

Design, Synthesis, and Biological Evaluation of Sphingosine 1-Phosphate Receptor (S1PR) Agonists

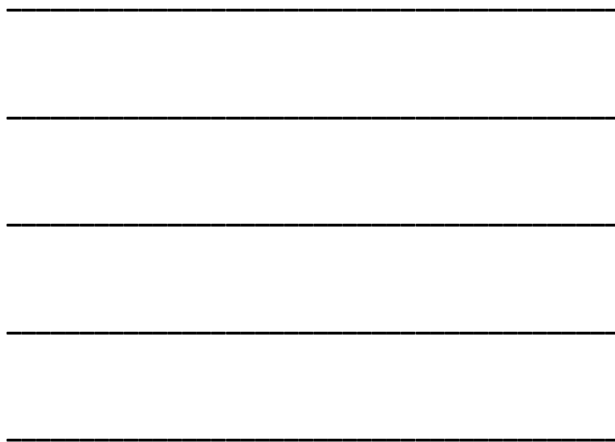
Michael James Ferracane  
Marshfield, Wisconsin

B.S. Chemistry, Biochemistry and Molecular Biophysics, University of Arizona, 2008

A Dissertation Presented to the Graduate Faculty of the University of Virginia  
in Candidacy for the Degree of Doctor of Philosophy

Department of Chemistry

University of Virginia  
June, 2014



## Abstract

Sphingosine 1-phosphate (S1P) is produced endogenously in response to a variety of environmental stimuli and serves as a critical signal in the processes of cell proliferation, vascularization, and inflammation. S1P elicits its effects primarily via modulation of its corresponding G protein-coupled receptors: the five sphingosine 1-phosphate receptors (S1PR<sub>1-5</sub>). This receptor family (especially S1PR<sub>1</sub>) has proven to be a particularly useful target for the treatment of autoimmune disease; fingolimod (Gilenya™, Novartis) is FDA-approved for the treatment of relapsing-remitting multiple sclerosis (RRMS), and it exhibits its activity via modulation of S1PR<sub>1,3,4,5</sub>. Unfortunately, this broad activity profile also contributes to its deleterious effects, which span from reversible macular edema to potentially lethal bradycardia.

The Macdonald and Lynch Labs are interested in developing a second-generation compound that more selectively modulates the S1PRs. Such a compound would be safer and more efficacious than fingolimod for treating those afflicted with RRMS or other autoimmune diseases. We sought to achieve this in two different ways: 1) through synthesis and biological evaluation of a library of rigid analogs of fingolimod and 2) through rational design, synthesis, and biological evaluation of novel compounds. The work described herein outlines present and past attempts to achieve this goal.

Based upon previous work, an indane-based fingolimod rigid analog was hypothesized to exhibit more selective agonism of the S1PRs. The indane-based analog was synthesized as a racemic mixture, which demonstrated S1PR<sub>1</sub>-S1PR<sub>5</sub> dual agonist activity *in vitro* and immunosuppressive activity similar to fingolimod *in vivo*. Though this rigid analog obtained encouraging preliminary results, it lacks therapeutic potential as a result of limitations in other chemical and biological properties.

While the indane-based analog was being evaluated, the 2.8 Å resolution crystal structure of an antagonist-bound S1PR<sub>1</sub>-T4L chimeric receptor was published. This new knowledge led us to ultimately change our strategy and use computational methods to rationally design new S1PR modulators rather

than our previous guess and check method. A computational model of S1PR<sub>1</sub> was developed using the Molecular Operating Environment software suite and crystal structure of the antagonist-bound S1PR<sub>1</sub>-T4L chimeric receptor. Docking studies were then performed using this model and S1PR<sub>1</sub> agonists previously reported in the literature; the compounds' *in silico* docking study results were compared to their *in vitro* assay results to create a method that reasonably predicts the activity and binding conformation of the chosen S1PR<sub>1</sub> agonists.

Using this method, a series of novel compounds were designed and are currently undergoing synthesis. Once finished, the compounds will undergo biochemical evaluation to test the validity and predictability of our model receptor. The validated model will then be used to better understand S1PR function and can be used to develop improved S1PR modulators to treat RRMS.

## Acknowledgments

To Professor Macdonald, for providing his insight, his encouragement, and the freedom to pursue both my scientific and nonscientific interests

To the Department of Chemistry, for providing me with the education, training, and resources needed to become a scientist and to conduct the work herein

To the current and past members of the Macdonald Lab, for furthering my education

To my family, for nurturing me into the individual that I am today  
and for constantly pushing and supporting that individual

and

To Victoria, for her love, sacrifice, and belief in me

# Table of Contents

Abstract.....	i
Acknowledgments .....	iii
Table of Contents .....	iv
List of Figures .....	v
List of Schemes .....	vii
List of Tables .....	viii
List of Abbreviations .....	ix
<b>Chapter 1: Sphingosine 1-Phosphate and its Receptors in Immunity .....</b>	<b>1</b>
1.1 Biology of Sphingosine 1-Phosphate .....	2
1.2 Structure and Signaling of the Sphingosine 1-Phosphate Receptors .....	4
1.3 Fingolimod and Other Modulators of the S1PRs .....	10
1.4 The Role of S1PR Modulation in Immune Cell Trafficking .....	13
1.5 Fingolimod for the Treatment of Relapsing-Remitting Multiple Sclerosis .....	16
1.6 Properties of Ideal Second-Generation S1PR Modulators .....	20
<b>Chapter 2: Rigid Analogs of Fingolimod as Improved Therapeutics .....</b>	<b>23</b>
2.1 Previous Rigid Analogs of Fingolimod .....	24
2.2 Synthesis & Evaluation of Racemic Indane-Based Rigid Analog .....	27
2.3 Attempts toward the Synthesis of an Indole-Based Rigid Analog .....	32
2.4 Conclusions .....	33
<b>Chapter 3: Molecular Modeling of S1PR<sub>1</sub> and its Agonists .....</b>	<b>34</b>
3.1 Overall Goals and Design of the Computational Study .....	35
3.2 Generation of the S1PR <sub>1</sub> Model Receptor .....	36
3.3 Design of the Docking Study and Choice of Ligand Set .....	38
3.4 Results from Docking Studies with Head Groups .....	42
3.5 Results from Docking Studies with Complete Ligands .....	44
3.6 Extension to Biology and Medicinal Chemistry and Future Work .....	54
3.7 Conclusions .....	55
<b>Chapter 4: Direct-Acting Second-Generation S1PR<sub>1</sub> Agonists .....</b>	<b>56</b>
4.1 Goals and Properties of Second-Generation S1PR <sub>1</sub> Modulators .....	57
4.2 Attempts at Designing Compounds to Elucidate the Binding Mode .....	59
4.3 Rational Design of Second-Generation S1PR <sub>1</sub> Agonists .....	62
4.4 Progress Toward the Synthesis of Second-Generation S1PR <sub>1</sub> Agonists .....	65
4.5 Conclusions .....	68
<b>Chapter 5: Experimental Details .....</b>	<b>69</b>
5.1 Biochemical & Biological Methods .....	69
5.2 Computational Methods .....	71
5.3 Synthetic Methods .....	80
<b>References .....</b>	<b>102</b>

## List of Figures

<b>1.1.</b> Biosynthesis and structure of representative sphingolipids .....	2
<b>1.2.</b> Pathways of S1P production and degradation .....	3
<b>1.3.</b> Crystal structure of the antagonist-bound S1PR <sub>1</sub> -T4L chimeric receptor .....	4
<b>1.4.</b> Ligand-receptor interaction diagram for ML056 in the S1PR <sub>1</sub> -T4L chimeric receptor .....	6
<b>1.5.</b> Sequence alignment of all human S1PRs .....	8
<b>1.6.</b> Properties of the S1PRs .....	9
<b>1.7.</b> Structure of myriocin, fingolimod, and (S)-fingolimod phosphate .....	10
<b>1.8.</b> Notable fingolimod analogs .....	11
<b>1.9.</b> Examples of some azetidinium carboxylate-containing S1PR modulators .....	11
<b>1.10.</b> Diagram of lymphatic system and lymph node .....	13
<b>1.11.</b> Egress of effector T cells from secondary lymphoid tissue .....	14
<b>1.12.</b> Types of MS and their relapse rate and progression .....	16
<b>1.13.</b> Comparison of signal transmission in healthy and demyelinated neurons .....	17
<b>2.1.</b> Fingolimod, AAL, and synthesized rigid analogs .....	25
<b>2.2.</b> Sphingosine kinase assay results .....	28
<b>2.3.</b> Lymphopenia assay results .....	30
<b>2.4.</b> Structure of bicyclic fingolimod rigid analogs .....	31
<b>3.1.</b> Docking study protocol .....	35
<b>3.2.</b> Comparison of the S1PR <sub>1</sub> -T4L crystal structure with the prepared model structure .....	37
<b>3.3.</b> SAR of some S1PR <sub>1</sub> modulators .....	38
<b>3.4.</b> Optimized head groups and ML056 inside the S1PR <sub>1</sub> model receptor pocket .....	42
<b>3.5.</b> Results from docking studies .....	45

<b>3.6.</b> Ligand-receptor interaction diagram for select compounds in model receptor .....	50
<b>3.7.</b> Compounds docked into the side cavity of S1PR <sub>1</sub> , highlighting important residues .....	52
<b>3.8.</b> Analogs of interest .....	54
<b>4.1.</b> Structure of fused polyaromatic compounds .....	59
<b>4.2.</b> Structure of indole- and benzimidazole-based analogs .....	60
<b>4.3.</b> Docked indole possessing the ( <i>R,S</i> )-cyclobutyl rigid linker compared with F-11t .....	61
<b>4.4.</b> Pharmacophore of second-generation agonists .....	62
<b>4.5.</b> Ligand-receptor interaction diagrams for rationally-designed compounds .....	63
<b>5.1.</b> Issues fixed using “Structure Preparation” program .....	71
<b>5.2.</b> “Protonate 3D” program settings .....	72
<b>5.3.</b> Head group minimization settings for azetidinium carboxylate .....	73
<b>5.4.</b> Initial conformational search settings for A-32 .....	74
<b>5.5.</b> Initial virtual screen settings for A-32 .....	75
<b>5.6.</b> Refined conformational search settings for A-32 .....	76
<b>5.7.</b> Docking settings for A-32 .....	78

## List of Schemes

<b>2.1.</b> Synthesis of VPC142072 and VPC142097 .....	27
<b>2.2.</b> Attempts toward the synthesis of the indole-based rigid analog .....	32
<b>4.1.</b> Progress toward the synthesis of 4.3 .....	65
<b>4.2.</b> Progress toward the synthesis of 4.7 .....	65
<b>4.3.</b> Progress toward the synthesis of 4.11 .....	66
<b>4.4.</b> Progress toward the synthesis of 4.16 .....	67



## List of Tables

<b>1.1.</b> Important features of the S1PR <sub>1</sub> -T4L crystal structure and their possible implications .....	5
<b>1.2.</b> Key properties and <i>in vivo</i> effects of fingolimod .....	20
<b>1.3.</b> Effects of S1PR <sub>1</sub> modulators <i>in vivo</i> .....	21
<b>2.1.</b> Nucleotide exchange assay results .....	29
<b>3.1.</b> Structure and biochemical data for compounds used in docking studies .....	41
<b>3.2.</b> Full numerical results from compounds used in docking studies .....	46
<b>4.1.</b> Ideal properties of second-generation S1PR <sub>1</sub> modulators .....	57
<b>4.2.</b> Results of docking studies with ethyl-and rigid ethyl-containing compounds .....	60

## List of Abbreviations

9-BBN	9-borabicyclo[3.3.1]nonane
[ $\gamma$ - <sup>32</sup> P]ATP	radiolabeled adenosine triphosphate
[ $\gamma$ - <sup>35</sup> S]GTP	radiolabeled, nonhydrolyzable guanosine triphosphate
[ <sup>33</sup> P]S1P	radiolabeled sphingosine 1-phosphate
ADME	absorption-distribution-metabolism-extraction
AMBER	assisted model building with energy refinement
AMBER12:EHT	forcefield for docking
APC	antigen-presenting cell
BBB	blood-brain barrier
BLAST	Basic Local Alignment Search Tool
Boc	<i>t</i> -butylcarbonate
cAMP	cyclic adenosine monophosphate
CBz	carboxybenzyl
CHO K1	chinese hamster ovary cell line K1
CNS	central nervous system
DIAD	diisopropyl azodicarboxylate
DIBALH	diisobutylaluminum hydride
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethylsulfoxide
EC <sub>50</sub>	concentration at which half the maximum efficacy is achieved
ECL	extracellular loop
EDTA	ethylenediaminetetraacetic acid
EHT	extended Hückel theory
GBVI/WSA dG	generalized-Born volume integral/weighted surface area, scoring function for docking
GDP	guanosine diphosphate
GIRK	G protein-gated inward-rectifying potassium channel
GPCR	G protein-coupled receptor
GRK	G protein-coupled receptor kinase
HEK293T	human embryonic kidney cell line 293T
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
IC <sub>50</sub>	concentration at which half the maximum inhibition is achieved
ICL	intracellular loop
MAPK	mitogen-activated protein kinase
mCPBA	<i>meta</i> -chloroperoxybenzoic acid
MMFF	Merck molecular force field
MMFF94x	forcefield for conformational
MOE	Molecular Operating Environment
MS	multiple sclerosis
pEC <sub>50</sub>	$-\log(\text{EC}_{50})$
pIC <sub>50</sub>	$-\log(\text{IC}_{50})$
RP-HPLC	reverse-phase high performance liquid chromatography
RMS	root mean square
RMSD	root mean square distance
RRMS	relapsing-remitting multiple sclerosis

S1P	sphingosine 1-phosphate
S1PR	sphingosine 1-phosphate receptor
SAR	structure-activity relationship
SK	sphingosine kinase
SPL	sphingosine 1-phosphate lyase
SPP	sphingosine 1-phosphate phosphatase
T4L	T4 bacteriophage lysozyme
TBAI	tetrabutylammonium iodide
TFA	trifluoroacetic acid
Tf	triflyl
THF	tetrahydrofuran
TLC	thin-layer chromatography
TMS	trimethylsilyl
TRAF2	tumor necrosis factor receptor-associated factor 2
Ts	toluenesulfonyl
WWP	WW domain-containing protein

---

# 1

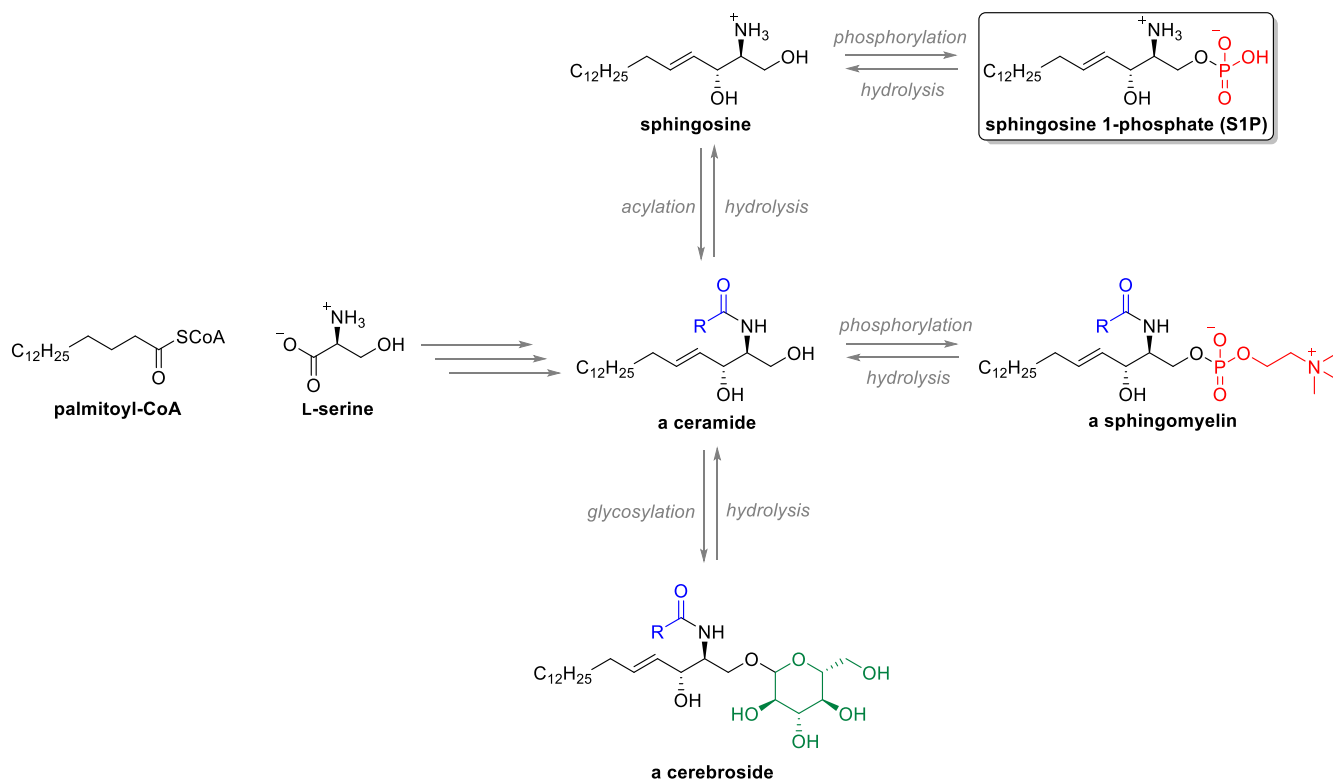
---

## Sphingosine 1-Phosphate and its Receptors in Immunity

**Sphingosine 1-phosphate (S1P)** and its family of five G protein-coupled receptors (GPCRs) - the sphingosine 1-phosphate receptors (S1PR<sub>1-5</sub>) - are involved in the signaling of many essential cellular processes. One such process involves the regulation of immune cell migration; **S1P** is produced during inflammation, stimulates S1PR<sub>1</sub> on the surface of B and T cells, and allows them to egress from lymph nodes into circulation, where they can elicit their effects. This egress is critical for immune system homeostasis, but becomes problematic for those afflicted with autoimmune diseases like relapsing-remitting multiple sclerosis (RRMS). Thus, compounds like  **fingolimod**  (Gilenya<sup>™</sup>, Novartis) that disrupt the **S1P**-S1PR signaling axis serve as promising treatments for autoimmune diseases like RRMS.

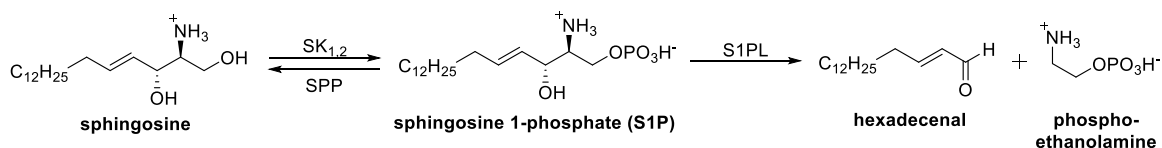
## 1.1 Biology of Sphingosine 1-Phosphate

**Sphingosine 1-phosphate (S1P)** is an important member of the class of lipids known as the sphingolipids. Unlike glycerol-based lipids, sphingolipids possess a sphingoid backbone, the product of a series of reactions beginning with the substrates palmitoyl CoA and serine. The sphingoid backbone can be acylated, glycosylated, or phosphorylated by a variety of different species, yielding an array of sphingolipids that vary widely in chemical structure and biological function [1, 2].



**Figure 1.1. Biosynthesis and structure of representative sphingolipids.** The sphingoid backbone (black) is initially constructed in several steps starting with palmitoyl-CoA and L-serine to yield a ceramide. The primary alcohol and the amine of the sphingoid backbone can be decorated with different groups to affect its structure and function.

**S1P** itself is produced via phosphorylation of **sphingosine** by two different kinase isoforms, sphingosine kinase 1 (SK<sub>1</sub>) and sphingosine kinase 2 (SK<sub>2</sub>). **S1P** can be hydrolyzed back to **sphingosine** via sphingosine 1-phosphate phosphatases (SPPs) or irreversibly degraded to hexadecenal and phosphoethanolamine by action by sphingosine 1-phosphate lyase (S1PL) [3].



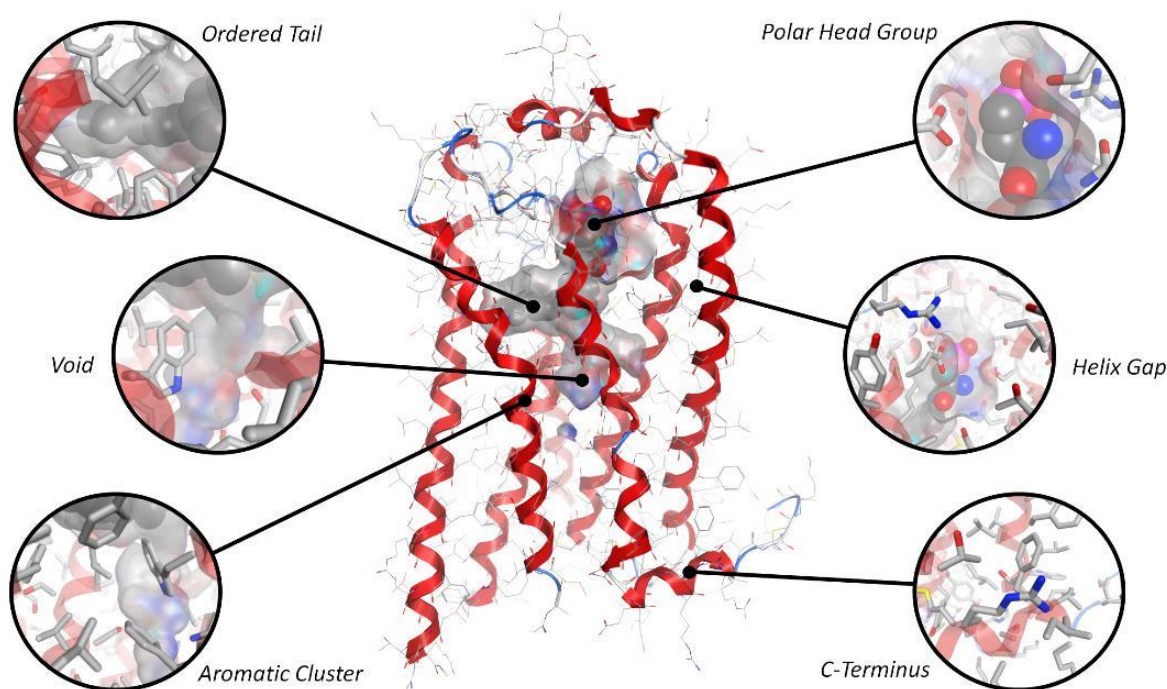
**Figure 1.2. Pathways of S1P production and degradation.** The combination of SK and SPP enzymes allows for reversible production of S1P depending on biological stimuli. S1PL irreversibly degrades S1P to products used in other pathways.

