University of Texas Rio Grande Valley [ScholarWorks @ UTRGV](https://scholarworks.utrgv.edu/)

[Theses and Dissertations](https://scholarworks.utrgv.edu/etd)

5-2021

Virtual Screening for Potential New Chemotherapeutic Agents for the GPR119 Receptor, a Target for Type II Diabetes

Jennifer Lizeth Bravo The University of Texas Rio Grande Valley

Follow this and additional works at: [https://scholarworks.utrgv.edu/etd](https://scholarworks.utrgv.edu/etd?utm_source=scholarworks.utrgv.edu%2Fetd%2F834&utm_medium=PDF&utm_campaign=PDFCoverPages)

Part of the Chemistry Commons

Recommended Citation

Bravo, Jennifer Lizeth, "Virtual Screening for Potential New Chemotherapeutic Agents for the GPR119 Receptor, a Target for Type II Diabetes" (2021). Theses and Dissertations. 834. [https://scholarworks.utrgv.edu/etd/834](https://scholarworks.utrgv.edu/etd/834?utm_source=scholarworks.utrgv.edu%2Fetd%2F834&utm_medium=PDF&utm_campaign=PDFCoverPages)

This Thesis is brought to you for free and open access by ScholarWorks @ UTRGV. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact [justin.white@utrgv.edu, william.flores01@utrgv.edu.](mailto:justin.white@utrgv.edu,%20william.flores01@utrgv.edu)

VIRTUAL SCREENING FOR POTENTIAL NEW CHEMOTHERAPEUTIC AGENTS FOR THE GPR119 RECEPTOR, A TARGET FOR TYPE II DIABETES

A Thesis

by

JENNIFER LIZETH BRAVO

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

MASTER OF SCIENCE

May 2021

Major Subject: Chemistry

VIRTUAL SCREENING FOR POTENTIAL NEW CHEMOTHERAPEUTIC AGENTS FOR THE GPR119 RECEPTOR, A TARGET FOR TYPE II DIABETES

A Thesis by JENNIFER LIZETH BRAVO

COMMITTEE MEMBERS

Dr. Evangelia Kotsikorou Chair of Committee

Dr. Frank Dean Co-Chair of Committee

Dr. Jose Gutierrez Committee member

Dr. Jason Parsons Committee member

May 2021

Copyright 2021 Jennifer Lizeth Bravo All rights reserved

ABSTRACT

Bravo, Jennifer L., Virtual Screening for Potential New Chemotherapeutic Agents for the GPR119 Receptor, a Target for Type II Diabetes. Master of Science (MS), May, 2021, 103 pp., 21 tables, 22 figures, references, 20 titles.

Three frames from a molecular dynamics simulation run of a GPR119 receptor homology model were used for this study. The homology model was validated by virtually screening 76 known GPR119 receptor. 85% of these agonists bound to the receptor. Following the validation, 21,000 molecules were selected for the virtual screening study. 862 ligands came from the GPCR Selleckchem/Prestwick library, 42 compounds from the Prestwick Phytochemical library, 20,000 compounds from the ZINC library, plus four molecules from the literature. All ligands were built, geometry-optimized, and docked in the GPR119 models using a protocol combining High Throughput Virtual Screening, Standard Precision, and Extra Precision Glide docking. 2,100 compounds fit inside the GPR119 model binding pocket. The agonists AR231453, AR437735, and oleoyl serinol as well as compound SRT1720 were tested for activation of GPR119 using an ELISA cAMP assay. The results agreed with values in the literature and with the computational results.

DEDICATION

I dedicate this work to the man I love, my always supportive and loving husband, Dennis K. Kim, this journey would not have been possible without your encouragement and unconditional love, thank you for always believing in me even when I doubted myself.

I also want to dedicate this thesis to my parents and my sisters who cheered for me since day one, thank you guys for being my pillars and my balance. Finally, I want to thank God for not leaving my side even in my darkest hours, for all honor and glory to him.

ACKNOWLEDGEMENTS

My eternal gratitude goes to Dr. Evangelia Kotsikorou, my mentor and research advisor. Thank you for taking me into your group and for seeing the potential in me, your guidance and teachings will come with me wherever life takes me, and I promise to put them into work in my career; I also would like to thank Dr. Frank Dean for all his help and imparted knowledge.

I want to give special thanks to my friend and lab peer Matthew D. Rosales, you have been an incredible mentor and friend, thank you for all the help in the computational aspects of the project and in the wet lab techniques. Without you, I would not have completed my project in time, thank you for always being willing to come and help no matter the time of the day; you are awesome.

Lastly, I would like to thank the rest of the lab members, Naila Bravo (my sister) thank you for being there for me and making sure my mental and physical health was always in check, I love you dearly. Thank you, Ashley, Solomon, and Jesus for putting up with me, even when I was super annoying or cranky due to the lack of caffeine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

G-protein coupled receptors (GPCRs) are a subtype of proteins that are found in great numbers in the cell membrane of eukaryotic cells¹. Conformational changes in the GPCR receptor caused by ligands binding, induce activation or deactivation of the receptor, which in turn allows and/or restricts the transduction of information to the inside of the cell for it to respond accordingly to the messages (secretion of hormones, regulation signals, etc). Their role is an essential part for the sustaining of life especially for humans; therefore, the study of the function of these types of receptors is key to discover how human physiology works, and how to treat targeted-receptor diseases that afflict people².

GPCRs are transmembrane receptors with one end exposed to the extracellular matrix and the other to the cytosol of the cell³⁻⁴. The transmembrane helices are connected by loops that are flexible and move to aid the binding of a molecule into the receptor pocket or to obstruct the entrance of the binding pocket making the ligand stay inside the receptor for added stability⁴. **Figure 1** shows a simple representation of the receptor; the nitrogen terminal on the extracellular part of the cell is followed by seven transmembrane α-helices (TMHs 1-7) that loop in and out of the cell membrane (extracellular loops ECL 1, ECL 2, ECL 3 and intracellular loops ICL 1, ICL 2 and ICL 3) ending by the carboxyl terminus in the inside of the cell⁵.

1

Figure 1. General structure of a G protein-coupled receptor.

Literature reports that more than a thousand GPCRs have been identified in the vertebrate genome, which tend to be highly specific towards their possible activators⁶⁻⁷; this specificity allows for ligands to activate a targeted receptor without affecting the rest. The process of activation by a ligand occurs when the molecule in the bloodstream interacts with the extracellular part of the receptor and then enters the hydrophobic binding pocket. Once inside, the ligand causes a series of rearrangements of the binding pocket that lead to changes in the intracellular end of the receptor where the heterotrimeric G protein binds and gets activated. Upon activation the G protein decouples from the receptor and the Ga subunit dissociates from the Gβ and Gγ subunits and it exchanges the GDP molecule with a GTP^{6, 8}. Once the Gα subunit is done with the communication process, GTPase activity converts GTP back into a GDP, and it rejoins the receptor and reforms the G-protein by binding with a Gβ and a Gγ subunit.³⁻⁴

GPCRs ligands can be separated into three types depending on their effect on the receptor. If a ligand is able to activate the receptor and increase its activity, it is considered an agonist. Within the spectrum of agonists there are full agonists and partial agonists. A molecule is called a full agonist if it increases the receptor activity close to 100%, otherwise the molecule

is categorized as a partial agonist. An antagonist is a molecule that binds to the receptor and it prevents an agonist from binding and activating it. A ligand considered as an antagonist does not alter the receptor activity; therefore, it does not affect the basal activity of the receptor. Basal activity, also referred as constitutive activity, is the signaling of a receptor in the absence of an agonist. The last class, the inverse agonists, not only prevent increased activity by blocking agonists, but also decrease basal activity levels^{4, 9-10}. All these examples can be seen in a graphical representation in **Figure 2**.

Figure 2. GPCR receptor activity based on drug interaction.

1.1 G Protein-Coupled Receptor GPR119

The GPR119 receptor, a type A GPCR, is also known as glucose-dependent insulinotropic receptor¹¹⁻¹³. The GPR119 receptor is one of the most common transmembrane receptors found in the gastrointestinal L-cells and K-cells as well as in the pancreatic β-cells of mammalian organisms¹⁰. Upon activation, the levels of cyclic adenosine monophosphate (cAMP) in the cell rise, which can cause L-cells and K-cells to secrete hormones like

glucagonlike peptide 1 (GLP1), glucose dependent insulinotropic peptide (GIP) and peptide YY (PYY). Similarly, activation of the GPR119 receptor in the pancreatic β-cells promotes the release of insulin into the bloodstream¹⁴. The ability of the GPR119 receptor to influence the secretion of insulin makes it an attractive target for the treatment of diabetes mellitus type 2 since people with this condition struggle to keep glucose homeostasis in their body.

