5.42 (Cyan). Toggle Switch Residues (M7.42 and F6.48) are Contoured at Their VdW Radius in Blue. Color Codes for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green) and TMH7 (Dark Blue).

RvD2 also has a carboxylic head group, which easily finds R2.60 for hydrogen binding, but it also has two hydroxyl groups which makes it more possible for RvD2 to have hydrogen binding with the other positively charged residue in the binding crevice, R5.42. However, although the carbonyl chain tail of RvD2 is smaller than NAGly, it is still used for the intervening interaction of M7.42 and the F6.48 toggle switch (see Figure 29).

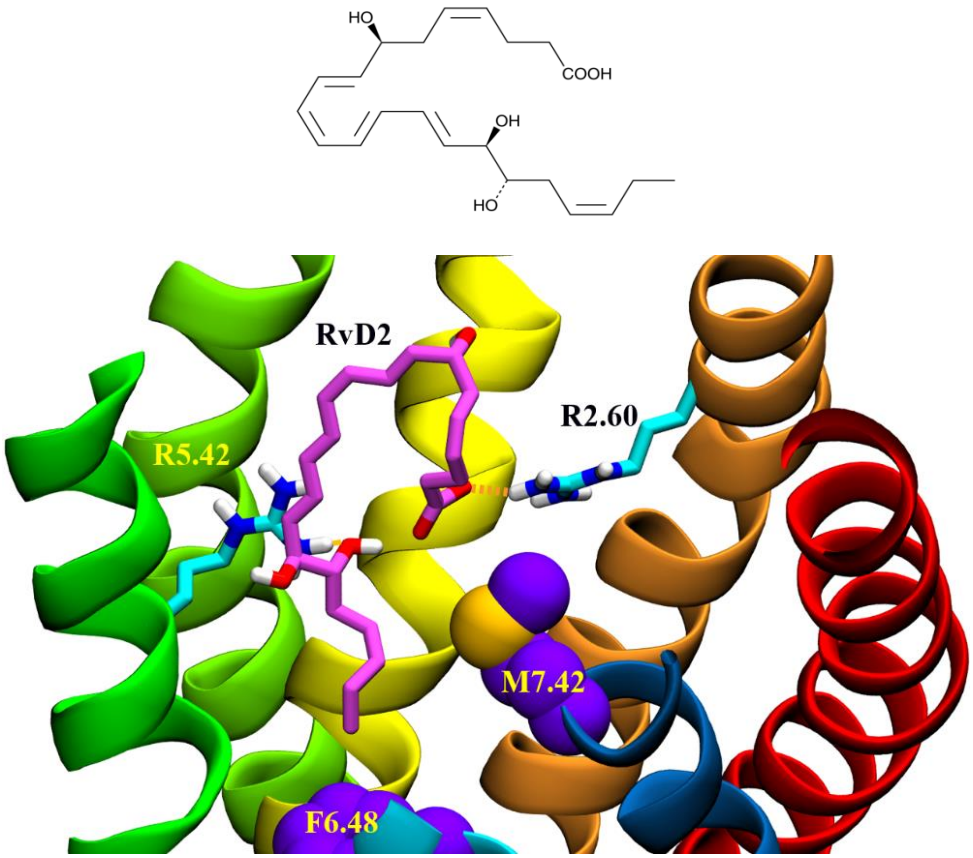


Figure 29. Docking of RvD2 (Pink) Inside the GPR18. Orange Dashed Line is Showing Hydrogen Bonding with R2.60 and R5.42 (Cyan). Toggle Switch Residues (M7.42 and F6.48) are Contoured at Their VdW Radius in Blue. Color Codes for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green) and TMH7 (Dark Blue).

Unlike the global minimum energy conformer of S4, the global min of S5 has no internal hydrogen bond, because the benzyl group is attached to the pyrazole nitrogen closest to a resorcinol hydroxyl group. Because this H bond interaction is lacking, the most stable arrangement for the resorcinol and pyrazole rings is perpendicular to each other. The benzyl group points down from the pyrazole ring, with its phenyl ring adopting an angle that mimics that of the resorcinol ring. Unlike S4, which adopts an elongated conformation, S5 is relatively compact, occupying less space in the binding

pocket. Figure 32 illustrates the S5/GPR18 R* complex. One resorcinol hydroxyl oxygen in S5 hydrogen bonds with R5.42 (O-N distance = 3.3 Å; NHO angle =161°). R2.60 hydrogen bonds with the pyrazole ring N2 position in S5 (N-N distance = 3.3 Å; NHN angle =148°). The S5 benzyl ring has an aromatic stacking interaction with F6.55 (centroid to centroid distance = 3.6 Å; angle of 19°). T3.32 has strong Van der Waals interactions with the benzyl and pyrazole rings, while the S5 alkyl chain has a VdW interaction with L5.38.

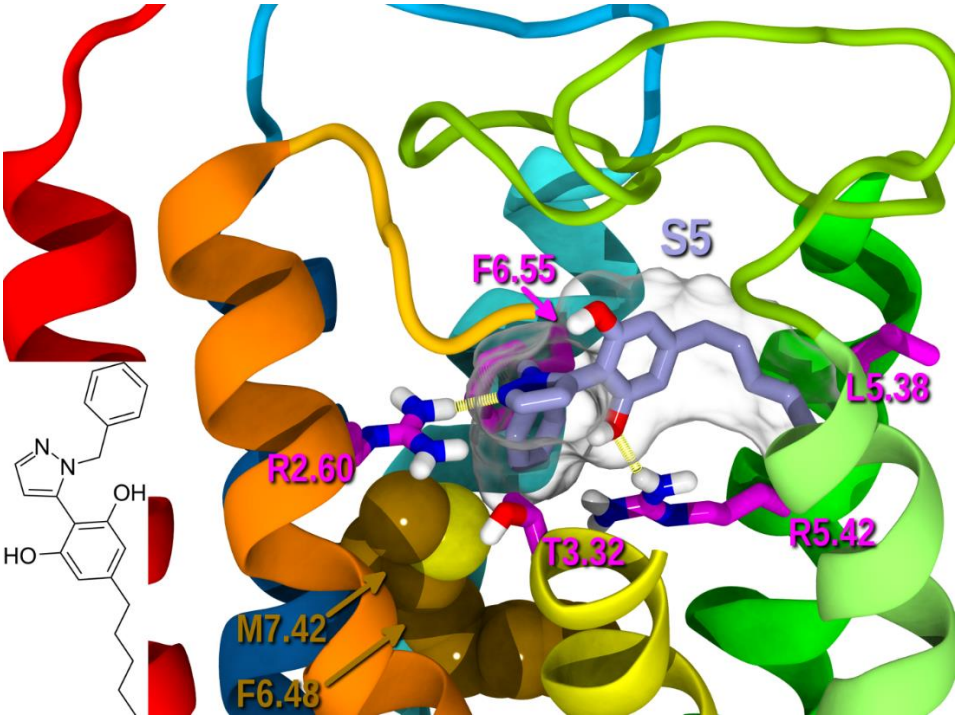


Figure 30. S5 (Light Blue Structure) Docked in R* Model of GPR18. The Hydrogen Bond between Resorcinol Ring and R5.42 and R2.60 are Shown with the Yellow Lines. The Putative Interaction Residues Highlighted in Pink. Toggle Switch Residues (M7.42 and F6.48) are Contoured at Their VdW Radius in Brown Color. Color Codes for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (in Cyan) and TMH7 (Dark Blue). 2D Structure of S4 Inserted in Far Left of the above Picture beside TMH1.

CHAPTER IV

TOWARDS A MOLECULAR UNDERSTANDING OF THE CANNABINOID RELATED ORPHAN RECEPTOR GPR18: A FOCUS ON ITS CONSTITUTIVE ACTIVITY

Noori Sotudeh [1], Paula Morales [2], Dow P. Hurst [1], Diane L. Lynch [1] and Patricia H. Reggio [1,*]

[1] Department of Chemistry and Biochemistry, University of North Carolina at Greensboro, Greensboro, NC 27412, United States; n_sotoud@uncg.edu (N.S.), lynchdl@gmail.com (D.L.L.), dphurst@uncg.edu (D.P.H.)
[2] Instituto de Química Médica, Consejo Superior de Investigaciones Científicas, Madrid 28006, Spain; paula.morales@iqm.csic.es (P.M.)

***** Correspondence: phreggio@uncg.edu (P.H.R.))

Received: 7 April 2019; Accepted: 8 May 2019; Published: date
Int. J. Mol. Sci. 2019, 20(9), 2300; https://doi.org/10.3390/ijms20092300

Abstract

The orphan G-protein coupled receptor (GPCR), GPR18, has been recently proposed as a potential member of the cannabinoid family as it recognizes several endogenous, phytogenic, and synthetic cannabinoids. Potential therapeutic applications for GPR18 include intraocular pressure, metabolic disorders, and cancer. GPR18 has been reported to have high constitutive activity, i.e., activation/signaling occurs in the absence of an agonist. This activity can be reduced significantly by the A3.39N mutation. At the intracellular (IC) ends of transmembrane helices TMH3 and TMH6 in GPCRs, typically, a pair of oppositely charged amino acids form a salt bridge called the "ionic lock." Breaking of this salt bridge creates an IC opening for coupling with the G protein.

The GPR18 “ionic lock” residues (R3.50/S6.33) can form only a hydrogen bond. In this paper, we test the hypothesis that the high constitutive activity of GPR18 is due to the weakness of its “ionic lock” and that the A3.39N mutation strengthens this lock. To this end, we report molecular dynamics simulations of wild-type GPR18 and the A3.39N mutant in fully hydrated phosphatidylcholine lipid bilayers. Results suggest that in the A3.39N mutant, TMH6 rotates and brings R3.50 and S6.33 closer together, thus strengthening the GPR18 “ionic lock.”

Keywords: GPR18; cannabinoids; homology model; G protein-coupled receptor; constitutive activity; molecular dynamics

Introduction

GPR18 is an orphan cannabinoid-like receptor that belongs to the class A family of G-protein coupled receptors (GPCRs). GPR18 is highly expressed in the central nervous system and lymphoid tissues, and it displays high constitutive activity in cells. Its modulation has been associated with numerous physiopathological processes, such as cellular migration, immunomodulation, sperm physiology, cardiac physiology, obesity, intraocular pressure, pain, and cancer, among others [32,33,37,43,89].

The lipid derivative, N-arachidonoyl glycine (NAGly), has been proposed as the putative endogenous ligand for this receptor [26,31,50]. Likewise, the endogenous polyunsaturated fatty acid metabolite, Resolvin D2 (RvD2), implicated in inflammatory processes, has also been reported to target GPR18 [60]. Nonetheless, the International Union of Pharmacology (IUPHAR) still categorizes GPR18 as an orphan GPCR due to

the lack of consistency amongst reports [45,46] and the existence of limited in vivo evidence [90–92].

GPR18 shows relatively low sequence homology with the cannabinoid receptors type 1 and type 2, CB1 and CB2 (~13% and 8%), yet a subset of endogenous, phytogenic, and synthetic cannabinoid ligands have been shown to modulate GPR18 [92]. This fact has prompted numerous studies to understand the relationship between GPR18 and the endocannabinoid system, which comprises the CB1 and CB2 receptors; endogenous ligands, anandamide, and 2-arachidonoylglycerol (2-AG); their associated enzymes; and transporter.

GPR18 signaling pathways are still not fully explored due to the lack of pharmacological tools, challenging heterologous expression, and high intracellular receptor expression [32,47]. Coupling to Gαi/o and Gαq has been reported for GPR18 [26,31,33,50,63], whereas GPR18 did not couple to other G proteins tested, including Gαs, Gαz, and Gα15 [46]. Several research groups have found that GPR18 has high constitutive activity [32,47]. For instance, Glass and coworkers observed that the introduction of an A3.39N mutation triggered a reduction in constitutive activity, as well as a higher steady-state surface receptor expression [47]. The role of the residue located at position 3.39 has been suggested to be relevant at the sodium binding pocket and to be critical to the control of constitutive activation at specific GPCRs, such as the chemokine receptors [93–95].

At the intracellular (IC) end of class A GPCRs, an “ionic lock” formed typically by R3.50 and D/E6.30 forms a salt bridge that holds the receptor in its inactive state.

Breaking of this lock creates an opening to allow coupling with the G protein, i.e., signaling. While GPR18 has an arginine at 3.50, it lacks a negatively charged residue at the end of TMH6. Instead, GPR18 has a hydrogen bonding residue, S6.33, at this position that interacts with R3.50. In this paper, we test the hypothesis that the high constitutive activity of GPR18 is due to the weakness of the GPR18 "ionic lock" and that the A3.39N mutation strengthens this lock and thereby reduces constitutive activity. To this end, we have constructed homology models of wild-type (WT) GPR18 and the A3.39N mutant in their inactive states. These models were used to conduct molecular dynamics simulations of WT GPR18 and the A3.39N mutant in fully hydrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayers. Our results suggest that the presence of N3.39 can lead (via changes in the sodium binding pocket) to a TMH6 rotation that brings R3.50 and S6.33 closer together, thus strengthening the "ionic lock" in the A3.39N mutant and reducing constitutive activity.

These molecular modeling studies provide structural insights into GPR18 and its activation. Knowledge gained in the present study should provide a structural basis for mutation studies of GPR18 and the development of novel ligands for GPR18.

Results and Discussion

GPR18 Model Development

Our GPR18 homology model of GPR18 was constructed using the X-ray crystal structure of the delta opioid receptor (DOR) at a 1.8 Å resolution as a template (PDB identifier: 4N6H). The rationale for the choice of template is provided in the Methods section. Upon mutation of the crystal template, the initial model was refined based on

sequence divergences between GPR18 (Figure 31) and the DOR crystal structure. Conformational differences in specific transmembrane helices were investigated using the simulated annealing/Monte Carlo (MC) conformational analysis method of conformational memories (CM) as described in the Methods.

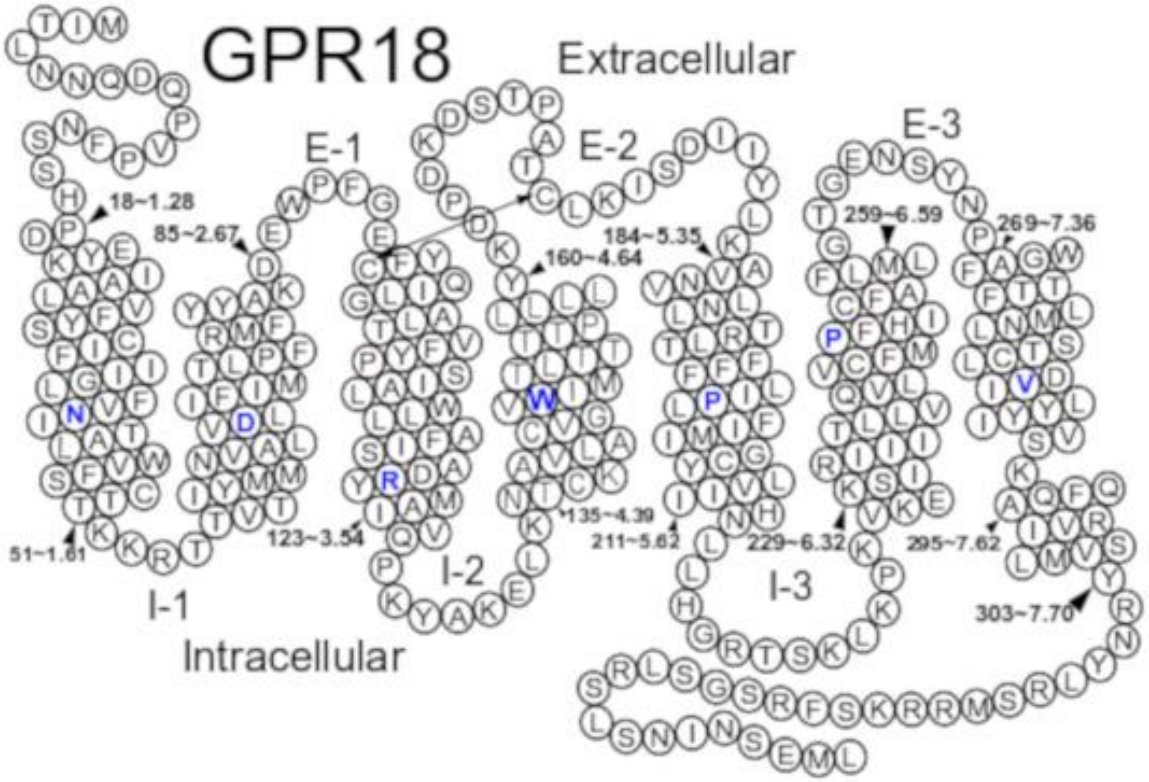


Figure 31. Helix Net of the Human GPR18 Sequence. Among Class A GPCRs, the Most Highly Conserved Residue in Each Helix is Colored Blue. GPR18 Has a Proline at P3.36, it Lacks the Highly Conserved Proline at Position 7.50, Bearing a Valine at That Position Instead. A Key Disulfide Bridge between the Extracellular Loop 2 (EC-2) and the top of TMH3 is Indicated by a Double Headed Arrow. Black Arrowheads Indicate Specific Residue Numbers in the GPR18 Sequence (Absolute and Ballesteros-Weinstein Numbering).

TMH3. In the GPR18 sequence, there is a P3.36 in TMH3 (transmembrane helix 3) that is unique. Since the presence of a proline introduces a kink in the helix (well

documented in previous reports) [75], the CM method was used to explore the possible conformational space of TMH3. To achieve this goal, an initial regular α-helix was built, and then variations around the proline kink region were introduced. This region includes the residues from P3.36 to T3.32. The range of variations was $-120° < \varphi < 40°$, $-70° < \psi < 0°$, and $-160° < \omega < 160°$. These values were extracted from the Protein Data Bank [66]. Finally, the 103 structures generated by CM were superimposed from P3.36 to Q3.56. Results were compared with the template. Although TMH3 in the DOR crystal structure lacks a P3.36 (M3.36 instead), it is bent in this region with an average bend angle of 16.5°, calculated using the ProKink program [96]. This bend is structurally explained by the presence of T3.38 in a gauche minus ($g^- = 0° < \chi1 < 120°$) dihedral. Average Ser/Thr kinks measured in alpha helices are usually around 12° [76]. CM outputs suggested that GPR18 TMH3 with its P3.36 has a proline kink angle of 16.8°, similar in magnitude to the Ser/Thr induced bend in TMH7 of DOR.

TMH7. GPR18 lacks the TMH7 NPXXY motif. Instead, it has DVXXY. Prolines are hinge residues capable of bending helices at specific positions [96]. It may seem that the lack of P7.50 will lead to a straight helix. However, our conformational studies led us to the observations detailed below. On one hand, incorporating a straight TMH7 in the TMH bundle created an empty space on the extracellular (EC) side of the bundle that cannot be filled by EC loops (specifically EC2 and 3), allowing more water to enter the orthosteric binding pocket than would be expected for a GPCR that binds lipid-like ligands. Second, a straight TMH7 changes the position of some key residues. For example, N7.43 is typically a binding pocket residue, but if TMH7 is straight, its Cα

position will face the lipid. Residue 7.39 (T7.39) has been found in several GPCRs to be a ligand interaction site, and therefore, normally, 7.39 is more accessible to the binding crevice. However, in a straight TMH7 helix, T7.39 faces towards TMH1 and away from the typical ligand binding pocket. Finally, in a straight TMH7, Y7.53 deviates from its normal position in the sodium binding pocket [95] (the region with polar residues, including N1.50, D2.50, S3.39, N7.49, and Y7.53, that are coordinating directly with sodium ion or water molecules).

While TMH7 lacks a proline at 7.50, there are other residues in TMH7 that may induce a helix distortion/hinge similar to a proline. Ballesteros et al. [76] showed that serines and threonines in alpha helices could act as hinge residues. The bending produced by each of these residues is produced by an intrahelical hydrogen bond between the backbone carbonyl oxygen (at position i-4 for an alpha helix and i-3 for a 3/10 helix) and the O atom of the Ser or Thr in a $\chi 1$ g- dihedral. In GPR18 TMH7, residues S7.45 and T7.46 deserve particular attention because of their close location to the common P7.50 in other GPCRs. Therefore, we used the CM method to explore the effect of S7.45 and T7.46 on TMH7 conformations. This study explored g^-/g^+, g^+/g^-, and g^-/g^- combinations of S7.45 and T7.46, respectively, by variation of the $\chi 1$ dihedral of these two residues to the chosen value. The variation range was the same as that used for the proline kink in TMH3 for the φ, ψ, and ω dihedrals. Altogether, the g^-/g^- combination mimicked best the kink produced by P7.50 in DOR. Among the low-free energy TMH7 conformers obtained using CM, the TMH7 helix that fit in the bundle without van der Waals overlaps with residues on other TMHs was selected for the GPR18 model. This TMH7 has a helix bend

angle of 14.4°, a wobble angle of −16.2°, and a face shift of 26.4° at S7.45 and has a helix bend angle of 15.7°, a wobble angle of −106.5°, and a face shift of 32.6° at T7.46.

GPR18 Activation. The binding of an agonist and subsequent receptor activation is accompanied by significant conformational changes in the receptor inactive state structure. Constitutively activated GPCRs do not need the presence of an agonist to achieve the active state (R*). The main structural changes that define the R* bundle occur in the cytoplasmic end of the receptor. One of the crucial features is the outward movement of TMH6 from the receptor core. This movement causes the breakage of the "ionic lock" (R3.50 and S6.33 interaction, further detailed below), increasing the distance between the IC ends of TMH6 and TMH3. This movement generates an intracellular opening that allows G-protein interaction. P6.50 determines the flexible hinge region that mediates the transition from R to R* [97]. To create our activated bundle from the GPR18 inactive state R model, a CM analysis was performed in the hinge region of P6.50 to V6.46. From the obtained outputs, a less kinked TMH6 conformer (bend angle: 22°) was selected for TMH6 in the R* state model. This TMH6 conformer pulls away from the TMH bundle on the IC side, ensuring that the TMH3-TMH6 "ionic lock" is broken. The resultant bundle was then minimized as described in the Methods section.