In its active form, SK<sub>1</sub> is localized on the inner leaflet of the cytosol, whereas SK<sub>2</sub> is membrane-bound in the nucleus and on the endoplasmic reticulum [4]. It is believed that intracellular S1P is important in a variety of functions, though only a few downstream targets have been identified [5-7]. Most importantly, S1P residing on the inner leaflet of the cytosol can be transported to the extracellular matrix via the Spns2 transporter [8, 9].

Once outside the cell, S1P acts as an agonist of the five sphingosine 1-phosphate G protein-coupled receptors (S1PR<sub>1-5</sub>), with K<sub>D</sub> values in the low nanomolar range [10]. The S1PRs belong to the largest subclass of GPCRs: class A, which includes around 700 of the roughly 800 known / predicted human GPCRs [11]. The S1PRs contain the common seven-transmembrane architecture and possess the conserved D(E)RY, CWxP, and NPxxY functional motifs [12]. Slight differences in their sequence and expression pattern give the S1PRs a wide range of activities; they have been found to regulate morphogenesis, cytoskeletal arrangement, cell proliferation, cellular migration, cellular adherence, and tight-junction formation [13].

## 1.2 Structure and Signaling of the Sphingosine 1-Phosphate Receptors

The best characterized and most pharmacologically interesting receptor subtype is S1PR<sub>1</sub>. The X-ray crystal structure of an antagonist-bound chimeric S1PR<sub>1</sub>-T4L receptor was determined to a resolution of 2.8 Å [14]. The receptor was engineered to minimize flexibility – truncating the N and C termini and replacing the third intracellular loop with the lysozyme of a T4 bacteriophage (T4L) – in order to facilitate crystallization and subsequent X-ray crystallographic study [15]. While these modifications destroy the receptor's biological activity, precluding study to assess the engineered receptor's viability, they provide a structure that gives insight into the structure and function of native S1PR<sub>1</sub>.



**Figure 1.3. Crystal structure of the antagonist-bound S1PR<sub>1</sub>-T4L chimeric receptor.** The antagonist ML056 (gray spheres) in the pocket of the receptor (colored surface) is shown. Residues corresponding to the T4 lysozyme are omitted for clarity. Structures that are unique or are involved in ligand binding, receptor activation, or receptor posttranslational modification are highlighted.

Like the crystal structures of other class A GPCRS, the S1PR<sub>1</sub>-T4L crystal structure possesses the common architecture (seven transmembrane helices and one intracellular helix towards the C-terminus) and structural motifs (D(E)RY, CWxP, and NPxxY motifs) that are involved in the receptor activation

process [12]. However, there are several unique features both of the antagonist ligand, **ML056**, and the S1PR<sub>1</sub>-T4L chimeric receptor itself (Table 1.1).

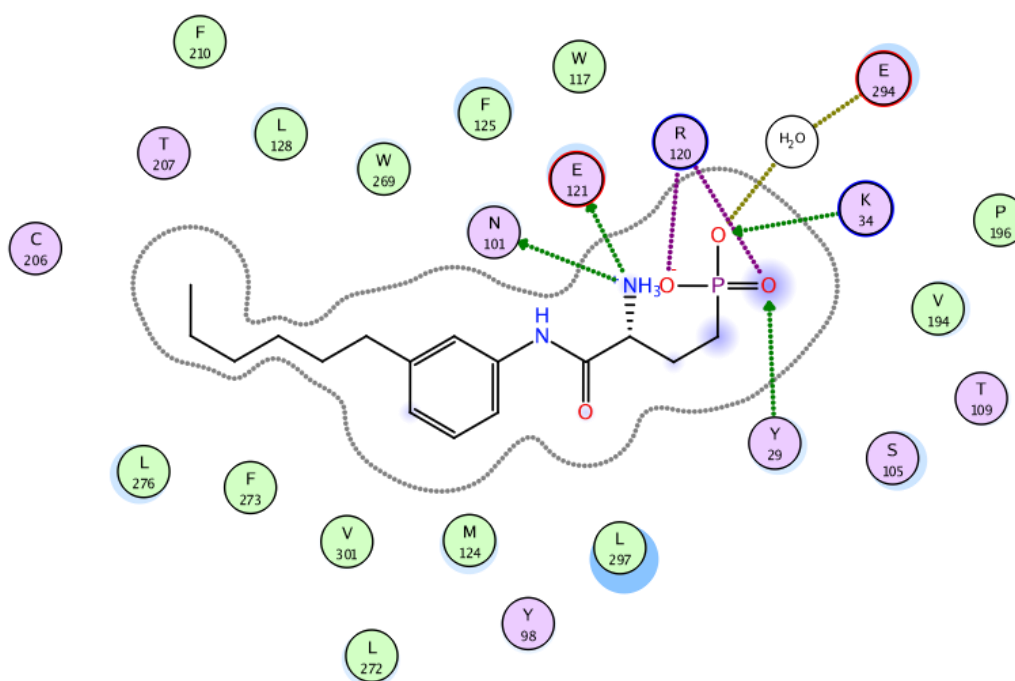
Structural Feature	Implication
Helical N-terminus caps the pocket	<i>Ligands are likely prevented from entering the pocket via the extracellular face</i>
Gap between helices I and VII is more pronounced than in other GPCR structures	<i>Ligands likely enter the pocket via the gap</i>
Receptor pocket is amphipathic	<i>Receptor binds charged, lipid-like molecules</i>
Portions of the pocket are unoccupied	<i>Other ligands may occupy different portions of the pocket OR solvent molecules are not resolved in crystal structure</i>
Hexyl tail of <b>ML056</b> is resolved and constrained to one conformation	<i>Ligand tail forms hydrophobic interactions that contribute to binding energy</i>

Table 1.1. Important features of the S1PR<sub>1</sub>-T4L crystal structure and their possible implications.

The first unique feature is the extracellular helix towards the N-terminus, which caps the receptor pocket and likely prevents ligand access from the extracellular matrix. Ligands are believed instead to enter the pocket directly from the membrane via the gap between helices I and VII, which is more pronounced than the same gap in structures of other class A GPCRs. The residue R292<sup>7.35</sup> lies along this gap and is positioned to interact with negatively charged groups like the phosphonate of **ML056** and the phosphate of **S1P** as they enter the receptor pocket. Not surprisingly, this residue was demonstrated to be necessary for **S1P** to saturate and activate S1PR<sub>1</sub> even though it is positioned nearly 12 Å from the phosphonate of **ML056** in the crystal structure [16]. This distance is far above the range of any noncovalent force, supporting the idea that R292<sup>7.35</sup> acts to “lure” ligands into the pocket but is not bind them once localized deep in the pocket [17]. However, the discrepancy between structural and biochemical results could also be the result of differences in the structure of S1PR<sub>1</sub> when bound by an antagonist like **ML056** versus an agonist like **S1P**.



Crystallography and mutagenesis are in better agreement about the other polar sites that bind zwitterionic ligands. The mutagenesis study that initially identified the importance of R292<sup>7,35</sup> also showed that R120<sup>3,28</sup> and E121<sup>3,29</sup> are essential for **S1P** to saturate and activate S1PR<sub>1</sub> [16]. Not surprisingly, these were found to bind the phosphonate and ammonium groups of **ML056** in the crystal structure. The crystal structure also identified a set of secondary residues that interact with **ML056** - Y29 and K34 of the N-terminal helix, N101<sup>2,60</sup>, and a bound H<sub>2</sub>O. These residues were also found to be involved in binding (**S**)-**ingolimod phosphate**, an ammonium phosphate-containing S1PR<sub>1</sub> agonist, during a subsequent site mutagenesis study [14]. These specific polar interactions confer the receptor's selectivity for **S1P** over structurally similar sphingolipids and lysophosphatidic acids, lipids that possess a glycerol backbone.



**Figure 1.4. Ligand-receptor interaction diagram for ML056 in the S1PR<sub>1</sub>-T4L chimeric receptor.** Residues lining the pocket (gray outline) are colored according to their polarity and interactions between them and **ML056** are highlighted.

The hexyl tail of **ML056** was found to exist in a single conformation in the crystal structure, an unexpected result given its conformational flexibility. Though it does not form specific interactions with residues lining the pocket, it does form van der Waals forces that compensate for the loss of entropy that

occurs with conformation restriction. Point mutations that changed the shape of this pocket affected both ligand binding and receptor activation, with F125<sup>3.33</sup>, L128<sup>3.36</sup>, F210<sup>5.48</sup>, W269<sup>6.48</sup>, and L276<sup>6.55</sup> even demonstrating ligand-dependent effects [14, 18-20]. These residues flank the hexyl tail of **ML056** in the crystal structure and demonstrate the contribution of van der Waals forces to the overall binding energy.

S1PR<sub>1</sub> has been shown to couple selectively to heterotrimeric G protein complexes containing the GTPase G<sub>i/o</sub>α [21]. Agonists like **S1P** produce conformational changes in S1PR<sub>1</sub> that activate this GTPase, leading to inhibition of adenylyl cyclase. Stimulation of the heterotrimeric G protein complex also leads to induction of Gβγ, which has been shown to activate GIRKs [22]. The effects of these actions vary based upon cell type and will be discussed in further detail later.

**S1P** also triggers internalization of S1PR<sub>1</sub>, with the majority of S1PR<sub>1</sub> quickly becoming internalized at high **S1P** concentration. **S1P** binding causes several serine residues along the C-terminus to become phosphorylated by GRK-2, which increases the receptor's affinity for β-arrestin [23]. β-Arrestin then mediates internalization of the receptor into intracellular vesicles [24]. Interestingly, it was shown that G protein signaling occurred even following internalization of S1PR<sub>1</sub> [25]. Once internalized, several specific lysine residues along the C-terminus can become ubiquitinated by the ubiquitin ligase WWP2. Monoubiquitinylation leads to recycling of the receptor back to the membrane whereas polyubiquitinylation leads to degradation of the receptor [26]. The preference for mono- or polyubiquitinylation depends upon the identity of the ligand: **S1P** predominately induces monoubiquitinylation whereas **(S)-fingolimod phosphate** primarily causes polyubiquitinylation [23].

The remaining S1PRs have not been crystallized and have had less mutagenesis studies conducted to evaluate residues involved in ligand binding and receptor activation. A simple sequence alignment demonstrates that all of the S1PRs have conserved residues at the positions needed to establish identical polar interactions with the zwitterionic head group of **S1P** and discriminate against similar lipids. The human receptors share reasonable overall homology to S1PR<sub>1</sub>: S1PR<sub>2</sub> = 51% identical, S1PR<sub>3</sub> = 56%, S1PR<sub>4</sub>

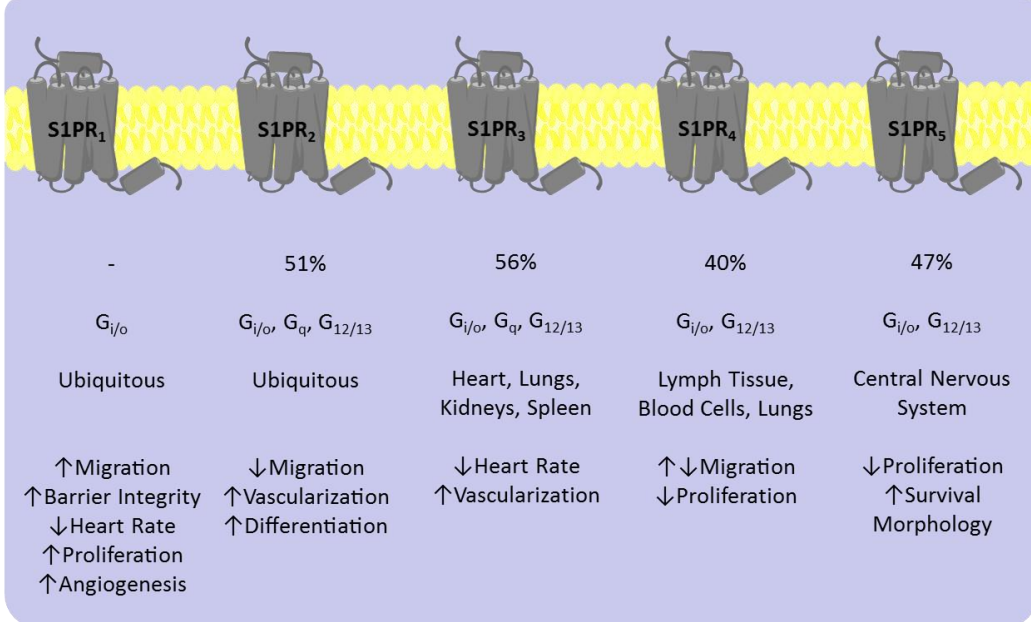
= 40%, S1PR<sub>5</sub> = 47%. The majority of the nonidentical residues are similar, and the few major differences occur at the intracellular and extracellular regions of these receptors (Figure 1.5).

1	MA-----TALPPRLQFVRGNETLREHYQYVGRKLAGRLKEASEGSLTTLVFLVICSPF	53	Q99500	S1PR3_HUMAN
1	MGPTSVPFLVKAHRSSVSDYVNYDIIVRHYNYPGRKLNISA-DKENSIKLTSVVVFLICCFI	59	P21453	S1PR1_HUMAN
1	-----MGSLYSEYLNPKNKVOEHYNYTKETLETO--ETTSROVASAFIVILCCAI	47	O95136	S1PR2_HUMAN
1	MNATGTFVAPESCCQLAAGGHSRLIVLHYNHSGRLAGRGGPEDGGLGALRGLSVAASCLV	60	O95977	S1PR4_HUMAN
1	-----MESGLLRPAPVSEVIVLHYNYPGRKLRGARYQPGAGLRADAVVCLAVCAPI	50	Q9H228	S1PR5_HUMAN
	: **:: . . . . : .. :			
54	VLENLMVLLAIWKNNRPHNMYFFIGNLALCDLLAGIAYKVNILMSGKTFSELSPTVWFL	113	Q99500	S1PR3_HUMAN
60	ILENIFVLLTIWTKKFKHRPMYFFIGNLALSDLLAGVAYTANLLSGATTYKLTFAQWFL	119	P21453	S1PR1_HUMAN
48	VVENLLVLLIAVARNRSHFSAHYLFLGNLAAASDLLAGVAFVANTLLSGSVTLRLTEVQWFA	107	O95136	S1PR2_HUMAN
61	VLENLLVLLAAITSHMRSRRWVYVYCLVNITLSDLLTGAAYLANVLLSGARTFRLAFAQWFL	120	O95977	S1PR4_HUMAN
51	VLENLAVLLVLRGRHPRFHAPMFLLLGSLTSLDLAGAAYAANILLSGPLTLKLSFALWFA	110	Q9H228	S1PR5_HUMAN
	::**:: ** .: : : : : : .: : .***: * : . * :*** * * :*. **			
114	REGSMFVALGASTCSLLAIAIERHLMIKMRP-YDANKRHRVFLIIGNCMLIAFTLGALP	172	Q99500	S1PR3_HUMAN
120	REGSMFVALSASVFSLLAIAIERYITMLKMKL-HNGSMNFRLLISACWVI SLILGGLE	178	P21453	S1PR1_HUMAN
108	REGSAFITLSASVFSLLAIAIERHVAIAKVKL-YGSDKSCRMLLIGASWLI SLVGGLE	166	O95136	S1PR2_HUMAN
121	REGLLPTALAASTFSLFTAGERFATMVRPVAESGATTSRVYGFIGLWLLAALLGMLE	180	O95977	S1PR4_HUMAN
111	REGGVFVALTASVLSLLAIALERSLTMARRGP-APVSSRGRTLAMAAAANGVSLLLGLE	169	Q9H228	S1PR5_HUMAN
	*** * : * ** . *** * ** : : : . * : . * : : ** **			
173	ILGMNCLHNLPCDSTILPLYSKKYIAPFCISIFTAILEVTIVELYARYFLVKSSSRKVANH	232	Q99500	S1PR3_HUMAN
179	IMGMNCISALSSCSTVLEPLYHKHYILPCTTVPTLLLSIVILYCRYSLVTRSRRLTFR	238	P21453	S1PR1_HUMAN
167	ILGMNCLGHLRACSTVLEPLYAKHYVLCVVITPFIILLAIAVAIVRYICVVRSSHADMA--	224	O95136	S1PR2_HUMAN
181	LLGMNCLCAFDRCSLLEPLYSKRYILPCLVIFAGVLAITMGLYGAIFPLVQASGQKAPR	240	O95977	S1PR4_HUMAN
170	ALGMNCLGRLDACSTVLEPLYAKAYVLPCLVAFVGLAAICALYARTYCCVRANARLEPAR	229	Q9H228	S1PR5_HUMAN
	:****: : **::**** * * : * : * * ** * : * : :			
233	-----NNSERSMALLRRTVVIIVVSVFIACWSPLFILFLIDVAC-RVOACPILFPA	280	Q99500	S1PR3_HUMAN
239	KNI----SKASRSSEKSLALKTVIIIVLSVFIACWAPLFIILLLDVGC-KVKTCDIAPRA	293	P21453	S1PR1_HUMAN
225	-----APQTLALKTVITVVLGVFVFCWLPAFSILLLDYAC-PVHSCPIEYKA	270	O95136	S1PR2_HUMAN
241	-----AARRKARREKTVLMILLAFVLCWGPLFGLLIADVFGSNLWAEYLRGM	289	O95977	S1PR4_HUMAN
230	PGTAGTTSTRARRKSLALLRSLVLLAFVACWGPLFLLLLLDVAC-PARTCEVELQA	288	Q9H228	S1PR5_HUMAN
	: : **:: : : : . : .** * * : : * : *			
281	QWFIVLAVLNSAMNPIVYITLASKEMRRAPFRVLCNCLVGRGA-----R---	324	Q99500	S1PR3_HUMAN
294	EYFLVLAVLNSGTNPIIYTLTKEMRRAFIRIMSCCKCPGDS-----AGK-	339	P21453	S1PR1_HUMAN
271	HYFFAVSTLNSLLNPIVYITWRSRDLRREVLRLQCNR-PGVGV-----QGR-	315	O95136	S1PR2_HUMAN
290	DWILALAVLNSAVNPIIYIFRSREVCRAVLSFLCCGCLRIGMRGPGDCLARAVEAHSAS	349	O95977	S1PR4_HUMAN
289	DFFLGLAMANSLNPIIYITLNRDLRHALRLVCCGRHSQGRDFSGSQSASAAEASGG-	347	Q9H228	S1PR5_HUMAN
	. : : : ** **::** : : : : : : : :			
325	--ASPIQPA--LDPSPKSSSSNNSHSPKVKEDLPHTAPSSC-----IMDKNAALONG	374	Q99500	S1PR3_HUMAN
340	-FKRPIAG--MEFSPK--SDNSMHPQKDEGDNPETIMSSG-----NVNSS--	382	P21453	S1PR1_HUMAN
316	-RRG-TPGHLLPLRSSSL-ERGMHMPST--PT-----FLEGNTVV--	353	O95136	S1PR2_HUMAN
350	TTDSBLRPR--DFRGR--SLRFR-----MREPLSSISSVR-----SI-----	384	O95977	S1PR4_HUMAN
348	-LRRCLPFGLDGSPGSRSS-----PQRDGLDTSGSTGSPGAPTAARTLVSEPAAD--	398	Q9H228	S1PR5_HUMAN
	* : *			
375	IFCN	378	Q99500	S1PR3_HUMAN
383	----	382	P21453	S1PR1_HUMAN
354	----	353	O95136	S1PR2_HUMAN
385	----	384	O95977	S1PR4_HUMAN
399	----	398	Q9H228	S1PR5_HUMAN

Figure 1.5. Sequence alignment of all human S1PRs. Alignment showing conserved residues (dark gray), similar residues (gray), and unique residues (light gray). Alignment was performed using BLAST.

Differences in the sequences of the S1PRs has a significant effect on each receptor's G protein coupling and ability to bind different synthetic ligands. Most of the S1PRs are capable of coupling to several heterotrimeric G protein complexes, which allows for a variety of downstream effects. To further

complicate this picture, the receptors have differing susceptibility to internalization, a property that is often ligand dependent [27, 28]. Luckily though, antagonists, agonists, and full agonists (agonists that induce receptor internalization) of many S1PR subtypes have been developed, which has led to a better understanding of their function and potential role as therapeutic targets.

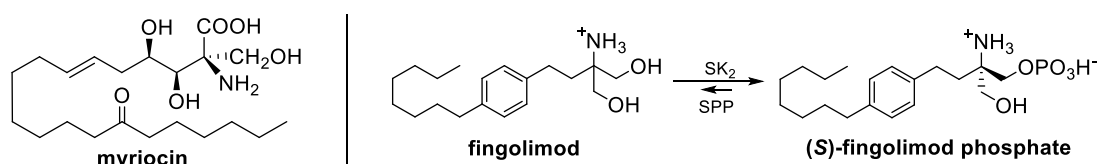


	S1PR <sub>1</sub>	S1PR <sub>2</sub>	S1PR <sub>3</sub>	S1PR <sub>4</sub>	S1PR <sub>5</sub>
<b>Homology</b>	-	51%	56%	40%	47%
<b>Coupling</b>	G <sub>i/o</sub>	G <sub>i/o</sub> , G <sub>q</sub> , G <sub>12/13</sub>	G <sub>i/o</sub> , G <sub>q</sub> , G <sub>12/13</sub>	G <sub>i/o</sub> , G <sub>12/13</sub>	G <sub>i/o</sub> , G <sub>12/13</sub>
<b>Expression Pattern</b>	Ubiquitous	Ubiquitous	Heart, Lungs, Kidneys, Spleen	Lymph Tissue, Blood Cells, Lungs	Central Nervous System
<b>Known Effects</b>	<ul style="list-style-type: none"> <li>↑ Migration</li> <li>↑ Barrier Integrity</li> <li>↓ Heart Rate</li> <li>↑ Proliferation</li> <li>↑ Angiogenesis</li> </ul>	<ul style="list-style-type: none"> <li>↓ Migration</li> <li>↑ Vascularization</li> <li>↑ Differentiation</li> </ul>	<ul style="list-style-type: none"> <li>↓ Heart Rate</li> <li>↑ Vascularization</li> </ul>	<ul style="list-style-type: none"> <li>↑ ↓ Migration</li> <li>↓ Proliferation</li> </ul>	<ul style="list-style-type: none"> <li>↓ Proliferation</li> <li>↑ Survival Morphology</li> </ul>

**Figure 1.6. Properties of the S1PRs.** The different receptors vary in their sequence and G protein coupling patterns. These properties, in combination with their differing expression patterns, give rise to the variety of effects observed *in vitro* and *in vivo* [29, 30].