In-vivo experiments and *in-vitro* experiments such as ELISA cAMP colorimetric analysis, immunofluorescence northern hybridization blot, liquid chromatography-mass spectrometry (LC-MS) among others^{11-12, 15} can provide information about the activation of the GPR119 receptor. However, these studies do not provide insight into conformational response of the receptor based on the ligand used, nor information of the chemical interactions happening inside the binding pocket $16-17$. The advances in computational methods have opened the door to visualization of molecules and modeling of their behavior under controlled environments.

There is not yet an x-ray crystal structure for the GPR119 receptor^{10, 18}. Research groups and pharmaceutical companies therefore have had to construct their own 3-dimensional homology models to study how molecules bind and interact with the binding pocket of the GPR119 receptor. Depending on how the homology model was developed (template used, conformation of helices, etc), there could be differences in how a ligand can fit in the binding pocket. Therefore, the homology model needs to be validated. If the *in vitro* results cannot be explained by using the homology model, then that model would have to be reworded until similar results are obtained. **Figure 3** shows a cartoon representation of the in-house GPR119 receptor model used for this study.

4

Figure 3. 3D in-house homology model of the GPR119 receptor.

1.2 G Protein-Coupled Receptor Ligands and Libraries

A compound library is a collection of chemical substances. These chemical substances are either theoretical molecules or molecules that have been synthesized, which can be used in diverse types of research¹⁹. These databases provide from thousands to billions of molecules which could be pre-screened for different purposes such as in the medical field to analyze how ligands fit a targeted receptor, in industry to give information about light activated chemicals, or in pharmaceutical industry to discover natural organic compounds for plant-based drug applications $19-20$.

Chemical libraries are usually subsets of larger databases of compounds that have been selected based on properties such as similar molecular structures or comparable chemical properties. The information provided about the molecules by a chemical library varies, basic libraries include information such as compound name (UPAC or other names by which the

molecules are also known), chemical structure, and sometimes chirality. Advanced libraries like those provided by chemical companies often also include information about solubility, receptor targets and a small summary about their properties.

The GPCR chemical libraries contain a great number of molecules that bind to the orthosteric site of different known G coupled receptors; however, not all ligands fit the same receptors. The ligands for the GPR119 receptor, the receptor of interest for this study, seem to have some general characteristics which were used as criteria in the selection of ligands that may fit in the binding pocket and possibly activate the receptor. These agonists are typically long with slender profile and are composed of rings (aromatic and unsaturated rings) and often have a polar group in one end $5, 9, 21-22$. However, even if the criteria for the selection of ligands are met, there is no guarantee that a ligand that fits into the binding pocket would be able to activate the receptor. Ultimately, the ligands identified from the chemical libraries will need to be tested experimentally to confirm if they bind and activate the receptor.

1.3 High Throughput Virtual Screening, Standard Precision and Extra Precision Screening

In order to ensure the fit of ligands into GPCRs, a molecular mechanics (MM) docking technique called high throughput virtual screening is done to eliminate those molecules that do not fulfil the criteria of the specific receptor to be studied²³. High throughput virtual screening (HTVS) eliminates molecules that are too polar, bulky, wrongly oriented or chiral in some cases. High throughput analysis is a compatibility assay for data processing that screens thousands of molecules in parallel and scores them to minimize the number of intermediate confirmations²⁴. This type of screening to be the first step in finding compounds that will eventually produce lead compounds for further studies^{16-17, 19, 25}.

6

Standard precision screening (SP) uses the same docking algorithm as the HTVS method but does a more thorough torsional refinement and sampling eliminating more compounds that do not fit the binding pocket.

Extra precision screening (XP) is much more thorough method of screening compared to HTVS and SP analysis. Extra precision performs more extensive sampling that only allow a small number of ligands to pass to the last steps of the analysis. XP screening docks ligands flexibly and uses an anchor-and-grow procedure that considers the position of the molecules as well as how they fit into the receptor's binding pocket²⁶⁻²⁸. The XP screening method penalizes and disregards double positives caused by the different ways a ligand fit in the binding pocket 28 .

1.4 Drug Design and Ligands

Drug design requires an enormous amount of work and resources: select and test vast numbers of compounds, identify lead compounds, then design derivatives of those compounds, synthesize them, purify them, verify their structure of using spectroscopic methods such as nuclear magnetic resonance (NMR), test them *in vitro* and *in vivo* experiments to confirm or disprove the suitability of that compound to be moved further in the drug development process. In short, for pharmaceutical companies and independent research groups, drug design is an intense and expensive process. Molecular modeling and calculations such as Glide HTVS, SP and XP screening can expedite the drug discovery process and save companies and research groups' invaluable time and money.

1.5 *In Vitro* **Studies and the Role of cAMP in Assessing GPR119 Receptor Activity**

In vitro experiments employ cultured cells to study the effect of external stimuli such as chemical compounds, on cell signaling pathways. The ligand-induced cAMP production assay, a

secondary messenger assay, was selected for the study of the GPR119 receptor activation since it has been reported that upon activation of the receptor, cAMP accumulates inside the cell, making it a perfect quantitative analysis assay to understand the effect of varying drug concentrations in the cells^{4, 12, 29}.

The effect of compounds on a receptor cannot be precisely determined using only computational methods. *In vitro* experiments can be employed to confirm or disprove the proposed mechanism of action of these compounds. Certain compounds that might serve as an agonist for a target receptor might behave as an antagonist for others $30-31$. In vitro experiments can help elucidate the mechanism of action of different compounds and the signaling pathways activated by binding to different receptors in the cells to avoid undesired effects by compounds that are drug candidates.

CHAPTER II

LITERATURE REVIEW

2.1 Background of GPR119 Receptor

The GPR119 receptor was discovered in the early 2000s and was identified as a rhodopsin-like, class A GPCR 13 . Since then, the activation of this receptor has been studied through *in vivo* and *in vitro* experiments as well as thought computational methods. In 2009, Brubaker *et al.* monitored GLP-1 secretion after GPR119 activation with the endogenous ligand oleoylethanolamide (OEA) in mGLUtag cells and *in vivo* studies treating euglycemic rats. OEA is the endogenous ligand that naturally activates the GPR119 receptor in the body. The *in vitro* results showed a GLP-1 increase of 2.1 \pm 0.2–fold from basal levels at 10 µmol/L, while the intraluminal *in vivo* results yielded 1.5 ± 0.2 increased fold at 20 nmol/rat that lasted for the duration of the experiment $(60 \text{ minutes})^{32}$.

In 2014, Engelstoft *et al*. applied computational modeling to complement *in vitro* experiments in order to explain the mechanics behind the activation of the receptor. OEA and the synthetic compound AR231453 were used as agonists for his studies. The *in vitro* experiments were performed on COS7 cells (monkey kidney cells) using wild type and mutated (FLAG tagged) cells. The computational studies were performed using their in-house homology model of the GPR119 receptor that was created using a hybrid method. They combined the method proposed by Mobarec *et al.* 2009 that used multiple structures of other type A GPCRs as templates for homology model development³³; and Worth *et al.*, 2011 that

used a fragment-based approach to create homology models of GPCRs from readily available type A GPCR crystal structures³⁴.

The computational results of Engelstoft *et al*., indicated that TMHs III,V, VI and VII (TMSs 3, 5, 6 and 7), as well as ECL 2 are of great importance for ligand interaction and signaling since these transmembrane helices encompass the binding pocket and give an insight to the interaction of the agonist with the amino acids of the toggle switch $(W^{6.48}$ and $F^{3.36})$ located in the center of the transmembrane region of the receptor. The mutations made in the ECL 2 also provided information for its role in keeping ligands inside the binding pocket. The *in vitro* analysis showed that the EC_{50} of OEA was 250 nM and that of AR231453 was 1.9 nM, proving that both compounds are strong agonists. Even though the wet lab experiments were in agreement with the results of the computations (OEA and AR231453 fit in the binding pocket and interact with the amino acids that shown to be important in the *in vitro* experiments), there was still uncertainty regarding the mode of binding of the agonists in the receptor pocket (43 possible different confirmations were found after performing docking calculations 1000 times). It was concluded that more computational studies applying MD simulations are necessary to fully understand the conformational interactions of the receptor with the ligands 4 .

.

Figure 4. Chemical structures of the agonists AR231453 and OEA.

2.2 Ligand Interactions with the GPR119 Receptor

Ligands approach receptors extracellularly, bind to the receptor binding pocket causing it to activate signaling pathways thus transmitting information to the cell. The GPR119 receptor increases the production of cAMP upon activation if a full agonist or partial agonist interacts with the binding pocket. In 2015, Ritter *et al.* provided a list of full agonists from leading pharmaceutical companies such as Arena, AztraZeneca, Pfizer, Prosidion, Merck and Astellas. The EC_{50} s of leading agonist varied from Arena's AR231453 at 1.355 nM to Astellas's AS1269574 at 2.6 µM. The higher the concentration of a compound the receptor needs to have the desired biological effect, the least likely it is that the company will pursue if for drug d evelopment²¹.