Ionic Lock and Toggle Switch. TMH6 plays an essential role in the transition between the inactive and activated states. When the receptor is activated, the intracellular end of TMH6 moves away from the transmembrane helix bundle, opening an intracellular space for coupling with the G protein. This mechanism is mediated by two important switches that are common between TMH3 and 6. The first one is the "ionic

lock" at the IC ends of TMH3 and TMH6. In most class A GPCRs, the "ionic lock" is a salt bridge between R3.50 and D/E6.30. In GPR18, the ionic lock is actually a stable hydrogen bond formed between R3.50 and S6.33 (O-N distance = 2.9 Å; NH-O angle of 164°). The role of the "ionic lock" is to limit the relative mobility of the cytoplasmic ends of TMH3/6 in the inactive state, and by disruption in the active state, leads to the formation of an opening at the bottom of the receptor into which the G-protein inserts its α-5 helix to couple with the receptor.

A second key structural feature is the "toggle switch." The toggle switch is a pair (trio or quartet) of residues, one of which is W/F6.48 (of the CWXP flexible hinge motif) and the other(s), is (are) typically an aromatic or bulky residue that forms an interaction or stack with W/F6.48, stabilizing it in a $\chi 1$ = g+ conformation. In GPR18, the toggle switch is formed by F6.48, M7.42, and H6.52 and Y3.55 (see Figure 32). When an agonist enters the GPR18 binding pocket, it triggers a F6.48 $\chi 1$ change from g+ to trans. This conformational change straightens TMH6 and breaks the ionic lock. Since antagonists/inverse agonists stabilize the inactive state of the receptor, an antagonist/ inverse agonist should block the F6.48 $\chi 1$ transition from g+ to trans, thus keeping GPR18 in its inactive state. The highly conserved F6.44XXCWXPXF6.52 motif is replaced by V6.44XXCFXPXH6.52 in GPR18. Moreover, the lack of a bulky residue, like M3.36 (in GPR55), or a F3.36 (like CB1 and CB2) [78] directly across from F6.48 makes it difficult to control the toggle switch in GPR18. Since there is a P3.36 at this position, there is much room for F6.48 to change its $\chi 1$ conformation from g+ to trans.

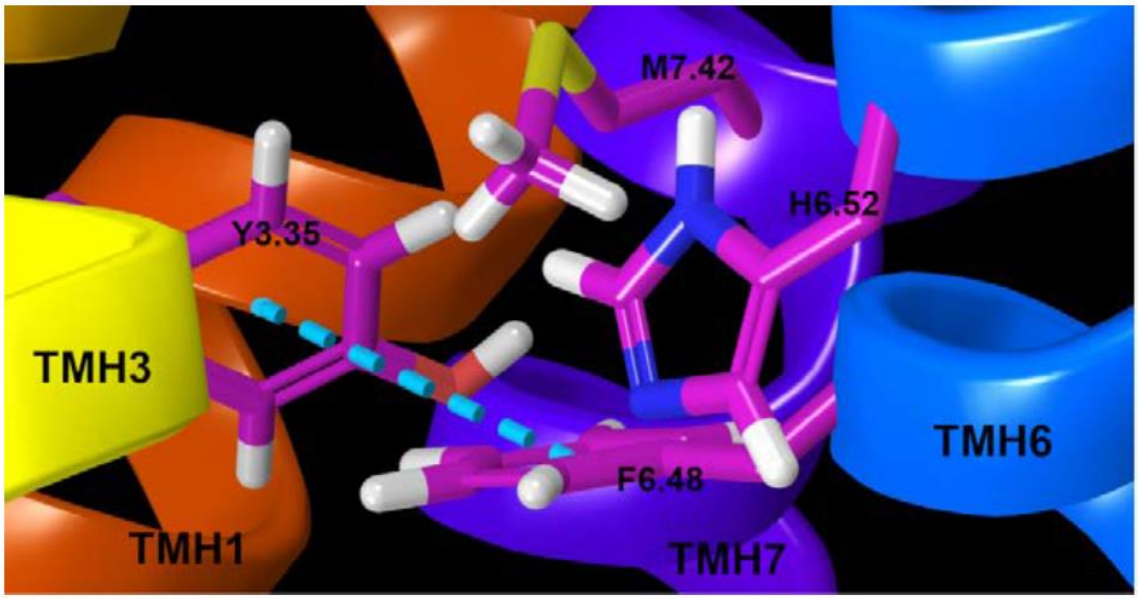


Figure 32. The Toggle Switch in GPR18 R State. M7.42/F6.48 and Y3.35/F6.48 Interactions (Blue Dashed Line) Keep the F6.48 χ1 Angle in a g+ Rotamer, Which Retains the H6.52 χ1 in g+ in the Inactive State (R) Model. Key Residues are Highlighted in Magenta Thin Tubes.

Our GPR18 model suggests that Y3.35 and M7.42 directly, and H6.52 indirectly, participate in the GPR18 toggle switch. The current model shows that M7.42 packs the terminal CD methyl carbon against the aromatic ring face of F6.48 by interacting through a favorable Met/aromatic interaction to hold F6.48 in its χ1 = g+ dihedral (M7.42 CG, SD, and CE atom distances to the centroid of F6.48 are 4.26 Å, 5.01 Å, and 3.93 Å, respectively). Also, Y3.35 presents an aromatic face to the aromatic edge of F6.48 in a tight tilted-T aromatic stack (centroid to centroid distance = 5.2 Å; angle of 58°). H6.52 is i + 4 to F6.48 and is sterically required to switch trans synergistically when F6.48 undergoes the χ1 g+ to trans rotameric change (Figure 33). These residues are also in a tilted-T aromatic stack (centroid to centroid distance = 5.99 Å; angle of 75.9°). Taken together, our model predicts that an antagonist should promote M7.42 to pack against

F6.48. An antagonist could also hold H6.52 in a χ1 = g+ to promote antagonism. However, an agonist could insert between Y3.35 and F6.48, while moving M7.42 away from F6.48.

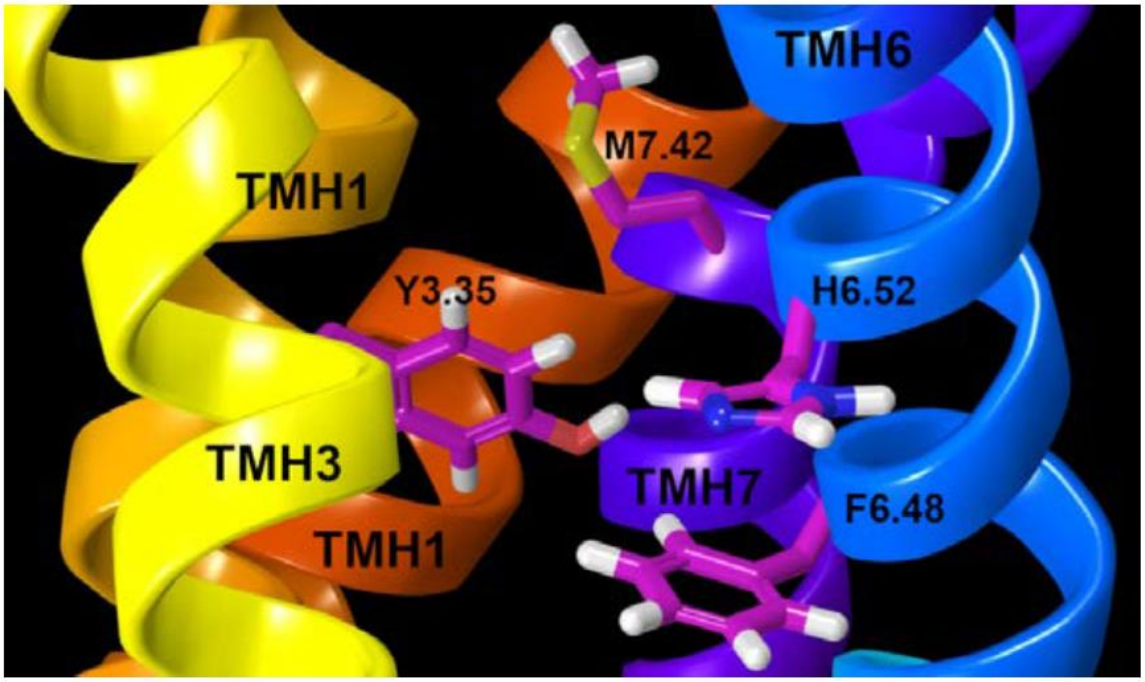


Figure 33. The Toggle Switch in the Activated State. The M7.42/F6.48 and Y3.35/F6.48 interactions have broken, causing the F6.48 χ1 to switch to a trans rotamer. This synergistically leads H6.52 to adopt a χ1 trans conformation in the R* model. Key residues are highlighted in magenta thin tubes.

Molecular Dynamics: Structural Insights into Its Constitutive Activation

GPR18 is a highly constitutively active receptor that produces low cell surface expression [47]. A GPR18 A3.39(108)N mutant has been reported to reduce this constitutively cycling phenotype, resulting in higher steady-state surface receptor expression [47]. As a test of our newly developed GPR18 models, we undertook a study of the effects of an A3.39(108)N mutation on the ability of the mutant to activate. To do

this, we performed molecular dynamics simulations of (1) the wild-type (WT) GPR18 inactive state and (2) the mutant A3.93(102)N inactive state. Each inactive state model was embedded in a fully hydrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer, and each was simulated for 700 ns, with the first 100 ns required for equilibration. Full details of MD simulations are provided in the Molecular Dynamics (MD) section. In these simulations, helix8 was placed parallel to the plane of the membrane at the lipid/water interface. Also, a sodium ion was positioned in the putative sodium ion binding site to stabilize the inactive state of the receptor. The role of sodium ions inside the GPCRs has been well documented [95,97–99].

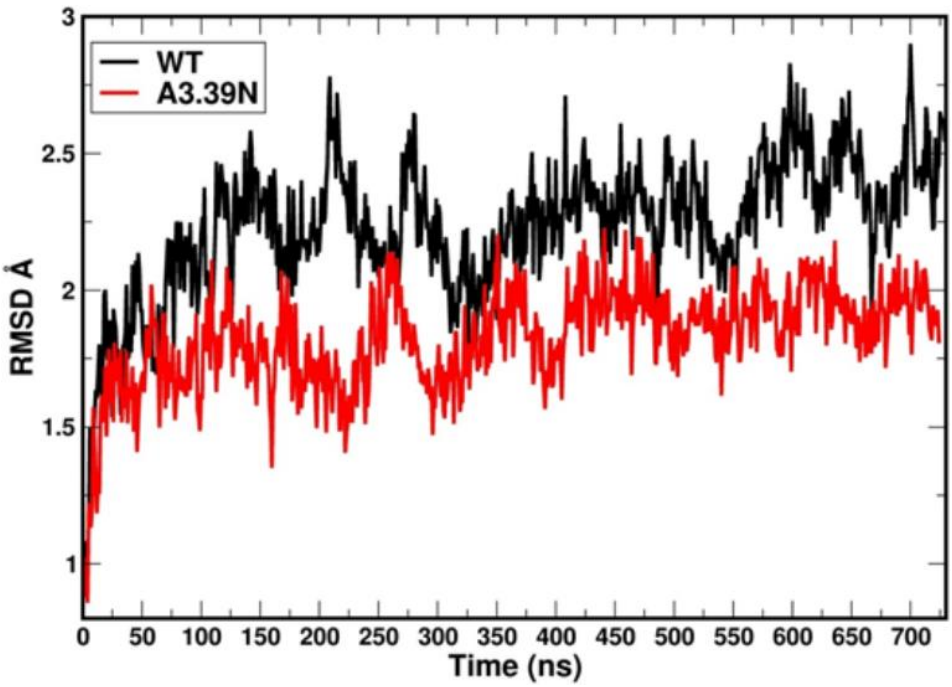


Figure 34. RMSD vs. Simulation Time Plots for the WT GPR18 R State (Black) and the A3.93N Mutant R State (Red) [72]. The RMSD Levels out by 100 ns in Both Simulations. So, the first 100 ns Correspond to the Equilibration Phase and the Next 600 ns are Considered Production.

Ability to Maintain "Ionic Lock" and Arginine Cage. The RMSD (root mean square deviation) plots for the WT (black) and A3.39N mutant [72] are shown in Figure 3. The WT system shows gradual fluctuations and quick stabilization with a final RMSD less than an average of 2.5 Å, while the A3.39N mutant also shows gradual fluctuations and quick stabilization with a final RMSD less than an average of 2.0 Å. Figure 35 illustrates the heteroatom-to-heteroatom distance for the R3.50-S6.33 "ionic lock" pair for the WT (black) and A3.93N [72] trajectories. The green line in Figure 4 marks the maximum heteroatom-to-heteroatom distance for a hydrogen bond to exist (3.2 Å). Points along the trajectory where the plot falls under the green lines are regions where the R3.50/S6.33 hydrogen bond is formed. As the molecular dynamics (MD) results in Figure 4 show, the "ionic lock" in WT breaks immediately and never re-forms.

In contrast, the "ionic lock" hydrogen bond for the A3.39N mutant first forms around 100 ns and persists to 225 ns; it then "flickers" on and off during the rest of the trajectory until very near the end at ~670 ns. The fact that the R3.50-S6.33 hydrogen bond in the WT trajectory never re-forms once broken, but the same hydrogen bond in the A3.39N mutant forms, breaks, and re-forms over the course of the trajectory is a reflection of the strength of the hydrogen bond between a charged and uncharged partner, which is not as strong as the R3.50-D/E6.30 salt bridge "ionic lock" found in many GPCRs. However, there clearly must be a difference between the IC domains in the WT vs. mutant since the mutant manages to form the "ionic lock" at least some of the time. Overall, the MD simulations suggest that the WT receptor is extremely constitutively active, while the mutant is less so. They also suggest that the A3.39N mutant should

show a decrease in basal activity. These two results are consistent with the GPR18 literature [47].

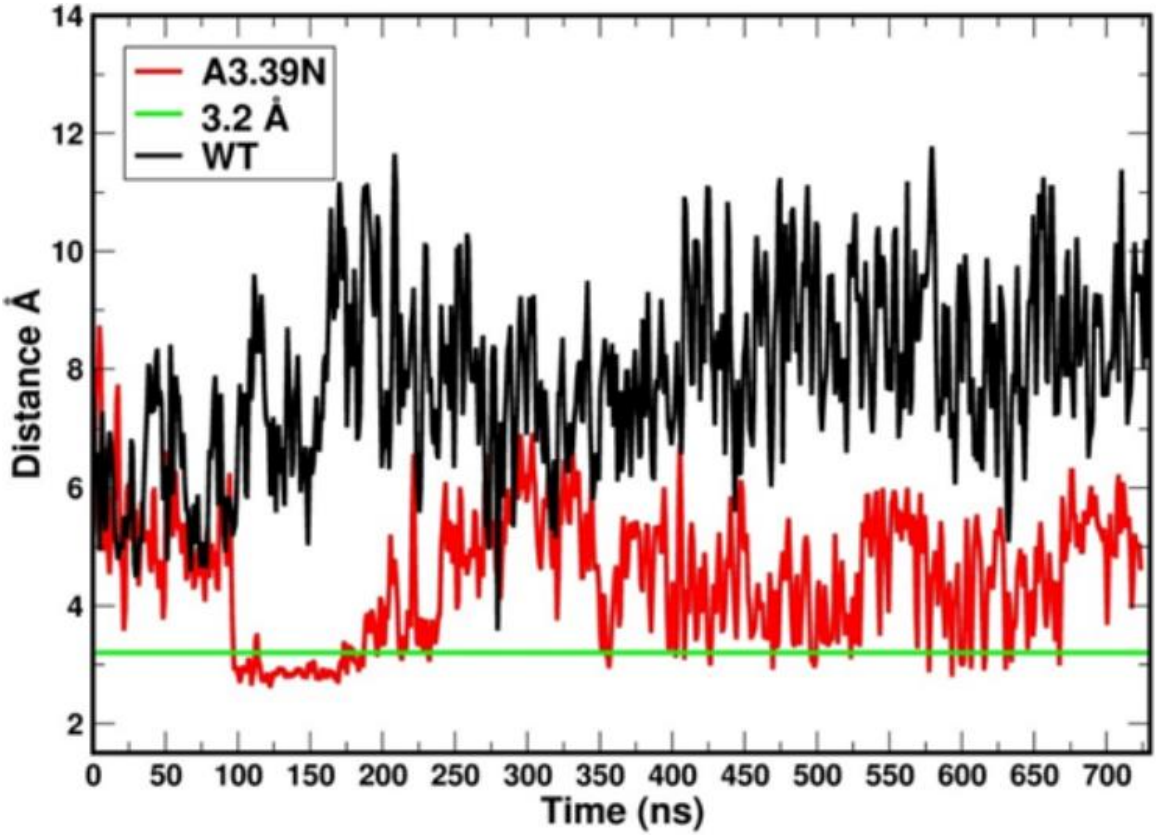


Figure 35. Stability of the "Ionic Lock" Over the Course of the WT (Black) and A3.93N (Red) Trajectories. The Green Line Shows the Maximum Distance between Heteroatoms for a Hydrogen Bond to Exist (3.2 Å). The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase.

It has been proposed that in the inactive state of class A GPCRs, R3.50 is located in an "arginine cage," in which it forms an intra-helical salt bridge interaction with D3.49 [100]. Figure 36 illustrates plots of the heteroatom to heteroatom distance between R3.50 and D3.49 for WT GPR18 (black) and the A3.39N mutant (red) [72]. The green line at 4.0 Å marks the maximum heteroatom-to heteroatom salt bridge distance. This plot

shows that an R3.50-D3.49 salt bridge exists for a significant amount of time for the mutant, but this salt bridge exists for WT until 200 ns and is afterward broken.

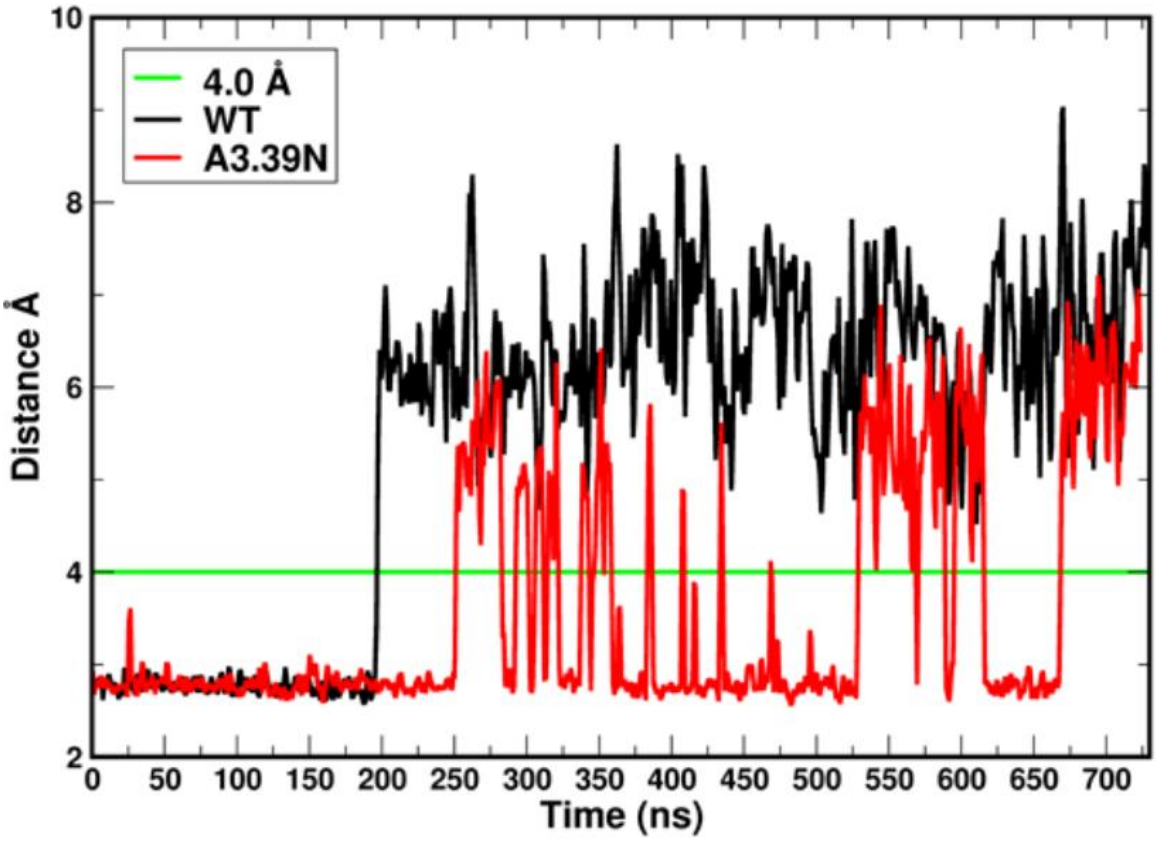


Figure 36. Heteroatom Distance between R3.50 and D3.49 in the WT GPR18 (Black) and the A3.39N Mutant (Red). The Green Line at 4 Å Marks the Maximum Salt Bridge Distance between Heteroatoms. The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase.