### 1.3 Fingolimod and other Modulators of the S1PRs

**Fingolimod** (Gilenya™, Novartis) was first described in 1995 and represents a breakthrough in the development of synthetic S1PR ligands [31]. **Fingolimod** was designed to be a simpler analog of **myriocin**, a fungal natural product found to inhibit serine palmitoyltransferase, the first enzyme involved in *de novo* synthesis of sphingolipids. While **fingolimod** did not inhibit serine palmitoyltransferase, the corresponding phosphate was later found to be an agonist of S1PR<sub>1,3,4,5</sub> [32]. **Fingolimod** is phosphorylated specifically in the *pro*-S position by sphingosine kinase 2 (SK<sub>2</sub>) to generate the active species, **(S)-fingolimod phosphate** [33, 34]. The corresponding sphingosine 1-phosphate phosphatases (SPP) likely catalyze the hydrolysis reaction to regenerate **fingolimod**, but this has not yet been confirmed experimentally. Regardless, both species are observed *in vivo* and are highly protein bound, with the alcohol to phosphate ratio being around 1:3 [35].

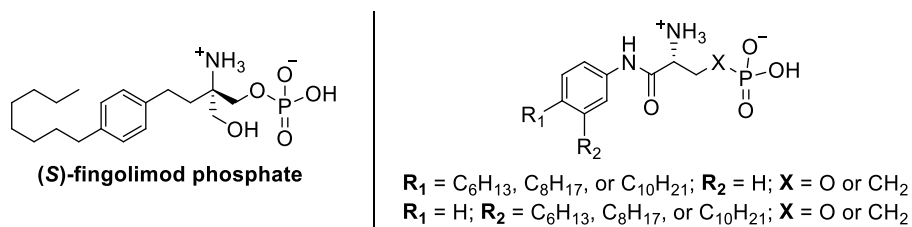


**Figure 1.7.** Structure of myriocin, fingolimod, and (S)-fingolimod phosphate. Myriocin was used as a template for the design of fingolimod, which is converted to its active form upon phosphorylation.

Like **S1P**, **(S)-fingolimod phosphate** induces activation of the G protein for S1PR<sub>1,3,4,5</sub>. It also causes rapid internalization of both S1PR<sub>1</sub> and S1PR<sub>5</sub>, and it is **(S)-fingolimod phosphate**'s ability to induce S1PR<sub>1</sub> internalization that gives rise to its beneficial and adverse effects, which will be discussed in further detail later [27]. The clinical success and eventual FDA-approval of **fingolimod** for the treatment of relapsing-remitting multiple sclerosis (RRMS) led to heavy interest in the development of new S1PR modulators.

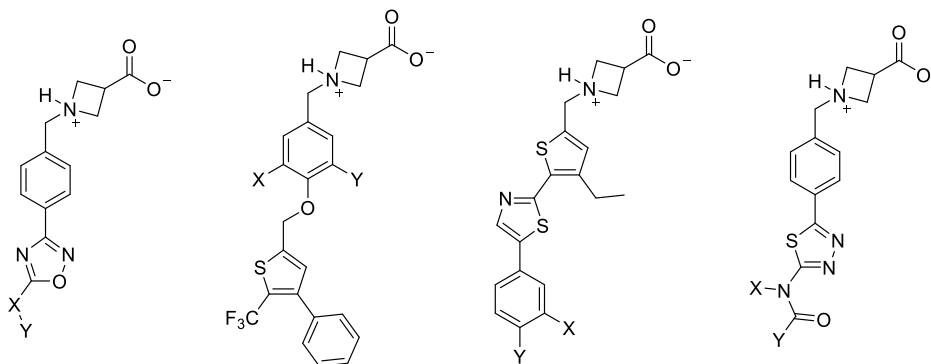
The Macdonald and Lynch labs made significant contributions to this early work, synthesizing a series of ammonium alcohols / ammonium phosphates that differed from **fingolimod** / **(S)-fingolimod phosphate** in their linker identity and tail substitution pattern and length [36, 37]. These compounds were demonstrated to be either low potency agonists or antagonists of the S1PRs *in vitro*, but upon biological

evaluation, it was realized that many of these compounds were unable to be phosphorylated to the active compound *in vivo*. To circumvent this problem, non-hydrolyzable ammonium phosphate derivatives were synthesized, including ammonium phosphonates and ammonium phosphothioates [38, 39]. Unfortunately, these compounds were far less potent than their ammonium phosphate analogs.



**Figure 1.8. Notable fingolimod analogs.** Ammonium phosphate, phosphonate, and thiophosphate (not shown) analogs of **fingolimod** and were tested *in vitro* and *in vivo*. None of the analogs had dramatically improved properties over **fingolimod**.

While we and others continued to synthesize similar ammonium phosphates and their analogs, these compounds fell out of interest due to their requirement for bioactivation. Instead, focus was shifted to ammonium carboxylate S1PR modulators (especially those containing the azetidinium carboxylate), as they act directly upon the S1PRs and are easy to synthesize from amino acid precursors [40-43]. These properties are amenable to high-throughput screen, which helped the S1PR modulator literature explode with new compounds [44-46].



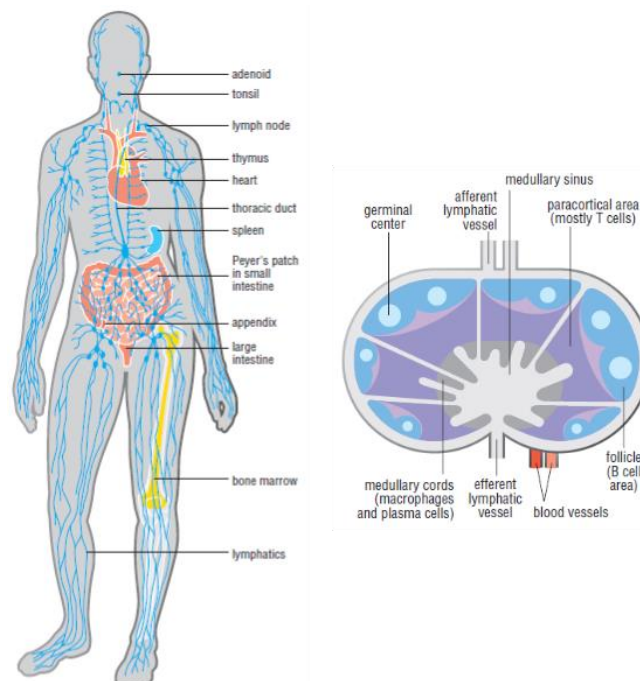
**Figure 1.9. Examples of some azetidinium carboxylate-containing S1PR modulators.** While other amino acid derivatives have been explored, azetidinium carboxylates found the most use. The X and Y groups indicate sites of substitution along the scaffolds.

Most importantly, the X-ray crystal structure of the S1PR<sub>1</sub>-T4L chimeric receptor was published and has led to the development of nonionic S1PR modulators [14]. While this development is only recent,

it will likely lead to significantly new innovation in the field. Overall, this entire body of work has led to the development of many subtype selective agonists and antagonists of the S1PRs, which has allowed for further study and better understanding of S1PR structure and function.

### 1.4 The Role of S1PR Modulation in Immune Cell Trafficking

The wealth of S1PR modulators, as well as biochemical and biological techniques, has allowed for extensive and systematic study of S1PR involvement in immune cell migration and trafficking [47]. The adaptive immune system consists primarily of B cells and T cells and serves to protect organisms from constantly evolving pathogens. These cells stem from a single progenitor cell and mature in primary lymphoid organs, either bone marrow (B cells) or the thymus (T cells) [48, 49]. Both cell types undergo selection processes that ensure that they are fully functional but are incapable of recognizing self-antigens. Functioning cells are allowed to mature, whereas disfunctioning cells are triggered to undergo apoptosis. Imperfections in these processes results in either infection or autoimmune disease [50, 51]. Cells that survive this process are referred to as naïve B or T cells, as they have not yet encountered a pathogen. Naïve cells are released from primary lymphoid organs and then cycle between the periphery and secondary lymphoid organs, which include the lymph nodes and spleen [52].

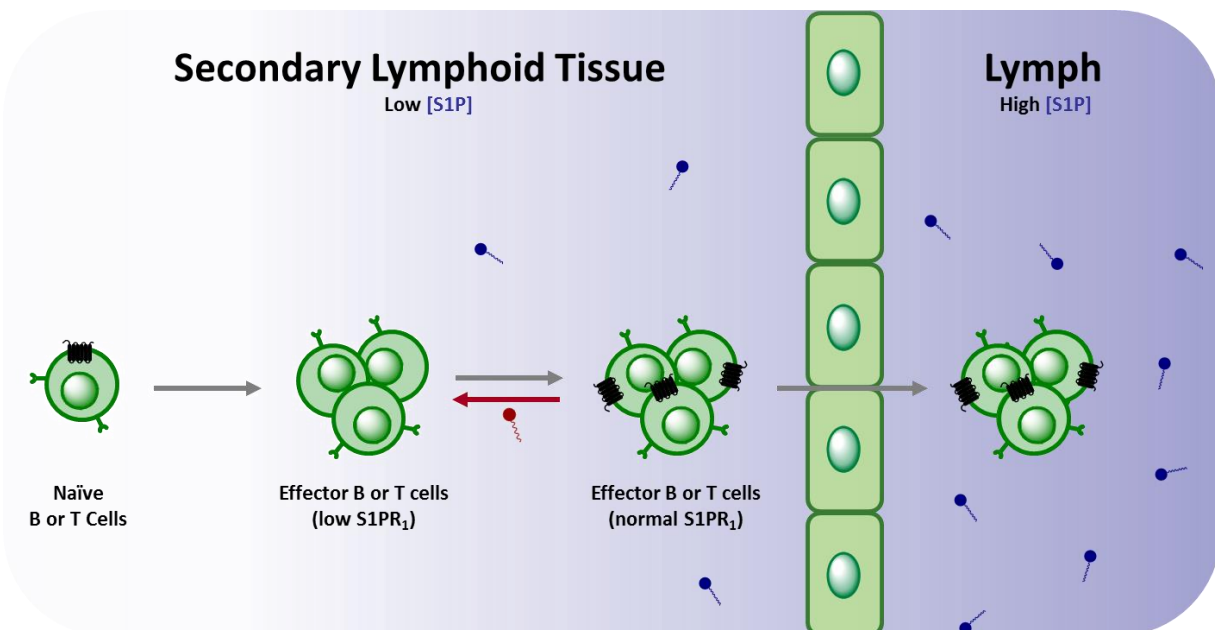


**Figure 1.10. Diagram of lymphatic system and lymph node.** Immune cells mature in primary lymphoid tissue. Once mature, they migrate to the periphery and cycle between the lymphatic / circulatory system and secondary lymphoid organs like the lymph nodes.



When an infection or autoimmune response occurs, antigen-presenting cells (APCs) migrate through the vasculature of the lymphatic system to secondary lymphoid tissue, where they encounter large numbers of naïve T cells. The APCs will then activate complementary T cells, eventually yielding populations of effector B and T cells capable of recognizing the original antigen and any cells that bear it [53]. The effector cells then egress from secondary lymphoid tissue and mediate their immune effects to combat infection or perpetuate autoimmune disease.

This cellular egress is mediated by the **S1P**-S1PR<sub>1</sub> signaling network. Again, **S1P** is produced during the process of inflammation and is subsequently exported outside the cell, establishing a gradient in which **S1P** is highest in blood and lymph (0.1 - 1.1  $\mu\text{M}$ ), modest in interstitial fluid (0.5 - 0.75 pmol / mg), and lowest in cytosol. The exact concentrations are difficult to determine, as **S1P** is highly protein bound in blood and collects in cellular membranes [54, 55]. Lymphocytes express high levels of S1PR<sub>1</sub> on their cell surface. They utilize this receptor to detect **S1P** and undergo S1PR<sub>1</sub>-mediated chemotaxis, migrating along this gradient out of secondary lymphoid tissue [47].



**Figure 1.11. Egress of effector T cells from secondary lymphoid tissue.** B and T cells undergo activation, proliferation, and differentiation to yield a population of effector cells that have low levels of S1PR<sub>1</sub> expression. These levels are restored over time and the B and T cells detect the **S1P** gradient (blue) using S1PR<sub>1</sub> and exit secondary lymphoid tissue. However, the **S1P**-S1PR<sub>1</sub>-mediated migration can be disrupted by **fingolimod** (red), which induces internalization of S1PR<sub>1</sub> *in vivo*.

Lymphocyte migration has been shown to be disrupted in S1PR<sub>1</sub> conditional knockout cells, cells with downregulated expression of S1PR<sub>1</sub>, and cells treated with **(S)-fingolimod phosphate** [56]. Disruption of the **S1P**-S1PR<sub>1</sub> signaling pathway prevents cellular migration and causes lymphocytes to become sequestered within secondary lymphoid tissues. These cells are incapable of reaching peripheral cells, a property that, though problematic for people battling cancer or infection, is desirable for treating those suffering from autoimmune disease. Thus **fingolimod**, once phosphorylated *in vivo* to the active species **(S)-fingolimod phosphate**, serves as a novel immunosuppressant capable of treating autoimmune disease.

### 1.5 Fingolimod for the Treatment of Relapsing-Remitting Multiple Sclerosis

Multiple sclerosis (MS) is an autoimmune disease that is estimated to affect approximately 2.3 million people worldwide [57]. Those afflicted with MS experience chronic deterioration of their motor and cognitive skills, which is often accompanied by acute episodes of increased debilitation. At this point, there is no cure for multiple sclerosis, and patients can only be given disease-modifying drugs that decrease their symptoms and/or disease progression.

Multiple sclerosis is categorized into four subtypes based upon the frequency of these episodes and the nature of disease progression; the most common of these forms is relapsing-remitting multiple sclerosis (RRMS), which occurs in roughly 85% of patients [58].

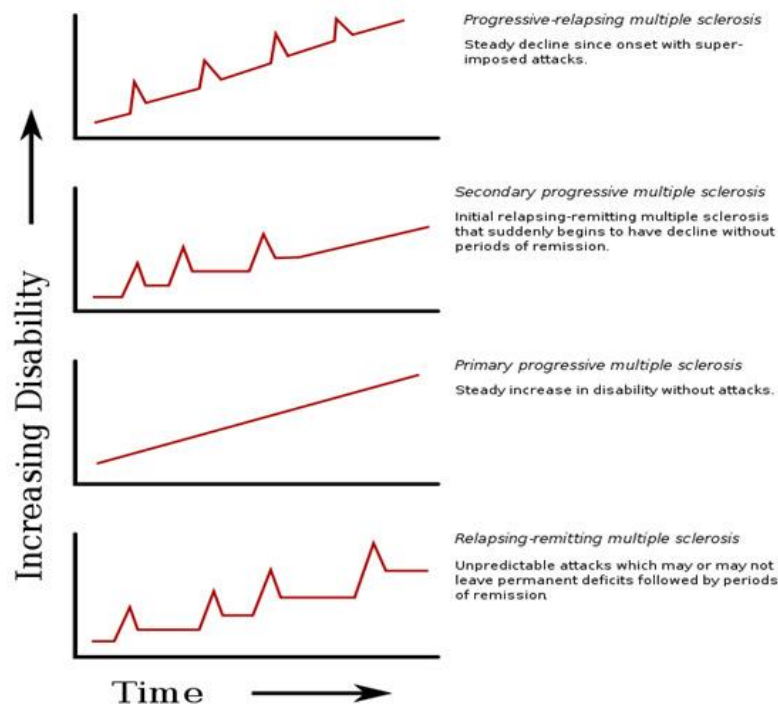


Figure 1.12. Types of MS and their relapse rate and progression. Most patients are diagnosed with relapsing-remitting multiple sclerosis (RRMS).

While B cells have increasingly been shown to play a role in RRMS, T cells are the primary drivers of its pathophysiology [47, 59]. Under most circumstances, T cells are unable to gain access to the central nervous system (CNS) as a result of the blood-brain barrier (BBB), a combination physical and chemical barrier that protects the CNS [60]. In multiple sclerosis, autoreactive T cells become capable of permeating

the BBB and gain entry to the CNS [61]. There they mediate damage to neurons, specifically at axons and the myelin sheath that covers them [62]. Damage to either causes rapid deterioration of a firing neuron's action potential, the traversing electrical signal required to stimulate muscle movement.

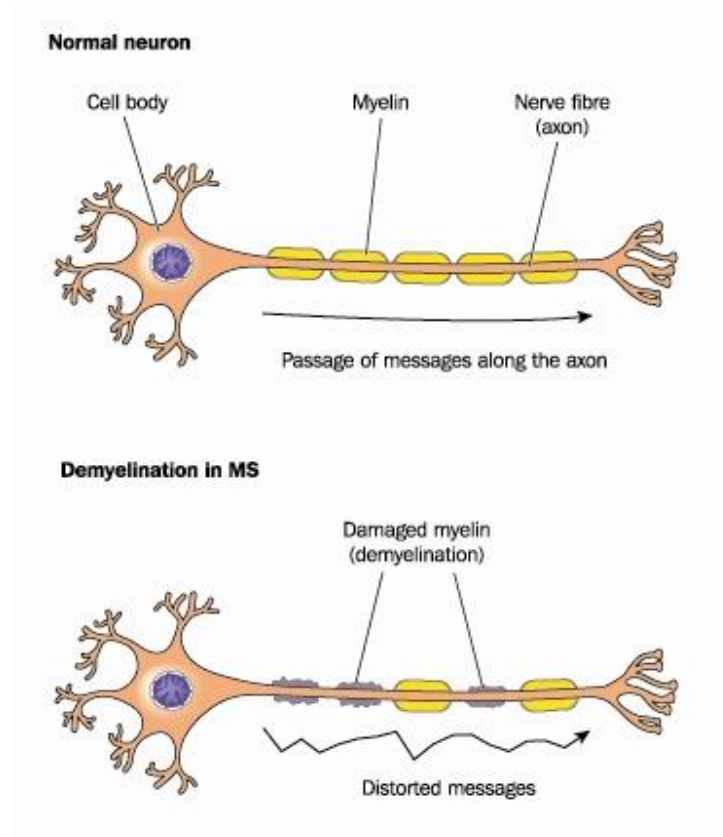


Figure 1.13. Comparison of signal transmission in healthy and demyelinated neurons. A wave of polarization rushes through the axon of healthy neurons. This is insulated by the fatty myelin sheath, which becomes destroyed in multiple sclerosis.

**Fingolimod** was approved by the FDA for the treatment of RRMS in the fall of 2010. It was considered a breakthrough drug, operating via a novel mechanism of action and serving as the first oral RRMS therapy. Until that point, most patients would take injectable biologics (interferon  $\beta$  or antibody treatment) to slow disease progression and treat acute relapses with anti-inflammatory corticosteroids. If these agents fail to slow disease progression, more dramatic treatments can be used to either prevent lymphocyte activation and proliferation or kill lymphocytes altogether [57].

**Fingolimod** is more nearly ideal as a therapeutic because it is orally bioavailable and not cytotoxic. More importantly, it has shown greater efficacy than interferon  $\beta$ , the previous standard of care treatment, in treating patients with RRMS. At its therapeutic dosage, **fingolimod** reduced patient annualized relapse rate by 73.4% versus placebo (interferon  $\beta$  = 67.4% versus placebo). That dosage also reduced the mean number of new/newly enlarged T2 MRI-active brain lesions to 66% of interferon  $\beta$  and 38% of placebo. It also reduced the number of gadolinium-enhancing T1 MRI-active lesions to 27% of placebo [63]. These are the standard clinical measurements used to determine episode occurrence and disease progression. Following these studies, **fingolimod** became the new standard of care, though it is currently competing with another novel compound, **dimethyl fumarate** (Tecfidera™, Biogen Idec) [64].

Again, **fingolimod** administration leads to internalization of S1PR<sub>1</sub> from the surface of lymphocytes, preventing their egress from secondary lymphoid organs. As a result, lymphocytes are unable to enter the bloodstream and perpetuate the autoimmune disease. Patients taking **fingolimod** demonstrate decreased numbers of peripheral blood lymphocytes (20 - 30% of normal levels) – a state referred to as lymphopenia [65]. This decrease is more pronounced in the Th1 and Th17 cell lineage that have been associated with autoimmune disease [66, 67]. By disproportionately affecting those T cell lineages, complications with infection and cancer are reduced, though they are still present and problematic.

Of course, lymphocytes are not the only cells bearing S1PR<sub>1</sub>, and S1PR<sub>1</sub> modulation in other cell types has been linked to side effects of **fingolimod**. S1PR<sub>1</sub>, via activation of the inhibitory G protein  $G\alpha_{i/o}$ , is responsible for tight junction formation and barrier maintenance [68]. Antagonism or internalization of S1PR<sub>1</sub> on endothelial cells compromises these barriers and causes vascular leak [23, 69]. This likely gives rise to minor adverse respiratory effects and macular edema, a reversible condition in which protein deposits in the eye affect vision, and diminishes the integrity of the BBB [70]. These effects were found to be dose-dependent but were observed in those taking the approved dosage.

More importantly, triggering of S1PR<sub>1</sub> induces Gβγ-mediated activation of GIRKS, membrane channels that hyperpolarize the cell with potassium ions [22, 71]. Agonism of S1PR<sub>1</sub> on cardiomyocytes alters ion levels and leads to a decrease in heart rate – a condition known as bradycardia. Luckily, this bradycardia is alleviated over time, likely the result of G protein uncoupling following receptor internalization. This effect is only observed in patients 1 – 24 hours following **fingolimod** administration, though it is substantial enough to require clinical observation of first-time patients for six hours following initial dosing [72]. Fortunately, a recent study has shown that this side effect can be ameliorated by using dose titration [73].

Bradycardia was originally believed to be solely the result of agonism at S1PR<sub>3</sub> based on early results in rodents [74]. Consequently, most research groups focused on the development of S1PR<sub>1</sub>-selective compounds, as they were believed to prevent bradycardia like that observed in rodents administered **fingolimod** or other nonselective S1PR agonists. Recently, it was demonstrated that the cause of bradycardia is actually species-dependent: agonism of S1PR<sub>1</sub> (and potentially agonism of S1PR<sub>3</sub>) is responsible for bradycardia in humans, whereas agonism of S1PR<sub>3</sub> is single-handedly responsible for bradycardia in rodents [75, 76].