An inverse agonist shuts down the constitutive activity of the receptor by stabilizing its inactive confirmation. Engelstoft *et al.* in 2014 reported that compound TM43718 is an inverse agonist of the GPR119 receptor in a dose-dependent manner where its IC_{50} was determined to be1.5 µM⁴ . Similarly, Norn *et al.* in 2015, claimed that AR437948, a third generation ligand derived from of AR231453, presented inverse agonistic behavior. The IC_{50} of that inverse agonist was not provided¹⁰. Antagonists on the other hand, are molecules that block agonistinduced activity without affecting the basal activity of the receptor. Syed *et al.* in 2012, performed cAMP colorimetric analysis and western blot experiments where he concluded that oleoyl serotonin (0.7 μ M), SB-366791 (10 μ M) and arvanil (50 μ M) antagonized the AR231453induced cAMP activity of the GPR119 receptor via competitive inhibition³⁰.

2.3 Database Screening

In the past, pharmaceutical companies would search for new drugs by performing high throughput screening of thousands of compounds using biological assays and other quantitative and qualitative methods of analysis. Once computational methods were developed, companies would first perform high throughput virtual screening of libraries of compounds to remove compounds that may not fit in the binding pocket and thus reduce the number of compounds to be tested physically and make the drug discovery process faster and more cost effective. In 2004, Shoichet discussed the advantages and disadvantages of using virtual analysis of molecules for their interactions with cellular receptors. He mentioned that the use of molecular modeling has proven beneficial for the screening of large databases since it relies on ligand receptor interactions that are governed by thermodynamic and quantum mechanical forces that permit the calculation of the energies of the ligand-receptor complexes. The downside of this type of analysis however was the introduction of compounds that are false negatives and false positives and would not be able to be identified until the compounds were tested experimentally³⁵.

Frienser *et al.* reported that before the introduction of Glide docking (grid-based ligand docking with energetics), the most commonly used virtual screening platforms in the 2000s were GOLD, FlexX and DOCK. While docking software prior to Glide treated the ligands as rigid when binding them into the PBD receptor, Glide considered the orientational, positional and conformational space of the ligands and their interaction with the binding pocket which doubled the docking accuracy from the former models¹⁷. Glide also allowed the refinement of compound analysis by introducing algorithms such as standard precision (SP), high throughput virtual screening (HTVS) and extra precision virtual screening (XP).

12

In 2006, Halgren *et al.* used Glide employing the SP algorithm to test 15,000 ligands from 15 different libraries. The ligands were energy-minimized and geometry-optimized using the Merc Molecular force field 94 (MMFF94) and were subsequently docked to nine different receptor types. From this analysis it was concluded that 70% of the enriched molecules preselected to bind to the receptors actually matched the experimental data and even though the results were promising, more work and understanding about ligand flexibility was required for a conclusive analysis 27 .

High throughput virtual screening has a similar scoring algorithm as SP; however, its use has been reported to give more detailed information about conformational analysis by providing extra details about docking orientations and reducing the toughness of the final torsional refinement and sampling. In a paper published in 2010 by Sciabola *et al.* it was mentioned that the use of GRID-MIFs (molecular interaction fields) and the FLAP (fingerprints for ligands and proteins) method in HTVS studies facilitated the ability to find unique ligands for target-receptor validation and for hit/lead identification (ligand-protein relationships) by using ligands that were structurally similar to drugs which have been previously reported to work in *in silico* or *in vitro* experiments.

The use of XP docking was first reported in 2006 by Friesner *et al.* as a novel addition to Glide 4.0 which incorporated new quantification terms like desolvation energy, hydrophobic enclosures, neutral- neutral and charged-charged interaction between hydrogen bonds , and also very importantly it excluded false positives 26. Triphati *et al.* (2013) used Glide XP docking to study the effect of 27 small 3,5-diaminoindazoles, imidazo(1,2-b)pyridazines, and triazolo(1,5 a)pyridazine derivatives for the inhibition of cyclin-dependent kinases (CDK2). The calculation results provided three possible CDK2 inhibitors from the initial 27 kinds and selected several

main binding postures based on their high-scoring binding affinity. The results were further corroborated by the use of MM/GBSA rescoring numbers that were consistent with literature²⁸.

CHAPTER III

EXPERIMENTAL METHODS

3.1 In-House GPR119 Receptor Homology Model Validation

The in-house GPR119 receptor homology model was validated through the docking of 76 GPR119 agonist ligands reported by Ritter *et al.* (2015). The ligands were constructed using Schrödinger Maestro simulation software considering the chirality and charge (if any) reported in the paper ²¹. The minimization was performed using force field OPLS3e and restraining bonds to metals around the input geometry. The calculation was carried on in no solvent using a distant dependent dielectric of 2(hydrophobic environment) and using charges from the force field. Extended cutoffs were used (8 Å van der Waals cutoff, 20 Å electrostatic cutoff and 4 Å hydrogen bond cutoff) and no constrains were applied on the ligand. The Polac-Ribier Conjugate Gradient (PRCG) minimization method was used. The maximum number of iterations was set to 5000, and the calculation convergence method was set gradient with a threshold of 0.05. The library of 76 agonists were further subjected to a LigPrep analysis that creates all the different chemical and structural possibilities a ligand can sample, before performing the virtual screening.

The GPR119 receptor homology model was extracted from four frames (frames 150, 200, 250, 300 and 350) of a 7-ns NAMD molecular dynamics (MD) simulation. These frames correspond to different conformations of the receptor, frame 150 where the transmembrane region is more compact to frame 350 where the receptor transmembrane region is more relaxed and the binding pocket is hydrate. The docking/virtual screening calculation was performed for

each one of the receptor conformations corresponding to the four frames. The virtual screening parameters were the following: pH was set to 7.4 ± 0.02 using Epik (to simulate the pH in the human body) and the high energy ionization/tautomer states were removed retaining up to four unspecified stereocenters and generate only one low energy conformer per ligand. The docking grid encompassing the receptor binding pocket was generated by Glide using the center of mass of the amino acids Phe157 (ECL2), Trp6.48²³⁸, Arg7.36²⁶², Trp7.39²⁶⁵, Cys155 (ECL2) and Arg3.28⁸¹ that line the pocket. A combination of HTVS, SP, and XP docking algorithms was used keeping compounds with a Glide score of -3 kcal/mol and above.

3.2 Library Preparations and Computational Screening

Four databases were selected for the purpose of this study. These databases were Prestwick Phytochemicals, SelleckChem GPCRs, Prestwick GPCR library and Zinc Naturals (ZINC15) library. The Prestwick Phytochemicals library contained 320 compounds from which 41 were manually selected for the docking experiments. Charged compounds and compounds with a MW<200 g/mol were excluded. The SelleckChem GPCRs (738 compounds) and Prestwick GPCR library (265 compounds) totaling 1003 compounds were used as well. However, upon inspection it was concluded that 141 compounds repeated between these libraries and were excluded which reduced the number of testing ligands to 862. The Zinc Naturals library contains 120 million compounds, but only the first $20,000$ were selected for docking³⁶.

The approximately 21,000 compounds were visually examined for missing hydrogens which were added; they were desalted and then they were energy minimized using the parameters covered in section 3.1. Once the ligands were prepared, they were used as input for docking on the GPR119 receptor homology model conformers from frames 150, 200, and 250. The docking parameters used were those mentioned in section 3.1. The docking calculations were done in two way: a) the automated sequential virtual screening protocol that takes the results of HTVS step and uses them as input for the SP step, and the SP results are used as input for the XP step and b) docking all the ligands using each one of the three docking methods separately (HTVS, SP and XP).

In addition to the libraries virtual screening calculations, four more molecules were built and docked in the GPR119 receptor because we had a certain amount of each in the lab and we can test them using a cAMP colorimetric assay. Of the four molecules, oleoyl serinol (OS), SRT1720, and AR231453 were purchased from Sigma-Aldrich and AR437735 was synthesized by a collaborator. AR231453, AR437735 and oleoyl serinol are known GPR119 receptor agonists whereas SRT1720 has not been tested for activity against the GPR119 receptor.

3.3 Maintaining of Eukaryotic HEK293 Cells

Human embryonic kidney cells (HEK293) were used to measure activation of the GPR119 receptor. The HEK293 cells were allowed to grow in 10 cm plates containing Dulbecco's Modified Eagle Medium (DMEM) and 10% fetal bovine serum (FBS), and 1% penicillin/streptomycin and 1% antibiotic-antimitotic additives were added in order to prevent bacterial and fungal contamination. Subsequently, the cells were placed in an incubator (5% CO2, humidified at 37°C) and left to grow for 3-4 days or until they reached a 90% confidence. Once this point was reached they were ready to be passaged to a new plate.