These results suggest that the A3.39N mutant receptor has a more stabilized arginine cage than WT. Since R3.50 has to be in a specific position in the arginine cage to interact with D3.49, this position may promote R3.50's ability to also hydrogen bond to S6.33. Figure 37 shows a plot of the average distance between Cα atoms of the last five residues on the cytoplasmic side of TMH3 and TMH6 for WT (black) and the

A3.39N mutant [72]. As depicted in Figure 6, the distance of the centers of mass of the last five cytoplasmic end residues of TMH3 (residues Y3.51 to V(124)) and TMH6 (residues K6.30 to I6.34) clearly reveals that in the A3.39N mutant receptor, the cytoplasmic ends of TMH3 and TMH6 are closer during the course of the trajectory than those of the WT receptor. In fact, after 200 ns, the bottom of the WT receptor appears to be opened, suggesting that it has activated. The tendency of the WT receptor to have its ionic lock broken and the TMH3 and TMH6 IC ends to move away from one another is consistent with the very high constitutive activity reported for GPR18 [47].

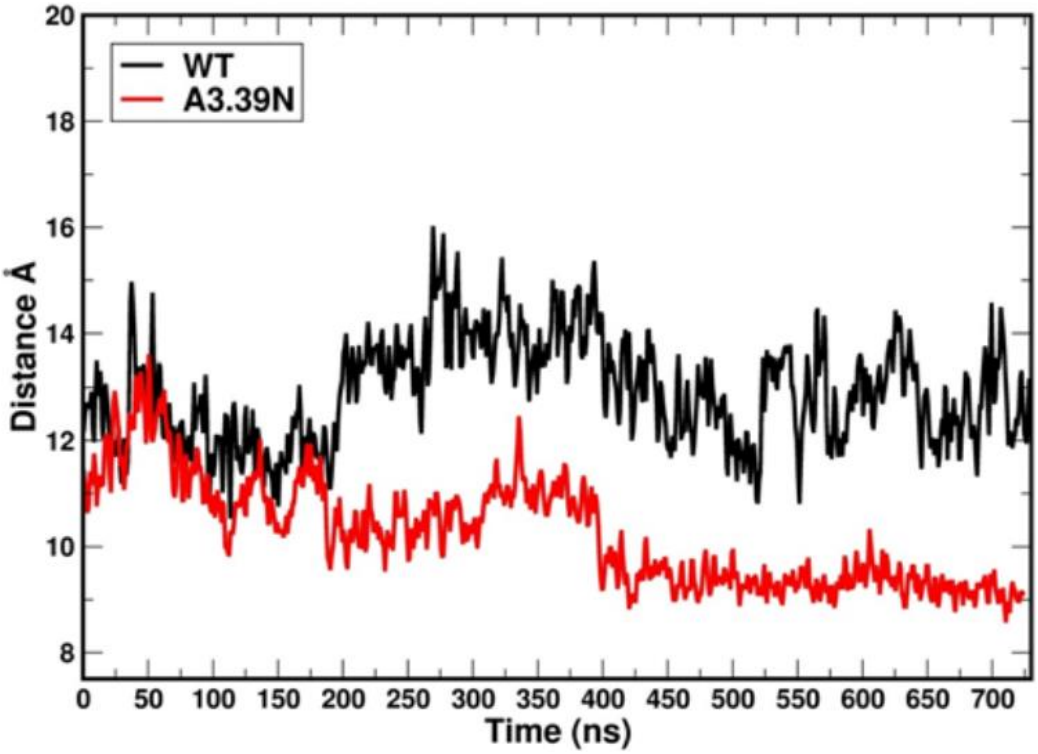


Figure 37. Average Distance between Cα Atoms of the Last Five Cytoplasmic Residues in TMH3 (Y3.51 to V(124) and TMH6 (K6.30 to I6.34) in Wild Type GPR18 WT (Black) and the A3.39N Mutant (Red). The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase.

The Sodium Binding Pocket. To explain the relationship between the A3.39N mutation and GPR18 constitutive activity, the position of A3.39 in the two simulations previously described were analyzed. Results from prior studies have shown that in many GPCRs, sodium ions stabilize the inactive state of the receptor by interacting with two highly conserved residues, D2.50 and S3.39, and three water molecules [95]. Interestingly, the GPR18 sequence is different from most class A GPCRs in this region as it lacks the highly conserved polar residue, S3.39 (WT GPR18 has an A3.39 instead). The second important difference in the GPR18 sequence is that unlike 86% of class A GPCRs, GPR18 lacks the NPXXY motif in TMH7 (N7.49-Y7.53). Instead, it has DVXXY (D7.49-Y7.53). The presence of a negatively charged residue, D7.49, instead of a neutral one, N7.49, and the lack of a polar residue, like S3.39, might cause a shift of the sodium ion position toward the cytoplasmic end of the receptor by coordination with two aspartic acid residues, D2.50 and D7.49 (instead of D2.50 and S3.39). These drastic sequence differences in the sodium binding pocket region indicate that it is more likely that the sodium ion is located in a lower place than typically observed in other receptors. Our MD results support this hypothesis. The distance between the sodium ion and the Cα atom of N3.39 in the A3.39N mutant is shorter than the distance between the sodium ion and the Cα atom of A3.39 in WT (Figures 38 and 39).

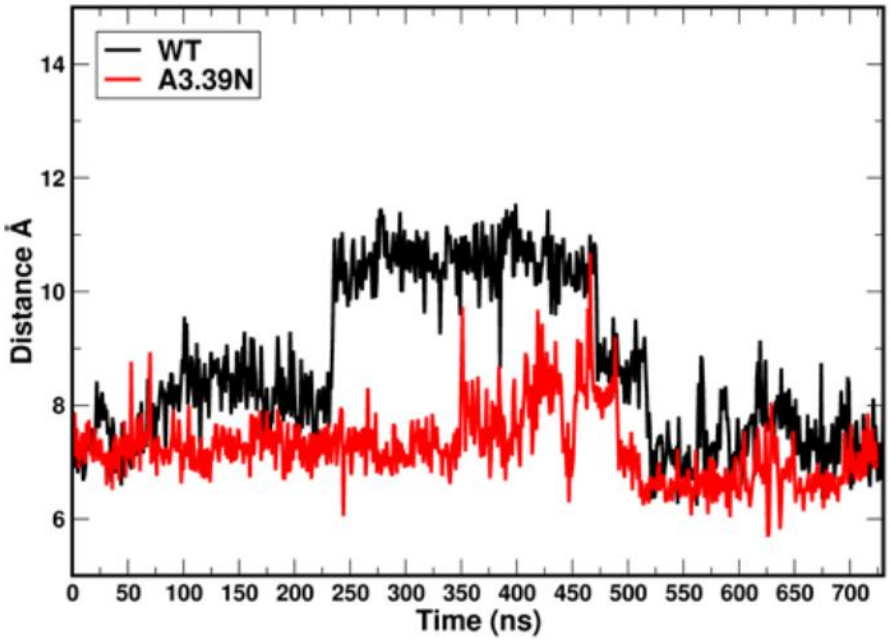


Figure 38. The Distance between the Sodium Ion and N3.39 in the A3.39N Mutant (Red) [72] and A3.39 in WT (Black). The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase.

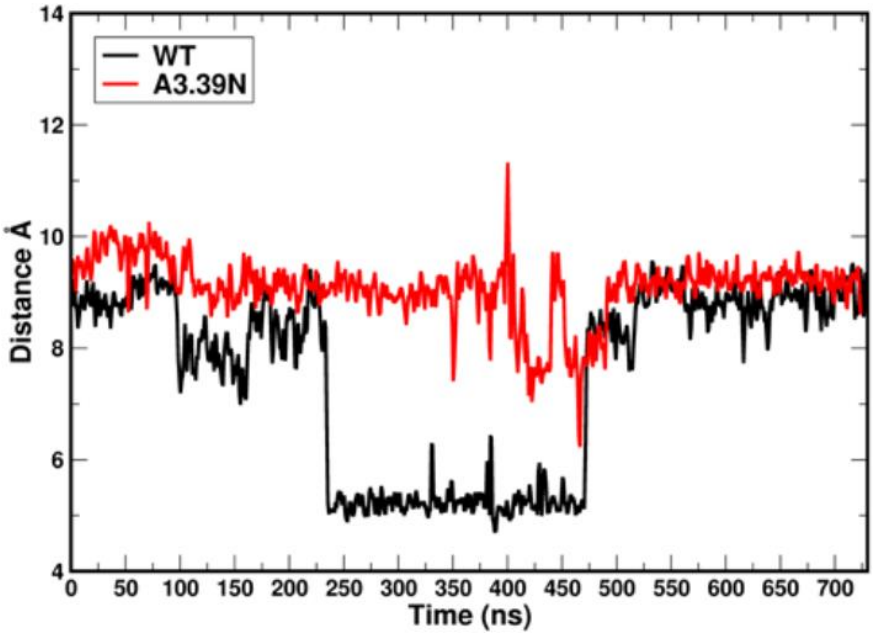


Figure 39. The Distance between the Sodium Ion and N1.50 in the A3.39N Mutant (Red) [72] and WT GPR18 (Black). The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase.

On the other hand, the distance between the Cα atom of N1.50 and the sodium ion increased in the A3.39N mutant (Figures 7B, 8). Therefore, the sodium ion has moved toward the EC side in the A3.39N mutant; this is due to the hydrophilicity of residue N3.39, which contributes more to binding to the sodium ion than the WT hydrophobic residue (A3.39) would contribute. The sodium ion is pulled approximately 3.0 Å up towards the EC region of the receptor (Figure 40).

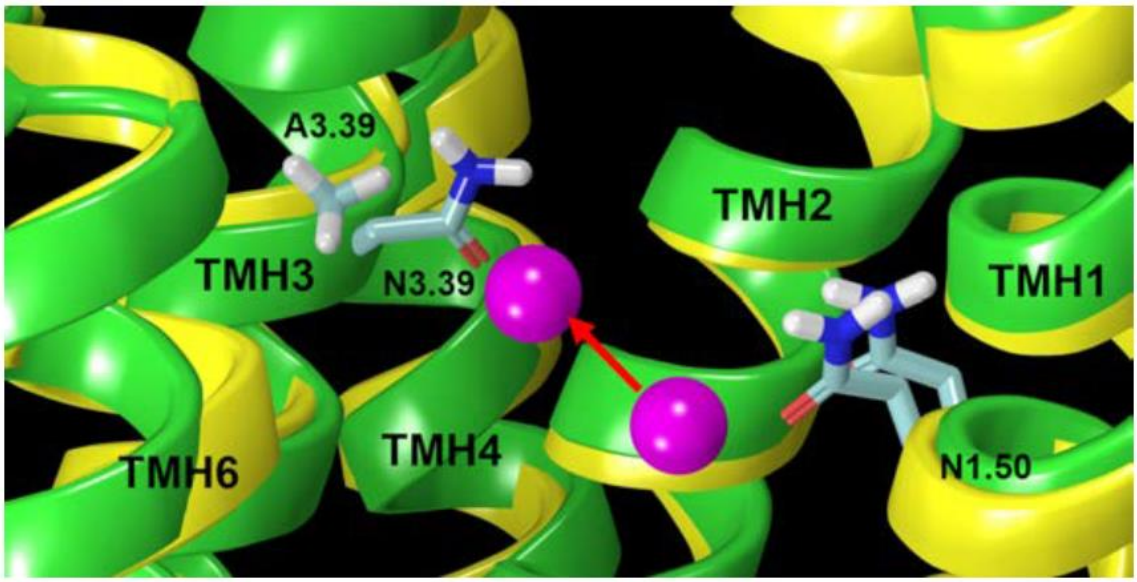


Figure 40. Comparison of the Sodium Ion (Magenta Sphere) Position in the WT Bundle (Yellow) and A3.39N Bundle (Green). The Lower Sodium Ion is Its Position in the WT Receptor (Yellow Bundle); While the Upper Sodium Ion Position is That in the A3.39N Mutant (Green Bundle). The Red Arrow Shows the Shift of the Sodium Ion Position towards the EC Side of the Receptor. Key Residues are Highlighted in Light Blue.

The other important aspect of the unique sodium ion binding pocket in GPR18 is the conformation of D7.49. MD results after 100 ns show that in the WT receptor, the D7.49 χ1 torsion tends to adopt a g^- rotamer conformation (+55°, average), instead of a typical g+ rotamer conformation (Figure 41). This is because in the WT system, two

negatively charged residues, D2.50 and D7.49, are strong enough to pull the sodium ion down below the D7.49 carboxylic group side chain. Thus, the ideal conformation in the WT receptor for D7.49 to easily access the sodium ion and directly interact with it is a $\chi1 = g^-$ conformation (Figure 42, Movie S1). Conversely, the A3.39N mutant simulations show that D7.49 retains its g+ conformation (−55° on average) (Figure 41). Since in the A3.39N mutant the sodium moves up (because of the presence of N3.39, Figure 40), the carboxylic acid group of D7.49 follows the ion upward so that they can interact, with the D7.49 assuming a g+ $\chi1$ value (Figure 43, Movie S3). The detailed analysis shows that D7.49 can form an additional important inter-helical hydrogen bond with Q6.43 in the A3.39N mutant (Figure 43, Movie S3). In contrast, this inter-helical hydrogen bond is broken in the WT bundle (Figure 42, Movie S1).

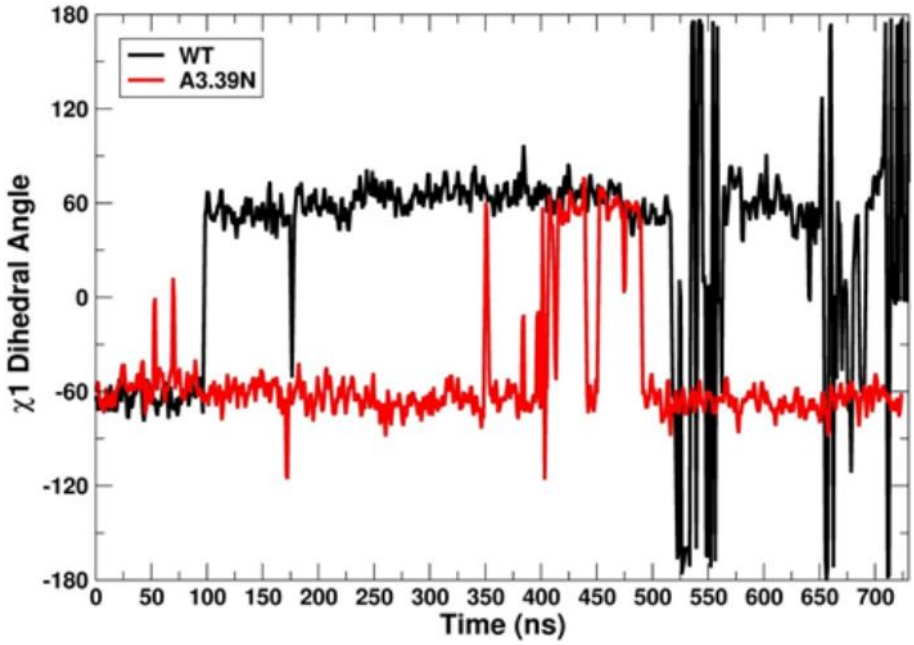


Figure 41. $\chi1$ Rotamer Conformational Change in D7.49 in WT GPR18 (Black) and the A3.39N Mutant (Red) [72] Due to Sodium ion Movement. The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase.

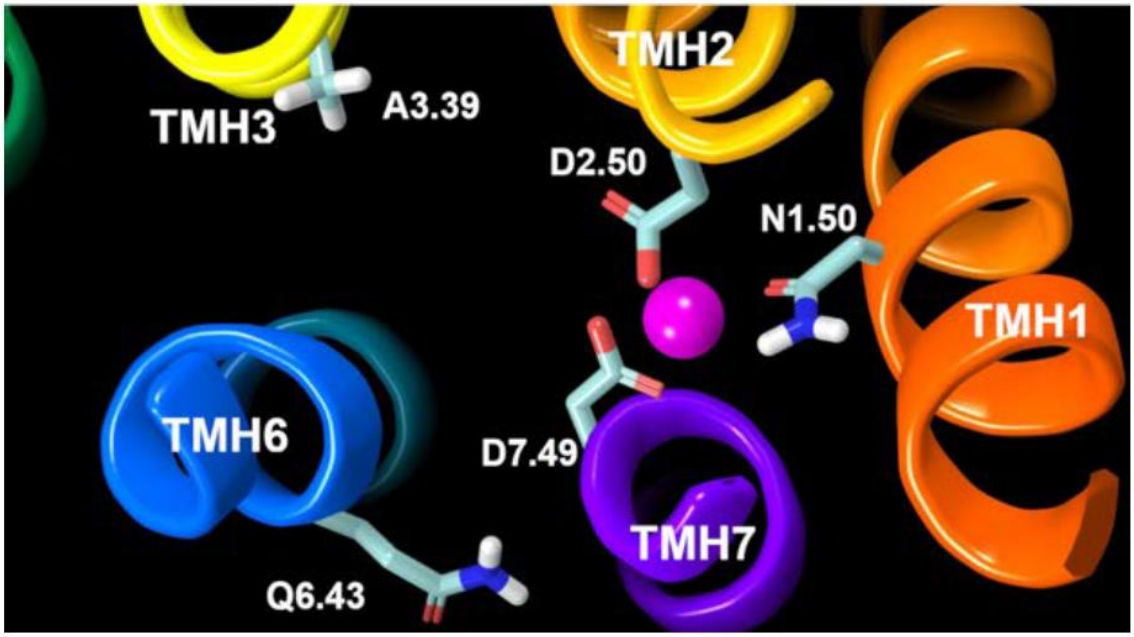


Figure 42. WT GPR18 Bundle. The Sodium Ion (Magenta Sphere) is Coordinated Between D7.49, D2.50, and N1.50 (Highlighted in Light Blue Tubes). To Achieve This, D7.49 Needs to Break its Hydrogen Bond with Q6.43.

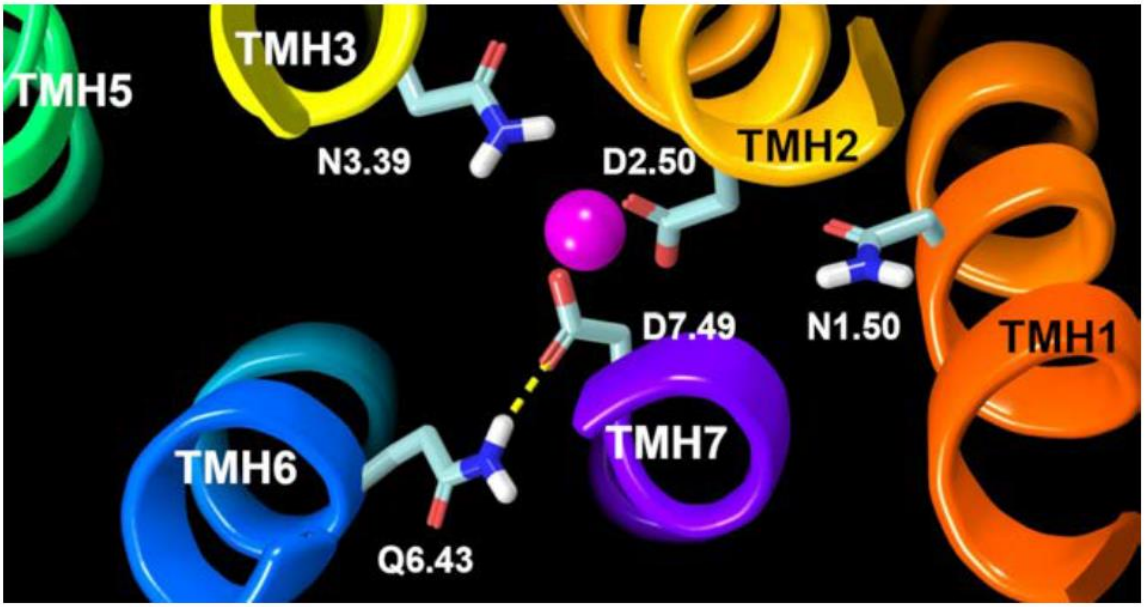


Figure 43. A3.39N Mutant Bundle. The Sodium Ion (Magenta Sphere) is Coordinated Between D7.49, D2.50, N1.50, and N3.39 (Highlighted in Light Blue Tubes) so That it is Possible for D7.49 to Keep the Hydrogen Bond with Q6.43 (Yellow Dashed Line).

The graphs in Figure 44 compare the heteroatom to heteroatom distances for the Q6.43-D7.49 pair. As shown in Figure 44, during the equilibration (first 100 ns), the distance is small enough for a hydrogen bond to exist between the two residues in the WT receptor (black line), which corresponds to the time frame in which D7.49 changes its conformation (Figure 41, black line). Past 100 ns, the WT distance is too great for a hydrogen bond to exist. In contrast, this distance is frequently below the maximum for a hydrogen bond in the A3.39N mutant.

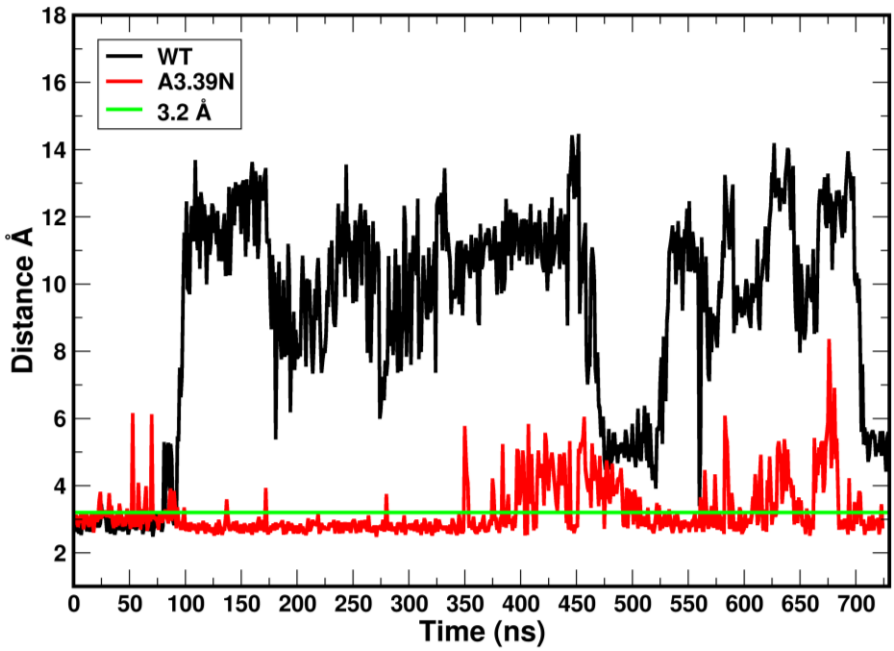


Figure 44. Heteroatom to Heteroatom Distance Between Residues Q6.43 and D7.49 in WT GPR18 (Black) and the A3.39N Mutant (Red) [72]. The First 100 ns Correspond to the Equilibration Phase. The Next 600 ns are Considered the Production Phase. The Green Line Corresponds to 3.2 Å, the Maximum Distance to be Considered a Hydrogen Bond.