Finally, **fingolimod** also stimulates S1PR<sub>5</sub>, which is highly expressed on oligodendrocytes (cells responsible for maintaining the protective myelin sheath covering axons of neuronal cells) and possibly on brain endothelial cells, though this finding is disputed. The agonism of S1PR<sub>5</sub> on oligodendrocytes leads to changes in shape or increased survival depending upon their state of maturity [77]. In the disputed study, agonism of S1PR<sub>5</sub> on brain endothelial cells was shown to enhance barrier integrity, which may improve BBB function [78]. These data, along with the positive results from clinical trials involving the S1PR<sub>1</sub>-S1PR<sub>5</sub> dual agonist **BAF312**, provide preliminary evidence that S1PR<sub>5</sub> agonism likely is beneficial for RRMS treatment [79].

## 1.6 Properties of Ideal Second-Generation S1PR Modulators

Again,  **fingolimod**  functions via a novel mechanism of action to provide improved treatment for RRMS. Unlike other immunosuppressants, it is neither cytotoxic nor disruptive to the activation, differentiation, or proliferation of lymphocytes. As a result, it is safer and more efficacious than the aforementioned treatments. At the same time, it still produces adverse effects as a result of its on-target (S1PR<sub>1</sub>) and off-target (S1PR<sub>3</sub>) activity and pharmacokinetic properties. A second-generation S1PR modulator must demonstrate improved activity and pharmacokinetic profiles to be therapeutically useful.

Property	Effect
Operates via novel mechanism of action	<i>(+) immunosuppression is reversible</i>
Acts at multiple S1PR subtypes	<i>(+) lymphopenia via S1PR<sub>1</sub></i> <i>(-) bradycardia via S1PR<sub>1</sub> and S1PR<sub>3</sub></i> <i>(-) vascular leak via S1PR<sub>1</sub></i> <i>(+) oligodendrocyte survival via S1PR<sub>5</sub></i>
Behaves as a prodrug	<i>(+) orally bioavailable</i> <i>(-) bioactivation required</i> <i>(-) increased bioaccumulation</i>

Table 1.2. Key properties and *in vivo* effects of fingolimod.

With regards to activity, the ideal S1PR modulator would attenuate activity at S1PR<sub>1</sub>, be devoid of activity at S1PR<sub>3</sub>, and stimulate S1PR<sub>5</sub>; such a compound would produce lymphopenia (via S1PR<sub>1</sub>) and promote oligodendrocyte survival (via S1PR<sub>5</sub>) while minimizing bradycardia (via S1PR<sub>3</sub>). Development of subtype-selective compounds is not trivial, but has been achieved before with the S1PRs [80]. However, the mechanism by which such a compound would attenuate S1PR<sub>1</sub> activity can vary and has a significant effect on the side effect profile.

S1PR<sub>1</sub> *antagonists* prevent G protein activation, producing both the desired lymphopenia and undesired vascular leak [69]. The extent of (and ability to balance) these effects is still poorly understood, as very few antagonists have been characterized *in vivo*. Antagonists are attractive however, because they

do not induce S1PR<sub>1</sub>-mediated bradycardia and further study of S1PR<sub>1</sub> antagonists is needed to fully appreciate their utility.

Unlike the antagonists, S1PR<sub>1</sub> *agonists* do not cause lymphopenia and instead only produce bradycardia, as G protein activation does always cause receptor internalization. Only agonists that induce receptor internalization, referred to as *functional antagonists*, produce lymphopenia. This effect would be accompanied initially by bradycardia and subsequently by vascular leak, as the G protein becomes slowly uncoupled from the receptor over time [25]. The balance between receptor recycling and degradation would likely affect the extent and reversibility of these effects, with recycling being preferred to degradation.

When dosed properly, a *functional antagonist*, would likely produce lymphopenia that would only be accompanied by vascular leak. Dose titration was demonstrated to attenuate the initial bradycardia of a S1PR<sub>1</sub> functional antagonist [73]. *Biased agonists*, compounds that elicit receptor internalization without G protein activation, would likely have the same side effect profile. Unfortunately, these compounds have been difficult to develop and are rarely described in the literature [81]. Together, it is likely that some combination of agonist design and dosing could produce an ideal therapeutic, but this requires significant insight into how agonist structure affects receptor signaling and processing.

	Antagonist	Agonist	Functional Antagonist	Titrated Functional Antagonist	Biased Agonist
<b>Lymphocyte Migration</b>	↓	↑	↓	↓	↓
<b>Heart Rate</b>	-	↓	↓	-	-
<b>Barrier Maintenance</b>	↓	↑	↓	↓	↓

**Table 1.3. Effects of S1PR<sub>1</sub> modulators *in vivo*.** S1PR<sub>1</sub> is involved in the signaling of many processes and can either increase or decrease the magnitude of these processes depending upon the type of modulation. Positive effects are shown in green, whereas negative effects are shown in red.



Before a compound can demonstrate any activity *in vivo*, it must first be absorbed by the body. **Fingolimod** is administered as its hydrochloride salt and is > 90% orally bioavailable [82]. Here, its behavior as a prodrug is advantageous, as the amino diol is likely better absorbed than the active phosphate. At the same time, the requirement for bioactivation limits the concentration of the active species *in vivo*, causing **fingolimod** to require higher dosage and potentially bioaccumulate. This same property also makes the drug discovery process significantly more tedious. Thus, amino alcohols like **fingolimod** fell out of favor once it was discovered that direct-acting, amino acid-containing compounds were also orally bioavailable [80, 83, 84]. The ideal second-generation compound would likely possess an amino acid group as a result.

**Fingolimod** and its phosphate have a tendency to bioaccumulate as a result of their long half-lives, which are between six to nine days [82]. Ideally, a therapeutic should have a half-life of ~ 12 - 16 hours, allowing daily administration. Thus, an ideal S1PR modulator should also have a scaffold that permits faster metabolism. Additionally, **fingolimod** has been shown to lead to elevation of transaminases in the liver [85]. A new scaffold would likely have differing effects on liver enzyme expression levels, which would need to be studied to determine its potential toxicity.

---

# 2

---

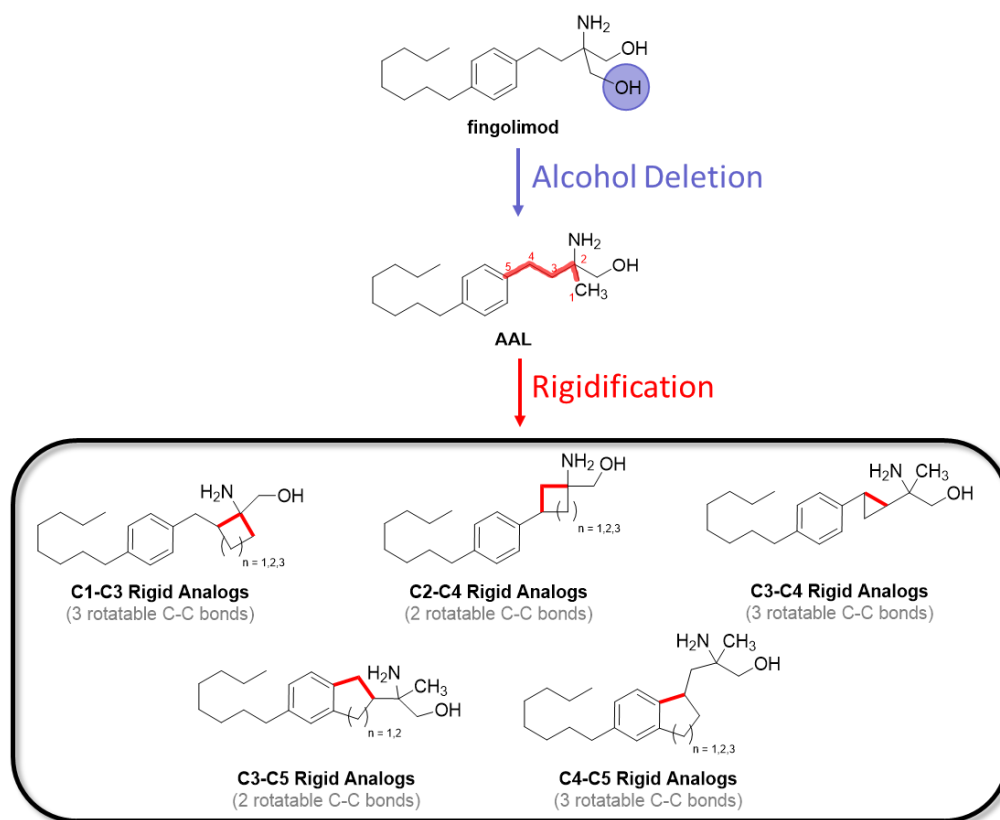
## Rigid Analogs of Fingolimod as Improved Therapeutics

The sphingosine 1-phosphate receptors (S1PR<sub>1-5</sub>) play an important role in a variety of different biological processes, with the effects depending upon both cell type and receptor subtype. **Fingolimod** (Gilenya™, Novartis) is FDA-approved for the treatment of relapsing-remitting multiple sclerosis (RRMS); both the beneficial and the adverse effects of this drug are the result of activity across this receptor family (S1PR<sub>1,3,4,5</sub>). It was believed that more selective compounds would likely show improved efficacy and safety in the treatment of RRMS. We hypothesized that limiting the conformational freedom of the **fingolimod** scaffold would likely alter its S1PR selectivity, potentially resulting in a second-generation compound with the optimal activity profile: functional antagonism of S1PR<sub>1</sub> and agonism of S1PR<sub>5</sub>.

## 2.1 Previous Rigid Analogs of Fingolimod

In medicinal chemistry, molecular scaffolds are commonly rigidified in hopes of modulating their potency and selectivity across different biomolecular targets [86, 87]. Rigidifying a scaffold restricts the number of rotatable bonds and conformations that it can possibly adopt. This decreases the entropic cost paid for confining a ligand to one active conformation. At the same time, this process also changes the shape of the ligand scaffold, affecting the degree of strain and intermolecular contacts that are added or removed upon protein binding. Of course, the extent of these effects will vary depending upon the mode of ligand rigidification and the biomolecular target. Ideally, one rigid analog utilizes the above effects to produce a second-generation compound with the desired activity profile. Again, the ideal **fingolimod** analog would be a dual S1PR<sub>1</sub>-S1PR<sub>5</sub> full or partial agonist; this compound would produce S1PR<sub>1</sub>-mediated lymphopenia and S1PR<sub>5</sub>-mediated oligodendrocyte modulation while minimizing S1PR<sub>3</sub>-mediated bradycardia.

The amino diol of **fingolimod** enables its biological activity. However, it was found that deletion of the non-phosphorylated alcohol actually increased activity *in vivo* [88]. We hypothesized that the resulting compound, **AAL**, could be rigidified in the alkyl linker that connects the two important functionalities: the amino alcohol and the phenyl ring. Rigidifying this region will likely produce the largest effect on binding and can be achieved in a variety of different ways to yield a library of rigid analogs (Figure 2.1). These compounds can be synthesized and tested as racemic mixtures to quickly assess their utility and then as the individual stereoisomers if needed.



**Figure 2.1. Fingolimod, AAL, and synthesized rigid analogs.** Removal of the non-phosphorylated alcohol in **fingolimod** was demonstrated to increase activity. A library of rigid analogs is proposed using this parent compound, **AAL**.

The C1-C3 rigid analogs were previously synthesized as racemates by Dr. Tao Huang, a previous member of the Macdonald Lab [89]. Of the analogs, only the cyclopentyl and cyclohexyl derivatives (where  $n = 2, 3$ ) were acceptable substrates of the mouse and human sphingosine kinases, required for generation of the active compound. These analogs were able to affect potent and long-lasting lymphopenia when administered to mice. The individual stereoisomers of the cyclopentyl derivative were synthesized and demonstrated significant differences in the extent and duration of lymphopenia they produced. Unfortunately, the corresponding phosphates of these analogs were never synthesized and assayed to precisely determine their S1PR activity profile.

The C2-C4 rigid analogs were synthesized as racemates by Dr. Ran Zhu, another member of the Macdonald Lab [90, 91]. Interestingly, these compounds were substrates of mouse sphingosine kinases but not human sphingosine kinases, precluding their utility as therapeutics. Despite this, the

corresponding phosphates were synthesized and demonstrated to be partial agonists of S1PR<sub>1</sub> and S1PR<sub>5</sub> and to be antagonists of S1PR<sub>3</sub>. This interesting result prompted the stereoselective synthesis of the individual stereoisomers of the cyclopentyl derivative (where  $n = 2$ ). Not surprisingly, the stereoisomers showed dramatic differences in their ability to serve as mouse sphingosine kinase substrates and S1PR<sub>1</sub> agonists as well as the extent and duration of the lymphopenia they elicited in mice.

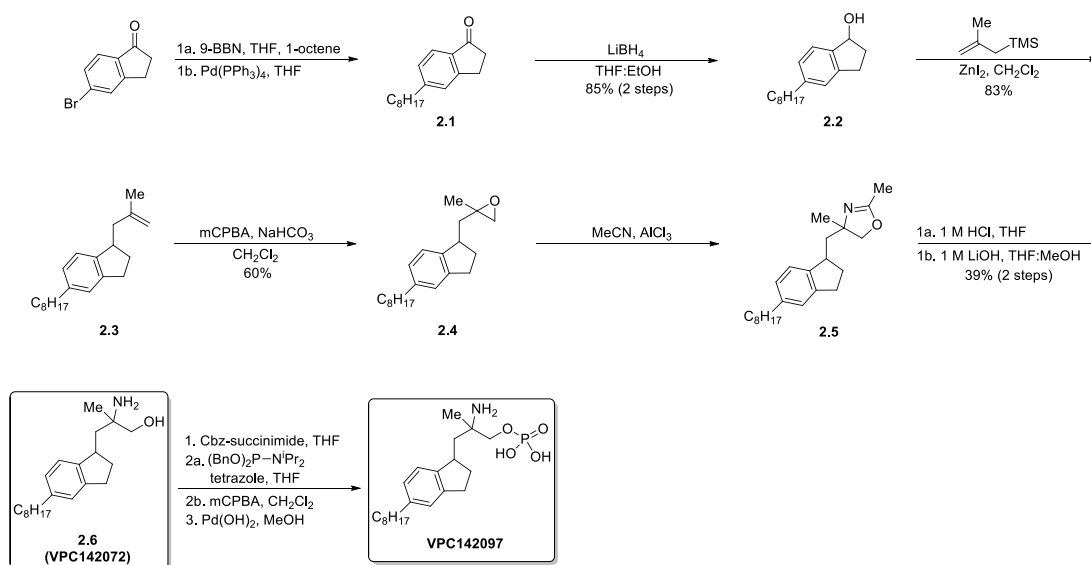
Both the *cis*- and *trans*-cyclopropyl C3-C4 rigid analogs were also synthesized and described by Dr. Tao Huang [89]. Neither of these compounds were substrates of the mouse or human sphingosine kinases. It was determined that any larger derivatives would also likely be devoid of sphingosine kinase activity, similarly preventing their therapeutic utility. As a result, these compounds were not synthesized.

The C3-C5 rigid analogs were synthesized and described by Dr. Tao Huang and through collaboration with BiogenIdec Inc. [89, 92]. The cyclopentyl derivative (where  $n = 1$ ) was synthesized as the racemate, but was not demonstrated to be a substrate of the sphingosine kinases. Surprisingly, two of the individual stereoisomers of the cyclohexyl derivative (where  $n = 2$ ) were substrates of both mouse and human sphingosine kinases. The corresponding phosphates of these derivatives were synthesized and displayed activities similar to that of **(S)-fingolimod phosphate**. One isomer showed potent and long-lasting lymphopenia in mice.

Finally, the C4-C5 rigid analogs were synthesized and described by Dr. Andrew Kennedy, Dr. Tao Huang, and myself [89]. The cycloheptyl derivative (where  $n = 3$ ) was synthesized as the racemate, but was not a substrate of the kinases. The smaller cyclohexyl derivative (where  $n = 2$ ) was also synthesized as a racemic mixture and showed limited sphingosine kinase and S1PR activity. This mixture was also able to produce lymphopenia when administered to mice. These results prompted interest in the smaller cyclopentyl analog ( $n = 1$ ), which was believed to have better activity and is described herein.

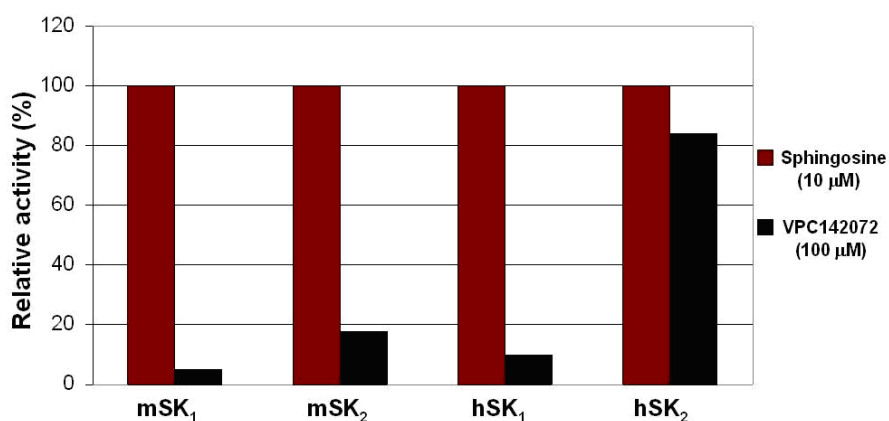
## 2.2 Synthesis & Evaluation of Racemic Indane-Based Rigid Analog (VPC142072 / VPC142097)

Previous work done to synthesize the tetralin- and benzocycloheptane-based analogs determined the synthetic route needed to synthesize the indane-based analog [89]. Synthesis of the analog began with a Suzuki coupling between 5-bromo-1-indanone and the hydroboration product of 9-BBN and 1-octene. Ketone **2.1** was reduced with  $\text{LiBH}_4$  in ethanol to yield **2.2** in 85% over the first two steps. The resulting alcohol was then alkylated in the benzylic position with methallyltrimethylsilane in the presence of the Lewis acid  $\text{ZnI}_2$  in 83% yield. Alkene **2.3** was epoxidized using mCPBA under basic conditions to yield **2.4** in 60% yield. The epoxide was then cleaved with acetonitrile under Lewis acidic conditions to yield oxazoline **2.5**, which was not isolable or characterizable due to degradation during purification. The oxazoline was hydrolyzed using sequential acidic and basic conditions to yield a racemic mixture of amino alcohol **2.6** or **VPC142072** in 39% over two steps. The corresponding amino phosphate, **VPC142097**, was synthesized by Dr. Andrew Kennedy. Its synthesis required protection of the amino group prior to formation of the phosphite ester and then oxidation to yield the corresponding phosphate ester. Global deprotection through hydrogenation yielded the product as a racemic mixture of four stereoisomers, as the stereocenters were previously set during the synthesis of **VPC142072**.



**Scheme 2.1. Synthesis of VPC142072 and VPC142097.** Both compounds were synthesized as racemic mixtures. They were then subjected to *in vitro* and *in vivo* testing to assess their therapeutic potential.

Like its parent compound,  **fingolimod**, **VPC142072** is a prodrug that does not become biologically active until converted to the corresponding phosphate, **VPC142097**, *in vivo*. An assay was conducted to confirm that this bulkier rigid analog is similarly phosphorylated by the mouse and human sphingosine kinase isoforms. Briefly, this was achieved by incubating the **VPC142072** racemate or **sphingosine**, [ $\gamma$ - $^{32}$ P]ATP, and extracts of cells overexpressing each respective kinase. The solutions were then extracted and purified to obtain the radiolabeled reaction products, which were quantified via scintillation counting. The results for **VPC142072** were standardized to results obtained when the endogenous ligand, **sphingosine**, was used as the substrate (Figure 2.2) [34].



**Figure 2.2. Sphingosine kinase assay results.** Results demonstrate that **VPC142072** is phosphorylated *in vitro* by both mouse and human sphingosine kinase isoforms (mSK<sub>2</sub> and hSK<sub>2</sub>), though it is less active than the endogenous ligand, **sphingosine**.

The results demonstrate that both mouse and human sphingosine kinase 2 (mSK<sub>2</sub> and hSK<sub>2</sub>) particularly catalyze the phosphorylation of **VPC142072** much like **fingolimod**, though it is likely that two stereoisomers within the mixture – those possessing *S* stereochemistry at C2 – are unable to be recognized by the kinase and persist as the alcohol. It is important to note that this assay does not measure the rate of the reaction, the rate and extent of the reverse reaction (catalyzed by a corresponding phosphatase), or the ratio of alcohol:phosphate that exists *in vivo*. These will be unique for each stereoisomer and are essential to truly understand the behavior of this compound *in vivo*.

Before beginning *in vivo* study, we wanted to measure the activity of the racemic phosphate, **VPC142097**, at the important S1PRs to determine the effect of rigidification on S1PR subtype selectivity. Agonist activity is determined by measuring the racemate's ability to induce receptor-mediated nucleotide exchange (GTP for GDP) at the corresponding G protein using a nonhydrolyzable, radiolabeled derivative of GTP, specifically [ $\gamma$ - $^{35}$ S]GTP. This was achieved by incubating the **VPC142097** racemate (or other indicated compound) with [ $\gamma$ - $^{35}$ S]GTP and membrane fractions containing a specific S1PR. Following this, the membrane fractions were isolated and [ $\gamma$ - $^{35}$ S]GTP binding was quantified using scintillation counting. The potency ( $pEC_{50}$ ) and efficacy (% of exchange induced by **S1P**) of nucleotide exchange induced by **VPC142097** and **(S)-fingolimod phosphate** were calculated (Table 2.1) [37].

Receptor	<b>VPC142097</b>	<b>(S)-fingolimod Phosphate</b>
<b>S1PR<sub>1</sub></b>	8.2 (70 %)	8.5 (60 %)
<b>S1PR<sub>3</sub></b>	8.4 (30 %)	8.1 (60 %)
<b>S1PR<sub>5</sub></b>	8.5 (70 %)	9.0 (20 %)

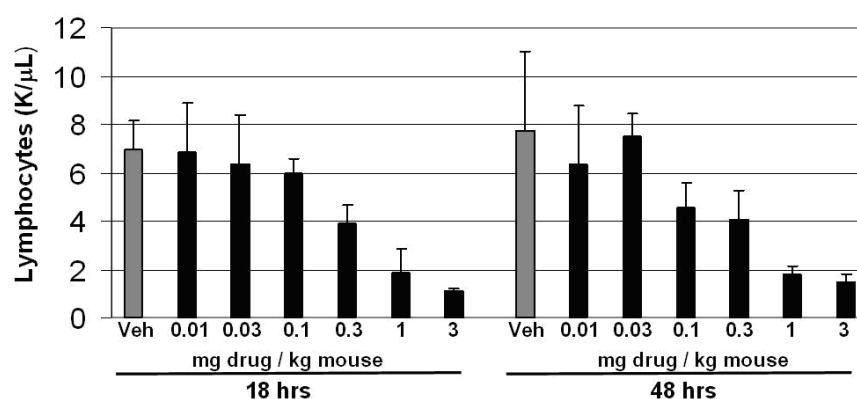
**Table 2.1. Nucleotide exchange assay results.** The potency ( $pEC_{50}$  – black) and efficacy (% of exchange induced by **S1P** – gray) of the ligands are shown. **VPC142097** demonstrates a more preferable S1PR profile than **(S)-fingolimod phosphate**.