The passaging consisted of carefully aspirating the medium from the plate while avoiding disruption of the cells attached to the plate. Then 5 mL of trypsin solution (0.05% trypsin, 0.5m MEDTA) to lift the cells from the plate. Using a 5 mL pipette the trypsin was pipetted up and
down about 12 times until the cell agglomerations were broken up and well dispersed in the solution. From this mixture 1-1.2 mL was added to a new plate that contained 10 mL of DMEM with the respective additives 10% FBS, 1% pen/strep, 1% anti-anti. The plate was then rocked front and back as well as side to side to assure a good dispersion of the cells in the new medium.

3.4 Preparing HEK293 Cells for Transfection

The medium was aspirated from the plate and 5 mL of trypsin was added to homogeneously suspend the cells. The mixture of trypsin and cells was then transferred to a 50 mL centrifuge tube. To prevent the further digestion of the HEK293 cells by the trypsin, 7.5 mL of DMEM containing 10% FBS was added to the tube. The 50 mL tube was centrifuged at 1000 RPM at 25 °C for 5 minutes to form a cell pellet at the bottom of the tube. After centrifugation the medium was aspirated from the tube and the cells were resuspended in 10 ml of DMEM (10% FBS).

A 20 µL micropipette was used to acquire 15 µL of the resuspended cells and dispense it to a hemacytometer. The hemacytometer was used to estimate the concentration of the cells in the centrifuge tube. The cells were counted from 4 base quadrants each containing 4 smaller ones within. The cells directly on top of the divisor lines between the major 4 quadrants were not considered. Once the total cells were accounted they were averaged to get the number of cells in 100 nL per quadrant. The averaged number was then multiplied by 10,000 to convert the the nanoliters to cells per milliliter. The resulting number was then multiplied by the total milliliters in the 50 mL centrifuge tube and divided by 250,000 (amount of cells desired per mL). This

provided the amount of DMEM required to deliver the desired amount of cells per mL of medium (dilution formula $M_1V_1=M_2V_2$).

Once the new volume of DMEM was added to the 50 mL test tube, often to an approximate volume of \sim 42 mL, 2 mL of the homogeneously mixed cells were pipetted into each well of three 6-well plates for a final concentration of 500,000 cells/well. The completed 6-well plates were rocked gently to evenly disperse the cells around the wells and then they were placed in the incubator (5% CO_2 , humidified at 37 $^{\circ}$ C) and allowed them to grow for 2 days.

Figure 5. Pre-transfection procedure of HEK293 cells.

3.5 Transfection of HEK293 Cells with Wild Type GPR119 DNA

This procedure was based on the protocol provided by the Invitrogen LipofectamineTM transfection kit. Eight tubes were labeled A and four were labeled B. Tubes A contained 125 μL of lipofectamineTM and 3.75 μL lipofectamine 3000 reagent (mixed on the vortex) and tubes B contained 250 μL of lipofectamineTM, 5 μL of p3000 reagent and 2.5 ng of DNA of the wild type GPR119 receptor. Tubes B were mixed gently and no more than 5 times due to the fragility of the DNA. DNA from tubes $B(125 \mu L)$ was then added to each tube A and mixed by gentle pipetting. Subsequently, the lipofectamine-DNA complex was left to form for 30 minutes.

The 6-well plates were removed from the incubator and 16 wells were transfected with the wild type GPR119 receptor DNA while two wells were left as controls; the plates were then placed in the incubator for four hours. The medium was aspirated from the plates and 2 mL of DMEM (10% charcoal-stripped fetal bovine serum) was added to each well. The switch to charcoal-stripped FBS at this point is important to remove undesired activating molecules from the cell medium. Charcoal-stripped FBS is made by passing the serum through a bed of charcoal, which absorbs hydrophobic compounds. This removes compounds such as steroids and lipids which can act as activating ligands for GPCRs or nuclear receptors.

Figure 6. Transfection of HEK293 cells with wild type GPR119 DNA.

3.6 Drug Treatment of GPR119 Transfected HEK293 Cells

The compounds AR231453, AR437735, SRT1720 and oleoyl serinol were each dissolved

in dimethyl sulfoxide (DMSO) to a final concentration of 10 mM. Since it is known that

AR231453 is a potent and recognized agonist for the GPR119 receptor, it was selected as a positive control and the performance of the other ligands were compared to the AR231453 activity in the cAMP production assay. The stock solutions for AR231453 and the compounds of interest were made by mixing 1 mL DMEM, 65.2μ of charcoal-stripped FBS and 10 μ L of specific ligand in a 50 mL centrifuge tube. The non-stock solution on the other hand was made combining 14 mL of DMEM, 350 μ l of charcoal-stripped FBS and 250 μ L of DMSO in a 50 mL tube. The 'stock' and 'non-stock' vials are shown in **Figure 7**.

 An aliquot from the stock of AR231453 was serially diluted to form the following concentrations: 10000 nM, 1000 nM, 100 nM, 10 nM, 1.0 nM, 0.1 nM, and 0.01 nM. After the serial dilutions, the 6-well plates were removed from the incubator and the medium was removed by aspiration. The cells were then washed with 1.0 mL of Hank's balanced salt solution (HBSS) followed by the careful addition 1.0 mL of DMEM to each well. Finally, 1 mL of the desired ligand concentration was pipetted into the wells and the plates were placed in the incubator for 30 minutes.

Once the incubation time passed, the cells were scraped from the plates and deposited in accordingly labeled 15 mL tubes. The tubes were then centrifuged at a temperature of 21 °C for 5 minutes at 1000 RPM. Upon completion of the cycle, the medium from the tubes was aspirated and the formed cell pellet was resuspended in 2.0 mL of a phosphate buffer solution (PBS). The cells were then centrifuged again under the same conditions used previously. The PBS was aspirated from all the 15 mL tubes and 286 µL of cold lysis buffer (thimerosal 0.01%, Triton X-100 2.0%) was added to the cells in a way that the pellet was not disturbed. The sample tubes were kept on an ice to prevent degradation of cAMP by cellular enzymes.

Figure 7. Drug treatment of transfected HEK293 cells.

3.7 cAMP ELISA Colorimetric Assay

The contents of the 15 mL centrifuge tubes (286 μ L lysis buffer + samples) were transferred to smaller scale 1.5-mL microfuge tubes for easier handling. The samples were then thoroughly mixed using a vortex and they were subjected to two cycles of freeze (dry ice/ethanol) and thaw (37°C water bath) that lasted 3 minutes each. The cells were then centrifuged at 13,000 RPM at 4°C for 10 minutes. Subsequently, 200 μL of the supernatant was transferred to a new microfuge tube that was placed in an ice bath to minimize degradation of cAMP. The 200 µl of cell lysate sample was meticulously mixed 6 times and 100 μL was transferred to a third and final 1.5-mL microfuge tube.

A dilution (1:100) of peroxidase cAMP tracer conjugate was made by using 12 μ l of tracer solution and 1188 µl of assay diluent (Thimerosal 0.01%) and 50 µL of this dilution was added to each of the sample tubes. A set of cAMP standards were prepared based on the

manufacturer's protocols and they were added in duplicate directly to the goat anti-rabbit antibody-coated plate (columns 1 & 2 in **Figure 8**) followed by the addition of 25 uL of the diluted peroxidase cAMP tracer conjugate to each cell.

The samples were then added in duplicates to the goat anti-rabbit antibody-coated plate. Columns 3 and 4 were reserved for the AR231453 treated cell extracts and columns 5 and 6 were for the cell extracts treated with the comparator ligand (refer to **Figure 8** for identification). A $(1:500)$ dilution of rabbit anti-cAMP polyclonal antibody was made using 4.8 µl of rabbit anticAMP polyclonal antibody stock and 2395.2 μ L of assay diluent. 50 μ L of this rabbit antibody dilution was then transferred to each of the testing wells.

Figure 8. Procedure before tracer/rabbit cAMP assay reaction.

The goat anti-rabbit antibody-coated plate was then placed on an orbital shaker for 2 hours. After the reaction, the contents of the wells were decanted and the wells were washed a total of six times with 200 μL of wash buffer (Thimerosal 0.02%). Subsequently, 100 μL of the substrate solution was added to each well. If the experiment went correctly, the contents of the cells containing low concentrations of cAMP would turn blue. The plate was returned to the orbital shaker to mix for an additional 15 minutes. The reaction was stopped by adding 100 μ L of the stop solution (0.5 N sulfuric acid) to each of the wells which would turn the contents yellow. The plate was then spectroscopically analyzed using a BioRad 480 micro-plate reader at a primary wavelength of 450 nm.