Our MD simulations show that there is a strong relation between the Q6.43 to D7.49 hydrogen bond and the R3.50 to S6.33 “ionic lock” interaction. Even though

glutamines are neutral, polar residues, based on our GPR18 model, Q6.43 is positioned so that it faces the lipid side of TMH6 (Figure 45).

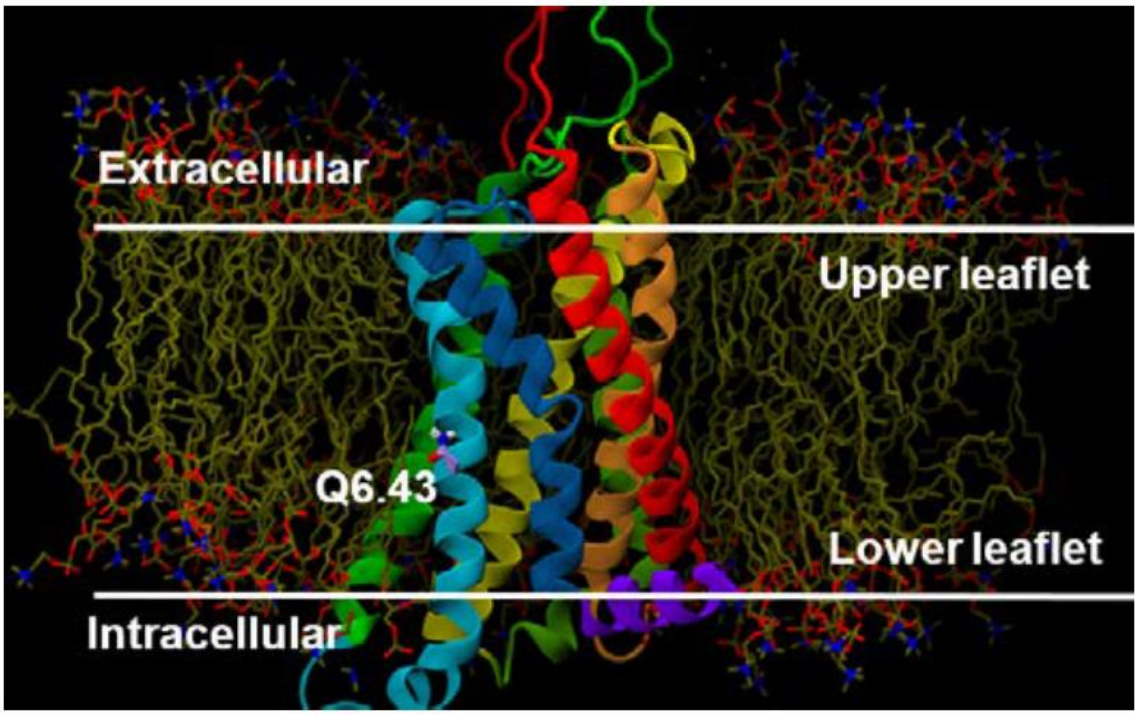


Figure 45. Position of Q6.43 in the GPR18 R Model (Highlighted in Purple Thin Tubes).

Comparison of the Q6.43 sequence position with the corresponding residue in other class A GPCRs shows that the amino acids at this position are either non-polar residues (such as Val in CB1 and CB2, Ala in rhodopsin, Leu in adenosine A2A receptor) or small polar residues (like Thr in the β2-adrenergic receptor, or Ser in the formyl peptide receptor 2). Thus, having a long polar residue, like the glutamine, along with D7.49, is unique in the GPR18 sequence. MD trajectories reveal that Q6.43 plays an essential role in the activation of the receptor in the absence of an agonist (i.e., constitutive activation). The driving force is the presence of the amide hydrophilic polar

side chain of Q6.43 inside the hydrophobic non-polar acyl chain region of the phospholipid bilayer. As it can be seen in the trajectories, right after starting the simulation, Q6.43 tends to interact with nearby polar residues. Hydrogen bonding with D7.49 is the best and shortest way to do this. However, in the wild type system, this interaction is broken soon after the conformational change of the D7.49 $\chi 1$ from g+ to g^{-} (Figure 13A, Movie S1). This change leads to an active state bundle, which can hardly be seen in empty GPCR simulations (Movie S2). Our results suggest that the rotameric change of the D7.49 $\chi 1$ from g+ to g^{-} in the WT receptor causes the hydrogen bond between D7.49 and Q6.43 to break, which ultimately leads to a clockwise movement of TMH6 (from a top view) which breaks the weak ionic lock between R3.50/S6.33 (Figure 46).

Conversely, in the A3.39N mutant receptor, the formation of the Q6.43/D7.49 hydrogen bond causes a counter-clockwise rotation (from the top view) of TMH6. This rotation brings S6.33 close enough to R3.50 to form an “ionic lock” hydrogen bond (Figure 47, Movie S3). These results are in agreement with the experimental data that show unusually high constitutive activity for GPR18 and reduced constitutive activity for the A3.39N mutant [47].

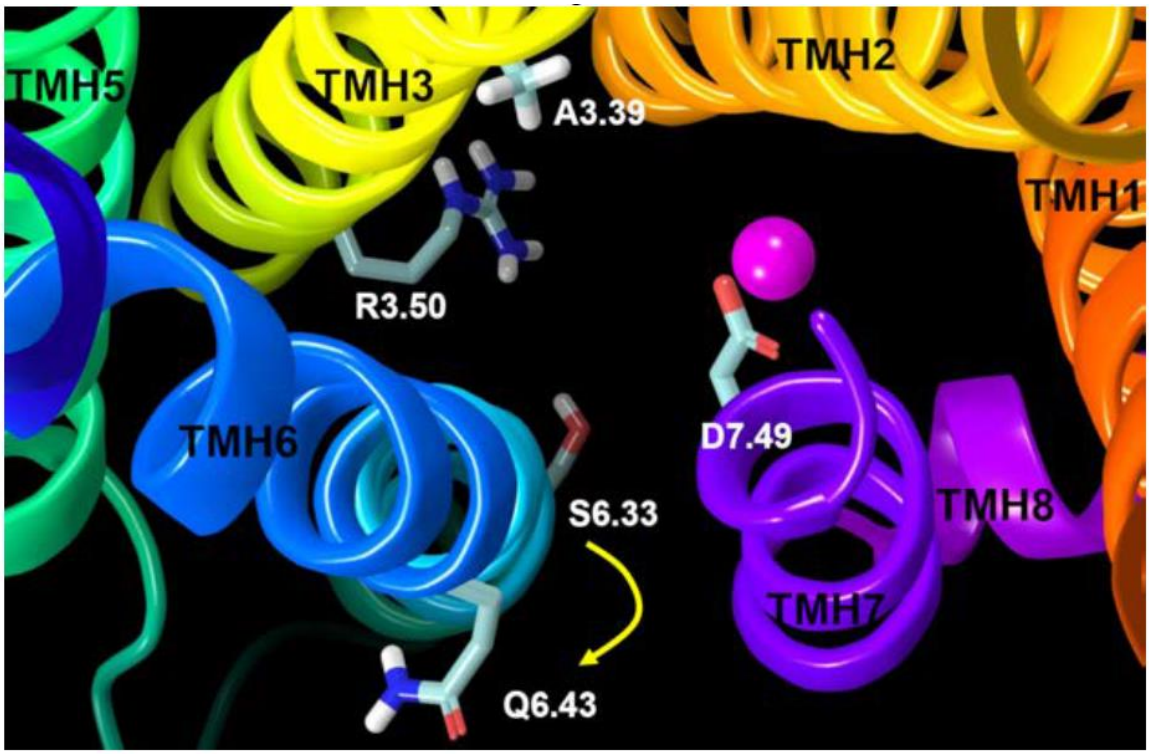


Figure 46. Configuration of the Important Residues (Highlighted in Light Blue Tubes) Involved in Our Proposed Activation Mechanism of GPR18. As D7.49 Moves Toward the Sodium Ion (Magenta Sphere) by Shifting its χ1 from g^+ to g^-, the Hydrogen Bond Between Q6.43 and D7.49 is Broken. This Leads to a Clockwise Movement (Yellow Arrow) of TMH6, Which Pulls it Away from TMH3 and Prevents the Formation of the R3.50/S6.33 "Ionic Lock" in the WT Receptor.

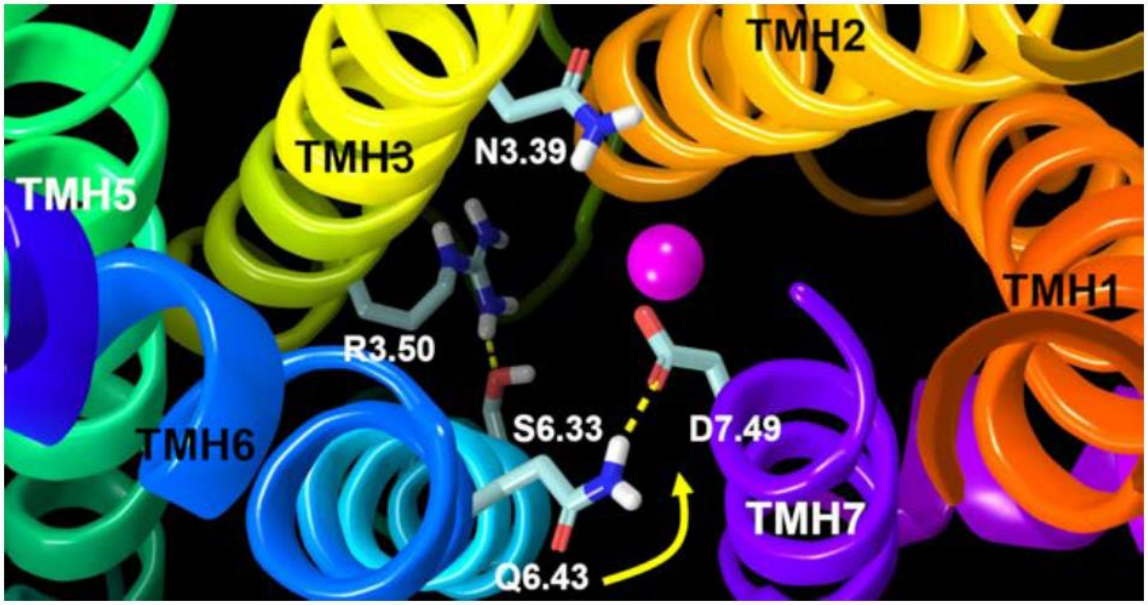


Figure 47. Configuration of the Important Residues (Highlighted in Light Blue Tubes) Involved in the Proposed Activation Mechanism in the A3.39N Mutant. Hydrogen Bonding Occurs Between Q6.43 and D7.49 (Yellow Dashed Line) Due to the Sodium Ion Position (Magenta Sphere). This is Facilitated by a Counter-clockwise Movement (Yellow Arrow) of TMH6. This Movement Facilitates the Formation of the "Ionic Lock" Hydrogen Bond Between R3.50/S6.33 (Yellow Dashed Line) in the A3.39N Mutant Receptor.

Other Structural Features That May Contribute to High Constitutive Activity of WT GPR18. Further evidence of GPR18′s tendency to activate can be observed in the activated state model we developed. The first evidence is the quality of the ionic lock in GPR18, which is formed between R3.50 and S6.33. The ionic lock keeps TMH3 and TMH6 close at their IC ends and precludes G-protein interaction. In many other class A GPCRs, this interaction is formed between two charged amino acids, R3.50 and D/E6.30. The interaction between a positively charged residue (arginine) and a negatively charged aspartic/glutamic acid is stronger than the interaction between the arginine and a polar neutral residue, like serine. Indeed, the "ionic lock" interaction in

GPR18 is a hydrogen bond. Thus, it could be expected that this interaction in GPR18 could be easily broken. The second reason is the presence of bulky residues at the right top of the ionic lock. In the current GPR18 model, two isoleucines are located at the same level on TMH3 and TMH6 in the GPR18 bundle (see Figure 48). These residues may play an important role in the movement of TMH6. Several studies have suggested that residues at the relative positions of 3.46 and 6.37 play critical roles in GPCR activation [101]. Since these two residues are positioned above the ionic lock, the contacts between them might help to break the weak R3.50/S6.33 interaction in GPR18.

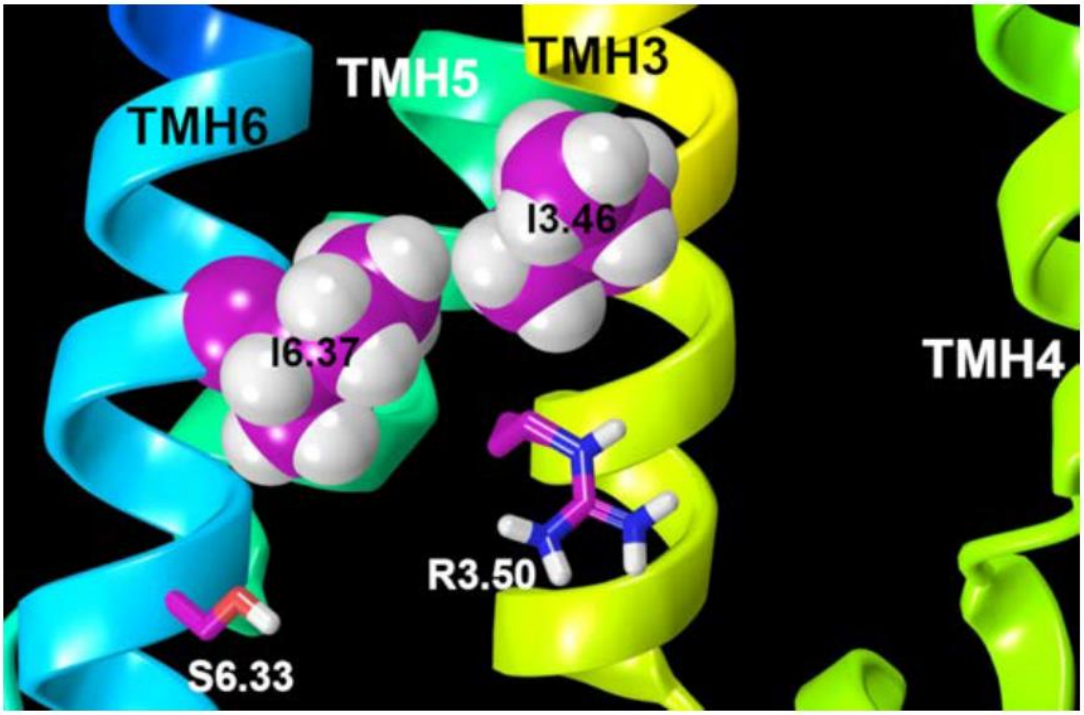


Figure 48. Presence of Two Bulky Residues, I3.46 and I6.37 (Magenta Residues in Van Der Waals), Above the Ionic Lock Residues, R3.50/S6.33 (in Magenta Tubes), Makes it More Difficult to Form a Hydrogen Bond in the WT Receptor.

Similar structural features have been previously observed in other highly constitutively active receptors. The constitutive activity of GPCRs can be intensely affected by the sodium concentration, as well as by mutations in sodium-coordinating residues. For instance, in the chemokine receptors, mutation of key conserved residues related to the sodium pocket, including 2.50, 3.39, 7.45, and 7.49, have been shown to reduce the constitutive activity [93–95]. In this way, an A3.39S mutation in GPR18 may provide a broader mechanistic perspective of this orphan receptor.

Materials and Methods

Amino Acid Numbering System

The Ballesteros–Weinstein numbering system for GPCR residues is used in this paper. This system is based on assigning the label, 0.50, to the most highly conserved amino acid in class A GPCRs for each TMH [67]. The TMH number precedes the label. In this system, for example, the most highly conserved residue in TMH4 is 4.50. The residue before this would be labeled 4.49, and the one placed immediately after this would be labeled 4.51. A residue in loops or the N- and C-termini are designated with its absolute sequence number.

Receptor Model Development

A homology model of GPR18 was constructed using the X-ray crystal structure of the delta opioid receptor (DOR) at a 1.8 A° resolution as a template (PDB identifier: 4N6H). The DOR was chosen as the template because it has 26% sequence homology (similarity) with GPR18, has an analogous sodium binding pocket region, and has a key disulfide bridge between TMH3 and the EC-2 loop. Sequence similarities include key

cysteines in the EC-2 loop (C(172) in GPR18; C(198) in DOR) that form a disulfide bridge with C3.25 on TMH3 (residues shown by the double-headed arrow in Figure 31). This disulfide bridge helps limit the conformational flexibility of the EC-2 loop and stabilizes its structure.

The GPR18 sequence was first aligned with the sequences of the Rho, CB1, CB2, DOR, β2-AR, and S1P1 receptors using the highly conserved class A GPCR residues/motifs as alignment guides (TMH 1 N1.50, TMH2 D2.50, TMH3 DRY motif from D3.49 to Y3.51, TMH4 W4.50, TMH5 P5.50, TMH 6 CWXP motif from C6.47 to P6.50 (blue residues in Figure 31), and TMH7 NPXXY motif from N7.49 to Y7.53). Based on this alignment, the DOR model was then mutated to the human GPR18 sequence. Two important motif variations in the human GPR18 sequence compared to other GPCRs in the alignment are (1) the presence of an F at position 6.48 in the TMH6 CWXP motif instead of W6.48, and (2) TMH7 motif variations with the GPR18 sequence being DVXXY vs. the common TMH7 motif, NPXXY. Also, one critical residue difference in the GPR18 sequence is the presence of a proline in TMH3 at residue 3.36. The presence of this proline may bend TMH3. To explore these possible sequence-dictated conformational deviations, we employed the combined Monte Carlo (MC) and simulated annealing (SA) technique, conformational memories (CM). CM allows an exploration of the flexibility introduced by helix bending residues, such as proline, and outputs a set of low free energy helix conformations from which a new helix conformer can be chosen.

Conformational Memories Technique for Calculating TMH Conformations

Conformational memories is a simulated annealing/Monte Carlo (MC) conformational analysis method that can be used to explore the sequence dictated range of conformations available at biological temperatures (310 K) for a TMH. Key sequence divergences between the template and GPR18 were explored using CM. This technique has been previously validated in the development of GPCR models [68–70]. In this method, the conformational consequences of helix-bending residues were extensively and effectively explored in two separate phases. In the exploratory phase, the starting temperature for the simulated annealing (SA) was 3000 K, with cooling to 310 K in 18 temperature steps. At each temperature, 50,000 MC steps were applied. With each step, two dihedral angles and one bond angle were varied in the range of ±180° for dihedrals and of ±8° for bond angles. In the biased sampling phase, CM explores only regions that were highly populated in the first phase. The starting temperature was 749.4 K, with cooling in nine temperature steps to 310 K. At each temperature, 50,000 MC steps were applied. All calculations were performed using a distance-dependent dielectric and the PARAM22 force field [71–73] without CMAP corrections [74].

Modeling of the EC and IC Loop Conformations of the Receptor and the N- and C-Termini

The DOR crystal structure has its EC-2 loop splayed out, providing water molecules with free access to the binding crevice. Since the GPR18 endogenous ligands are lipid-derived, GPR18 ligands likely enter via the lipid bilayer [102]. As a result, the EC loops and the N-terminus likely shield the receptor binding pocket from water. Therefore, the GPR18 EC loops and N-terminus were modeled so that they would shield

the binding site from water. As Figure 31 shows, the EC-2 is the longest GPR18 loop (L(159)-V(184)), and it includes C(172), which likely forms a disulfide bridge with C3.25, providing some structural stability. MODELLER 8.2 [103–105] was used to generate the N terminus (Met(1) to Pro(18)) and the IC and EC loops (IC-1: Thr(51) to Thr(55), IC-2: Pro(126) to Lys(133), IC-3: His(216) to Val(226), EC-1: Glu(86) to Gly(90), EC-2: Leu(159) to Val(184), EC-3: Thr(259) to Tyr(264)). A limited distance between Cys(172) and Cys3.25 was placed in the EC-2 loop calculation so that the highly conserved disulfide bridge was preserved. The final loops were energy minimized to a gradient of 0.05 kJ/mol.Å with 2500 maximum iterations using the generalized born/surface area (GB/SA) continuum solvation model for water and the OPLS3 [81–83] force field in MacroModel 10.7.014 (Schrödinger Release 2017-4: MacroModel, Schrödinger, LLC, New York, NY, USA, 2017), while the TMH region was frozen.

Receptor Minimization

The resultant model was energy minimized using the protocol previously reported in our lab [45]. The minimization protocol used the OPLS3 [81–83] all-atom force field in MacroModel (Schrödinger 2017-4: MacroModel, Schrödinger, LLC, New York, NY, USA, 2017). An extended Van der Waals cutoff (8.0 Å, updated every 10 steps), a 20.0 Å electrostatic cutoff, and a 4.0 Å hydrogen bond cutoff were used in each stage of the calculation.