The potency of racemic **VPC142097** is very similar to that of **(S)-fingolimod phosphate** across the three important S1PRs, demonstrating that rigidification minimally affects the affinity of these ligands for the individual S1PRs. Conversely, the efficacy of the **VPC142097** racemate is dramatically different than that of **(S)-fingolimod phosphate** across the tested S1PRs; racemic **VPC142097** shows the ideal S1PR<sub>1</sub>-S1PR<sub>5</sub> dual agonist behavior, whereas **(S)-fingolimod phosphate** displays less desirable S1PR<sub>1</sub>-S1PR<sub>3</sub> dual agonist behavior. Though the individual stereoisomers of the racemate will have different activity at the S1PRs, it can be reasonably assumed that the active isomers – those possessing *R* stereochemistry at C2 - will be roughly twice that of the racemate. It should be noted that this assay does not measure receptor



internalization, which is required at S1PR<sub>1</sub> (and may be important at other S1PRs) for a compound's biological activity.

Following these encouraging *in vitro* results, we sought to test the racemic amino alcohol, **VPC142072**, *in vivo* to measure its ability to produce lymphopenia. To do this, healthy mice were dosed various amounts of racemic **VPC142072** via oral gavage. The mice were lightly anesthetized, peripheral blood was collected from the orbital sinus, and the lymphocytes were quantified using a blood analyzer at various time points (Figure 2.3) [91].



**Figure 2.3. Lymphopenia assay results.** Mice were dosed **VPC142072** and blood samples were collected over time and analyzed. Lymphopenia was observed after 18 and 48 hours with doses greater than 0.1 mg drug / kg mouse.

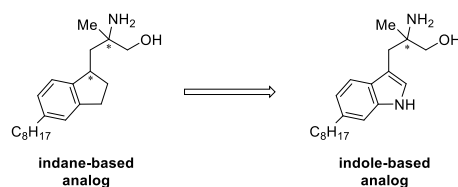
The results of this assay show that the **VPC142072** racemate was able to induce long-lasting lymphopenia at doses above 0.1 milligram drug / kilogram mouse. Inherent chemical and biological limitations precluded direct assessment of the racemate's half-life, though it is expected that the metabolic properties of the individual stereoisomers will likely vary, as was the case with other rigid analogs [91, 92]. Still, the assay does demonstrate that the racemate is orally bioavailable, is phosphorylated *in vivo* to the active species, and induces internalization of S1PR<sub>1</sub>.

The animals' heart rates were not measured during the evaluation period, preventing assessment of the racemate's S1PR<sub>3</sub> activity *in vivo*. However, it is expected that the mixture would produce substantially less bradycardia in mice based upon its *in vitro* properties. It is important to note that measuring heart rate in this assay actually provides little insight into the compound's potential safety in

humans. This is because the effect the S1PRs have on heart rate is species-dependent: agonism of S1PR<sub>1</sub> is primarily responsible for bradycardia in humans, whereas agonism of S1PR<sub>3</sub> is solely responsible for bradycardia in rodents [74-76].

Racemic **VPC142072** showed desirable *in vitro* and *in vivo* activity, which of course will vary for each individual stereoisomer. The individual stereoisomers possessing the **VPC142072** scaffold require relatively difficult, expensive, and time-intensive syntheses. Additionally, the individual stereoisomers would need to undergo the same preliminary screen prior to a more complete characterization of each compound's pharmacokinetic and pharmacodynamic properties, which will likely be dramatically different and potentially undesirable [91, 92]. These synthetic, biochemical, and biological limitations prevent the **VPC142072** scaffold from being therapeutically useful.

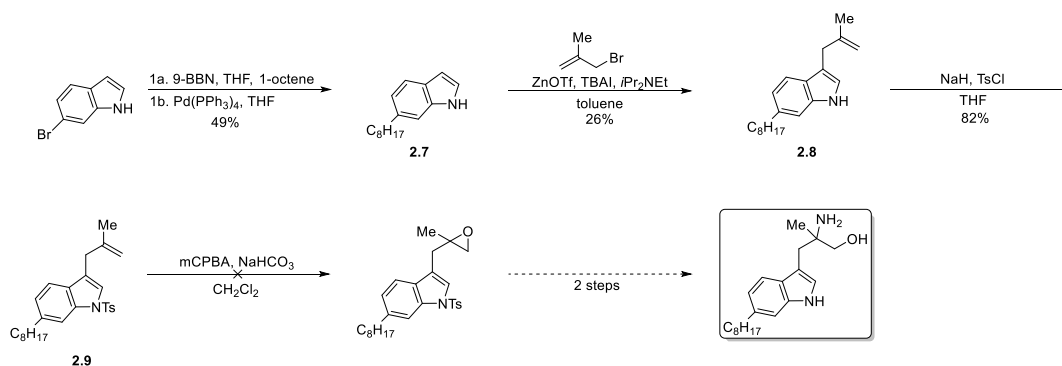
Rather than undertake these complicated studies, we decided to investigate the corresponding indole-based rigid analog (Figure 2.4). It has a very similar shape to the indane-based analog, which would likely give the two compounds similar S1PR subtype selectivities. At the same time, the indole-based analog has one less stereocenter, which allows for more straightforward synthesis and evaluation.



**Figure 2.4. Structure of bicyclic fingolimod rigid analogs.** Two desired **fingolimod** rigid analogs, with stereocenters denoted by an asterisk.

### 2.3 Attempts toward the Synthesis of an Indole-Based Rigid Analog

We chose to synthesize the racemic material first, as many of the reactions were previously parameterized during synthesis of the alkyl rigid analogs. The synthesis began with a Suzuki coupling between 6-bromoindole and the hydroboration product of 9-BBN and 1-octene to afford **2.7** in 49% yield. The product was then alkylated at C3 using 3-bromo-2-methylpropene under Lewis acidic conditions to give alkene **2.8** in 26% yield. The low yield is primarily the result of difficulty in cleanly isolating the final product, as both **2.7** and **2.8** are essentially equal in polarity. Interestingly, this reaction was unsuccessful when the tosyl group was installed prior to alkylation, as the tosyl group is electron withdrawing and impairs the reactivity of the indole ring. Alkene **2.8** was tosylated to afford **2.9** in 82% yield. The subsequent epoxidation using mCPBA was unsuccessful and yielded a variety of different products, which were likely hydroxyindoles.



**Scheme 2.2.** Attempts toward the synthesis of the indole-based rigid analog. The analog was to be synthesized as a racemic mixture and then tested *in vitro* and *in vivo* to assess its therapeutic potential.

Following this negative result, another route to the product was determined and pursued. During that synthesis, the 2.8 Å resolution crystal structure of an antagonist-bound S1PR<sub>1</sub> chimeric receptor was published [14], which caused us to reevaluate our strategy. While the rigid analog approach did yield some valuable information about the structure-activity relationship (SAR) of the  **fingolimod**  scaffold, it was too often marred by limitations in the scaffold's chemistry and biology. As a result, we decided to abandon this approach and transition to rational design of S1PR modulators using information from the crystal structure and computational methods.

## 2.4 Conclusions

The indane-based analog showed the ideal S1PR<sub>1</sub>-S1PR<sub>5</sub> dual agonist activity *in vitro* and produced significant and long-lasting lymphopenia *in vivo*. Unfortunately, the presence of multiple stereocenters in the scaffold limits its therapeutic utility, as the individual stereoisomers will have differing activities / metabolic properties and require difficult syntheses in order to fully characterize them. We began the synthesis of the indole-based analog thinking that it would likely possess the similar S1PR<sub>1</sub>-S1PR<sub>5</sub> activity but would lack the second complicating stereocenter. However, this molecule was never synthesized nor evaluated, as new developments in the field caused us to change our approach to developing second-generation S1PR modulators.

---

# 3

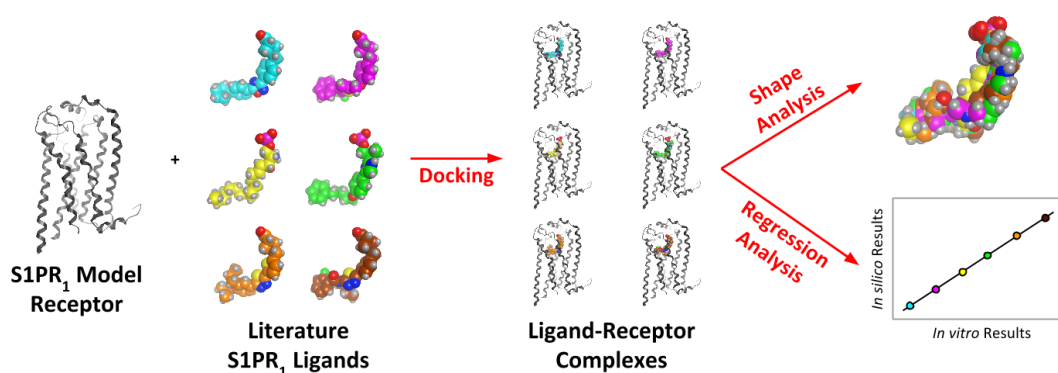
---

## Molecular Modeling of S1PR<sub>1</sub> and its Agonists

The previous studies aimed at synthesizing and evaluating rigid analogs of  **fingolimod**  were met with modest success; most of these molecules possessed multiple chiral centers and suffered from poor kinase activity and receptor selectivity, limiting their utility as either probe compounds or therapeutics. In 2012, the 2.8 Å resolution crystal structure of the antagonist-bound S1PR<sub>1</sub>-T4L chimeric receptor was published. We believed that we could use this structure and the wealth of biochemical data available in the literature to develop a computational model to rationally design new molecules that avoid the problems of our  **fingolimod**  rigid analogs. Such a model could be developed using molecular modeling software, docking compounds previously described in the literature and by comparing their *in silico* and *in vitro* results.

### 3.1 Overall Goals and Design of the Computational Study

Computational chemistry is a unique tool that aids the drug discovery process, giving insight into which molecules should be synthesized and how they potentially behave *in vitro* and *in vivo*. There are a variety of different computational techniques commonly employed: docking studies, pharmacophore identification, dynamics simulations, etc. Each method has its own unique strengths and weaknesses and serves to answer different questions [93, 94]. Regardless of the method chosen, the overall strength of any *in silico* computational study depends upon the amount of *in vitro* and *in vivo* data used to develop and validate the study; the *in silico* environment is merely an approximation that may or may not reflect the actual biological environment.



**Figure 3.1. Docking study protocol.** A series of agonists were docked into a model receptor to obtain qualitative (molecular shape) and quantitative (docking score) results.

In designing our study, we wanted to incorporate information from previous biochemical studies into a model that rationalizes this prior data and that aid in the development of improved S1PR modulators. We felt that this would be best accomplished by developing a S1PR<sub>1</sub> model receptor and incorporating that receptor and previously described S1PR<sub>1</sub> agonists into docking studies. The agonists would be docked into the model receptor, and each resulting agonist-receptor complex would be analyzed to identify common shapes or interactions. Finally, linear regression would be used to compare *in silico* docking scores with *in vitro* biochemical results (Figure 3.1). The overall predictability and utility of the model would be determined by designing, synthesizing, and evaluating novel compounds.

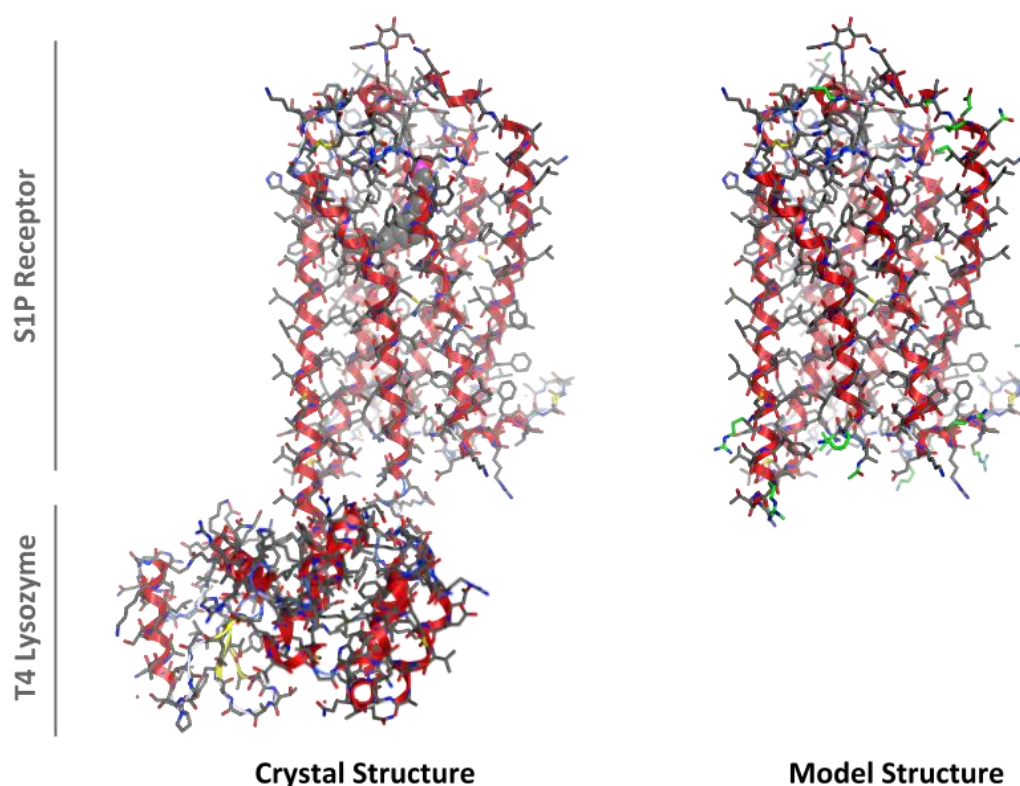
### 3.2 Generation of the S1PR<sub>1</sub> Model Receptor

Previous work identified sites in the S1PR<sub>1</sub> pocket that bind zwitterionic ligands. Site mutagenesis studies demonstrated that R120<sup>3,28</sup>, E121<sup>3,29</sup>, and R292<sup>7,34</sup> were involved in binding the ammonium phosphate of the agonist **S1P** [16]. X-ray crystallography demonstrated that only R120<sup>3,28</sup> and E121<sup>3,29</sup> were directly involved in binding the ammonium phosphonate of the antagonist **ML056** (X = O, R = C<sub>6</sub>H<sub>13</sub>) and that Y29, K34, N101<sup>2,60</sup>, and a bound H<sub>2</sub>O molecule also interacted with the zwitterion [14]. Again, this finding led to multiple theories for the role of R292<sup>7,34</sup> in S1PR<sub>1</sub> function: R292<sup>7,34</sup> either binds differently to antagonists and agonists or it serves as a “cationic lure” that facilitates entry of charged ligands from the membrane [17]. A second, more extensive series of mutagenesis studies identified several residues involved in receptor activation and signal transduction [18]. Not surprisingly, these were found along helices III, V, and VI, which are known to be involved in the activation of other Class A GPCRs [95].

Our model was developed to be in agreement with these previous studies and to not deviate significantly from the X-ray crystal structure of the S1PR<sub>1</sub>-T4L chimeric receptor (PDB ID: 3V2Y). While S1PR<sub>1</sub> undergoes both covalent and noncovalent modification of its structure, we felt that structural changes made *in silico* would be speculative and difficult to interpret. This has two important implications. First, by not repositioning R292<sup>7,34</sup>, we assumed that it serves solely as a “cationic lure” and that it does not interact with ligands localized deep within the receptor pocket. More importantly, by not solvating the unoccupied portions of the receptor pocket (which may not exist as a void *in vivo*), we assumed that solvation and solvent displacement are essentially equal for S1PR<sub>1</sub> ligands regardless of their orientation in the pocket, though this is unlikely to be the case *in vivo*.

The S1PR<sub>1</sub> model receptor was generated following minimal processing of the crystal structure. First, those residues corresponding to the T4 lysozyme were removed, those side chains unresolved in the crystal structure were added, and the gaps and truncated C and N termini were capped. These alterations primarily affected the intracellular and extracellular faces of the receptor; they were performed to simplify

any future dynamics simulations and likely had no impact on the present docking studies. Most importantly, hydrogen atoms were titrated into the model receptor so that it attained the proper valency and charge, ensuring that K34, R120<sup>3,28</sup>, and E121<sup>3,29</sup> were charged and that Y29, N101<sup>2,60</sup>, and the bound water molecule were neutral and that these groups were properly oriented to optimize interactions inside in the pocket. This work yielded the S1PR<sub>1</sub> model receptor, which was utilized in all subsequent study (Figure 3.2).

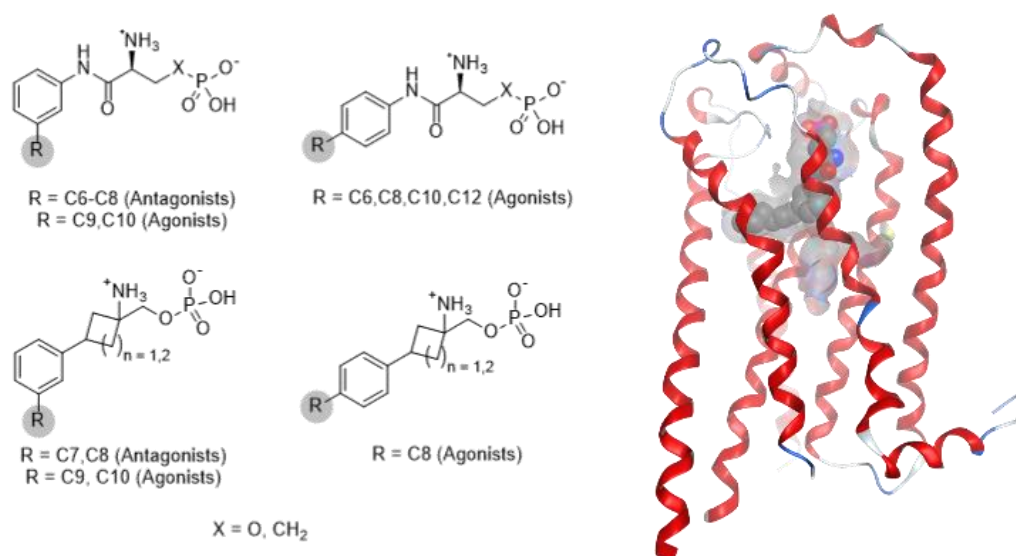


**Figure 3.2. Comparison of the S1PR<sub>1</sub>-T4L crystal structure with the prepared model structure.** Residues corresponding to the T4 lysozyme were deleted and unresolved atoms (green) and hydrogens (not shown) were added so that the side chains were complete and all atoms were of proper valency and charge.



### 3.3 Design of the Docking Study and Choice of Ligand Set

In the crystal structure, the tail of the antagonist, **ML056**, is directed into the side cavity of the pocket, leaving the linear cavity unoccupied. Interestingly, we previously observed a relationship between ligand shape and activity at S1PR<sub>1</sub>: bent, *meta*-substituted molecules like **ML056** tended to behave as antagonists whereas linear, *para*-substituted molecules like **S1P** and **(S)-fingolimod phosphate** tended to behave as agonists, though tail length also affected ligand activity (Figure 3.3) [37, 39, 96]. Interestingly enough, previous *in silico* studies using homology models developed from the rhodopsin crystal structure predicted that agonists would be oriented linearly in the S1PR<sub>1</sub> pocket [18, 19].



**Figure 3.3.** SAR of some S1PR<sub>1</sub> modulators. Unlike *para*-substituted compounds, short, *meta*-substituted compounds like **ML056** (X = O, R = C<sub>6</sub>H<sub>13</sub>) behave as antagonists (left). The bent structure of **ML056** positions its tail into the side cavity of the overall S1PR<sub>1</sub> pocket, outlined in gray (right).

We wanted to investigate this phenomenon to see if we could differentiate the ligand requirements for antagonism and agonism of S1PR<sub>1</sub>. This could be achieved by docking a series of ligands into the side and linear cavities of the model receptor, with the ligands' *in silico* results (both when oriented into the side cavity and into the linear cavity) being compared to their *in vitro* results to elucidate the differences between antagonism and agonism. This would be the first step in developing an overall model that explains how ligands bind, induce G protein activation, cause receptor internalization, and

affect the subsequent receptor trafficking or degradation. A model that provides that level of insight would be invaluable to the design and development new S1PR modulators to treat RRMS.

We initially sought to conduct our study using the above *meta*- and *para*-substituted compounds. Most of the compounds only weakly interact with S1PR<sub>1</sub> (micromolar to millimolar activities) and often yielded inconsistent results that were not conducive to quantitative modeling studies. Following this negative result, we decided to limit our study to only those compounds that are potent modulators of S1PR<sub>1</sub> (low and sub-nanomolar activities). Unfortunately, this requirement precluded us from studying the *meta*- and *para*-substituted compounds and all other S1PR<sub>1</sub> antagonists reported at the time of study.

Following this, we looked to dock compounds from our library of **fingolimod** rigid analogs, believing that the rigid head groups would cause them to occupy either the side or linear cavity of the receptor, but not both. As discussed previously, the active amino phosphates were only assayed for a few of these scaffolds, with even fewer of them having had their individual stereoisomers characterized. Thus, docking studies conducted with the **fingolimod** rigid analogs were similarly fruitless and were subsequently abandoned.

Finally, we settled on docking agonists that contained either an ammonium phosphate [97, 98] or azetidinium carboxylate [40-43] head group. While both are present in compounds that have shown clinical efficacy [75], the azetidinium carboxylate head group is particularly desirable due to its rigidity, a property that limits conformational mobility and facilitates the docking process. Overall, these zwitterionic ligands are easier to initially place within the receptor pocket than more recent S1PR<sub>1</sub> agonists that lack a charged head group [44].

The chosen molecules belong to one of seven unique scaffolds, each comprised of several compounds that vary by the identity of their substituents (Table 3.1). This ensures that the test set shows diversity in chemical structure, allowing the docking study to explore a variety of ligand shapes and be broadly applicable. Ultimately, only compounds possessing minimal heteroatoms were chosen, as

heteroatoms may produce favorable or unfavorable electrostatic interactions that would complicate interpretation of the results. Additionally, compounds that did not follow their scaffold's general SAR trend were eliminated, as these outliers would likely impair data analysis. Finally, only compounds that obtained *in vitro* potency values ( $\text{pIC}_{50}$  of [ $^{33}\text{P}$ ]S1P displacement or  $\text{pEC}_{50}$  of [ $^{35}\text{S}$ ]GTP $\gamma$ S exchange) greater than 8.0 were utilized, as less potent compounds yielded poor docking results.