Figure 9. cAMP assay after tracer/rabbit reaction

3.8 Data Analysis of cAMP ELISA

Microsoft excel was used to create an absorbance vs concentration plot obtained from the cAMP standards. Beer's law was then employed to use the absorbance of the wells in order to determine the approximate concentration of cAMP produced in each lysate sample. SigmaPlot 11 (Systat Software Inc., San Jose, CA) was the graphical software employed to create a sigmoidal dose-response curve considering the production of cAMP in nM vs. the log concentration of the test compounds. The cAMP levels produced by the controls were used to normalize the values of cAMP produced by GPR119 activation. The cAMP values of the

samples were also scaled and normalized to the plateau of cAMP achieved by the ligand (100%). The sigmoidal dose-response curve was set to produce a hillslope of 2.0 and the values of log EC50 and EC50 produced by SigmaPlot 11 were converted to concentration in nM.

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Homology Model Validation

Due to the lack of a crystallographic structure for the GPR119 receptor, an in-house homology model was built using Schrödinger software by taking fragments of several type A GPCR receptors such as the cannabinoid 1, the adenosine A_{2A} and β 2 adrenergic receptors that have similar sequence and motifs, as well as structurally verifying the model with those provided by the literature 4, 8. In general, homology models tend to vary in some degree based on the templates used for modeling of the structures, sequence alignment, model building and model refinement³⁷. Therefore, it was imperative to validate the in-house built homology model through the use docking results before applying it to more extensive experiments.

The 76 molecules obtained from Ritter *et al.* paper were all found in the published literature from pharmaceutical companies like Arena, Pfizer, Roche, Bristol Meyer Squibb (BMS), Merck and AstraZeneca. The molecules varied in size and composition and among them there were agonists, inverse agonists and antagonists²¹. From a 7-ns MD simulation of the GPR119 receptor homology model embedded in a hydrated lipid bilayer patch, 5 frames were obtained and designated Frame 150, 200, 250, 300 and 350. The receptor structure was extracted from each frame and used for the docking calculations. **Table 1** shows the docking results obtained for each of the frames. For the majority of the compounds, the proper name of the

structure was not supplied by the paper. Instead, an identity number was provided, which was used to locate their common and/or UPAC name which was used for easier identification and comparison.

Frame 150 had the tightest transmembrane region (or least relaxed) of the five GPR119 homology model structures. This is reflected by allowing the least number of ligands to fit the receptor, In contrast, the GPR119 homology model structure from frame 350 was the most relaxed and open. The percentage of ligands that fit in the receptors were 78%, 87%, 92%, and 97% for frames 150, 200, 250, and 350 respectively. It was evident that as the receptor relaxed and opened to a greater extent, it allowed more complex ligands to fit. However, frame 300 was an exception since only 64 out of 76 ligands fit (84%). A possible explanation of this discrepancy might be that, although the receptor structure from frame 300 appears to have relaxed and have a large binding pocket, the amino acids lining the binding pocket (such as the large $Trp7.39^{265}$) have enough room to swing into the binding pocket partially obstructing it.

From the docking results, it was verified that the GPR119 homology model met the criteria for a relatively effective model, since it allowed a high percentage of known agonists to fit into the binding pocket. It is important to mention that the receptor was able to fit ligands that were long, slender and aromatic, just like the most commonly used high-potency agonist on the market, AR231453. However, the binding pocket was also able to accommodate large and bulky molecules, which affords the possibility that a wider number of molecules may be tested and new scaffolds may be discovered. Since it is understood that the GPR119 receptor is not a static structure it was necessary to take the results of a combination of less open/relaxed and more open/relaxed frames to make the most robust study. Frames 150, 200 and 250 were selected to be used for the library analysis.

Table 1. Extra precision results for the GPR119 receptor model from the molecular dynamics simulations frames 150, 200, 250 and 350. The 76 tested ligands can be accessed through DOI:10.1021/acs.jmedchem.5b01198

4.2 Computational Analysis

Prestwick phytochemicals offered 320 total molecules, from which 42 were handpicked for the docking experiments. The criteria for the selection was based on size and polarity. If a ligand had a molecular weight less than 200 g/mol, the assumption was that it would not be big or long enough to interact with the conserved Trp6.48 residue found in TMH 6, which acts as a toggle switch for the activation of the GPR119 receptor⁹. The second point was charge; if a molecule was charged, the likelihood of the ligand entering the binding pocket was low, due to the overall non-polar environment created by the seven transmembrane helices. Therefore, these 42 molecules had the optimal size and charge and were docked to each of the GPR119 receptor frames.

For clarity in the display of the results, the top ten best Glide scoring molecules were selected for each frame analyzed; the rest of the molecules are shown in the Appendix. Out of the approximately 21,000 molecules analyzed through the sequential virtual screening protocol (automated) using HTVS, SP and XP, 1,632 compounds docked in the frame 150 GPR119 receptor structure (7.8%); 1,735 compounds docked in the frame 200 GPR119 receptor structure (8.3%) and 1,701 compounds docked in the frame 250 GPR119 receptor structure (8.1%).

A second computational study was done in parallel to the automated sequential virtual screening study where the ligands were docked using HTVS by itself, the results obtained from that algorithm were used as input for a SP docking calculation, and subsequently, the results from the SP docking where used as input for XP docking.

This stepwise (manual) set of calculations were done to see if the resulting molecules were the same for both analyses and to compare the time it took to perform the automated and the stepwise calculations. The results of the stepwise analysis for frames 150, 200 and 250 were

1,135 (5.4%), 1,217 (5.8%), and 1,179 (5.6%) molecules, respectively. The difference in the results between the two computational methods might be due to the variation in parameters in the algorithms compared to when they ran using the sequential virtual screening protocol. The variation is most probably due to the number of conformations retained for each ligand. When the docking algorithms are run through the automatic protocols, some of the options are preset compared to when using just one of the docking methods. Also, since HTVS and SP use the same algorithm, but different docking criteria, it is safe to say that the results would not vary much between these steps. However, when run together with more specific restrictions, like those imposed by XP, the calculation might not have been able to eliminate double positives or penalize unfavorable interactions^{20, 34, 38}.

For the time analysis comparison, the stepwise experiment took longer than the automated sequential virtual screening protocol. For example, for the 20,000 compounds of the Zinc library, the stepwise experiment the HTVS calculation took approximately 4 hours, the SP calculation took 6 hours and the XP calculation took 32 hours for a total of 42 hours. In the automated sequential virtual screening analysis, the complete calculation took approximately 34 hours. This time difference was not as noticeable with the Prestwick phytochemical library, which only contained 42 compounds, since both the stepwise and the automated docking protocols took around 5 hours. Even though the automated sequential virtual screening calculation might take less time than the stepwise calculations, it is important to understand that false positives or different docking poses of the same structure might be present.

The analysis of the results of the Prestwick phytochemicals for each of the GPR119 receptor structures from the chosen frames are represented in **Tables 2, 5** and **8**. Molecules that tended to repeat between frames makes them attractive candidates for in vitro analysis, because

they fit into the GPR119 receptor (Glide scores -12.434 to -7.940 kcal/mol), regardless of how closed or open it was. The molecules that repeated between the three frames were verbenalin (average Glide score -10.03 kcal/mol), 4,4'-(2,3-dimethyltetramethylene)dipyrocatechol (average Glide score -10.31 kcal/mol) and chlorogenic acid (average glide score -9.763 kcal/mol).

For the SelleckChem and Prestwick GPCR libraries, it can be seen in **Tables 3, 6** and **9** that the compounds that got the highest (more negative) Glide scores varied from frame to frame. However, just like for the Prestwick phytochemicals library, a trend was observed, where the top compounds tended to repeat more often between frame 150 and frame 200, and at a reduced frequency in frame 250. It may be that this trend is associated with the state of the receptor, since for frames 150 and 200 the receptor is in a more closed conformation, before the binding pocket becomes more hydrated as in frame 250. Therefore, similar intermolecular forces should be found among them. The highest Glide score achieved in the GPCR libraries was obtained for D-Glucitol, 1,5-anhydro-1-*C*-[3-[[5-(4-fluorophenyl)-2-thienyl]methyl]-4-methylphenyl]-, (1*S*)- (- 12.87 kcal/mol) and the lowest was by dobutamine (-8.477 kcal/mol), both from frame 250.

Tables 4, 7 and **10** show the results for the Zinc Naturals library. Most of the names for this library were found by searching databases and literature; nevertheless, some of them were not able to be identified. For the compounds for which names were not found the Zinc ID was provided. As shown in **Tables 4, 7** and **10,** the compounds varied in molecular weight from the mid 200 Da to the high 500 Da, which highlights the versatility of the GPR119 receptor for interacting with both simple and more complex molecules. The Glide scores were among the highest (-11.68 kcal/mol to -13.57 kcal/mol) for each of the frames, which signifies that the ligands from Zinc Naturals library are more likely to fit better in the binding pocket of the GPR119 receptor compared to the other libraries.