Molecular Dynamics (MD)

The wild type and mutant GPR18 models were oriented with respect to the membrane using the template, 4n6h, from the Orientation of Proteins in Membranes

database [107] and placed in a hydrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer using the CHARMM-GUI server [108,109]. The resulting system contained 155 POPC molecules, 15,177 water molecules, and an ionic strength of 0.15 M NaCl. The resulting simulation cell contained 71,363 atoms and was 80.3 Å x 80.3 Å 119.5 Å. All molecular dynamics runs employed the all-atom additive CHARMM36m force field for proteins [110], and CHARMM36 for lipid and ions [72]. The system was energy minimized and relaxed using the multistep schedule as described previously [109]. After this relaxation, 700 ns of unbiased MD was performed for each system. The first 100 ns of which were considered as equilibration (see Figure 34).

All MD simulations were performed with the GPU version of AMBER 16 [73,111,112] in the semi-isotropic NPT ensemble using a Langevin thermostat (manufacturer, city, state abbreviation if USA or Canada, country) (T = 310 K, using a friction coefficient of $1ps^{-1}$) and a Monte Carlo barostat (manufacturer, city, state abbreviation if USA or Canada, country) (P = 1 bar) as implemented in AMBER 16. Long range electrostatics were treated using PME (particle mesh ewald), and the Van der Waal force switching method was applied starting at 10 Å with a cutoff of 12 Å. High-frequency bonds to hydrogen were constrained using the SHAKE method, allowing the use of a 2-fs integration time step.

Finally, a sodium ion was positioned in the putative sodium ion binding site. This site includes the polar residues, N1.50, D2.50, S3.39, N7.49, and Y7.53, and directly coordinates a sodium ion to stabilize the inactive state of the receptor.

RMSD plots for the TMH bundle versus time were used to analyze the bundle geometry stability in the bilayer.

Conclusions

Our results suggest that the high constitutive activity of GPR18 is sequence-driven and dependent upon sodium binding pocket effects. To date, there has not been a systematic, comprehensive, and extended mutagenesis study of GPR18. Hopefully, by identifying the critical residues involved in the activation of GPR18, our model can aid future drug design studies. Furthermore, we suggest that ligand development for GPR18 should concentrate on the development of inverse agonists/antagonists. It is possible that the sequence-driven, high constitutive activity of GPR18 may preclude the need for agonist development.

Supplementary Materials

Supplementary materials (Movies S1, S2, and S3 in .avi format) can be found at www.mdpi.com/xxx/s1.

Author Contributions

Conceptualization, N.S. and P.H.R.; Data curation, N.S. and D.P.H.; Formal analysis, N.S., P.M. and D.P.H.; Investigation, N.S.; Methodology, N.S., P.M., D.L. and P.H.R.; Supervision, P.H.R.; Validation, N.S.; Visualization, D.P.H.; Writing—original draft, N.S., P.M.; Writing—review and editing, P.H.R.

Funding

This research was supported by a grant from the National Institutes on Drug Abuse KO5 DA021358 to PHR.

Acknowledgments

P.M. would like to thank CAM for her postdoctoral fellowship “Atracción de Talento” 2018-T2/BMD-10819.

Conflicts of Interest

The authors declare no conflict of interest.

Abbreviations

1.	GPCR	2.	G-Protein Coupled Receptor
3.	NAGly	4.	N-arachidonoyl glycine
5.	POPC	6.	Palmitoyloleoylphosphatidylcholine
7.	CM	8.	Conformational Memories
9.	SA	10.	Simulated Annealing
11.	TMH	12.	Transmembrane helix
13.	DOR	14.	delta opioid receptor
15.	MD	16.	Molecular dynamics
17.	RMSD	18.	root-mean-square deviation
19.	EC	20.	Extracellular
21.	IC	22.	Intracellular
23.	WT	24.	Wild type
25.	AMBER	26.	Assisted Model Building with Energy Refinement

CHAPTER V

PARAMETERIZATION OF GPR18 PROPOSED ANTAGONIST S4

Introduction

Since there are no experimental structural techniques data (e.g., x-ray crystallography, electron microscopy, nuclear magnetic resonance) concerning our proposed antagonist S4, the best and only tools for an in-depth dynamic study at an atomic resolution is to use Molecular dynamics (MD) simulations. The CHARMM potential energy function (Eq. (1), the so-called molecular mechanics (MM) force field) is the most commonly used force field for MD simulations.

$$
\begin{aligned}
E = & \sum_{bonds} K_b (b - b_0)^2 \\
& + \sum_{angles} K_\theta (\theta - \theta_0)^2 \\
& + \sum_{Urey-Bradley} K_{b^{1-3}} (b^{1-3} - b_0^{1-3})^2 \\
& + \sum_{dihedrals} K_\emptyset [1 + \cos(n\emptyset - \delta)] + \sum_{improper} K_\varphi (\varphi - \varphi_0)^2 \\
& + \sum_{nonbonded} \left\{ \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\in r_{ij}} \right\}
\end{aligned}
$$

The internal (bonded) part of the equation includes potential energy functions for bonds (b), angles (θ), Urey-Bradley ($b^{1\text{-}3}$), dihedrals (ϕ), and improper dihedral (φ). Subscript zero corresponds to the equilibrium value for each, and the K's are the corresponding force constants. The sinusoidal energy term also incorporates periodicity (n) and phase shift (δ). External (nonbonded) energy terms include Lennard-Jones 6-12 potential and Coulombic potential for electrostatic interactions. In the nonbonded energy terms ε_{ij} is the well depth calculated as the geometric mean ($\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$), $R_{min,ij}$ is the arithmetic mean, $(R_{min,ij} = (r_{min,i} + r_{min,j})/2)$, q_i, q_j are the partial atomic charges. After choosing a particular force field, the force parameters must be determined. These parameters can be produced from experimental data. Although there is a considerable amount of experimental data for some simple molecules such as alkanes, as discussed in the following sections, the CHARMM force field lacks a few new parameters (as well as dihedral and charge parameterization) that have been introduced only in our novel targeted molecule and explicitly are not covered in the current available force fields. The process of defining the new parameters and addition of them to the earlier force fields is very challenging and takes a long time; therefore, the inability to generate required data rapidly in complex large molecules could inhibit researchers in the application of MD and severely limit studies to obtain the most realistic and accurate results from ligand-receptor binding simulations. Since parameterization is a critical step before starting MD simulations and because of the lack of entirely empirical values to provide satisfactory results, researchers need another technique to resolve the failure of MM force fields.

In contrast to Molecular Mechanics, which use potential functions that are derived from classical physics to model molecular systems, Quantum mechanical (QM) models calculate precisely various properties of the electrons using Schroëdinger's equation,

$$H\Psi = E\Psi \qquad \text{(Eq. 2)}$$

where

Ψ = psi, the wave function

E = the energy of the system

H = The "Hamiltonian," a mathematical operator.

The Hamiltonian equation is shown below:

$$H^{elec} = -\frac{1}{2}\sum_i^{elec}\left(\frac{\partial^2}{\partial x_i^2}+\frac{\partial^2}{\partial y_i^2}+\frac{\partial^2}{\partial z_i^2}\right) - \sum_i^{elec}\sum_I^{nuclei}\left(\frac{Z_I}{|R_I - r_i|}\right) + \sum_i^{elec}\sum_{j<i}^{elec}\left(\frac{1}{|r_i - r_j|}\right) + \sum_I^{elec}\sum_{J<I}^{nuclei}\left(\frac{Z_I Z_J}{|R_i - R_j|}\right) \qquad \text{(Eq. 3)}$$

As can be concluded from Equations 1, 2, and 3 in the quantum mechanical models, electrons are not particles (unlike Bohr's model in classical physics) but have wavelike characteristics. Hence, these assumptions make QM calculations more accurate and also complicated and expensive than MM equations (compare Equation 1 with 2 and 3) from a computational perspective and in turn make it reliable, suitable data to refer to when MM calculations fail to provide the required data.

The Force Field Toolkit (ffTK) [113] is a plugin for the VMD molecular modeling software package explicitly designed to develop CHARMM36 parameters including charges, bonds, angles, and dihedrals for small or large molecules directly from *ab* initio calculations (a computational chemistry method based on quantum chemistry).

The ffTK parameterization method philosophy is to find the best fit between MM simulations and QM data at a modest level of theory HF/6-31G* (for charge optimization) and MP2/6-31G* level of theory (for bond, angle, and dihedral optimization) using mathematical optimization algorithms like “Simulated annealing” or “downhill.” The ffTK supports the use of Gaussian [114] software package for these calculations and compared with the previously suggested methods [115] ffTK readily performs many difficult and time-consuming steps such as missing force field parameters, creates Gaussian input files for the required *ab initio* calculations, reads the data from the Gaussian output files, and finally convergence of data between QM and MM. Since *ab initio* calculation is exceptionally computationally expensive depending on the size of the molecule, the trick is to split large molecules into several small fragments to ease parameterization then join them together. In this chapter, we first describe why and how we use ffTK to parameterize our proposed antagonist S4. Then, we apply the newly generated parameter for running MD simulations of the ligand in the GPR18’s inactive state model to describe the resulting essential interactions.

Why S4 Needs to be Parameterized

Before starting any parameterization step, at the first we need to make sure if our component requires a full parameterization or not. In this way, to identify possible minimum energy structures, a conformational search using the Spartan 08 molecular modeling program (Wavefunction, Inc., Irvine, CA) [116] was performed at the MP2/6-31G* level of theory. The resulting S4 global minimum was minimized by CHARMM to evaluate what dihedrals and angles need to be developed or parameterized.

As depicted in Figure 49, it seems that the resorcinol ring and carbon chain tail is completely fitted in the CHARMM minimized structure (S4 in orange Figure 49) with the QM global minima (S4 in grey Figure 49). The only unpleasant deviations from QM global minima are around the part of the pyrazole ring and phenyl ring. Table 2 shows the detailed investigation.

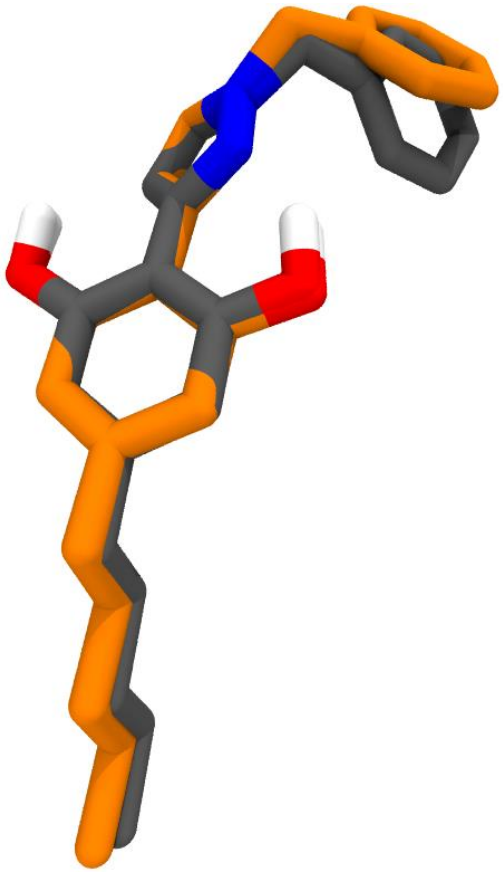


Figure 49. Superposition of the S4 Global Minima at MP2/6-31G* Level of Theory (Grey Structure) with CHARMM36 Minimized Structure (Orange) Before Parameterization.

Table 2

Comparison of Angles and Dihedrals between QM and MM

Dihedrals	**MP2**	**CHARMM**		**Δ < 8°**	
H13-C12-C5-C4	-105.3°	-137.2°	-106.2	31.9°	**0.9°**
H13-C12-C5-C6	73.1°	43.2°	69.4	29.9°	**3.7°**
H14-C12-C5-C6	-165.9°	157.0°	-170.2	37.1°	**4.3°**
H14-C12-C5-C4	15.7°	-23.4°	17.2	39.1°	**1.5°**
N15-C12-C5-C6	-45.4°	-80.1°	-46.4	34.7°	**1.0°**
N15-C12-C5-C4	136.2°	99.5°	132.3	36.7°	**3.9°**
Angles	**MP2**	**CHARMM**		**Δ < 3°**	
N15-C12-C5	111.7°	116.0°	113.3°	4.3°	**1.6°**
H13-C12-C5	112.0°	107.0°	111.4°	5.0°	**0.6°**
H13-C12-N15	105.8°	110.1°	103.8°	4.3°	**3.0°**

Note. QM is the column under MP2. The column under CHARMM shows the MM values before parameterization (red) and after parameterization (yellow). The column under Δ shows the differences of QM/MM before parameterization in red and after parameterization in yellow. Δ< 3° and Δ< 8° are the allowed criteria for angle and dihedral differences, respectively.

Previous reports' acceptance criteria for the bond, angle, and dihedral measurements are within 0.3Å, 3°, or 8°, respectively, of their QM values [117]. As Figure 51 and the angle differences probing in Table 2 show, the angle N15 C12 C5 plays a vital role in the deviation of MM from the QM global minima. The dihedrals differences investigation in Table 2 also shows that the deviated dihedrals are similar torsions around the C12 C5 bond, but with the different names. Thus, we guess that by parameterization of one angle and one dihedral, MM would fit the QM structure.

The previously famous published CHARMM General Force Field (CGenFF) database [118,119] for CHARMM-compatible parameters was used as the best resource

for obtaining the initial parameters here, as it includes hundreds of generated topology and parameter files of organic compounds. Since the CGenFF parameterization method is based on analogy, care must be taken in using those parameters together with the corresponding penalty scores due to the generalization of parameters. Penalties below 10 mean that the analogy is fair; penalties between 10 and 50 mean that validation of parameters is required and penalties over 50 mean that parameter optimization is required. For instance, in our case, after the submission of the S4 structure to www.paramchem.org, the output is shown in Figure 50. The output of ParamChem online tools reveals that what angles and dihedrals are new and missed in the existing general force field parameter file. It means that all of the missed angles and dihedrals parameters must be generated. For the starting point, ParamChem uses an analogy-based analysis and suggests an initial parameter from a similar (not identical) motif, as shown in Figure 52. This initial value might be an excellent or imperfect analogy; the penalties scores indicated its quality. However, as mentioned before, one should be careful about the analogy analysis, and each of the suggested values should be double-checked before using. Taken together, as Figure 50 and Table 2 show, our selected angle N15 C12 C5 and dihedral N15 C12 C5 C4 have high penalties, which is an indication that their proposed parameters are not meant to be used for production simulations and must be parameterized.

```
read param card flex append
* Parameters generated by analogy by
* CHARMM General Force Field (CGenFF) program version 1.0.0
! Penalties lower than 10 indicate the analogy is fair; penalties between 10
! and 50 mean some basic validation is recommended; penalties higher than
! 50 indicate poor analogy and mandate extensive validation/optimization.
ANGLES
CG2R51 CG2R52 CG2R61   45.80    130.00 ! M0071 , from CG2R51 CG2R51 CG321, penalty= 76
CG2R61 CG2R52 NG2R50   45.80    120.00 ! M0071 , from CG321 CG2R51 NG2R50, penalty= 76
CG2R52 CG2R61 CG2R61   36.00    120.00 ! M0071 , from CG2DC1 CG2R61 CG2R61, penalty= 22.5
CG2R61 CG321  NG2R51   53.00    115.20 ! M0071 , from CG2R61 CG321 NG2S1, penalty= 11
DIHEDRALS
CG2R51 CG2R51 CG2R52 CG2R61     2.0000  2   180.00 ! M0071 , from CG2R51 CG2R51 CG2R51 CG2RC0, penalty= 52.5
HGR51  CG2R51 CG2R52 CG2R61     2.8000  2   180.00 ! M0071 , from CG2RC0 CG2R51 CG2R51 HGR51, penalty= 52.5
CG2R51 CG2R52 CG2R61 CG2R61     1.0000  2   180.00 ! M0071 , from NG2S1 CG2O1 CG2R61 CG2R61, penalty= 228
NG2R50 CG2R52 CG2R61 CG2R61     1.0000  2   180.00 ! M0071 , from NG2S1 CG2O1 CG2R61 CG2R61, penalty= 201
CG2R61 CG2R52 NG2R50 NG2R51    13.5000  2   180.00 ! M0071 , from CG2RC0 CG2R52 NG2R50 NG2R51, penalty= 47.5
CG2R52 CG2R61 CG2R61 CG2R61     3.1000  2   180.00 ! M0071 , from CG2DC1 CG2R61 CG2R61 CG2R61, penalty= 22.5
CG2R52 CG2R61 CG2R61 OG311      3.1000  2   180.00 ! M0071 , from CG2R61 CG2R61 CG2R61 OG311, penalty= 28.5
CG2R61 CG2R61 CG321  NG2R51     0.1190  2     0.00 ! M0071 , from CG2R61 CG2R61 CG321 NG2S1, penalty= 11
CG2R61 CG321  NG2R51 CG2R51     0.4200  2     0.00 ! M0071 , from CG331 CG321 NG2R51 CG2R51, penalty= 71.9
CG2R61 CG321  NG2R51 NG2R50     0.4200  2     0.00 ! M0071 , from CG331 CG321 NG2R51 NG2R50, penalty= 71.9
CG2R52 NG2R50 NG2R51 CG321      2.5000  2   180.00 ! M0071 , from NG2R50 NG2R50 NG2R51 CG321, penalty= 53
END
```

Figure 50. The Output .str File of Submission of S4 to Paramchem Online Database. Yellow Sections Highlighted Important High Penalty New Angles and Dihedrals.

Parameter Optimization of Components by ffTK
Optimization of Bonds and Angles

Since ffTK only works with Gaussian files, the first required step of parameterization by ffTK is to run a geometry optimization by a Gaussian software package at MP2/6-31G* level of theory to find the optimum conformation of S4. However, before building the structure, we needed to decide about the fragment that included the targeted angles and dihedrals. Since our targeted dihedral and angle are localized between pyrazole and phenyl rings, as shown in Figure 51, we built a fragment of S4 of which the aliphatic chain was removed.

After finding the optimum conformation of the S4 fragment, the resulting low-energy conformation was used as the input conformation for the bond, angle, and in the next step for dihedral parameterizations. The approach employed by ffTK for bonds and angles parameterization in the target molecule is to derive a Potential Energy Distribution

(PED) from the MM Hessian calculated (a matrix containing the second derivative of the potential energy as function of the input coordinates) using the trial parameters, and compare it with the corresponding QM Hessian calculated PED at MP2/6-31G* level of theory [113]. This step is very straightforward, and after Hessian calculation and updating the parameter file, the new bond and angle parameters automatically would be added to the parameter file.

Optimization of Dihedrals

The first stage in dihedral optimization in ffTK is to generate a dihedral scan for the un-parameterized targeted dihedral. To specify the targeted dihedral manually, one can visualize the selected dihedral in the VMD main window (see Figure 51). The scanning strategy in ffTK, as shown in Figure 52, was used to generate two Gaussian input files.

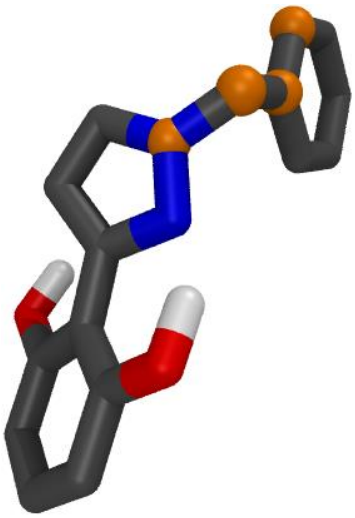


Figure 51. The Fragment Built from S4. The Orange Balls Show the Selected Dihedral Atoms for Parameterization.

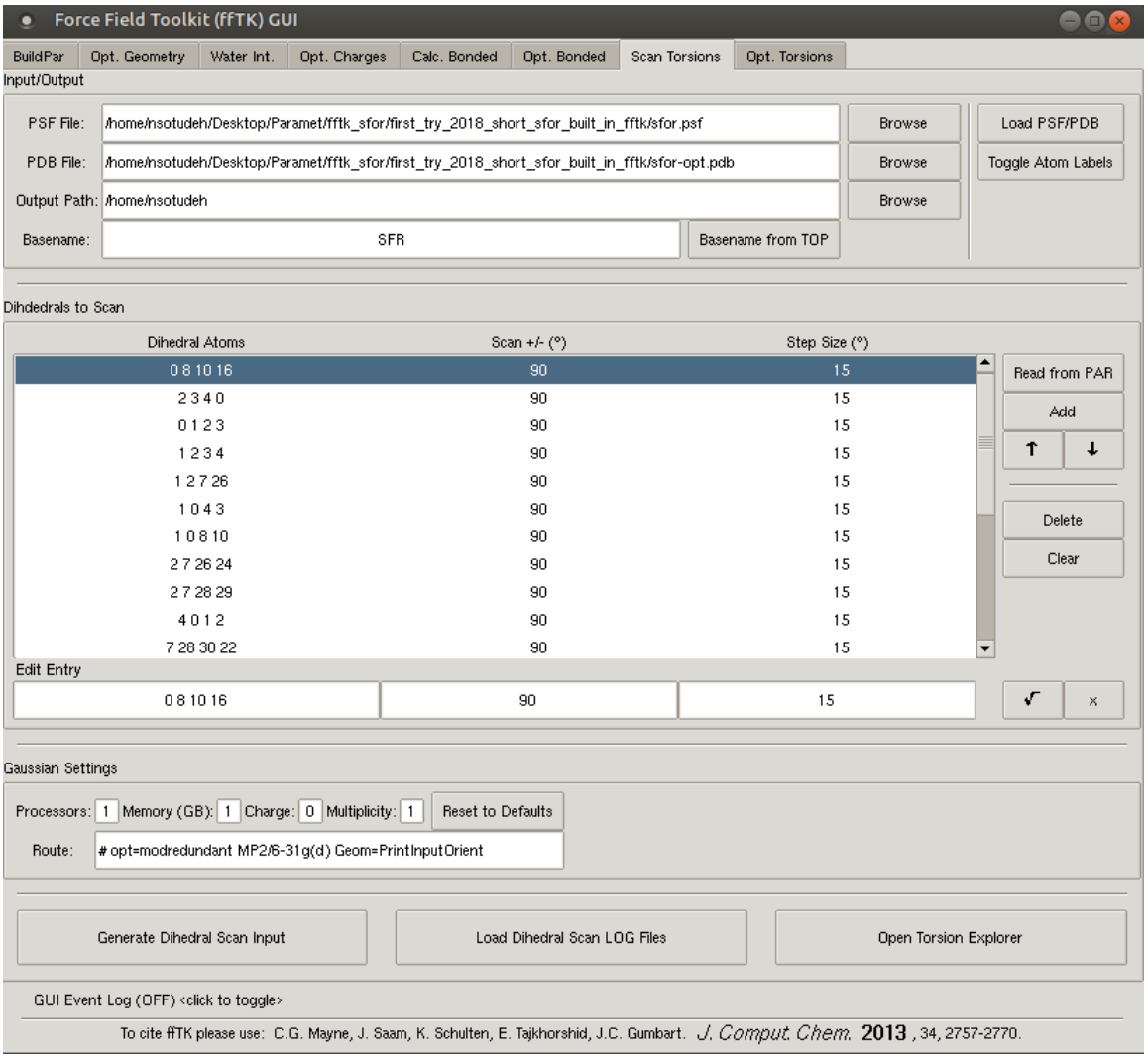


Figure 52. Dihedral Scanning Step. After Loading the Topology and In-progress Parameter File, By Choosing the Targeted Dihedral with the Favorite Degree and Step Size of Scanning, While VMD Main Window Shows the Picture of Selected Dihedral as Shown in Figure 51. Then, ffTK Will Generate the Dihedral Scan Gaussian Input Files.