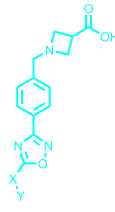
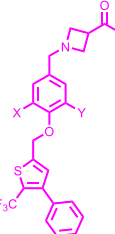
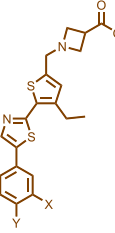
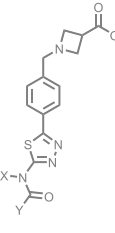
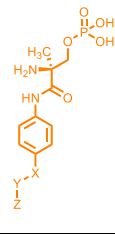
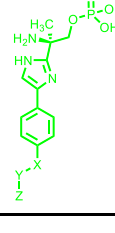
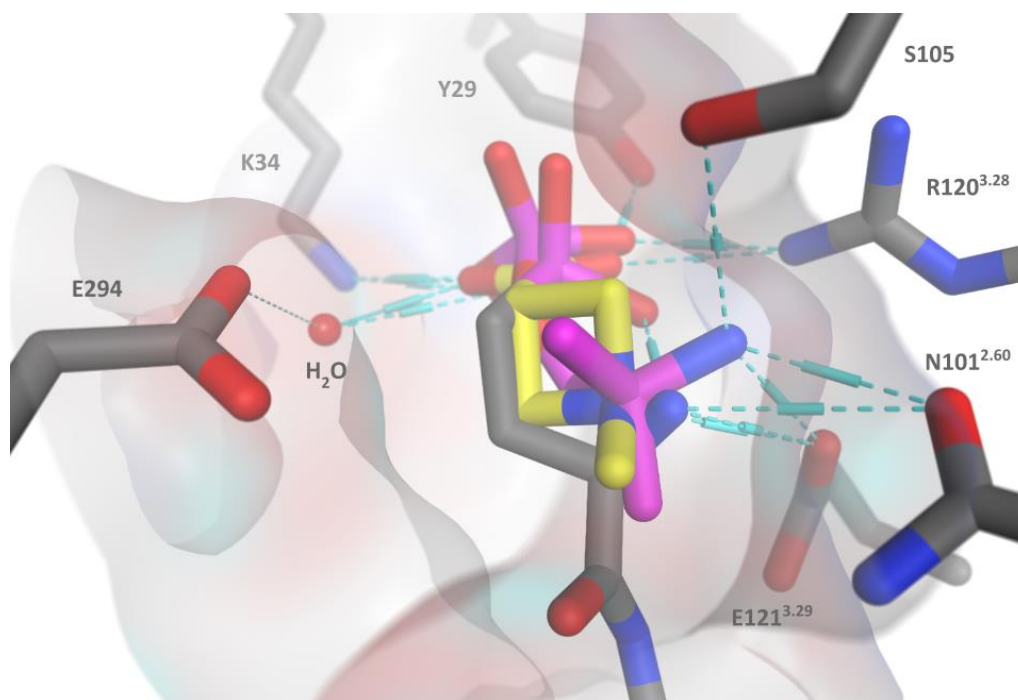
	pIC <sub>50</sub>	pEC <sub>50</sub>	X	Y	Z	
 <p><b>Scaffold A</b></p>	15	8.1	-	Thiophenyl	2-CF <sub>3</sub> , 3-Ph	-
	21	8.4	-	Ph	4-C(CH <sub>3</sub> ) <sub>3</sub>	-
	22	8.9	-	Ph	4-C(CH <sub>3</sub> ) <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	-
	23	8.9	-	Ph	4-CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	-
	24	8.5	-	Ph	4-CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	-
	26	9.2	-	Ph	4-CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	-
	27	8.4	-	Ph	4-CH <sub>2</sub> C(CH <sub>3</sub> ) <sub>3</sub>	-
	30	8.3	-	Ph	4-cyclopropyl	-
	31	8.7	-	Ph	4-cyclobutyl	-
	32	9.4	-	Ph	4-cyclopentyl	-
	33	8.9	-	Ph	4-cyclohexyl	-
34	8.8	-	Ph	4-Ph	-	
 <p><b>Scaffold B</b></p>	18	8.9	9.3	H	H	-
	20	8.1	8.2	CF <sub>3</sub>	H	-
	22	8.5	8.9	Cl	H	-
	23	8.6	8.9	F	H	-
	24	8.7	9.1	CH <sub>3</sub>	H	-
	25	8.5	8.9	CH <sub>2</sub> CH <sub>3</sub>	H	-
	27	8.4	8.9	CH <sub>3</sub>	CH <sub>3</sub>	-
	28	8.1	8.6	Cl	Cl	-
 <p><b>Scaffold C</b></p>	24a	-	8.3	H	OPh	-
	24b	-	8.0	H	OCH(CH <sub>3</sub> ) <sub>2</sub>	-
	24c	-	8.5	CH <sub>3</sub>	OCH(CH <sub>3</sub> ) <sub>2</sub>	-
	24d	-	8.5	CH <sub>2</sub> CH <sub>3</sub>	OCH(CH <sub>3</sub> ) <sub>2</sub>	-
	24e	-	9.0	CH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	OCH(CH <sub>3</sub> ) <sub>2</sub>	-
	24f	-	8.9	CH(CH <sub>3</sub> ) <sub>2</sub>	OCH(CH <sub>3</sub> ) <sub>2</sub>	-
 <p><b>Scaffold D</b></p>	18	-	9.5	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	2-F-Ph	-
	23	-	8.3	(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	2-F-Ph	-
	24	-	8.8	CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	2-F-Ph	-
	25	-	8.7	(CH <sub>2</sub> ) <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	2-F-Ph	-
	28	-	8.7	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	CH <sub>2</sub> Ph	-
	30	-	9.0	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	Ph	-
	31	-	9.1	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	2-Cl-Ph	-
 <p><b>Scaffold E</b></p>	11c	9.1	9.3	O(CH <sub>2</sub> ) <sub>2</sub>	Ph	4-Ph
	11i	8.8	7.4	O(CH <sub>2</sub> ) <sub>4</sub>	Ph	-
	11j	9.8	9.0	O(CH <sub>2</sub> ) <sub>5</sub>	Ph	-
	11m	9.0	8.6	O(CH <sub>2</sub> ) <sub>4</sub>	cyclohexyl	-
	11n	8.7	7.3	OCH <sub>2</sub>	Ph	4-Ph
	11o	9.1	8.5	O(CH <sub>2</sub> ) <sub>3</sub>	Ph	4-Ph
	11p	8.3	7.4	O(CH <sub>2</sub> ) <sub>4</sub>	Ph	4-Ph
	12a	9.0	8.6	O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	-	-
 <p><b>Scaffold F</b></p>	11q	8.2	7.4	O(CH <sub>2</sub> ) <sub>2</sub>	Ph	Ph
	11r	9.1	8.7	OCH <sub>2</sub>	Ph	Ph
	11s	9.1	8.1	O(CH <sub>2</sub> ) <sub>4</sub>	Ph	-
	11t	9.8	8.0	O(CH <sub>2</sub> ) <sub>5</sub>	Ph	-
	12d	9.5	9.3	O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	-	-
<p><b>Other</b></p>	S1P	9.2	8.3	-	-	-
	(S)-fingolimod phosphate	9.6	8.9	-	-	-

Table 3.1. Structure and biochemical data for compounds used in docking studies.

### 3.4 Results from Docking Studies with Head Groups

Before the chosen agonists were docked, their corresponding head groups were optimized within the receptor pocket. To do this, the azetidinium carboxylate (both *cis*- and *trans*-*N*-methylazetidinium-3-carboxylate) and the ammonium phosphate (2-ammonium-2-methylpropyl hydrogenphosphate) head groups were constructed and then positioned in various orientations in the appropriate region of the S1PR<sub>1</sub> model receptor pocket. Each resulting head group-receptor complex was allowed to minimize in the Amber12:EHT forcefield [99] and subsequently scored using the GBVI/WSA dG methodology [100] to estimate the binding energy, ensuring that contacts were established between the head groups and important sites of the pocket (Figure 3.4).



**Figure 3.4. Optimized head groups and ML056 inside the S1PR<sub>1</sub> model receptor pocket.** The ammonium phosphate (magenta) and *trans* azetidinium carboxylate (yellow) are shown atop **ML056** (gray), the antagonist of the crystal structure, within the pocket of the S1PR<sub>1</sub> model (shaded region). The important solvent molecule and side chains are identified, and their interactions with the different ligands are highlighted (cyan).

The substituents on the zwitterionic azetidinium carboxylate can be either *cis* or *trans* to each other, with both species being interconvertible under physiological conditions. While both *cis* and *trans* head groups were built and minimized in the receptor pocket, only the *trans* head group was capable of

establishing contacts with R120<sup>3.28</sup> and E121<sup>3.29</sup> simultaneously. In the top pose for the *trans* azetidinium carboxylate (calculated binding energy of -5.1 kcal / mol), contacts were made between the head group and all other important sites with the exception of N101<sup>2.60</sup>. This result was expected, as the lone ammonium hydrogen interacts more strongly with the charged carboxylate of E121<sup>3.29</sup> than the neutral amide of N101<sup>2.60</sup>.

Unlike the azetidinium carboxylate, the ammonium phosphate head group was able to interact with all of the important sites within the receptor pocket, owing to the decreased degree of substitution on its ammonium nitrogen (primary amine vs. tertiary amine). This is in agreement with biochemical studies showing that ammonium phosphates are more potent than azetidinium carboxylates of the same scaffold [45]. Interestingly, a previously undescribed hydrogen bond was observed between a hydrogen of the ammonium and the oxygen of the S105 side chain. The validity of this interaction could be confirmed using site mutagenesis studies.

It is important to note that the top pose for the ammonium phosphate (calculated binding energy of -5.5 kcal / mol) does not perfectly overlap with the ammonium phosphonate of **ML056** from the X-ray crystal structure. The main difference between the two is that the optimized ammonium phosphate does not form strong intramolecular ionic contacts and instead establishes a more direct contact with R120<sup>3.28</sup>. Both zwitterion conformations likely exist and interchange rapidly *in vivo*, and the two poses represent different snapshots of the mobile ligand-receptor complex. On average though, the heavy atoms only deviate by  $\sim 1$  Å between the two zwitterionic groups, with the greatest difference being 1.5 Å. This deviation might also be the result of the different stereochemistry of the chiral carbon of the ammonium phosphate head group and the ammonium phosphonate of **ML056**, which are *R* and *S* respectively.

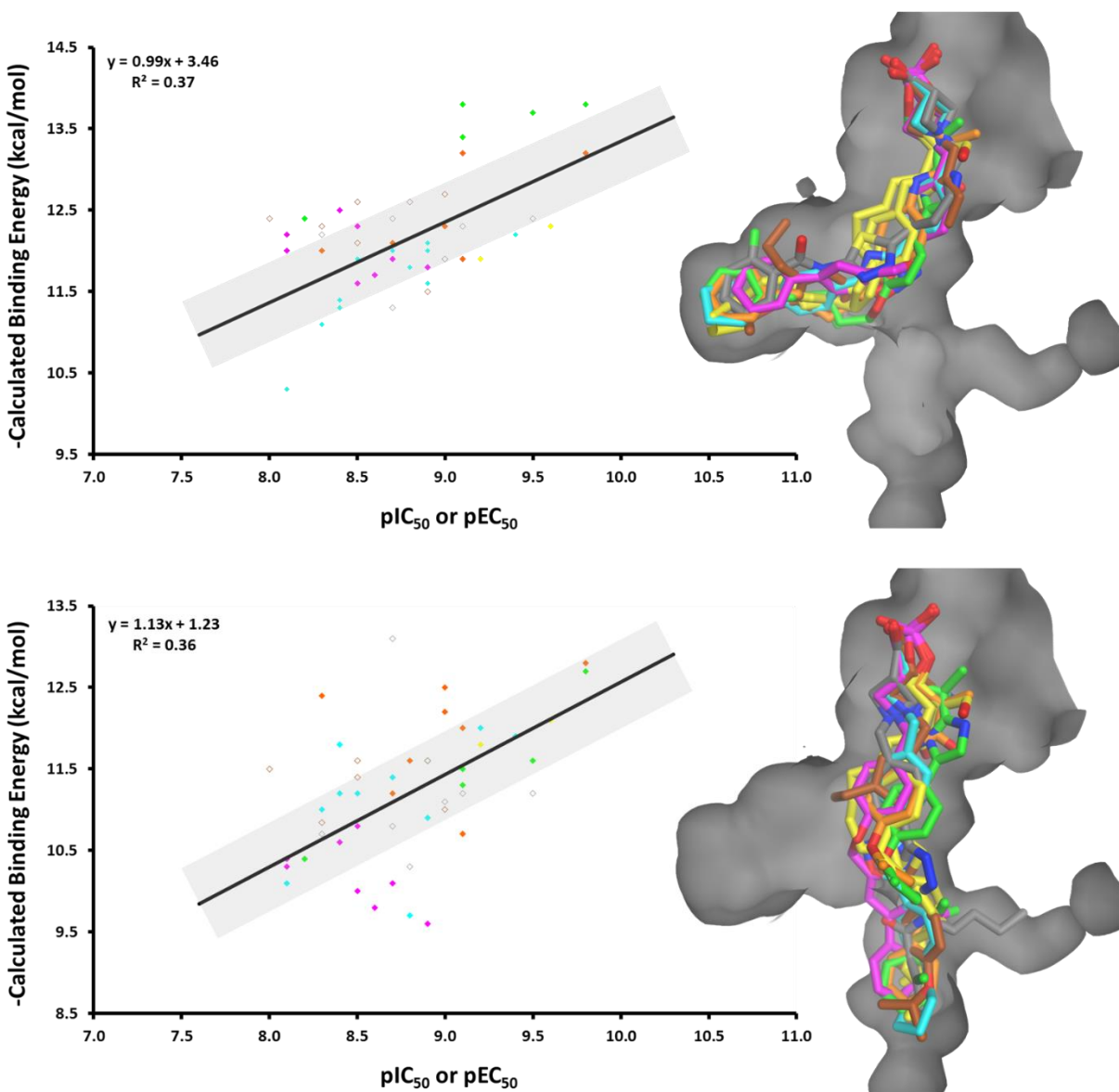
### 3.5 Results from Docking Studies with Complete Ligands

With the head groups placed, the whole ligands were constructed and subsequently docked into both the side and linear cavities of the model receptor to test our original hypothesis. Our initial induced fit docking studies achieved poor and irreproducible results: essential contacts between the receptor and the head groups were usually broken and the ligands obtained strained geometries and low binding energies. This is because the induced fit docking process is essentially a brief dynamics simulation, and the ligand-receptor complex can become kinetically trapped in local minima and never properly equilibrate and achieve its global minimum.

To address this issue, we developed a stepwise protocol that optimized the ligand pose prior to induced fit docking. Using the minimized head groups, the different compounds were constructed and underwent successive conformational searches and virtual screens. The conformational searches sampled all possible conformations of each ligand and the subsequent virtual screens determined top poses for the ligands docked into both the side and linear cavities of the model receptor. Following this, a second series of conformational searches and virtual screens were employed to yield more refined poses for the compounds docked again into both the side and linear cavities. These new top poses were added into the appropriate region of the pocket and each resulting ligand-receptor complex was minimized in the Amber12:EHT forcefield and scored using the GBVI/WSA dG methodology to estimate the binding energy. While time intensive, this process sampled all conformations of the ligands and yielded binding energies and docking poses that varied by less than 2 % with repeated trials (replicate data not shown).

The calculated binding energy for each orientation of each compound was then compared to the compound's corresponding *in vitro* potency value ( $pIC_{50}$  of [ $^{33}P$ ]S1P displacement or  $pEC_{50}$  of [ $^{35}S$ ]GTP $\gamma$ S exchange), and these results are compiled in Table 3.2. Linear regression analysis was performed on data from bent conformers docked into the side cavity and then again on linear conformers docked into the linear cavity (Figure 3.5). Compounds possessing only  $pEC_{50}$  values are shown but are excluded from

regression analysis; the assay used to obtain pEC<sub>50</sub> values measures both the ligand's affinity for the receptor and its ability to induce G protein activation, and the latter component complicates our investigation of binding orientation.



**Figure 3.5. Results from docking studies.** Poses from top members of each scaffold (right) and regression analysis (left) performed on *in silico* and *in vitro* results. Top poses and data points are colored according to their scaffold; points corresponding to pIC<sub>50</sub> values (shaded) are included in the regression, whereas points corresponding to pEC<sub>50</sub> values (hollow) are shown for comparison. Studies in which the ligands are docked into the side cavity (top) are more predictive than those with ligands docked into the linear cavity (bottom), with most compounds falling within  $\pm 0.5$  log order of the line (gray shaded area of plot).



			Calculated Binding Energy (kcal / mol)		
	pIC <sub>50</sub>	pEC <sub>50</sub>	Bent Orientation	Linear Orientation	
Scaffold A	15	8.1	-	-10.3	-10.1
	21	8.4	-	-11.3	-11.2
	22	8.9	-	-12.0	-10.9
	23	8.9	-	-11.6	-11.6
	24	8.5	-	-11.9	-11.2
	26	9.2	-	-11.9	-12.0
	27	8.4	-	-11.4	-11.8
	30	8.3	-	-11.1	-11.0
	31	8.7	-	-12.0	-11.4
	32	9.4	-	-12.2	-11.9
33	8.9	-	-12.1	-11.6	
34	8.8	-	-11.8	-9.7	
Scaffold B	18	8.9	9.3	-11.8	-9.6
	20	8.1	8.2	-12.2	-10.3
	22	8.5	8.9	-11.6	-10.0
	23	8.6	8.9	-11.7	-9.8
	24	8.7	9.1	-11.9	-10.1
	25	8.5	8.9	-12.3	-10.8
	27	8.4	8.9	-12.5	-10.6
	28	8.1	8.6	-12.0	-10.4
Scaffold C	24a	-	8.3	-12.3	-10.8
	24b	-	8.0	-12.4	-11.5
	24c	-	8.5	-12.1	-11.6
	24d	-	8.5	-12.6	-11.4
	24e	-	9.0	-12.7	-11.0
	24f	-	8.9	-11.5	-11.6
Scaffold D	18	-	9.5	-12.4	-11.2
	23	-	8.3	-12.2	-10.7
	24	-	8.8	-12.6	-10.3
	25	-	8.7	-12.4	-10.8
	28	-	8.7	-11.3	-13.1
	30	-	9.0	-11.9	-11.1
	31	-	9.1	-12.3	-11.2
Scaffold E	11c	9.1	9.3	-13.2	-12.0
	11i	8.8	7.4	-12.6	-11.6
	11j	9.8	9.0	-13.2	-12.8
	11m	9.0	8.6	-12.3	-12.2
	11n	8.7	7.3	-12.1	-11.2
	11o	9.1	8.5	-11.9	-10.7
	11p	8.3	7.4	-12.0	-12.4
	12a	9.0	8.6	-11.9	-12.5
Scaffold F	11q	8.2	7.4	-12.4	-10.4
	11r	9.1	8.7	-13.4	-11.5
	11s	9.1	8.1	-13.8	-11.3
	11t	9.8	8.0	-13.8	-12.7
	12d	9.5	9.3	-13.7	-11.6
Other	S1P	9.2	8.3	-11.9	-11.8
	(S)-fingolimod phosphate	9.6	8.9	-12.3	-12.1

Table 3.2. Full numerical results from compounds used in docking studies.

When docked into the side cavity (occupied by **ML056** in the X-ray crystal structure), ligands obtained predicted binding energies between -10.3 kcal / mol and -13.7 kcal / mol. These binding energies correlate well with the observed *in vitro* potency values. This demonstrates that the receptor model, docking protocol, and scoring method reasonably approximate the binding energetics of each ligand-receptor complex. Overall, the data set shows linearity ( $m = 0.99$  kcal / mol) between the calculated binding energy and *in vitro* activity. The correlation of this data ( $R^2 = 0.37$ ) is limited as a result of error in the biochemical data itself (estimated to be around 20 %) [43], in using biochemical data from multiple sources (though each data set was standardized using **S1P** as a reference), and in having few characterized compounds near the extremes of this spectrum of activity. More importantly, it is difficult to develop a precise yet comprehensive model to describe each specific ligand-receptor complex, even when using an induced fit docking protocol.

Compounds belonging to Scaffold A (cyan) established specific hydrophobic interactions with M124<sup>3.32</sup>, F125<sup>3.33</sup>, W269<sup>6.48</sup>, and L297<sup>7.39</sup> when docked into the side cavity. Interestingly, three of these residues showed ligand-dependent changes in activation of wild type and mutant S1PR<sub>1</sub> by **S1P** and **SEW2871** [18]. Compounds **A-21** through **A-34** differ only by the terminal substituent of the phenyl ring, and the model and docking method were able to differentiate the subtle differences in structure and correctly order the activity of these compounds. Compound **A-15** was correctly predicted to be the least potent, but its activity was underestimated by approximately a log order.

Like **A-15**, compounds belonging to Scaffold B (magenta) possess a terminal 2-trifluoromethyl-3-phenylthiophenyl group. These compounds were shown to also interact with W269<sup>6.48</sup> and L297<sup>7.39</sup> when oriented into the side cavity. The predicted binding energies for **B-18** through **B-28** are scattered about the expected region and do not correlate perfectly with activity. Here, the model and docking method were unable to explain how substituents on the phenyl ring modestly affect the activity of members of this scaffold.