The molecular structure of the ligands for each library is presented right after the corresponding tables. In frame 150, the molecules selected tended to be long and slender, similar to AR231453 and its derivates. As expected, as the receptor relaxed and opened up, as in frame 250, the ligands selected showed increased structural complexity. Accordingly, in the more open frames, the sizes of the ligands covered a wider range of molecular weight, with the smallest molecule being chrysin (MW 254 g/mol) and the largest one being naringin (MW 580 g/mol). The molecules selected were also increasingly aromatic and some structures even presented extended ring structures containing 7,8 or even 9 atoms.

Frame 150

Table 2. 10 best Glide scoring compounds for Prestwick phytochemicals in Frame 150

Figure 10. Best glide scoring Pretwick phytochemicals molecules for Frame 150. a) verbaline, b) 4,4'-(2,3-dimethyltetramethylene)dipyrocatechol), c) baicalin, d) glycocholic acid, e) chlorogenic acid, f) (+) s-camptothecine, g) menaquinone, h) (-) α -lobeline, i) glycyrrhetinic acid and j) capsaicin

Table 3. 10 best Glide scoring compounds for the GPCRs libraries in Frame 150

Figure 11. Best glide scoring for GPCR molecules for Frame 150. a) empagliflozin, b) dapagliflozin, c) phlorizin, d) shanzhiside, e) flibanserin, f) netupitant, g) formoterol, h) ketanserin, i) droperidol and j) terazosin

Table 4. 10 best Glide scoring compounds for the Zinc Naturals library in Frame 150

Figure 12. Best glide scoring for zinc library molecules for Frame 150. a) ZINC31155896, b) ZINC35457506, c) ZINC36728548, d) ZINC03842067, e) ZINC05414350, f) ZINC31155429, g) ZINC35442872, h) ZINC35442868, i) ZINC01667455, and j) ZINC35457485

Frame 200

Table 5. 10 best Glide scoring compounds for Prestwick phytochemicals in Frame 200

Figure 13. Best glide scoring Pretwick phytochemicals molecules for Frame 200. a) naringin, b) glycocholic acid, c) curcumin, d) baicalin, e) (+) s-camptothecine, f) 4,4'-(2,3 dimethyltetramethylene)dipyrocatechol, g) calciferol, h) verbenalin, i) chlorogenic acid and j) (-)-cinchonidine

Figure 14. Best glide scoring for GPCR molecules for Frame 200. a) azilsartan medoxomil, b) ipragliflozin, c) phlorizin, d) penfluridol, e) tropisetron, f) pirenzepine, g) formoterol, h) risperidone, i) 1-azoniabicyclo[2.2.2]octane, 4-[(2-hydroxy-2,2-diphenylacetyl)oxy]l and j) Fluphenazine.

Table 7. 10 best Glide scoring compounds for the Zinc Naturals library in Frame 200.

Figure 15. Best glide scoring for zinc molecules for Frame 200. a) ZINC33838191, b) ZINC31163744, c) ZINC04235989, d) ZINC20463632, e) ZINC35457506, f) ZINC06041521, g) ZINC08662732, h) ZINC35415777, i) ZINC31155532, and j) ZINC08662730.

Frame 250

Table 8. 10 best Glide scoring compounds for Prestwick phytochemicals in Frame 250.

Figure 16. Best glide scoring Prestwick phytochemicals molecules for Frame 250. a) 4,4'-(2,3 dimethyltetramethylene) dipyrocatechol, b) chlorogenic acid, c) verbenalin, d) capsaicin, e) abietic acid, f) berlambine, g) chrysin, h) ajmalicine, i) naringenine and j) olivacine.

Table 9. 10 best Glide scoring compounds for the GPCRs libraries in Frame 250.

Figure 17. Best glide scoring for GPCR molecules for Frame 250. a) d-Glucitol, 1,5-anhydro-1- *C*-[3-[[5-(4-fluorophenyl)-2-thienyl]methyl]-4-methylphenyl]-, (1*S*)-, b) 5-Heptenoic acid, 7- [(1*R*,2*R*,3*R*,5*S*)-3,5-dihydroxy-2-[(3*R*)-3-hydroxy-5-phenylpentyl]cyclopentyl]-, 1-methylethyl ester, (5*Z*)-, c) 1-Azoniabicyclo[2.2.2]octane, 4-[(2-hydroxy-2,2-diphenylacetyl)oxy], d) fluphenazine, e) cloperastine, f) ketanserin, g) darifenacin, h) carvedilol, i) Piperidinium, 3-[(2 hydroxy-2,2-diphenylacetyl)oxy]-1,1-dimethyl-, (3*R*)-, j) dobutamine.

Table 10. 10 best Glide scoring compounds for the Zinc Naturals library in Frame 250.

Figure 18. Best glide scoring for zinc molecules for Frame 250. a) ZINC36728548, b) ZINC31155902, c) ZINC04236655, d) ZINC35457506, e) ZINC31155664, f) ZINC04236552, g) ZINC35457671, h) ZINC08681833, i) ZINC04236634, j) ZINC04236575.

4.3 Wet Lab Ligand Selection and Computational Results

After the computational analysis performed on the chemical libraries, the procedure was implemented on 4 molecules that were already in wet lab. Three of these molecules AR231453, AR437735, and oleoyl serinol shown in **Figure 19** have been reported in the literature to be agonists for the GPR119 receptor^{4, 18, 21, 39}. AR437735 was made by a collaborator's organic

synthesis group, and this molecule is a variation of the consolidated agonist AR231453, both previously patented by Arena pharmaceuticals. Oleoyl serinol was chosen for the experiment due to its close resemblance to the natural body activator of GPR119, oleoylethanolamide (OEA), and lastly SRT1720 was a novel molecule chosen for the experiment even though it has never been studied for the interaction with this specific receptor. It has been reported that SRT1720, shown in **Figure 19,** is a selective activator of the SIRT1 receptor belonging to sirtuin family of proteins which are associated with diseases like aging, metabolism impairment and periapical peridionitis. Therefore, it would be interesting to see if this compound could be repurposed for a new receptor 40-41.

Figure 19. Molecular structures for the molecules used in the wet lab experiments

Table 11 presents the computational results of the automated sequential virtual screening protocol analysis based on the GPR119 homology model structure from frames 150, 200, 250 and 300. In Frame 150, the only molecules that fit were oleoyl serinol and SRT1720; oleoyol serinol is a lipid-like molecule that can fit in the receptor with no problem since the GPR119

tends to accept molecules with this type of structure. However, SRT1720, as seen in **Figure 12,** is bulky and elongated. Typical observations of the almost closed receptor at Frame 150 would suggest that SRT1720 would not be able to fit in the pocket but perhaps the intermolecular interactions with the α helical bundle are favorable for the structure in this stage of the homology model. Frame 200 allowed all molecules with exception of AR437735, and frames 250 and 300, which were more open and the binding pocket more hydrated, were able to fit the four molecules. To create some consistency between the results of the library and give a realistic image of the receptor in the body, frame 250 was selected for the automated sequential virtual screening protocol analysis.

Table 11. The automated sequential virtual screening protocol results for oleoyl serinol, AR231453, AR437735 and SRT1720.

The results of the computational analysis done in Frame 250 can be seen in **Table 12**. The highest glide score (-7.961 kcal/mol) was found to be for oleoyl serinol, which was not surprising since as previously mentioned, it has a very similar chemical structure to OEA, predicting its favorable fitting into the binding pocket of GPR119. AR231453 and AR437735 had very similar Glide scores, which could be attributed to AR147735 having a chemical core based on AR231453, as shown in **Figure 12**. SRT1720, on the other hand, had one of the lowest glide scores of the overall computational analysis $(\sim 21,000$ compounds) at -3.758 kcal/mol, which indicates that the molecule can fit in the receptor binding pocked, but that the fit would be more strained than for the other compounds analyzed.

Table 12. Drugs used for the ELISA cAMP colorimetric assay.

4.4 *In Vitro* **Studies**

Once the results from the computational analysis indicated that three of the four molecules in **Figure 19** (oleoyl serinol, AR231453, and AR437735) would more than likely fit into the receptor, cAMP ELISA colorimetric assays were performed for each compound. **Figure 20** shows the results of the oleoyl serinol vs the AR231453. As can be seen from the graph, both ligands present agonistic activity, varying only in the EC_{50} values (half maximal response achieved by concentration of drug). In the cAMP analysis for oleoyl serinol the calculated EC_{50} was 96 nM which correlates with the good Glide score, -7.96 kcal/mol, from the computational analysis. Also, the experimental EC_{50} of oleoyl serinol was comparable, though better, to 1.6 μ M, the EC₅₀ value reported for Cohen *et al.* for a cAMP assay using ACTOne HEK293 cells transfected with GPR119 receptor. The difference in values of the EC_{50} s for oleoyl serinol for the cAMP assay might be due to differences in the procedures. For example the Cohen *et al.* paper does not mention that they used charcoal-stripped FBS to prevent early activation of the receptor 42 .

Figure 20. Dose-response curve of wild type HEK293 cells treated with AR231453 and oleoyl serinol and ligand effects on cAMP production.