Taking the initial value of the aimed dihedral angle from optimized geometry and then evenly generation ±15° (or ±5° is the optional user-chosen step size) about it until ending at the chosen scan range. For instance, the dihedral angle N15 C12 C5 C4 with the starting value of 136.2° has been specified. The ffTK Scan Torsions tab would produce

two Gaussian files. The positive scanning file begins at 136.2° and after 18 steps (each step size 5.0° is chosen to cover 90° scanning range which is enough for all of the conformations about this dihedral) ends at -46.2°. In the same pattern, the negative Gaussian input file starts at 136.2° and ends at 46.2°. The next stage in dihedral optimization is to enter all of the scanned dihedrals. In this step, to minimize the difference between QM and MM, we can make use of the simulated annealing method [120]. This method has some advantages to the other method, which is called the downhill Simplex method [121, 122]. For instance, simulated annealing is faster than downhill, and it does not require many function evaluations. In the downhill method, it is possible to miss global minima in favor of local minima. Both methods run multiple cycles for convergence. Hence, we first use simulated annealing to explore parameters effectively and rapidly arrive at convergence; following simulated annealing, the downhill Simplex optimization method is used for the final optimization.

Based on Eq. 1, three variables of Force constant (k), Periodicity (n), and Phase Shift (δ) must be identified to parameterize a dihedral. Between those values, ffTK only optimizes force constant (k) and periodicity (n), leaving phase shift undetermined. The (δ) parameter can take any value from the CHARMM potential energy function; however, in most cases, they are 0° or 180°. The selection of the in-progress parameter file adds all of the missing parameters, along with the earlier parameterized dihedrals which can be found in CHARMM36 parameter files. For example, NG2R51 CG321 CG2R61 CG2R61 is the targeted dihedral that we aim to parameterize. However, there is also the NG2R50 NG2R51 CG321 HGA2 dihedral that already been parameterized, and

its parameters in the par_all36_cgenff.prm file (CGenFF: Parameters for the CHARMM General Force Field v. 3.0.1 for Small Molecule Drug Design) are 1.010, 3, and 180.00 for k, n, and δ, respectively (see Figure 53).

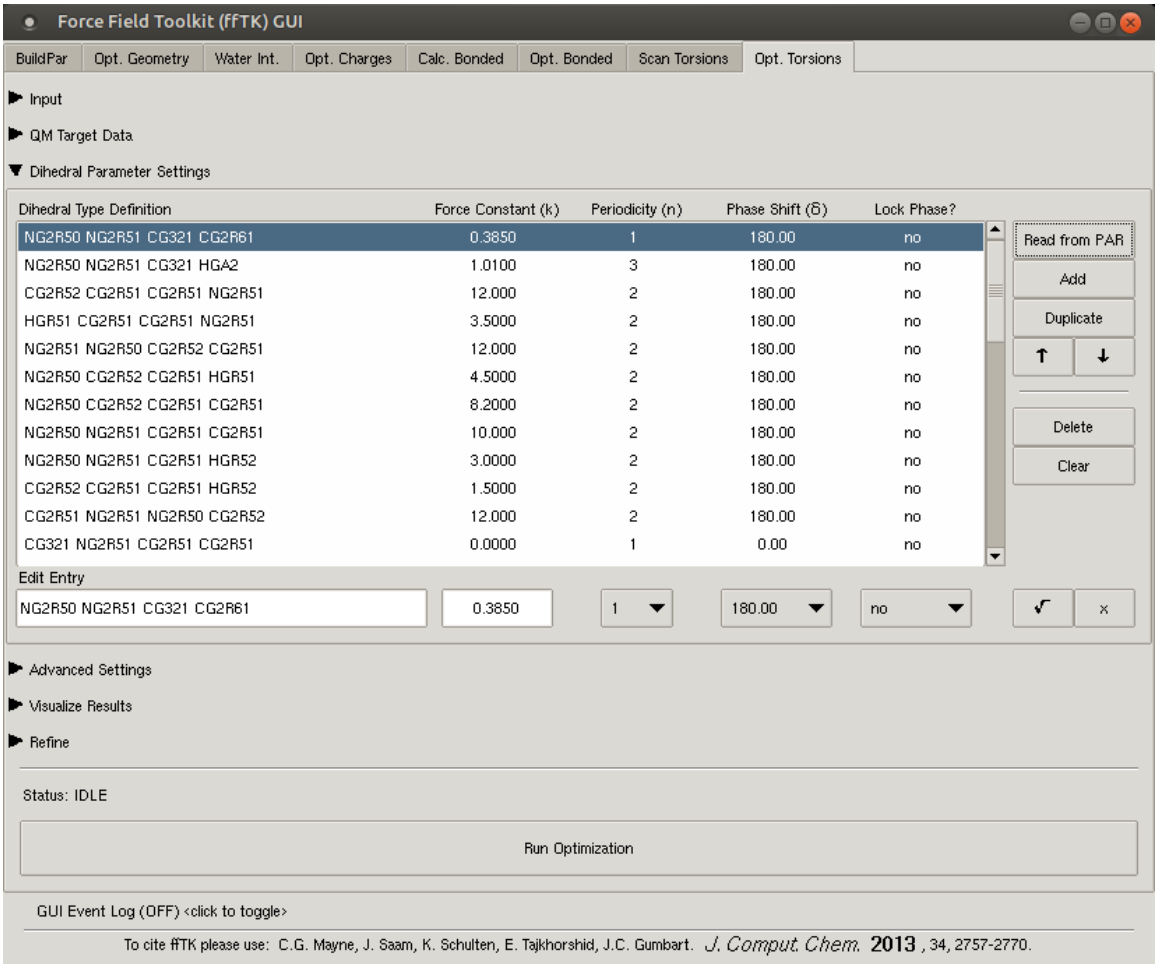


Figure 53. Dihedral Parametrization Step. After Loading the Topology and In-progress Parameter File, By Choosing the Targeted Dihedral, the Favorite Force Constant (k), Periodicity (n) and Phase Shift (δ) Can Be Entered Manually. After Checking the Comparison Graphs in “Visualize Results” Tab the Parametrization Values Can Be Modified in “Refine” Tab.

We can use the initial values prepared by ParamChem online database or par_all36_cgenff.prm file, but it is not necessary because all of those values are

manipulated during the refinement. The current author believes that this is a disadvantage in using ffTK because we prefer to keep fixed the existing parameters of unwanted dihedrals (provided in par_all36_cgenff.prm like NG2R50 NG2R51 CG321 HGA2) and only optimize the targeted new defined dihedrals (such as NG2R51 CG321 CG2R61 CG2R61), but it is impossible by the current version of ffTK plugin. Since, the change in unwanted dihedrals force constant is inevitable, we have tracked their changes to avoid a significant difference (with the values in par_all36_cgenff.prm) during refinement steps.

After each step of the refinement, a plot is generated to visualize QM energy and the corresponding MM energy for each optimized conformation, which makes it easier to compare their differences. The refinement process will continue as long as RMSE value, or the shape of the curves is not yet correct (see Figure 54). Some strategies help us to improve the best fit of the MM and QM curves. In this way, multiplicity and periodicity settings changes are the most effective techniques. For example, the multiplicity of a dihedral can be added (see Figure 53). By pressing the Add button, removed and modified during the refinement stage because it is not crucial to put all of the terms correctly at this step. When the developed parameters are in good agreement between the shape of the MM and QM curves, a final refinement was performed by switching to downhill mode with a lower tolerance yielded a rapid converge and improved acceptable fit. The final optimized dihedral parameters data were written to an ffTK log file, and then by using the BuildPar tab (see Figure 52 and 53), the parameter file was updated.

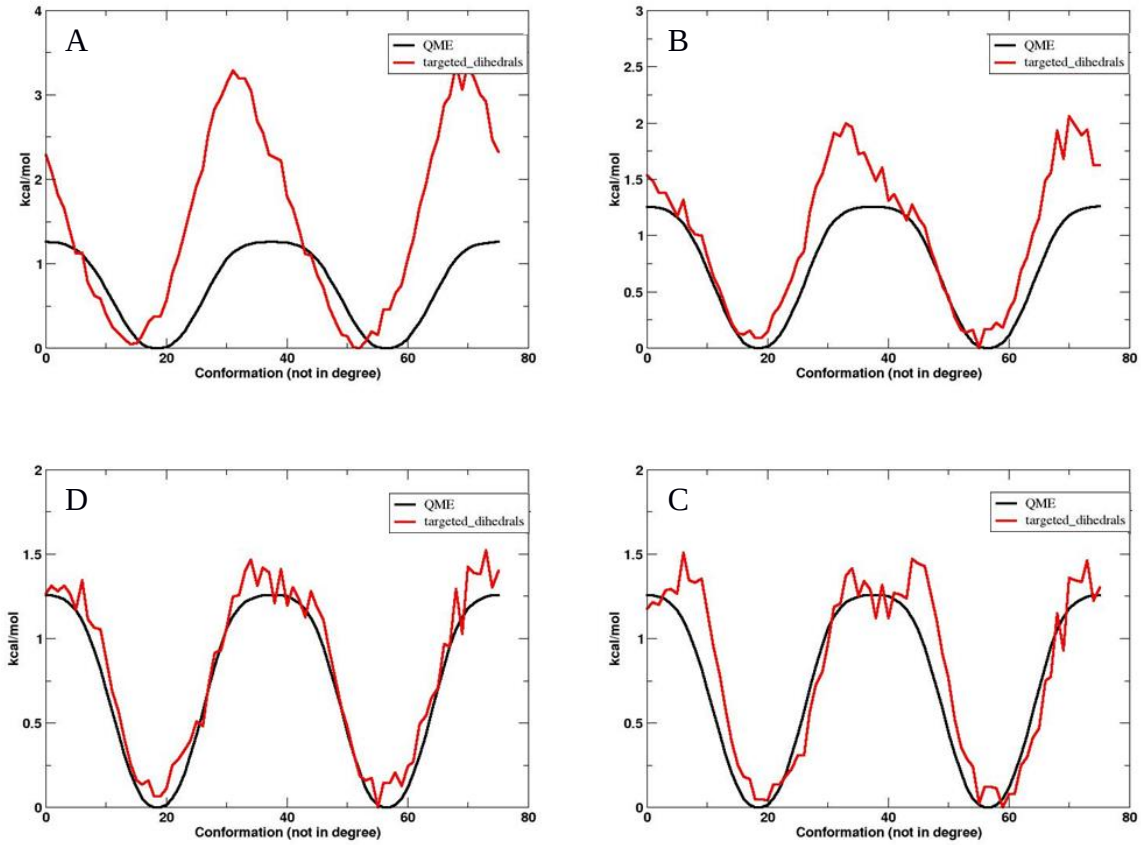


Figure 54. Energy-Conformation Plots; the Steps Need to be Taken Before Getting Into the Best Final Fit. The Curves in Red Show MM Values While the Constant Black Curves Show QM Values.

As shown in Figure 55, overlaying of the S4 structures after parameterization by CHARMM (green) and QM global minima (grey) has a stronger match than the un-parameterized S4 structure (orange). Also, comparing the measured dihedral and angles after parametrization with the QM values in Table 2 (the white column) confirms that all of them are within the accepted range of 3° for angles and 8° for dihedrals. The corresponding potential energy curves in Figure 56 also confirm this strong fit.

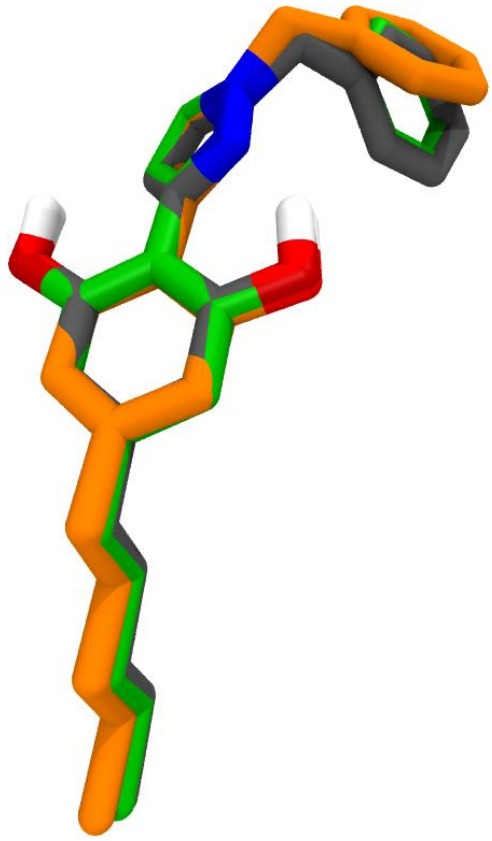
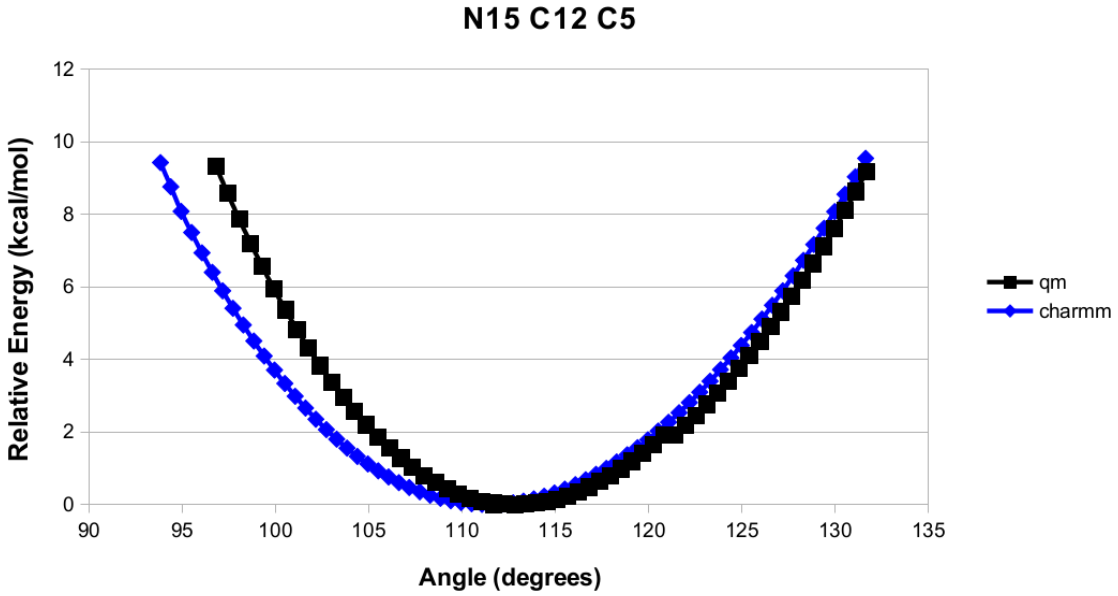


Figure 55. Superposition of the S4 Global Minima at MP2/6-31G* Level of Theory (Grey Structure) with CHARMM Minimized Structure Before Parameterization (Orange) and After Parameterization (Green).



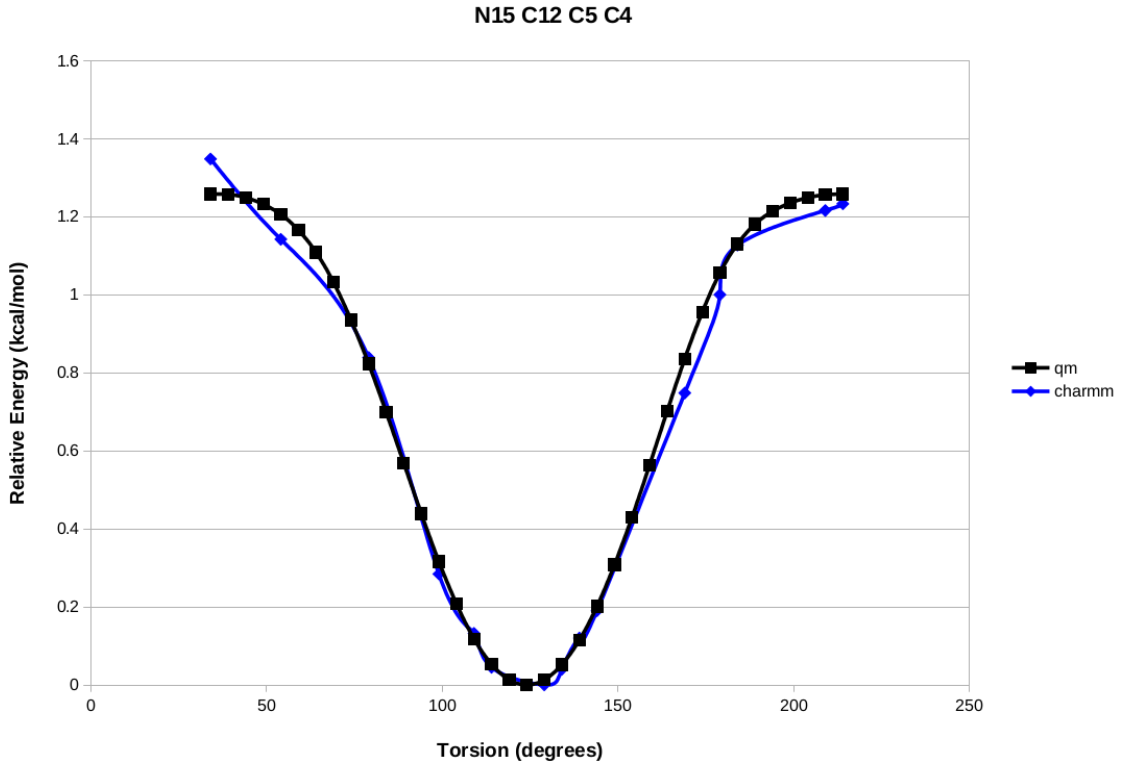


Figure 56. Potential Energy Scan for Targeted Angle N15 C12 C5 and Dihedral N15 C12 C5 C4. QM Curve in Black and MM Curve is in Blue.

MD Simulation

To investigate the quality of the S4 docking job in the static model, all-atom MD simulations were performed in a POPC bilayer. Since in this simulation we are not looking for the pathway of the ligand entry from the membrane bilayer to the GPR18, we used precisely the model shown in Figure 20 with no S4 molecule added to the leaflets. The resulting simulation contains 74721 atoms, and the simulation cell (80.1 Å × 80.1 Å ×126.1Å, Z-axis normal to the plane of the bilayer) was built using the freely available CHARMM-GUI input generator. The resulting system contains 153 POPC molecules (77/76 lipids in the upper/lower leaflet), solvated with 16347 water molecules and was neutralized at 0.15M NaCl. Standard CHARMM-GUI minimization and equilibration steps were performed, followed by a higher than 120ns production run. Isotropic (NPT) ensemble was performed using the Langevin thermostat (T=310K) and Monte Carlo barostat (P=1atm) with the GPU-enabled AMBER16 molecular dynamics package using the all-atom CHARMM36m force field for proteins and CHARMM36 for lipids and ions.

However, in the early frames, hydrogen bonding with R5.42 and cation-pi interaction with R2.60 both were formed (as shown in Figure 20), but after around 80 ns the interaction with R2.60 is broken and the phenyl ring of the S4 tends to rotate so that places the right top of the M7.42, which forces it to form Met-aromatic interaction with F6.48 (see Figure 57). On the other hand, a new stable hydrogen bonding is formed with Y180 in the EC2 loop, which is very stable and can be seen in most of the frames of the trajectory. The comparison between the results of MD in Figure 57 and the previous docking job in Figure 20 show that the only important difference which needs to be

considered is the rotation of the phenyl ring in the top of the S4. Since the structure was used for docking of S4 in the static R state model was the global minima that produced from an extensive conformational search at HF/6-31G* level of theory by the Spartan 08 molecular modeling program (Wavefunction, Inc., Irvine, CA) [116], hence the question is how much energy the current conformation of S4 (as shown in Figure 57) costs. Our investigation shows that the relative energy of the adopted conformation of S4 in MD simulation is 0.0 kcal, which means that this conformation is a global minimum of S4 (see Figure 58). The reason why S4 adopts a conformation like what is shown in Figure 57 is the importance of toggle switch residue interaction. As mentioned before in Chapter III, in GPR18 two critical residues, M7.42 and F6.48, by the formation of a Methionine Pi interaction, play the role of the toggle switch. In this way the toggle F6.48 cannot change its conformation from gauche plus to trans. The role of an antagonist is to keep F6.48 in its gauche plus conformation and as MD results show, S4 is doing this job pretty well. On the other hand, investigations of the ionic lock interaction between R3.50 and S6.33 show that this interaction is stable and strong. This also is consistent with the role of S4 as an antagonist (see Figure 59).