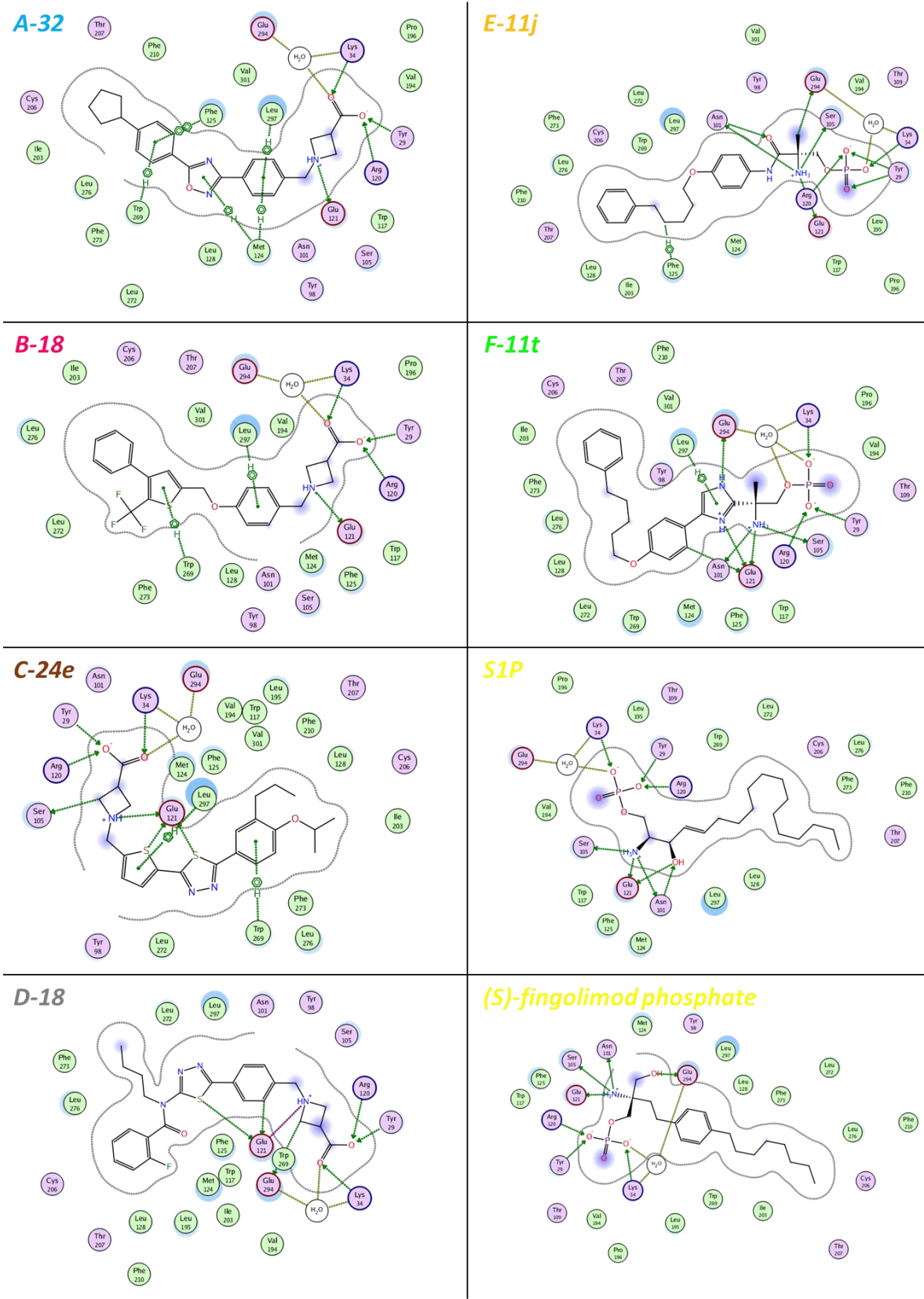
In their bent conformation, compounds possessing Scaffold C (brown) similarly established interactions with W269<sup>6,48</sup> and L297<sup>7,39</sup>. Additionally, the sulfur atoms of the thiophene and thiadiazole rings were positioned to interact with D121<sup>3,29</sup>. Compounds from Scaffold D (gray) established a similar interaction between their thiadiazole ring and D121<sup>3,29</sup>, though they did not form specific hydrophobic interactions with other sites along the side cavity. Not surprisingly, the calculated binding energies are not particularly predictive of *in vitro* activity for these scaffolds, as pEC<sub>50</sub> values assess both binding affinity and efficacy of receptor activation. However, it is important to note that the model and docking method predicted these scaffolds to be active and differentiated between the individual members of each scaffold. Interestingly, the pEC<sub>50</sub> values for compounds in Scaffold D range greatly depending upon the substituent on the amide nitrogen (pEC<sub>50</sub> = 6.4 when X = Me, 7.4 when X = Et, 8.3 when X = Pr, and 9.5 when X = Bu). This finding indicates that compounds belonging to Scaffold D activate S1PR<sub>1</sub> via size-induced changes of the pocket rather than specific interaction-induced changes, similar to results from the docking study.

The compounds possessing Scaffold E (orange) and Scaffold F (lime) differ only in the identity of their linker. In their bent orientations, the amide of Scaffold E compounds was shown to interact with N101<sup>2,60</sup> and the imidazole of Scaffold F compounds was demonstrated to interact with D121<sup>3,29</sup> and D294. For both scaffolds, the calculated binding energies show a linear relationship with the *in vitro* potency data, but the binding energies for compounds of Scaffold F are slightly overestimated.

We believe that water molecules likely occupy the vacancy adjacent to the zwitterionic head groups and that these solvent molecules are simply not resolved in the X-ray crystal structure; the corresponding electron density map shows electron density in this region (0.3 $\sigma$ ), but that it is not well defined, which is consistent with this idea. Docking studies were conducted using an aporeceptor consisting of the model receptor and any crystallized solvent molecules, with the receptor pocket and periphery existing as a vacuum. Thus the steric and energetic factors involved in (de)solvation are not captured during the docking process, biasing the receptor-ligand complex to maximize contacts *in silico*,

whether or not they are relevant *in vivo*. This error is likely consistent across the other ligand scaffolds and only becomes problematic with specific interactions at solvent-exposed residues. We speculate as to the validity of this interaction *in vivo* and believe it causes inflation of the binding energies estimated for compounds of Scaffold F.

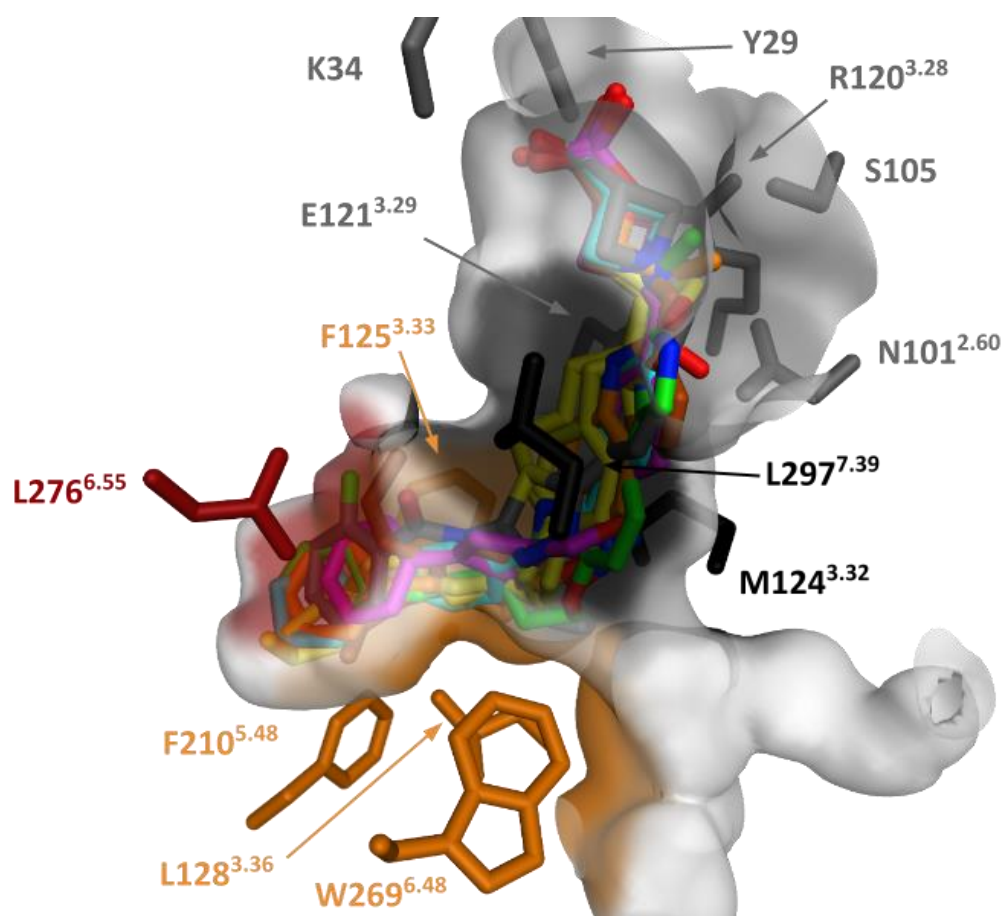
Like the previous ammonium phosphates, both **S1P** and **(S)-fingolimod phosphate** do not form hydrophobic interactions with specific residues when docked into the side cavity. Instead, the ligand tails appear to maximize van der Waals contacts throughout the cavity. The additional hydroxyl groups in the ligand head groups form interactions with D121<sup>3,29</sup> and D294 respectively. The predicted binding energies for these compounds correlate well with their *in vitro* potency data.



**Figure 3.6.** Ligand-receptor interaction diagram for select compounds in model receptor. Residues lining the pocket (gray outline) are colored according to their polarity. Interactions between residues and ligands are highlighted.

Overall, agonists possessing alkyl tails established van der Waals interactions with the residues lining the pocket and likely activate the receptor via nonspecific, size-induced rearrangement of the receptor. This hypothesis can also be used to explain why the original *meta*-substituted compounds switched from antagonism to agonism as they increased in size; the corresponding *para*-substituted compounds all behaved as agonists because that substitution pattern places the ligand tails significantly closer to the residues lining the side cavity, causing the receptor to rearrange no matter the ligand tail length. Additionally, it also may explain why potent S1PR<sub>1</sub> antagonists have remained elusive: addition of substituents to the scaffolds of weak antagonists likely confers agonism rather than increased potency of antagonism. Conversely, ligands with aromatic tails were found to form interactions with specific residues that comprise or flank the aromatic cluster, conserved residues implicated in the receptor activation process [95].

This multimodal mechanism of receptor activation was previously proposed based upon results of mutagenesis studies, which demonstrated ligand-dependent changes in agonists' pEC<sub>50</sub> values upon mutation of F125<sup>3,33</sup>, L128<sup>3,36</sup>, F210<sup>5,48</sup>, and W269<sup>6,48</sup> [14, 18, 19]. These residues line the side cavity of the receptor and support this proposed binding mode and the results from our *in silico* study. More importantly, this binding mode seems to be validated by another critical mutagenesis study involving L276<sup>6,55</sup> [20]. This study revealed ligand-dependent changes on both the pIC<sub>50</sub> and pEC<sub>50</sub> values for **S1P**, **(S)-fingolimod phosphate**, **A-26**, and **B-24** upon mutation of the residue to phenylalanine - an increase in side chain volume of approximately 25%. The mutation had a dramatic effect on the larger, aromatic ligands **A-26** and **B-24**, whereas it had only a minimal effect on the smaller, alkyl ligands **S1P** and **(S)-fingolimod phosphate**. The results of this study indicate that this residue must either directly interact with S1PR<sub>1</sub> ligands or lie in close proximity to them. It resides at the end of the side cavity and serves as the best evidence to support this binding mode and results from our *in silico* study.



**Figure 3.7. Compounds docked into the side cavity of S1PR<sub>1</sub>, highlighting important residues.** Top ligands from each scaffold docked into the side cavity of the model receptor (shaded region). Residues in gray were shown to be involved in binding zwitterionic head groups during previous site-mutagenesis studies. Residues in black were shown to be involved in binding ligands during docking studies. Finally, the remaining residues showed ligand-dependent effects on pIC<sub>50</sub> (red) or pEC<sub>50</sub> (orange) in prior mutagenesis studies.

So why then did agonists obtain similar docking scores (between -9.6 kcal / mol and -12.5 kcal / mol) and overall correlation ( $R^2 = 0.36$ ) when docked into the linear cavity? Are these poses energetically accessible, as is suggested from the results of this docking study? If so, can ligands interconvert between the two orientations, a process that would likely regulate the dynamics and activation state of the receptor? We find these scenarios to be unlikely, as this docking pose is inconsistent with *in vitro* results from mutagenesis studies and is likely an artifact of the docking process.

The majority of the ligands' calculated binding energy ( $\sim -7$  to  $-8$  kcal / mol) is contributed by their head group, linker, and first aromatic group. These groups have essentially the same orientation when

ligands are docked into either the side or linear cavities. As a result, it is not surprising that the ligands receive similar calculated binding energies when oriented into both the side and linear cavities.

This problem is further magnified by the method's inability to properly capture (de)solvation energy. The desolvation energy of the ligands is likely negligible, as these lipid-like molecules are typically protein-bound in blood and are believed to enter the receptor pocket via the membrane [101]. Solvation of the receptor, on the other hand, is much more complicated. Solvent molecules and ions have been shown to occupy the internal cavity of other Class A GPCRs and potentially play a role in their process of activation [102], and a recent set of dynamics simulations demonstrated this phenomenon for S1PR<sub>1</sub> [103]. While the electron density map in this region is not well defined, it does not preclude solvent molecules from occupying this region. For ligands to occupy the linear cavity, they would need to displace any solvent molecules from a relatively polar portion of the pocket. Additionally, new solvent molecules or, more likely, receptor rearrangement would be required to fill the vacuum that would be left in the hydrophobic side cavity. These processes are either energetically unfavorable or require changes that are inconsistent with current knowledge of receptor activation. Even if these processes were to occur, they are impossible to incorporate into docking studies utilizing an aporeceptor.

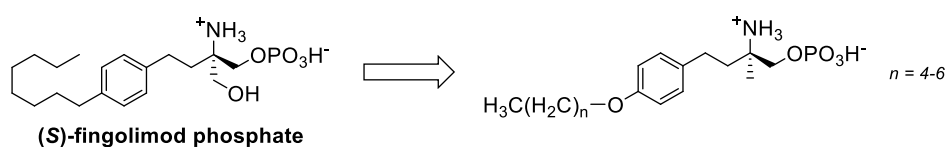
In summary, we used previous studies and an understanding of the limitations of molecular modeling to analyze our *in silico* results, concluding that all S1PR<sub>1</sub> ligands likely occupy the side cavity of the receptor. Using our model, docking methodology, and regression analysis, we can reasonably predict the pose and *in vitro* activity of novel S1PR<sub>1</sub> agonists. Additionally, our results support previous work demonstrating that activation of S1PR<sub>1</sub> is multimodal and depends upon agonist structure.



### 3.6 Extension to Biology and Medicinal Chemistry and Future Work

So what then do these studies tell us about the biology of S1PR<sub>1</sub>, and how can we then use this information to design improved S1PR therapeutics? As discussed previously, the model supports the idea of a multimodal mechanism for S1PR<sub>1</sub>-mediated G protein activation, with the pathway of activation dependent upon ligand structure. At this point, it has not yet been determined whether antagonism or functional antagonism of S1PR<sub>1</sub> is preferred for RRMS treatment [104], and information from this model will facilitate the development of novel S1PR<sub>1</sub> antagonists and agonists to investigate this question.

Unfortunately though, the model does not yet elucidate how ligands affect receptor internalization and subsequent trafficking, properties of functional antagonists like **(S)-fingolimod phosphate**. To date, only S1P and **(S)-fingolimod phosphate** have been characterized well enough to investigate all levels of receptor signaling and processing, which limited the scope of our modeling study. We hope to eventually reach this level of understanding, using *in silico* study to guide our inquiry and *in vitro* / *in vivo* studies to validate our hypotheses in a stepwise process. We are particularly interested in a series of **(S)-fingolimod phosphate** analogs that vary in their tail length (Figure 3.8).



**Figure 3.8. Analogs of interest.** Compounds show differing ability to induce S1PR<sub>1</sub> internalization.

These compounds demonstrated similar G protein activation but differed dramatically in their ability to cause receptor internalization [25]. These differences may arise from minor differences in how well the ligands are phosphorylated, how they modulate the structure of S1PR<sub>1</sub>, how quickly they enter and exit the receptor pocket, and how long they reside in the pocket.

Finally, we want to expand our study to include nonionic S1PR<sub>1</sub> modulators and develop homology models of the remaining S1PRs to aid in the development of S1PR subtype-selective compounds that minimize the potential for adverse effects.

### 3.7 Conclusions

Docking simulations were performed to assess the potential orientation of agonists within the S1PR<sub>1</sub> receptor pocket. The inherent limitations of the method prevented the incorporation of the binding pocket's (de)solvation energy during the docking process. By understanding this limitation and comparing the *in silico* results with *in vitro* results recorded using wild type and mutant S1PR<sub>1</sub>, we have demonstrated that all ligands likely occupy the side cavity of the S1PR<sub>1</sub> pocket similar to the antagonist **ML056** in the crystal structure. We intend to design new compounds that 1) validate these *in silico* results, 2) distinguish the ligand requirements for antagonism and agonism, 3) elucidate new information about the biology and pharmacology of S1PR<sub>1</sub>, and 4) lead to the development of improved therapeutics for the treatment of RRMS and other autoimmune diseases.

---

# 4

---

## Direct-Acting Second-Generation S1PR<sub>1</sub> Agonists

Compounds within the library of  **fingolimod**  rigid analogs were typically hindered by their poor bioactivation, lack of selectivity, and chemical properties. These properties limit their ease of synthesis and evaluation as well as their overall therapeutic utility. Rather than continue to explore the unpredictable activity of these scaffolds using trial and error, we utilized the S1PR<sub>1</sub> model to rationally design new compounds. These molecules are designed to be achiral and direct acting (requiring no bioactivation step) and will test the predictive capability of the S1PR<sub>1</sub> model. Once validated, homology models of the other S1PRs can be developed and utilized in the design of more potent and selective therapeutics.

#### 4.1 Goals and Properties of Second-Generation S1PR<sub>1</sub> Modulators

While the previous *in silico* results are interesting, they are not particularly useful unless they aid in the understanding of the receptor's pharmacology and the design of new drugs. To determine the predictability and utility of the model, a series of new compounds would need to be designed, synthesized, and evaluated biochemically. As before, the *in silico* docking results would be compared to the *in vitro* assay results to either validate the model or indicate that it needs refinement. Before work began, a set of standards were defined, and molecules were designed to meet all or most of those criteria (Table 4.1).

Compounds should provide clear definition of the binding mode for S1PR <sub>1</sub> ligands
Compounds should be designed to span a range of activities and test the predictability of the model
Compounds should be easy to synthesize and characterize
Compounds should have properties similar to those of molecules in the test set

Table 4.1. Ideal properties of second-generation S1PR<sub>1</sub> modulators.

Ideally, compounds would be able to resolve issues of the previous *in silico* study. While only docking studies in which ligands were oriented into the side cavity of the receptor actually agreed with *in vitro* results, it would be best to validate these results through design, synthesis, and characterization of novel S1PR<sub>1</sub> ligands.

Additionally, the series of compounds should be designed to span a range of activities. This will assess how well the model is able to differentiate the activity of various agonists. It is important that an effective model not just predict that compounds are active, but be able to predict their *relative* activity. This will facilitate the development of novel, highly potent compounds.

Furthermore, these molecules should also be simple to synthesize and evaluate biochemically. For this reason, we chose to only develop compounds that possess the azetidinium carboxylate head group. Ligands with this head group do not require phosphorylation by sphingosine kinases to become bioactive and instead act directly upon the S1PRs. Furthermore, they are achiral and synthesis of similar compounds have been reported extensively in the literature, which facilitates their development.

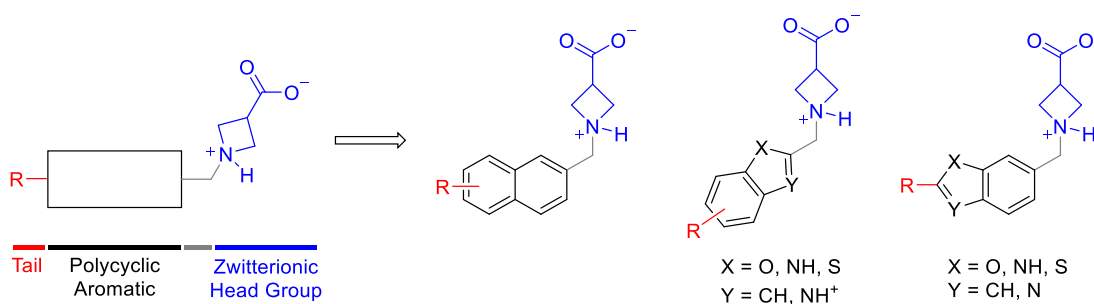
We intend to test other zwitterionic head groups, but incorporation of these groups into our initial study adds another variable that may complicate interpretation of the results. It is important that the initial compounds are structurally similar to compounds of the test set, as any deviation provides another variable that makes analysis more difficult. Thus, as was the case for compounds of the test set, second-generation compounds should be designed to possess minimal heteroatoms.

## 4.2 Attempts at Designing Compounds to Elucidate the Binding Mode

Initially, we sought to design compounds that would clearly define the orientation in which ligands bind to S1PR<sub>1</sub>, as investigated in our previous study. Ligands possessing the azetidinium carboxylate head group are particularly tractable to this study, as the rigid head group has minimal conformational flexibility. We believed that we could design ligands possessing more rigid tails than those of the test set that would show a strong preference for orientation into one cavity of the receptor pocket over another.

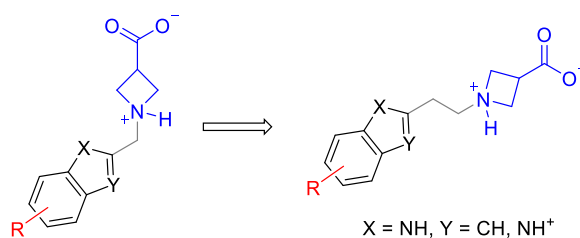
The study began by screening the 800,000 different scaffolds available in the MOE software suite database [105]. This allowed for thorough investigation of the 3-dimensional space of the pocket and would hopefully yield a few candidates that would warrant further study. Unfortunately, this work afforded very few reasonable candidates, all of which performed poorly during subsequent docking studies.

Following this, compounds were designed manually using trial and error, focusing on azetidinium carboxylates that possessed adjacent polycyclic aromatic groups in their tails (Figure 4.1). These groups are more rigid than the nonfused, polyaromatic compounds present in the test set and were thought to potentially bias the orientation of ligands in the pocket. The two groups were separated by a methylene spacer, as direct connection would attenuate the basicity of the nitrogen of the azetidinium carboxylate. A series of compounds were again designed and docked, but these compounds also did not show conformational preference during docking studies due to the flexibility of the methylene linker.



**Figure 4.1. Structure of fused polyaromatic compounds.** Structures were constructed, docked, and scored according to previously-described methods. The position of substitution and identity of the R groups varied, but did not largely affect the results from scaffold to scaffold. Rotation about the bonds of the methylene linker makes the compounds flexible enough to occupy both cavities.