Figure 21 demonstrates the agonistic activity of AR437735 and from the image it can be seen that there are slight variations between the sigmoidal plots of both tested ligands, with

AR437735 tending to have greater standard deviations and some stray points like the one found at a concentration of $1.0X10^{-5}$ M. Nonetheless, both compounds show a similar activation of the receptor; this is confirmed by comparing the Glide scores which were -7.286 kcal/mol for AR231453 and -7.186 kcal/mol for AR437735. Literature EC_{50} s were also in accordance with the results, providing a value of 10.5 nM for AR231453 and 0.1-1 nM for AR437735. More than likely, the difference in the results could be attributed to differences in the experimental methods.

Figure 21. Dose-response curve of wild type HEK293 cells treated with AR231453 and AR437735 and ligand effects on cAMP production.

The SRT1720 vs AR231453 results are shown in **Figure 22**. The SRT1720 ligand had never been tested with the GPR119 receptor even though it has been reported that its target receptor, SRIT1, could be involved in the activity of other GPCR receptors, like GPR30, which regulates the GPER pathway (breast cancer influencer)⁴³. From **Figure 22**, it can be seen that SRT1720 did not promote any receptor-dependent cAMP production, since its activity is in the

zero range. This statement is also shown in **Table 12** where no EC₅₀ was able to be calculated from the results of the ligand. The computational analysis for SRT1720 suggested that the ligand may fit into the receptor's binding pocket since a Glide score was reported. However, the Glide score is move positive, -3.758 kcal/mol, compared to the agonists suggesting that SRT1720 may not fit well in the binding pocket and so it was not able to activate the GPR119 receptor.

Figure 22. Dose-response curve of wild type HEK293 cells treated with AR231453 and SRT1720 and ligand effects on cAMP production.

CHAPTER V

CONCLUSION

GPR119 is a transmembrane receptor that belongs to the type A GPCR family and is involved in the regulation of insulin levels and other hormones. Due to its influence on insulin secretion, research has focused on finding possible drugs to regulate its activity. Pharmaceutical companies have the ability to computationally screen millions of drugs using homology models to find agonists like AR231453, which activate the receptor without major secondary effects. However, their search tends to be solely based on HTVS, a "quick-and-dirty" algorithm.

This study proposed the possibility of using three computational screening algorithms (high throughput virtual screening, standard precision and extra precision virtual screening) in sequentially, based on an in-house homology model of the GPR119 receptor for the efficient and cost-effective analysis of 21,000 chemical compounds from 4 chemical libraries. From the automated sequential virtual screening protocol results, approximately \sim 2,100 compounds fit the GPR119 receptor with promising Glide scores (likelihood of fitting in a binding pocket) varying from -3.20 kcal/mol to -13.80 kcal/mol.

The compounds AR231453, AR437735, oleoyl serinol and SRT1720 available in the lab, were tested both computationally (virtual screening using Glide docking) and experimentally using a cAMP ELISA colorimetric assay to test the correlation between virtual screening/Glide docking results of the ligands and their ability to induce GPR119 mediated cAMP production. The results showed that AR231453, AR437735 and oleoyl serinol, which had a more negative Glide score compared to SRT1720, promoted agonistic activity in the receptor and their EC_{50}

62

values were in close accordance with the literature. SRT1720 on the other hand, which had a more positive Glide score, did not promote any cAMP build up. This result might suggest either that SRT1720 does not fit well in the receptor binding pocket resulting in no activation, or that it could be an antagonist for the receptor. However, competitive inhibition experiments would need to be made to prove this characterization of SRT1720.

REFERENCES

- 1. Zhou, Q., et al., Common Activation Mechanism of Class a Gpcrs. *eLife* **2019**, *8*, e50279.
- 2. Latorraca, N. R.; Venkatakrishnan, A. J.; Dror, R. O., Gpcr Dynamics: Structures in Motion. *Chem Rev* **2017**, *117*, 139-155.
- 3. Lameh, J.; Cone, R. I.; Maeda, S.; Philip, M.; Corbani, M.; Nádasdi, L.; Ramachandran, J.; Smith, G. M.; Sadée, W., Structure and Function of G Protein Coupled Receptors. *Pharmaceutical Research* **1990**, *7*, 1213-1221.
- 4. Engelstoft, M. S.; Norn, C.; Hauge, M.; Holliday, N. D.; Elster, L.; Lehmann, J.; Jones, R. M.; Frimurer, T. M.; Schwartz, T. W., Structural Basis for Constitutive Activity and Agonist-Induced Activation of the Enteroendocrine Fat Sensor Gpr119. *British Journal of Pharmacology* **2014**, *171*, 5774-5789.
- 5. Tyurenkov, I. N.; Ozerov, A. A.; Kurkin, D. V.; Logvinova, E. O.; Bakulin, D. A.; Volotova, E. V.; Borodin, D. D., Structure and Biological Activity of Endogenous and Synthetic Agonists of Gpr119. *Russian Chemical Reviews* **2018**, *87*, 151-166.
- 6. Keshelava, A.; Solis, G. P.; Hersch, M.; Koval, A.; Kryuchkov, M.; Bergmann, S.; Katanaev, V. L., High Capacity in G Protein-Coupled Receptor Signaling. *Nature Communications* **2018**, *9*, 876.
- 7. Dunham, J. H.; Hall, R. A., Enhancement of the Surface Expression of G Protein-Coupled Receptors. *Trends in biotechnology* **2009**, *27*, 541-5.
- 8. Latorraca, N. R.; Venkatakrishnan, A. J.; Dror, R. O., Gpcr Dynamics: Structures in Motion. *Chemical Reviews* **2017**, *117*, 139-155.
- 9. Holst, B.; Nygaard, R.; Valentin-Hansen, L.; Bach, A.; Engelstoft, M. S.; Petersen, P. S.; Frimurer, T. M.; Schwartz, T. W., A Conserved Aromatic Lock for the Tryptophan Rotameric Switch in Tm-Vi of Seven-Transmembrane Receptors. *J Biol Chem* **2010**, *285*, 3973-3985.
- 10. Norn, C.; Hauge, M.; Engelstoft, M. S.; Kim, S. H.; Lehmann, J.; Jones, R. M.; Schwartz, T. W.; Frimurer, T. M., Mutation-Guided Unbiased Modeling of the Fat Sensor Gpr119 for High-Yield Agonist Screening. *Structure* **2015**, *23*, 2377-2386.
- 11. Chu, Z.-L., et al., N-Oleoyldopamine Enhances Glucose Homeostasis through the Activation of Gpr119. *Molecular Endocrinology* **2010**, *24*, 161-170.
- 12. Flock, G.; Holland, D.; Seino, Y.; Drucker, D. J., Gpr119 Regulates Murine Glucose Homeostasis through Incretin Receptor-Dependent and Independent Mechanisms. *Endocrinology* **2011**, *152*, 374-383.
- 13. Pamir, N., et al., Glucose-Dependent Insulinotropic Polypeptide Receptor Null Mice Exhibit Compensatory Changes in the Enteroinsular Axis. *American journal of physiology. Endocrinology and metabolism* **2003**, *284*, E931-9.
- 14. Odori, S., et al., Gpr119 Expression in Normal Human Tissues and Islet Cell Tumors: Evidence for Its Islet-Gastrointestinal Distribution, Expression in Pancreatic Beta and Alpha Cells, and Involvement in Islet Function. *Metabolism* **2013**, *62*, 70-78.
- 15. Tough, I. R.; Forbes, S.; Herzog, H.; Jones, R. M.; Schwartz, T. W.; Cox, H. M., Bidirectional Gpr119 Agonism Requires Peptide Yy and Glucose for Activity in Mouse and Human Colon Mucosa. *Endocrinology* **2018**, *159*, 1704-1717.
- 16. Zhang, M., et al., High-Throughput Screening for Gpr119 Modulators Identifies a Novel Compound with Anti-Diabetic Efficacy in Db/Db Mice. *PLOS ONE* **2013**, *8*, e63861
- 17. Friesner, R. A., et al., Glide:  A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **2004**, *47*, 1739-1749.
- 18. Shiri, F.; Teymoori, M., In Silico Approaches to Explore Structure of New Gpr 119 Agonists for Treatment of Type 2 Diabetes Mellitus. *Medicinal Chemistry Research* **2017**, *26*, 947-961.
- 19. Walters, W. P., Virtual Chemical Libraries. *Journal of Medicinal Chemistry* **2019**, *62*, 1116-1124.
- 20. Zhang, X.; Betzi, S.; Morelli, X.; Roche, P., Focused Chemical Libraries--Design and Enrichment: An Example of Protein-Protein Interaction Chemical Space. *Future medicinal chemistry* **2014**, *6*, 1291-307.
- 21. Ritter, K.; Buning, C.; Halland, N.; Pöverlein, C.; Schwink, L., G Protein-Coupled Receptor 119 (Gpr119) Agonists for the Treatment of Diabetes: Recent Progress and Prevailing Challenges. *Journal of Medicinal Chemistry* **2016**, *59*, 3579-3592.
- 22. Valentin-Hansen, L.; Holst, B.; Frimurer, T. M.; Schwartz, T. W., Phevi:09 (Phe6.44) as a Sliding Microswitch in Seven-Transmembrane (7tm) G Protein-Coupled Receptor Activation *. *Journal of Biological Chemistry* **2012**, *287*, 43516-43526.
- 23. Ghosh, S.; Nie, A.; An, J.; Huang, Z., Structure-Based Virtual Screening of Chemical Libraries for Drug Discovery. *Current Opinion in Chemical Biology* **2006**, *10*, 194-202.
- 24. Dragojlovic, V., Conformational Analysis of Cycloalkanes. *ChemTexts* **2015**, *1*, 14.
- 25. Mayr, L. M.; Bojanic, D., Novel Trends in High-Throughput Screening. *Current opinion in pharmacology* **2009**, *9*, 580-8.
- 26. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra Precision Glide:  Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *Journal of Medicinal Chemistry* **2006**, *49*, 6177-6196.
- 27. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide:  A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry* **2004**, *47*, 1750-1759.
- 28. Tripathi, S. K.; Muttineni, R.; Singh, S. K., Extra Precision Docking, Free Energy Calculation and Molecular Dynamics Simulation Studies of Cdk2 Inhibitors. *Journal of theoretical biology* **2013**, *334*, 87-100.
- 29. Kotsikorou, E.; Lynch, D. L.; Abood, M. E.; Reggio, P. H., Lipid Bilayer Molecular Dynamics Study of Lipid-Derived Agonists of the Putative Cannabinoid Receptor, Gpr55. *Chemistry and physics of lipids* **2011**, *164*, 131-43.
- 30. Syed, S. K.; Bui, H. H.; Beavers, L. S.; Farb, T. B.; Ficorilli, J.; Chesterfield, A. K.; Kuo, M.-S.; Bokvist, K.; Barrett, D. G.; Efanov, A. M., Regulation of Gpr119 Receptor Activity with Endocannabinoid-Like Lipids. *American Journal of Physiology-Endocrinology and Metabolism* **2012**, *303*, E1469-E1478.
- 31. Zhang, S.-y.; Li, J.; Xie, X., Discovery and Characterization of Novel Smallmolecule Agonists of G Protein-Coupled Receptor 119. *Acta Pharmacologica Sinica* **2014**, *35*, 540-548.
- 32. Lauffer, L. M.; Iakoubov, R.; Brubaker, P. L., Gpr119 Is Essential for Oleoylethanolamide-Induced Glucagon-Like Peptide-1 Secretion from the Intestinal Enteroendocrine L-Cell. *Diabetes* **2009**, *58*, 1058-66.
- 33. Mobarec, J. C.; Sanchez, R.; Filizola, M., Modern Homology Modeling of G-Protein Coupled Receptors: Which Structural Template to Use? *Journal of Medicinal Chemistry* **2009**, *52*, 5207-5216.
- 34. Worth, C. L.; Kreuchwig, A.; Kleinau, G.; Krause, G., Gpcr-Ssfe: A Comprehensive Database of G-Protein-Coupled Receptor Template Predictions and Homology Models. *BMC Bioinformatics* **2011**, *12*, 185.
- 35. Shoichet, B. K., Virtual Screening of Chemical Libraries. *Nature* **2004**, *432*, 862-5.
- 36. Sterling, T.; Irwin, J. J., Zinc 15 Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling* **2015**, *55*, 2324-2337.
- 37. Meier, A.; Söding, J., Automatic Prediction of Protein 3d Structures by Probabilistic Multi-Template Homology Modeling. *PLoS computational biology* **2015**, *11*, e1004343.
- 38. Wellenzohn, B.; Lessel, U.; Beller, A.; Isambert, T.; Hoenke, C.; Nosse, B., Identification of New Potent Gpr119 Agonists by Combining Virtual Screening and Combinatorial Chemistry. *Journal of Medicinal Chemistry* **2012**, *55*, 11031-11041.
- 39. Li, N. X., et al., Gpr119 Agonism Increases Glucagon Secretion During Insulin-Induced Hypoglycemia. *Diabetes* **2018**, *67*, 1401.
- 40. Mitchell, S. J., et al., The Sirt1 Activator Srt1720 Extends Lifespan and Improves Health of Mice Fed a Standard Diet. *Cell Rep* **2014**, *6*, 836-843.
- 41. Pacholec, M., et al., Srt1720, Srt2183, Srt1460, and Resveratrol Are Not Direct Activators of Sirt1. *The Journal of biological chemistry* **2010**, *285*, 8340-8351.
- 42. Cohen, L. J., et al., Commensal Bacteria Make Gpcr Ligands That Mimic Human Signalling Molecules. *Nature* **2017**, *549*, 48-53.
- 43. Liarte, S.; Alonso-Romero, J. L.; Nicolás, F. J., Sirt1 and Estrogen Signaling Cooperation for Breast Cancer Onset and Progression. *Front Endocrinol (Lausanne)* **2018**, *9*, 552-552.