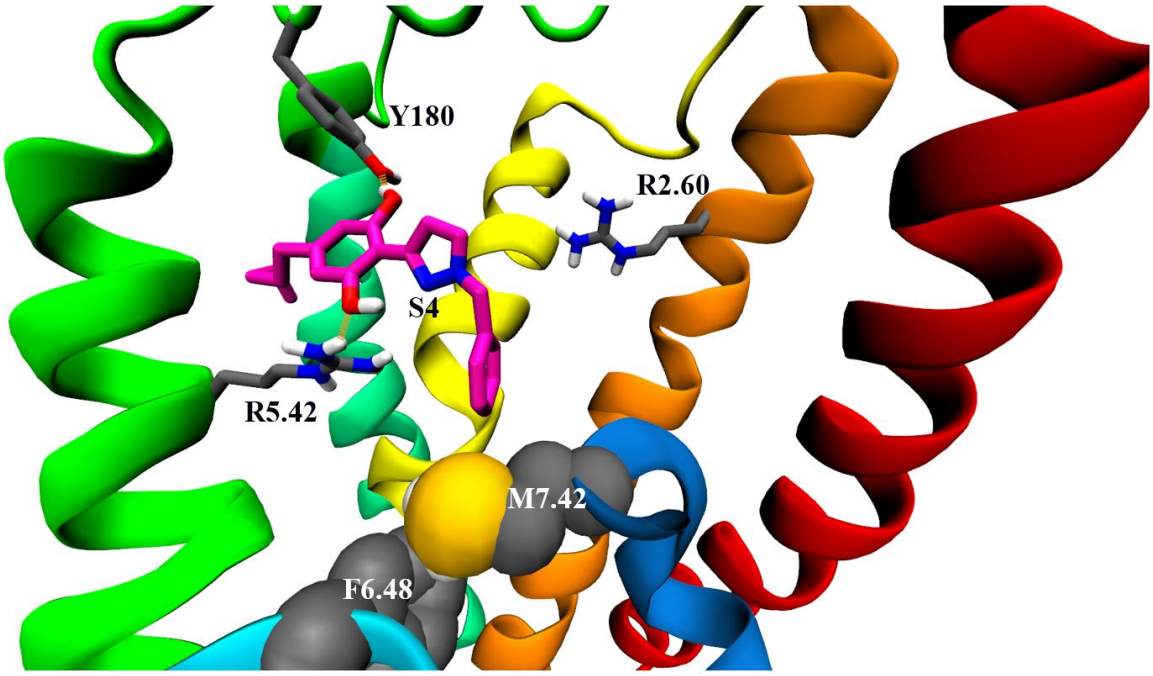


Figure 57. MD Simulations of Parameterized S4 Inside GPR18 R State Model. S4 is Hydrogen Bonding (Orange Dashed Line) with R5.42 and Y180 in EC2 Loop. Toggle Switch Residues (M7.42 and F6.48) are Contoured at Their VdW Radius in Brown Color. Color Code for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (in Cyan) and TMH7 (Dark Blue). 2D Structure of PSB-CB5 Inserted in Far Left of the above Picture Beside TMH1.

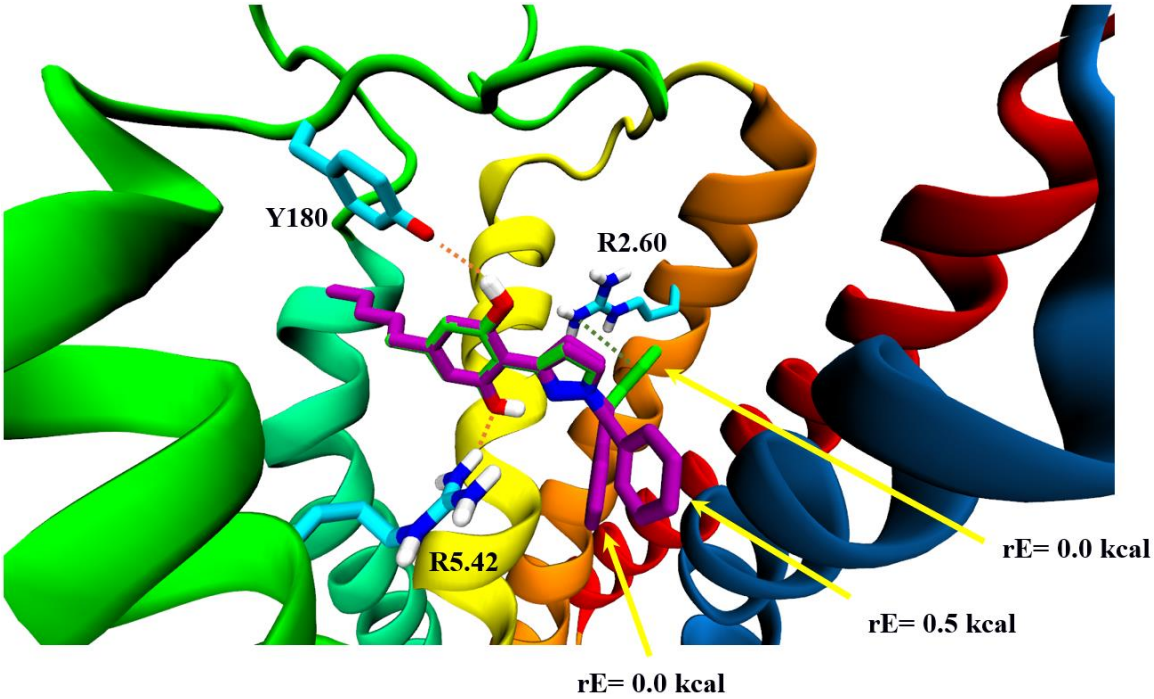


Figure 58. Comparison of Three Different Conformation of S4 Inside the Binding Site of GPR18. Their Relative Energy Values Inserted Beside Each of Them. The Structure in Green Was Used for Docking Job and is Able to Form Cation Pi Interaction (Green Dashed Line) with R2.60. The Structures in Pink Have Seen in MD Simulations and Only Form Hydrogen Bonding with R5.42 and Y180 (Orange Dashed Line). Color Code for the Helices are TMH1 (Red), TMH2 (Orange), TMH3 (Yellow), TMH4 (Light Green), TMH5 (Dark Green), TMH6 (in Cyan) and TMH7 (Dark Blue). 2D Structure of PSB-CB5 Inserted in Far Left of the Above Picture Beside TMH1.

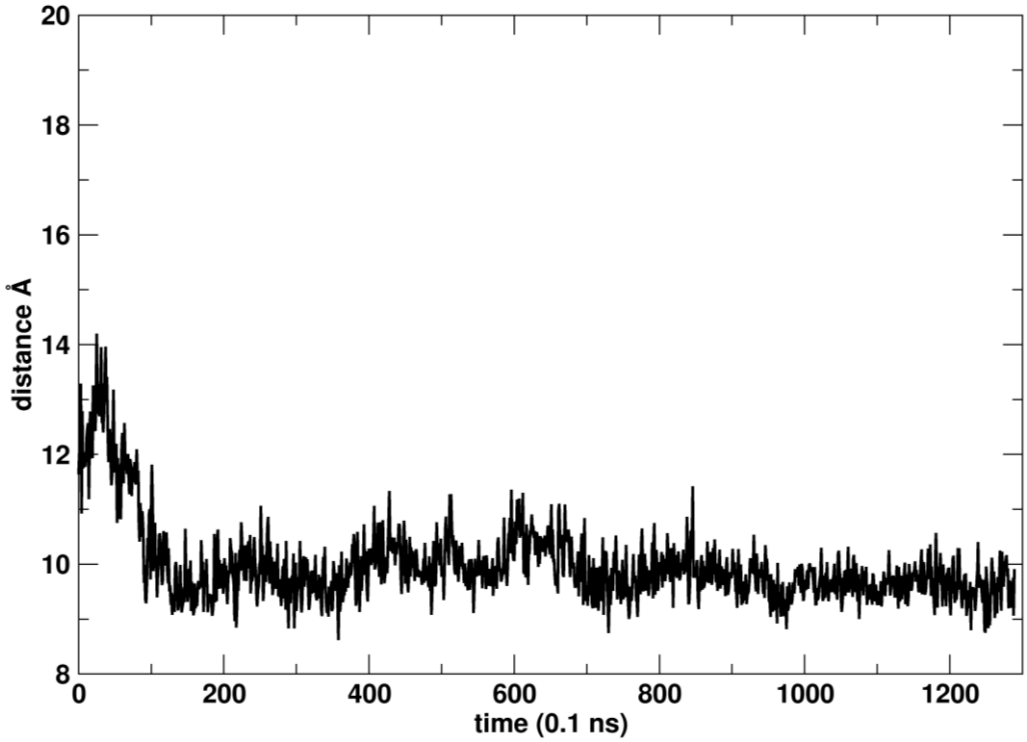


Figure 59. Distance of the Mass Center of the 5 Last Intracellular End Residues of TMH3 and 6.

CHAPTER VI

CONCLUSIONS AND FUTURE WORKS

In this project, a series of GPR18 agonists and antagonists synthesized, which resulted in two potent structures S4 and S5. The best antagonist S4 inhibited NAGly (the first reported agonist of GPR18) affects in calcium signaling assay completely at 10 μM, while the earlier reported antagonist PSBCB-5 can only reduce a NAGly-generated calcium signal by 40-50% at 10 μM. The best agonist in this series S5 showed efficacy comparable to NAGly.

We also presented an inactive and active states computer model of the GPR18 based on its homology with the crystal structure of the delta opioid receptor (DOR). These models used to dock the most important reported antagonist and agonists for GPR18 and identifying the crucial residues involved in the binding crevice of GPR18.

Furthermore, the constructed inactive model was used to explain mutation A3.39N by MD simulation. These simulations revealed the mechanism of activation of GPR18 by itself, which could shed light on future studies.

Finally, by parameterizing our proposed antagonist S4, we could investigate the quality of our static docking job in MD simulation. As the results show, this compound is able to make strong hydrogen bonding with critical residues R5.42 and Y180 in the GPR18 binding pocket and also keep toggle switch residue F6.48 in the off state while the ionic lock interaction between R3.50 and S6.33 is formed stably. To date, there has

not been a systematic, comprehensive, and extended mutagenesis study of GPR18. Hence, a mutagenesis study based on the results of the current project could be performed for future studies to improve the quality of the models constructed here.

Our docking studies suggest that R2.60, T7.39, and T3.32 are putative sites for interaction with GPR18 antagonist PSB-CB5 and S4. Therefore, we suggest R2.60M/A, T7.39S/A, and T3.32S/A mutations. Docking studies on GPR18 agonists, NAGly, RvD2, and S5 revealed that R5.42 plays a vital role in binding. Thus R5.42K/M/A mutant are suggested. Our inactive model of GPR18 suggests that the putative ionic lock in GPR18 can be formed between R3.50 and S6.33.

Therefore, the presence of a small residue like S6.33 instead of a highly conserved D/E6.30 in this region makes a weak ionic lock which may argue for high basal activity in GPR18. We would expect that the S6.33A mutation increases GPR18 constitutive activity. Another region that based upon the inactive model presented here could play an important role is at the cytoplasmic ends of TMH3 and 6. As mentioned before, GPR18 is reported to have high constitutive activity. Hence the contact between bulky residues Ile3.46 and Ile6.37 causes these amino acids to push each other, making the ionic lock more easily broken, leading to receptor activation in the basal state. Therefore, we would expect that I3.46A and I6.37A mutations affect the basal activity of the GPR18. One necessary switch that needs to be broken for activation of the receptor is the binding pocket toggle switch. The key residue here is 6.48, which in GPR18 is F6.48. It is well known that from the binding pocket, an agonist can activate the receptor by causing the 6.48 chi1 torsion angle to change from gauche plus to trans. Our model suggests that this

switch is fragile in GPR18. In this way, M7.42, H6.52, and Y3.35 help to confine F6.48 to its gauche plus inactive conformation. Thus, mutation of M7.42/A, H6.52A, and Y3.35A should increase the basal activity of GPR18.

Taking into consideration the results of the mutation study, another future aim could be to propose a synthesis plan whose goal is to generate a structure-activity relationship (SAR) based on our proposed antagonist S4 that should result in compounds with high potency. Compounds could be screened for their ability by the MAPK signaling, calcium mobilization, and recruitment of β-arrestin assays produced by the GPR18 agonists, NAGly, RvD2, and S5 in stably transfected CHO cells expressing GPR18.

REFERENCES

1. Manuel, I.; de San Román, E. G.; Giralt, M. T.; Ferrer, I.; Rodriguez-Puertas, R., Type-1 cannabinoid receptor activity during Alzheimer's disease progression. *Journal of Alzheimer's Disease* **2014**, 42, (3), 761-766.

2. Brotchie, J. M., CB1 cannabinoid receptor signalling in Parkinson's disease. *Current Opinion in Pharmacology* **2003**, 3, (1), 54-61.

3. More, S. V.; Choi, D.-K., Promising cannabinoid-based therapies for Parkinson's disease: motor symptoms to neuroprotection. *Molecular Neurodegeneration* **2015**, 10, (1), 17.

4. Pazos, M.; Sagredo, O.; Fernandez-Ruiz, J., The endocannabinoid system in Huntington's disease. *Current Pharmaceutical Design* **2008**, 14, (23), 2317-2325.

5. Lastres-Becker, I.; Miguel, R.; Fernandez-Ruiz, J., The endocannabinoid system and Huntington's disease. *Current Drug Targets-CNS & Neurological Disorders* **2003**, 2, (5), 335-347.

6. Leimuranta, P.; Khiroug, L.; Giniatullin, R., Emerging role of (Endo) cannabinoids in migraine. *Frontiers in Pharmacology* **2018**, 9, 420.

7. Manseau, M. W.; Goff, D. C., Cannabinoids and schizophrenia: risks and therapeutic potential. *Neurotherapeutics* **2015**, 12, (4), 816-824.

8. Matsuda, L. A.; Lolait, S. J.; Brownstein, M. J.; Young, A. C.; Bonner, T. I., Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature* **1990**, 346, (6284), 561.

9. Howlett, A.; Barth, F.; Bonner, T.; Cabral, G.; Casellas, P.; Devane, W.; Felder, C.; Herkenham, M.; Mackie, K.; Martin, B., International Union of Pharmacology. XXVII. Classification of cannabinoid receptors. *Pharmacological Reviews* **2002**, 54, (2), 161-202.

10. Herkenham, M.; Lynn, A. B.; Little, M. D.; Johnson, M. R.; Melvin, L. S.; De Costa, B. R.; Rice, K. C., Cannabinoid receptor localization in brain. *Proceedings of the National Academy of Sciences* **1990**, 87, (5), 1932-1936.

11. Gerard, C. M.; Mollereau, C.; Vassart, G.; Parmentier, M., Molecular cloning of a human cannabinoid receptor which is also expressed in testis. *Biochemical Journal* **1991**, 279, (1), 129-134.

12. Herkenham, M.; Lynn, A. B.; Johnson, M. R.; Melvin, L. S.; de Costa, B. R.; Rice, K. C., Characterization and localization of cannabinoid receptors in rat brain: a quantitative in vitro autoradiographic study. *Journal of Neuroscience* **1991**, 11, (2), 563-583.

13. Zou, S.; Kumar, U., Cannabinoid receptors and the endocannabinoid system: signaling and function in the central nervous system. *International Journal of Molecular Sciences* **2018**, 19, (3), 833.

14. Kruk-Slomka, M.; Dzik, A.; Budzynska, B.; Biala, G., Endocannabinoid system: the direct and indirect involvement in the memory and learning processes—a short review. *Molecular Neurobiology* **2017**, 54, (10), 8332-8347.

15. Parsons, L. H.; Hurd, Y. L., Endocannabinoid signalling in reward and addiction. Nature reviews. *Neuroscience* **2015**, 16, (10), 579-594.

16. Barrie, N.; Manolios, N., The endocannabinoid system in pain and inflammation: Its relevance to rheumatic disease. *European Journal of Rheumatology* **2017**, 4, (3), 210-218.

17. Bellocchio, L.; Cervino, C.; Pasquali, R.; Pagotto, U., The endocannabinoid system and energy metabolism. *Journal of Neuroendocrinology* **2008**, 20, (6), 850-857.

18. Koch, M., Cannabinoid Receptor Signaling in Central Regulation of Feeding Behavior: A Mini-Review. *Frontiers in Neuroscience* **2017**, 11, 293.

19. Munro, S.; Thomas, K. L.; Abu-Shaar, M., Molecular characterization of a peripheral receptor for cannabinoids. *Nature* **1993**, 365, (6441), 61-65.

20. Turcotte, C.; Blanchet, M. R.; Laviolette, M.; Flamand, N., The CB2 receptor and its role as a regulator of inflammation. *Cellular and Molecular Life Sciences: CMLS* **2016**, 73, (23), 4449-4470.

21. Maresz, K.; Pryce, G.; Ponomarev, E. D.; Marsicano, G.; Croxford, J. L.; Shriver, L. P.; Ledent, C.; Cheng, X.; Carrier, E. J.; Mann, M. K.; Giovannoni, G.; Pertwee, R. G.; Yamamura, T.; Buckley, N. E.; Hillard, C. J.; Lutz, B.; Baker, D.; Dittel, B. N., Direct suppression of CNS autoimmune inflammation via the cannabinoid receptor CB1 on neurons and CB2 on autoreactive T cells. *Nature Medicine* **2007**, 13, (4), 492-497.

22. McAllister, S. D.; Griffin, G.; Satin, L. S.; Abood, M. E., Cannabinoid receptors can activate and inhibit G protein-coupled inwardly rectifying potassium channels in a xenopus oocyte expression system. *The Journal of Pharmacology and Experimental Therapeutics* **1999**, 291, (2), 618-626.

23. Agler, H. L.; Evans, J.; Colecraft, H. M.; Yue, D. T., Custom distinctions in the interaction of G-protein beta subunits with N-type (CaV2.2) versus P/Q-type (CaV2.1) calcium channels. *Journal of General Physiology* **2003**, 121, (6), 495-510.

24. Howlett, A. C.; Blume, L. C.; Dalton, G. D., CB(1) cannabinoid receptors and their associated proteins. *Current Medicinal Chemistry* **2010**, 17, (14), 1382-1393.

25. Jarai, Z.; Wagner, J. A.; Varga, K.; Lake, K. D.; Compton, D. R.; Martin, B. R.; Zimmer, A. M.; Bonner, T. I.; Buckley, N. E.; Mezey, E.; Razdan, R. K.; Zimmer, A.; Kunos, G., Cannabinoid-induced mesenteric vasodilation through an endothelial site distinct from CB1 or CB2 receptors. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 96, (24), 14136-14141.

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

27. Gantz, I.; Muraoka, A.; Yang, Y. K.; Samuelson, L. C.; Zimmerman, E. M.; Cook, H.; Yamada, T., Cloning and chromosomal localization of a gene (GPR18) encoding a novel seven transmembrane receptor highly expressed in spleen and testis. *Genomics* **1997**, 42, (3), 462-466.

28. Wain, H. M.; Bruford, E. A.; Lovering, R. C.; Lush, M. J.; Wright, M. W.; Povey, S., Guidelines for human gene nomenclature. *Genomics* **2002**, 79, (4), 464-470.

29. Vassilatis, D. K.; Hohmann, J. G.; Zeng, H.; Li, F.; Ranchalis, J. E.; Mortrud, M. T.; Brown, A.; Rodriguez, S. S.; Weller, J. R.; Wright, A. C.; Bergmann, J. E.; Gaitanaris, G. A., The G protein-coupled receptor repertoires of human and mouse. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, 100, (8), 4903-4908.

30. Rosenkilde, M. M.; Benned-Jensen, T.; Andersen, H.; Holst, P. J.; Kledal, T. N.; Luttichau, H. R.; Larsen, J. K.; Christensen, J. P.; Schwartz, T. W., Molecular pharmacological phenotyping of EBI2. An orphan seven-transmembrane receptor with constitutive activity. *The Journal of Biological Chemistry* **2006**, 281, (19), 13199-13208.

31. Kohno, M.; Hasegawa, H.; Inoue, A.; Muraoka, M.; Miyazaki, T.; Oka, K.; Yasukawa, M., Identification of N-arachidonylglycine as the endogenous ligand for orphan G-protein-coupled receptor GPR18. *Biochemical and Biophysical Research Communications* **2006**, 347, (3), 827-832.

32. Qin, Y.; Verdegaal, E. M.; Siderius, M.; Bebelman, J. P.; Smit, M. J.; Leurs, R.; Willemze, R.; Tensen, C. P.; Osanto, S., Quantitative expression profiling of G-protein-coupled receptors (GPCRs) in metastatic melanoma: the constitutively active orphan GPCR GPR18 as novel drug target. *Pigment Cell & Melanoma Research* **2011**, 24, (1), 207-218.

33. Takenouchi, R.; Inoue, K.; Kambe, Y.; Miyata, A., N-arachidonoyl glycine induces macrophage apoptosis via GPR18. *Biochemical and Biophysical Research Communications* **2012**, 418, (2), 366-371.

34. Caldwell, M. D.; Hu, S. S.; Viswanathan, S.; Bradshaw, H.; Kelly, M. E.; Straiker, A., A GPR18-based signalling system regulates IOP in murine eye. *British Journal of Pharmacology* **2013**, 169, (4), 834-843.

35. Penumarti, A.; Abdel-Rahman, A. A., The novel endocannabinoid receptor GPR18 is expressed in the rostral ventrolateral medulla and exerts tonic restraining influence on blood pressure. *The Journal of Pharmacology and Experimental Therapeutics* **2014**, 349, (1), 29-38.

36. Flegel, C.; Vogel, F.; Hofreuter, A.; Wojcik, S.; Schoeder, C.; Kiec-Kononowicz, K.; Brockmeyer, N. H.; Muller, C. E.; Becker, C.; Altmuller, J.; Hatt, H.; Gisselmann, G., Characterization of non-olfactory GPCRs in human sperm with a focus on GPR18. *Scientific Reports* **2016**, 6, 32255.

37. Miller, S.; Leishman, E.; Oehler, O.; Daily, L.; Murataeva, N.; Wager-Miller, J.; Bradshaw, H.; Straiker, A., Evidence for a GPR18 Role in Diurnal Regulation of Intraocular Pressure. *Investigative Ophthalmology & Visual Science* **2016**, 57, (14), 6419-6426.

38. Matouk, A. I.; Taye, A.; El-Moselhy, M. A.; Heeba, G. H.; Abdel-Rahman, A. A., The Effect of Chronic Activation of the Novel Endocannabinoid Receptor GPR18 on Myocardial Function and Blood Pressure in Conscious Rats. *Journal of Cardiovascular Pharmacology* **2017**, 69, (1), 23-33.

39. Zhang, M. J.; Sansbury, B. E.; Hellmann, J.; Baker, J. F.; Guo, L.; Parmer, C. M.; Prenner, J. C.; Conklin, D. J.; Bhatnagar, A.; Creager, M. A.; Spite, M., Resolvin D2 Enhances Postischemic Revascularization While Resolving Inflammation. *Circulation* **2016**, 134, (9), 666-680.

40. Pridgeon, J. W.; Klesius, P. H., G-protein coupled receptor 18 (GPR18) in channel catfish: expression analysis and efficacy as immunostimulant against Aeromonas hydrophila infection. *Fish & Shellfish Immunology* **2013**, 35, (4), 1070-1078.

41. Pascoal, L. B.; Bombassaro, B.; Ramalho, A. F.; Coope, A.; Moura, R. F.; Correa-da-Silva, F.; Ignacio-Souza, L.; Razolli, D.; de Oliveira, D.; Catharino, R.; Velloso, L. A., Resolvin RvD2 reduces hypothalamic inflammation and rescues mice from diet-induced obesity. *Journal of Neuroinflammation* **2017**, 14, (1), 5.

42. Chiang, N.; de la Rosa, X.; Libreros, S.; Serhan, C. N., Novel Resolvin D2 Receptor Axis in Infectious Inflammation. *Journal of Immunology* **2017**, 198, (2), 842-851.

43. Rajaraman, G.; Simcocks, A.; Hryciw, D. H.; Hutchinson, D. S.; McAinch, A. J., G protein coupled receptor 18: A potential role for endocannabinoid signaling in metabolic dysfunction. *Molecular Nutrition & Food Research* **2016**, 60, (1), 92-102.

44. Malek, N.; Kostrzewa, M.; Makuch, W.; Pajak, A.; Kucharczyk, M.; Piscitelli, F.; Przewlocka, B.; Di Marzo, V.; Starowicz, K., The multiplicity of spinal AA-5-HT anti-nociceptive action in a rat model of neuropathic pain. *Pharmacological Research* **2016**, 111, 251-263.

45. Yin, H.; Chu, A.; Li, W.; Wang, B.; Shelton, F.; Otero, F.; Nguyen, D. G.; Caldwell, J. S.; Chen, Y. A., Lipid G protein-coupled receptor ligand identification using beta-arrestin PathHunter assay. *The Journal of Biological Chemistry* **2009**, 284, (18), 12328-12338.

46. Lu, V. B.; Puhl, H. L., 3rd; Ikeda, S. R., N-Arachidonyl glycine does not activate G protein-coupled receptor 18 signaling via canonical pathways. *Molecular Pharmacology* **2013**, 83, (1), 267-282.

47. Finlay, D. B.; Joseph, W. R.; Grimsey, N. L.; Glass, M., GPR18 undergoes a high degree of constitutive trafficking but is unresponsive to N-Arachidonoyl Glycine. *PeerJ* **2016**, 4, e1835.

48. Bradshaw, H. B.; Rimmerman, N.; Hu, S. S.; Benton, V. M.; Stuart, J. M.; Masuda, K.; Cravatt, B. F.; O'Dell, D. K.; Walker, J. M., The endocannabinoid anandamide is a precursor for the signaling lipid N-arachidonoyl glycine by two distinct pathways. *BMC Biochemistry* **2009**, 10, 14.

49. Bradshaw, H. B.; Lee, S. H.; McHugh, D., Orphan endogenous lipids and orphan GPCRs: a good match. *Prostaglandins & Other Lipid Mediators* **2009**, 89, (3-4), 131-134.

50. McHugh, D.; Page, J.; Dunn, E.; Bradshaw, H. B., Delta(9) -Tetrahydrocannabinol and N-arachidonyl glycine are full agonists at GPR18 receptors and induce migration in human endometrial HEC-1B cells. *British Journal of Pharmacology* **2012**, 165, (8), 2414-2424.

51. Sheskin, T.; Hanus, L.; Slager, J.; Vogel, Z.; Mechoulam, R., Structural requirements for binding of anandamide-type compounds to the brain cannabinoid receptor. *Journal of Medicinal Chemistry* **1997**, 40, (5), 659-667.

52. Huang, S. M.; Bisogno, T.; Petros, T. J.; Chang, S. Y.; Zavitsanos, P. A.; Zipkin, R. E.; Sivakumar, R.; Coop, A.; Maeda, D. Y.; De Petrocellis, L.; Burstein, S.; Di Marzo, V.; Walker, J. M., Identification of a new class of molecules, the arachidonyl amino acids, and characterization of one member that inhibits pain. *The Journal of Biological Chemistry* **2001**, 276, (46), 42639-42644.

53. Succar, R.; Mitchell, V. A.; Vaughan, C. W., Actions of N-arachidonyl-glycine in a rat inflammatory pain model. *Molecular Pain* **2007**, 3, 24.

54. Vuong, L. A.; Mitchell, V. A.; Vaughan, C. W., Actions of N-arachidonyl-glycine in a rat neuropathic pain model. *Neuropharmacology* **2008**, 54, (1), 189-193.

55. Burstein, S. H.; Huang, S. M.; Petros, T. J.; Rossetti, R. G.; Walker, J. M.; Zurier, R. B., Regulation of anandamide tissue levels by N-arachidonylglycine. *Biochemical Pharmacology* **2002**, 64, (7), 1147-1150.

56. McHugh, D.; Wager-Miller, J.; Page, J.; Bradshaw, H. B., siRNA knockdown of GPR18 receptors in BV-2 microglia attenuates N-arachidonoyl glycine-induced cell migration. *Journal of Molecular Signaling* **2012**, 7, (1), 10.

57. McHugh, D., GPR18 in microglia: implications for the CNS and endocannabinoid system signalling. *British Journal of Pharmacology* **2012**, 167, (8), 1575-1582.

58. McHugh, D.; Roskowski, D.; Xie, S.; Bradshaw, H. B., Delta(9)-THC and N-arachidonoyl glycine regulate BV-2 microglial morphology and cytokine release plasticity: implications for signaling at GPR18. *Front Pharmacology* **2014**, 4, 162.

59. Spite, M.; Norling, L. V.; Summers, L.; Yang, R.; Cooper, D.; Petasis, N. A.; Flower, R. J.; Perretti, M.; Serhan, C. N., Resolvin D2 is a potent regulator of leukocytes and controls microbial sepsis. *Nature* **2009**, 461, (7268), 1287-1291.

60. Chiang, N.; Dalli, J.; Colas, R. A.; Serhan, C. N., Identification of resolvin D2 receptor mediating resolution of infections and organ protection. *The Journal of Experimental Medicine* **2015**, 212, (8), 1203-1217.

61. Norling, L. V.; Dalli, J.; Flower, R. J.; Serhan, C. N.; Perretti, M., Resolvin D1 limits polymorphonuclear leukocyte recruitment to inflammatory loci: receptor-dependent actions. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2012**, 32, (8), 1970-1978.

62. Claria, J.; Dalli, J.; Yacoubian, S.; Gao, F.; Serhan, C. N., Resolvin D1 and resolvin D2 govern local inflammatory tone in obese fat. *Journal of Immunology* **2012**, 189, (5), 2597-2605.

63. Console-Bram, L.; Brailoiu, E.; Brailoiu, G. C.; Sharir, H.; Abood, M. E., Activation of GPR18 by cannabinoid compounds: a tale of biased agonism. *British Journal of Pharmacology* **2014**, 171, (16), 3908-3917.

64. Harms, H.; Rempel, V.; Kehraus, S.; Kaiser, M.; Hufendiek, P.; Muller, C. E.; Konig, G. M., Indoloditerpenes from a marine-derived fungal strain of Dichotomomyces cejpii with antagonistic activity at GPR18 and cannabinoid receptors. *Journal of Natural Products* **2014**, 77, (3), 673-677.

65. Rose, P. W.; Beran, B.; Bi, C.; Bluhm, W. F.; Dimitropoulos, D.; Goodsell, D. S.; Prlic, A.; Quesada, M.; Quinn, G. B.; Westbrook, J. D.; Young, J.; Yukich, B.; Zardecki, C.; Berman, H. M.; Bourne, P. E., The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Research* **2011**, 39, (Database issue), D392-401.

66. Richardson, J. S., The anatomy and taxonomy of protein structure. *Advances in Protein Chemistry* **1981**, 34, 167-339.

67. Ballesteros, J. A.; Weinstein, H., [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in neurosciences*, Elsevier: 1995; Vol. 25, pp 366-428.

68. Guarnieri, F.; Weinstein, H., Conformational memories and the exploration of biologically relevant peptide conformations: an illustration for the gonadotropin-releasing hormone. *Journal of the American Chemical Society* **1996**, 118, (24), 5580-5589.

69. Guarnieri, F.; Wilson, S. R., Conformational memories and a simulated annealing program that learns: application to LTB4. *Journal of Computational Chemistry* **1995**, 16, (5), 648-653.

70. Reggio, P. H., Computational methods in drug design: modeling G protein-coupled receptor monomers, dimers, and oligomers. *The AAPS Journal* **2006**, 8, (2), E322-336.

71. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-

atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B* **1998**, 102, (18), 3586-3616.

72. Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W., Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *The Journal of Physical Chemistry. B* **2010**, 114, (23), 7830-7843.

73. Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of Chemical Theory and Computation* **2013**, 9, (9), 3878-3888.

74. Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; MacKerell, A. D., Jr., Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophysical Journal* **2006**, 90, (4), L36-38.

75. Ballesteros, J. A.; Weinstein, H., Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains. *Biophysical Journal* **1992**, 62, (1), 107-109.

76. Ballesteros, J. A.; Deupi, X.; Olivella, M.; Haaksma, E. E.; Pardo, L., Serine and threonine residues bend alpha-helices in the chi(1) = g(-) conformation. *Biophysical Journal* **2000**, 79, (5), 2754-2760.

77. Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K., Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **2011**, 477, (7366), 549-555.

78. McAllister, S. D.; Hurst, D. P.; Barnett-Norris, J.; Lynch, D.; Reggio, P. H.; Abood, M. E., Structural mimicry in class A G protein-coupled receptor rotamer toggle switches: the importance of the F3.36(201)/W6.48(357) interaction in cannabinoid CB1 receptor activation. *The Journal of Biological Chemistry* **2004**, 279, (46), 48024-48037.

79. Webb, B.; Sali, A., Protein structure modeling with MODELLER. *Methods in Molecular Biology* **2014**, 1137, 1-15.

80. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics* **2014**, 47, 5 6 1-32.

81. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L., OPLS3: a force field providing broad

coverage of drug-like small molecules and proteins. *Journal of Chemical Theory and Computation* **2015**, 12, (1), 281-296.

82. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W., Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation* **2010**, 6, (5), 1509-1519.

83. Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **1988**, 110, (6), 1657-1666.

84. Kotsikorou, E.; Sharir, H.; Shore, D. M.; Hurst, D. P.; Lynch, D. L.; Madrigal, K. E.; Heynen-Genel, S.; Milan, L. B.; Chung, T. D.; Seltzman, H. H.; Bai, Y.; Caron, M. G.; Barak, L. S.; Croatt, M. P.; Abood, M. E.; Reggio, P. H., Identification of the GPR55 antagonist binding site using a novel set of high-potency GPR55 selective ligands. *Biochemistry* **2013**, 52, (52), 9456-9469.

85. Lynch, D. L.; Hurst, D. P.; Shore, D. M.; Pitman, M. C.; Reggio, P. H., Molecular Dynamics Methodologies for Probing Cannabinoid Ligand/Receptor Interaction. *Methods in Enzymology* **2017**, 593, 449-490.

86. Rempel, V.; Atzler, K.; Behrenswerth, A.; Karcz, T.; Schoeder, C.; Hinz, S.; Kaleta, M.; Thimm, D.; Kiec-Kononowicz, K.; Müller, C., Bicyclic imidazole-4-one derivatives: a new class of antagonists for the orphan G protein-coupled receptors GPR18 and GPR55. *MedChemComm* **2014**, 5, (5), 632-649.

87. Jensen, A. D.; Guarnieri, F.; Rasmussen, S. G.; Asmar, F.; Ballesteros, J. A.; Gether, U., Agonist-induced conformational changes at the cytoplasmic side of transmembrane segment 6 in the beta 2 adrenergic receptor mapped by site-selective fluorescent labeling. *The Journal of Biological Chemistry* **2001**, 276, (12), 9279-9290.

88. Rich, M. R., Conformational analysis of arachidonic and related fatty acids using molecular dynamics simulations. *Biochimica Biophysica Acta* **1993**, 1178, (1), 87-96.

89. Nourbakhsh, F.; Atabaki, R.; Roohbakhsh, A. The role of orphan G protein-coupled receptors in the modulation of pain: A review. *Life Sciences* **2018**, 212, 59–69.

90. Laschet, C.; Dupuis, N.; Hanson, J. The G Protein-Coupled Receptors deorphanization landscape. *Biochemical Pharmacology* **2018**, 153, 62–74.

91. Alexander, S.P.H.; Christopoulos, A.; Davenport, A.P.; Kelly, E.; Marrion, N.V.; Peters, J.A.; Faccenda, E.; Harding, S.D.; Pawson, A.J.; Sharman, J.L.; et al. The concise guide of pharmacology 2017/18: G protein-coupled receptors. *British Journal of Pharmacology* **2017**, 174, S17–S129.

92. Pertwee, R.G.; Howlett, A.C.; Abood, M.E.; Alexander, S.P.H.; Marzo, V.D.; Elphick, M.R.; Greasley, P.J.; Hansen, H.S.; Kunos, G. International Union of Basic and Clinical Pharmacology. LXXIX. Cannabinoid Receptors and Their Ligands: Beyond CB1 and CB2. *Pharmacological Reviews* **2010**, 62, 588–631.

93. Taddese, B.; Deniaud, M.; Garnier, A.; Tiss, A.; Guissouma, H.; Abdi, H.; Henrion, D.; Chabbert, M. Evolution of chemokine receptors is driven by mutations in the sodium binding site. *PLoS Computational Biology* **2018**, 14, 1–25.

94. Montaner, S.; Kufareva, I.; Abagyan, R.; Gutkind, J.S. Molecular Mechanisms Deployed by Virally Encoded G Protein–Coupled Receptors in Human Diseases. *Annual Review of Pharmacology and Toxicology* **2013**, 53, 331–354.

95. Katritch, V.; Fenalti, G.; Abola, E.E.; Roth, B.L.; Cherezov, V.; Stevens, R.C. Allosteric sodium in class A GPCR signaling. *Trends in Biochemical Sciences* **2014**, 39, 233–244.

96. Visiers, I.; Braunheim, B.B.; Weinstein, H. Prokink: A protocol for numerical evaluation of helix distortions by proline. *Protein Engineering* **2000**, 13, 603–606.

97. Shang, Y.; Lerouzic, V.; Schneider, S.; Bisignano, P.; Pasternak, G.W.; Filizola, M. Mechanistic insights into the allosteric modulation of opioid receptors by sodium ions. *Biochemistry* **2014**, 53, 5140–5149.

98. Miao, Y.; Caliman, A.D.; McCammon, J.A. Allosteric Effects of Sodium Ion Binding on Activation of the M3 Muscarinic G-Protein-Coupled Receptor. *Biophysical Journal* **2015**, 108, 1796–1806.

99. Gutiérrez-de-Terán, H.; Massink, A.; Rodríguez, D.; Liu, W.; Han, G.W.; Joseph, J.S.; Katritch, I.; Heitman, L.H.; Xia, L.; Ijzerman, A.P.; et al. The Role of a Sodium Ion Binding Site in the Allosteric Modulation of the A2A Adenosine G Protein-Coupled Receptor. *Structure* **2013**, 21, 2175–2185.

100. Ballesteros, J.; Kitanovic, S.; Guarnieri, F.; Davies, P.; Fromme, B.J.; Konvicka, K.; Chi, L.; Millar, R.P.; Davidson, J.S.; Weinstein, H.; et al. Functional microdomains in G-protein-coupled receptors: The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. *Journal of Biological Chemistry* **1998**, 273, 10445–10453.

101. Venkatakrishnan, A.J.; Deupi, X.; Lebon, G.; Heydenreich, F.M.; Flock, T.; Miljus, T.; Balaji, S.; Bouvier, M.; Veprintsev, D.B.; Tate, C.G.; et al. Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature* **2016**, 40, 383–388.

102. Hurst, D.P.; Grossfield, A.; Lynch, D.L.; Feller, S.; Romo, T.D.; Gawrisch, K.; Pitman, M.C.; Reggio, P.H. A lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. *Journal of Biological Chemistry* **2010**, 285, 17954–17964.

103. Eswar, N.; Webb, B.; Marti-Renom, M.A.; Madhusudhan, M.S.; Eramian, D.; Shen, M.-y.; Pieper, U.; Sali, A. Comparative Protein Structure Modeling Using Modeller. *Current Protocols in Bioinformatics* **2006**, 15, 1–46.

104. Fiser, A.; Kinh Gian Do, R.; Sali, A. Modeling Loops in Protein Structures. *Protein Science* **2000**, 9, 1753–1773.

105. Šali, A.; Blundell, T.L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology* **1993**, 234, 779–815.

106. Morales, P.; Hurst, D.P.; Reggio, P.H. Methods for the Development of In Silico GPCR Models. In *Methods in Enzymology* Ed., Reggio, P.H., Ed. Elsevier: Amsterdam, The Netherlands, 2017; pp. 405-448.

107. Lomize, M.A.; Pogozheva, I.D.; Joo, H.; Mosberg, H.I.; Lomize, A.L. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Research* **2011**, 40, D370–D376.

108. Jo, S.; Kim, T.; Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* **2007**, 2, e880.

109. Lee, J.; Cheng, X.; Swails, J.M.; Yeom, M.S.; Eastman, P.K.; Lemkul, J.A.; Wei, S.; Buckner, J.; Jeong, J.C.; Qi, Y. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of Chemical Theory and Computation* **2015**, 12, 405–413.

110. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B.L.; Grubmüller, H.; MacKerell, A.D. CHARMM36: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Biophysical Journal* **2017**, 112, 175a–176a.

111. Case, D.; Betz, R.; Cerutti, D.; Cheatham, T.; Darden, T.; Duke, R. AMBER16, 2016. San Francisco. http://ambermd.org.

112. Case, D.A.; Cheatham III, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **2005**, 26, 1668–1688.

113. Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C., Rapid parameterization of small molecules using the Force Field Toolkit. *Journal of Computational Chemistry* **2013**, 34, (32), 2757-2770.

114. Ortiz, J.; Cioslowski, J.; Fox, D., Gaussian 09, revision B. 01. Wallingford CT 2009.

115. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I., CHARMM general force field: A force field for drug- like molecules compatible with the CHARMM all- atom additive biological force fields. *Journal of Computational Chemistry* **2010**, 31, (4), 671-690.

116. Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C. P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., 3rd; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., 3rd; Kong, J.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M., Advances in methods and algorithms in a modern quantum chemistry program package. *Physical Chemistry Chemical Physics* **2006**, 8, (27), 3172-3191.

117. Iliff, H. A.; Lynch, D. L.; Kotsikorou, E.; Reggio, P. H., Parameterization of Org27569: an allosteric modulator of the cannabinoid CB1 G protein-coupled receptor. *Journal of Computational Chemistry* **2011**, 32, (10), 2119-2126.

118. Vanommeslaeghe, K.; MacKerell, A. D., Jr., Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of Chemical Information and Modeling* **2012**, 52, (12), 3144-3154.

119. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr., Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *Journal of Chemical Information and Modeling* **2012**, 52, (12), 3155-3168.

120. Guvench, O.; MacKerell, A. D., Jr., Automated conformational energy fitting for force-field development. *Journal of Molecular Modeling* **2008**, 14, (8), 667-679.

121. Nelder, J. A.; Mead, R., A simplex method for function minimization. *The Computer Journal* **1965**, 7, (4), 308-313.

122. Flannery, B. P.; Press, W. H.; Teukolsky, S. A.; Vetterling, W., *Numerical recipes* in C. Press Syndicate of the University of Cambridge, New York 1992, 24, 78.