APPENDIX

APPENDIX

Complete Tables for Frame 150

Table 13. 36 ligands out of the 41 Prestwick phytochemical compound library bound to the GPR119 receptor

Г

Table 14. 646 ligands out of the 862 Prestwick and Selleckchm GPCR compound libraries bound to the GPR119 receptor

Table 15. 949 ligands out of the 20,000 Zinc Naturals library compounds bound to the GPR119 receptor

Complete tables for Frame 200

Table 16. 38 ligands out of the 41 Prestwick phytochemical compound library bound to the GPR119 receptor

	trihexyphenidyl	Aceclidine	AM251	S5437 4 4 DD E	Bepotastine besil ate	Cytisine	IEM 1754 d ihydrobroMi de
	carbinoxamine	Diphenhydra mine	K Ras G12C inhibitor	unamed	S3706 Sarpogrel ate hydrochlorid	fty720	S4718 Acet ylcholine io dide
	Astemizole	Betaxolol	unamed	Maprotiline hy drochloride	Atomoxetine hy drochloride	VU 0361737	unamed
	1H-Pyrido[4,3-b]indole, 2,3,4,5-tetrahydro-2-methyl- 5-(phenylmethyl)	Oxprenolol	S4635 Cyproh eptadine hydr ochloride	S4696 Arbinox aMine Maleate	Azasetron hydro chloride	unamed	S3639 Tacri ne hydrochl oride hydrat
	1H-Indol-5-ol, 3-(2- aminoethyl)-	Tolazoline	Hexestrol	ABILIFY_aripi prazole	Endoxifen HCl	S3761 Eucalyp tol	Arecoline h ydrobromide
	Pheniramine	$2-(4-$ Hydroxyphe nethylamino) $-1-(4-$ hydroxyphen yl)propanol	S3953_L_Lysi ne hydrochlori de	Medetomidine hcl	Hexamethon ium bromide	unamed	unamed
	3469 LCZ696	S5341 Metr oprolol succ inate	S4700 4 Ami nobutyric acid	unamed	S5131 Homotaur ine	unamed	AMD3465
	NSC23766	S5066 Pram ipexole dihy drochloride	unamed				

Table 18. 741 ligands out of the 862 Prestwick and Selleckchm GPCR compound libraries bound to the GPR119 receptor

Table 18. 956 ligands out of the 20,000 Zinc Naturals library compounds bound to the GPR119 receptor

Complete tables for Frame 250

Table 19. 32 ligands out of the 41 Prestwick phytochemical compound library bound to the GPR119 receptor

Table 20. 698 out of the 862 Prestwick and Selleckchem GPCR compound libraries bound to the GPR119 receptor

Table 21. 968 ligands out of the 20,000 Zinc Naturals library compounds bound to the GPR119 receptor

BIOGRAPHICAL SKETCH

Jennifer Lizeth Bravo was born on February 12, 1991 in the city of Linares, Nuevo Leon Mexico. She started her early studies in Mexico and upon moving to the United States in 2006 she continued to pursue a career in science. For her undergraduate she attended The University of Texas Pan American and graduated as a magna cum laude with a BS in chemistry in 2015. In 2019, she joined Dr. Evangelia Kotsikorou's computational chemistry laboratory where she studied computational algorithms for the analysis of big chemical libraries with the purpose of finding possible ligands to fit the GPR119 receptor. Mid 2020, Jennifer started to be co-advised by biochemist Dr. Frank Dean, and she started performing ELISA cAMP analysis on some molecules of the library that produced good receptor fitting scores. She earned her MS in Chemistry degree in May 2021 from the University of Texas Rio Grande Valley. Her email is jennybrav2@gmail.com.