During these studies, several of the indole- and benzimidazole-containing compounds ( $X = \text{NH}$ ,  $Y = \text{CH}$  or  $\text{NH}^+$ ) were observed to establish weak interactions with E121<sup>3,29</sup>. It was found that these interactions were nearly equal between compounds possessing a methylene linker and those with an ethylene linker (Figure 4.2). Unlike a methylene linker, an ethylene linker can be rigidified to provide an additional level of conformational restraint that, along with the constraint imposed by the interaction with E121<sup>3,29</sup>, should bias the ligand to occupy only one cavity of the receptor pocket.



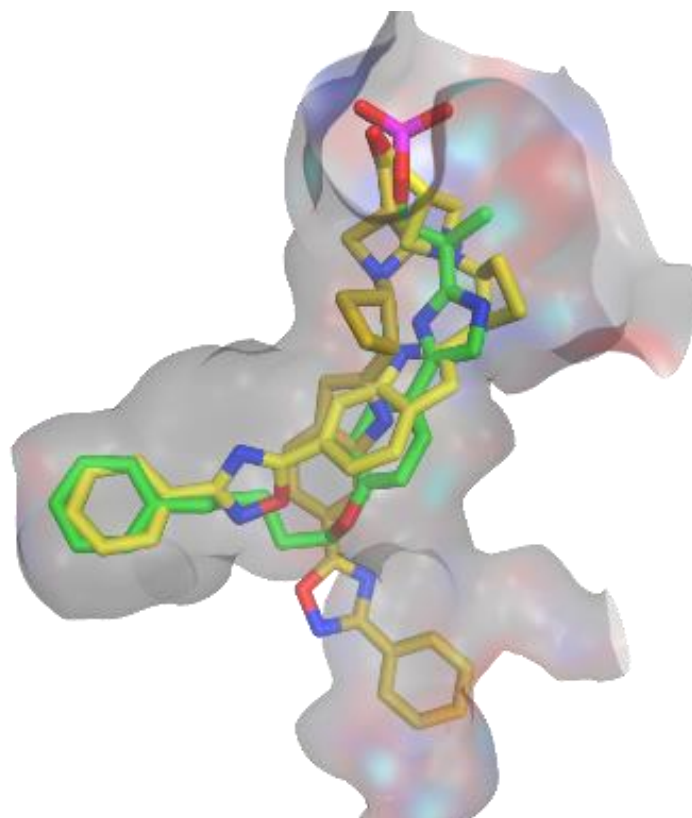
**Figure 4.2. Structure of indole- and benzimidazole-based analogs.** Both scaffolds were found to interact with E121<sup>3,29</sup>. Unlike the methylene linker, the ethylene linker can be rigidified to affect the scaffold's conformation.

Thus, a series of indole-based analogs possessing ethylene and rigid ethylene linkers were constructed, docked, and scored *in silico* as previously described. The compounds' docking scores were converted to predicted  $\text{pIC}_{50}$  values using formulas obtained from the regression analysis discussed previously and shown in Figure 3.5. Of the compounds, only the (*S,R*) stereoisomers of the compounds with cyclopropylene and cyclobutylene linkers are predicted to have significant differences in their activity depending upon their binding mode (Table 4.2).

	( <i>R,R</i> )	( <i>R,S</i> )	( <i>S,R</i> )	( <i>S,S</i> )	( <i>R,R</i> )	( <i>R,S</i> )	( <i>S,R</i> )	( <i>S,S</i> )
Calculated Binding Energy (Bent)	-11.6	-11.9	-11.0	-11.7	-12.1	-12.3	-11.2	-12.0
Predicted $\text{pIC}_{50}$ (Bent)	8.2	8.5	7.6	8.3	8.7	8.9	7.8	8.6
Calculated Binding Energy (Linear)	-10.6	-11.2	-10.4	-8.6	-10.7	-11.1	-10.7	-8.0
Predicted $\text{pIC}_{50}$ (Linear)	8.3	8.8	8.1	6.5	8.4	7.8	8.4	6.0

**Table 4.2. Results of docking studies with ethylene- and rigid ethylene-containing compounds.** Compounds were constructed, docked, and scored similar to previous studies. Their predicted  $\text{pIC}_{50}$  values were determined using the appropriate formula from previous linear regression. Compounds with significant differences are highlighted in red.

While the numerical results appeared to show promise, they ultimately had to be discounted. Unlike compounds docked into the linear cavity, compounds docked into the side cavity had their cyclopropyl and cyclobutyl groups project into the same problematic region of the pocket as the imidazole of compounds from Scaffold F (Figure 4.3). Again, the imidazole of Scaffold F compounds occupied an area of the receptor that is likely solvated in reality. Thus, compounds in this series received inflated docking scores (often greater than 1 kcal / mol), which incorporated false contacts with the receptor and did not properly account for desolvation energy of the receptor. These effects are likely more pronounced for the aliphatic cyclopropyl and cyclobutyl groups that project even further into this region. These molecules and this approach were subsequently abandoned as a result of the ambiguous *in silico* results and the difficulty required to synthesize these strained, chiral scaffolds.



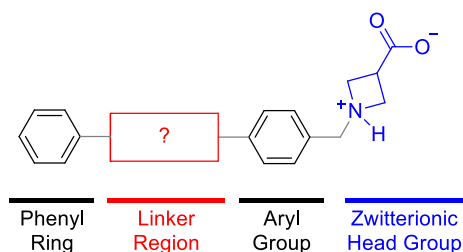
**Figure 4.3. Docked indole possessing the (*R,S*)-cyclobutyl rigid linker compared with F-11t.** The indole obtained dramatically different docking scores when docked into the side (yellow) and linear (gold) cavities. Overestimation of the binding energy for the compound in the side cavity is likely similar to that of a previously problematic compound, F-11t (lime).



### 4.3 Rational Design of Second-Generation S1PR<sub>1</sub> Agonists

Following these failed studies, we realized that we likely could not design and develop molecules that would definitively determine the binding mode of ligands in the S1PR<sub>1</sub> pocket. However, we still felt confident that ligands likely occupied the side cavity, as results obtained with ligands docked in this manner aligned very well with previous *in vitro* studies. Thus, we went about designing, synthesizing, and evaluating new agonists under the assumption that ligands were oriented into the side cavity of the receptor. Again, we wanted our second-generation agonists to span a range of activities and structures, but not deviate too significantly from molecules in the original test set.

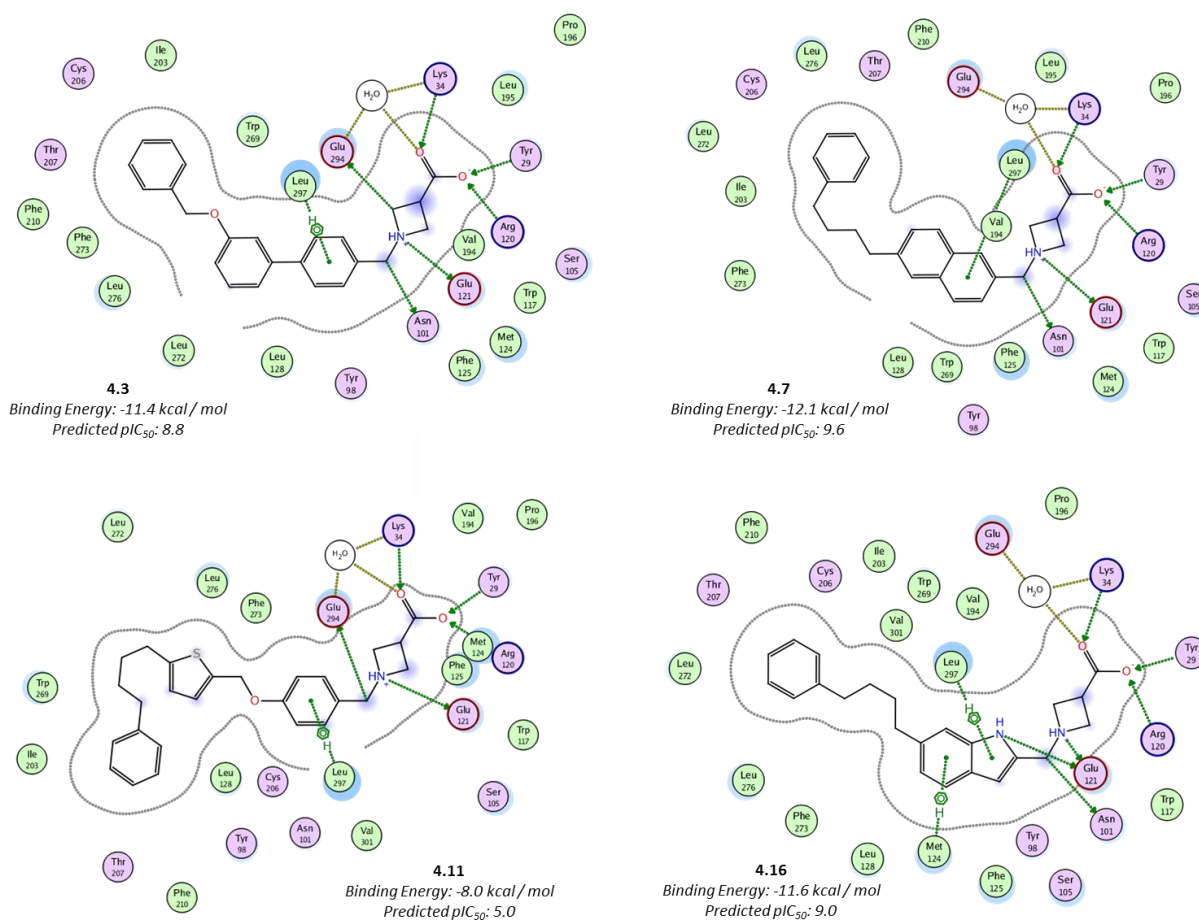
Overall, compounds of the test set that bind potently to S1PR<sub>1</sub> possess a zwitterionic head group and an adjacent aryl group, though **S1P** is an obvious exception to this statement. Agonists bearing a terminal phenyl group demonstrated particularly high *in vitro* potency and high *in silico* docking scores. Furthermore, these terminal phenyl groups were predicted to occupy a similar position in the receptor pocket no matter the agonist scaffold. We postulated that a zwitterionic head group and a terminal phenyl group could serve to anchor ligands in the receptor pocket, which could be connected using various linkers designed to probe different portions of the receptor pocket. This analysis yielded our general pharmacophore (Figure 4.4).



**Figure 4.4. Pharmacophore of second-generation agonists.** The compounds are designed to be anchored by the terminal zwitterionic and phenyl groups. The intermediate aryl and linker groups are variable among the compounds.

Again using trial and error, a series of four compounds possessing different aryl groups (biphenyl, naphthalyl, phenyl, and indolyl) were manually designed, docked, and scored in the model receptor

(Figure 4.5). The alkyl and alkoxy linkers of these scaffolds permit flexibility so that the terminal phenyl ring is capable of achieving the predicted arrangement in the pocket.



**Figure 4.5. Ligand-receptor interaction diagrams for rationally-designed compounds.** Compounds **4.3**, **4.7**, and **4.16** were designed to test the validity of the model; compound **4.11** served as the null. All compounds achieved interactions with the typical polar sites in the head groups, and the aryl groups also all interact with L297<sup>7,39</sup>.

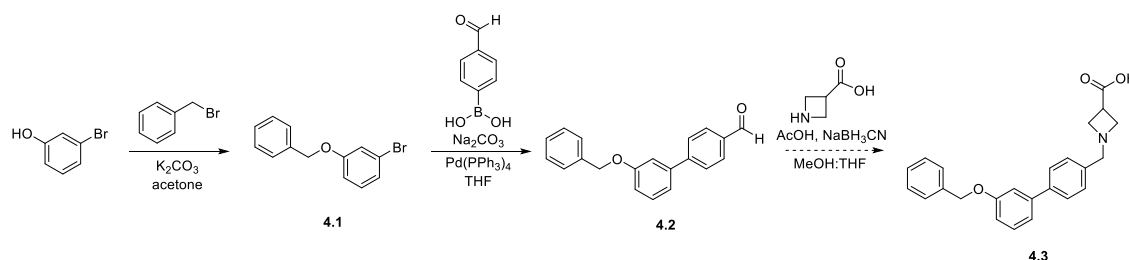
The biphenyl-based compound **4.3** achieved an *in silico* docking score of - 11.4 kcal / mol, which corresponds to a predicted  $pIC_{50}$  of 8.8. The first aryl ring is predicted to undergo a specific hydrogen-arene interaction with a methyl hydrogen on the side chain of L297<sup>7,39</sup>, an interaction common among all members of the previously-docked agonist scaffolds. The naphthalene-based compound **4.7** yielded a docking score of - 12.1 kcal / mol and is estimated to be the most potent compound with a predicted  $pIC_{50}$  of 9.6. It also is hypothesized to interact with L297<sup>7,39</sup> via a specific hydrogen-arene interaction. Indole-based compound **4.16** obtained a docking score of - 11.6 kcal / mol, corresponding to a predicted  $pIC_{50}$  of

9.0. This aryl group was particularly interesting, as it is able to form a hydrogen bond with E101<sup>3.29</sup> and have hydrogen-arene interactions with both L297<sup>7.39</sup> and M124<sup>3.32</sup>. Finally, the phenyl-based compound **4.11** yielded a docking score of - 8.0 kcal / mol, which corresponds to a predicted pIC<sub>50</sub> of 5.0. This molecule is predicted to interact with L297<sup>7.39</sup> via a hydrogen-arene interaction but, more importantly, is predicted to be too long to appropriately fit inside the receptor pocket. Compound **4.11** is ~ 18 atoms in length from the ammonium nitrogen to the end of the terminal phenyl ring, whereas the other compounds are ~ 15 - 16 atoms in length. It is designed to serve as a null compound, testing whether or not the model can differentiate between active and inactive compounds.

These compounds are also chosen based upon their ease of synthesis, as they can be obtained using well established chemistry in five steps or less. Furthermore, the simple building blocks required for their synthesis can be modified relatively easily to produce libraries of analogs that investigate the structure-activity relationship of each scaffold.

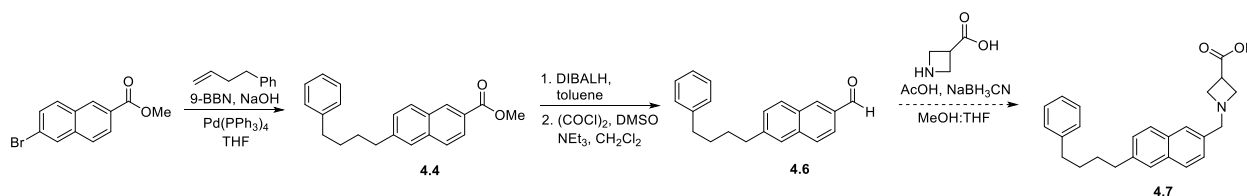
#### 4.4 Progress Toward the Synthesis of Second-Generation S1PR<sub>1</sub> Agonists

Synthesis of the biphenyl-based compound (Scheme 4.1) began with a Williamson ether synthesis between 3-bromophenol and benzyl bromide to yield **4.1** in 98% yield. A Suzuki coupling was performed between **4.1** and commercially-available 4-formylphenylboronic acid to yield **4.2** in 45% yield. Interestingly, this reaction did not run to completion even after prolonged reflux. The final product, **4.3**, is to be synthesized finally via reductive amination with azetidione-3-carboxylic acid.



**Scheme 4.1.** Progress toward the synthesis of **4.3**. The product is to be synthesized through a linear sequence and evaluated *in vitro* to determine its activity at S1PR<sub>1</sub>.

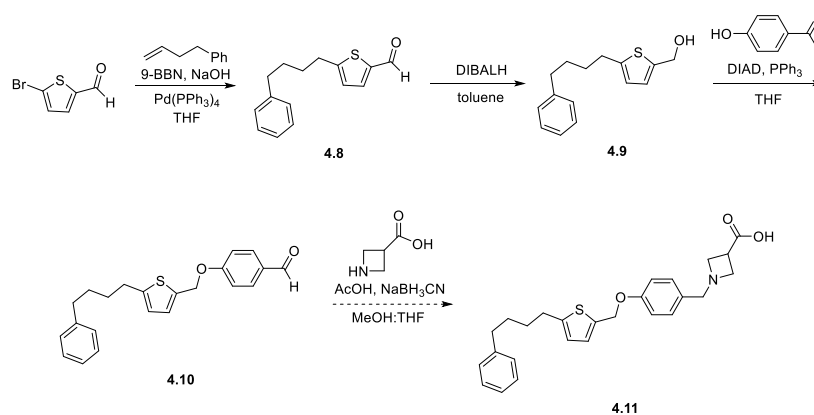
Synthesis of the naphthalene-based compound (Scheme 4.2) began with a Suzuki coupling using methyl 6-bromo-2-naphthoate and the hydroboration product between 9-BBN and 4-phenyl-1-butene to yield **4.4** in 83% yield. Corresponding alcohol **4.5** was obtained in 90% yield following reduction using DIBALH. Not surprisingly, aldehyde **4.6** was not observed by TLC during this reaction and was instead generated in 96% yield by performing a subsequent Swern oxidation. As before, the aldehyde is to be used as a substrate for a reductive amination to afford **4.7**.



**Scheme 4.2.** Progress toward the synthesis of **4.7**. The product is to be synthesized in four steps using a linear sequence and later evaluated *in vitro* to determine its activity at S1PR<sub>1</sub>.

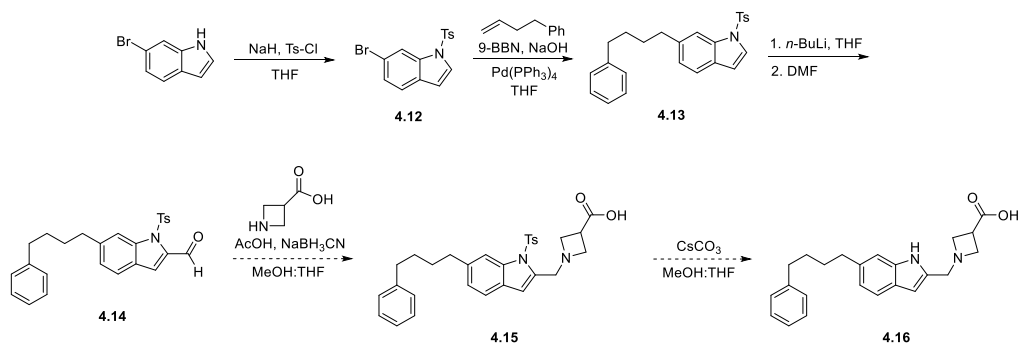
Synthesis of the phenyl-based compound (Scheme 4.3) began with a Suzuki coupling between 5-bromothiophene-2-carbaldehyde and the hydroboration product between 9-BBN and 4-phenyl-1-butene to yield **4.8** in 63% yield. The aldehyde was then reduced using DIBALH in 80% yield to afford **4.9**, which

underwent a subsequent Mitsunobu reaction with 4-hydroxybenzaldehyde to yield **4.10** in an abysmal 8.4% yield. The starting alcohol was consumed during this reaction, and the low yield is attributed to formation of the dithiophenyl ether and difficulties during purification. In future studies, an alternative route or method would be employed to achieve **4.10** in more acceptable yields. Finally, another reductive amination is to be utilized to form **4.11**.



**Scheme 4.3.** Progress toward the synthesis of **4.11**. The product is to be synthesized in five steps using a linear sequence and evaluated *in vitro* to determine its activity at S1PR<sub>1</sub>.

Finally, synthesis of the indole-based compound (Scheme 4.4) required protection of its nitrogen. The tosyl group was specifically chosen as a result of its ability to induce *ortho*-metallation at C2. Indoles are most nucleophilic at C3 as a result of their enamine character; however, this property is negated when alkylations are performed using a C2-metallated substrate. Thus, 6-bromoindole was tosyl protected to afford **4.12** in 61% yield. This yield was surprisingly low and is mainly the result of difficulty in isolating the product during flash chromatography; the starting material and product have nearly identical R<sub>f</sub> values. The protected indole then underwent a Suzuki coupling with the hydroboration product between 9-BBN and 4-phenyl-1-butene to afford **4.13** in 95% yield. This product was *ortho*-metallated using *n*-BuLi and then formylated using DMF as the electrophile to give aldehyde **4.14** in a modest 38% yield. As before, the aldehyde is to be used in a reductive amination with azetidine-3-carboxylic acid to give **4.15** and **4.16** following deprotection.



**Scheme 4.4.** Progress toward the synthesis of 4.16. The product is to be synthesized in five steps using a linear sequence and evaluated *in vitro* to determine its activity at S1PR<sub>1</sub>.

In general, most of the employed reactions were straightforward to perform and purify, yielding useful amounts of the intermediates; most of these reactions are staples of synthetic chemistry as a result of this utility. However, one common problem occurred throughout synthesis: the materials were difficult to isolate by flash chromatography. The scaffolds are very nonpolar and the individual transformations did not substantially alter the polarity of the molecules, a problem especially common with Suzuki couplings between aryl bromides and the hydroboration product of 4-phenyl-1-butene. The opposite is likely to be true once the zwitterionic head group is installed: the molecules will become extremely polar and even more difficult to purify, owing to solubility issues as a result of their detergent character. Nevertheless, the synthetic routes should afford the title compounds in reasonable yields and atom economies. More importantly, the routes allow for exploration of the structure-activity relationships of these scaffolds through use of derivatives of the above building blocks.

#### 4.5 Conclusions

A set of criteria were defined for our second-generation S1PR<sub>1</sub> agonists, and a series of ligands were designed with these standards in mind. Unfortunately, we were unable to validate our initial *in silico* study that investigated the orientation of agonists in the S1PR<sub>1</sub> pocket. Despite this, a series of novel compounds were designed to test the predictability of the model when compounds occupy the side cavity of the receptor. The intermediates required to form these compounds were synthesized with relative ease. The remaining step in the synthesis of all the second-generation agonists is a reductive amination, which has been reported throughout the literature for similar substrates. Once the synthesis of these molecules is complete, we intend to subject them to *in vitro* testing to finally determine the model's predictability and utility.

---

# 5

---

## Experimental Procedures

### 5.1 Biochemical & Biological Methods

All biochemical and biological work was conducted by Dr. Perry Kennedy and Dr. Yugesh Kharel in the laboratory of Dr. Kevin Lynch of the Department of Pharmacology at the University of Virginia and meets the strict guidelines outlined by the Animal Welfare Act and Animal Welfare Regulations.

#### [ $\gamma$ -<sup>35</sup>S]GTP Binding Assay [37]

A plasmid encoding a human S1P receptor was introduced to CHO K1 cells via DNA-mediated transfection, and a clonal line of cells overexpressing the receptor was then isolated. Crude membrane fractions of the cells with ca. 5  $\mu$ g of protein were incubated for 30 minutes at 30°C with 0.1 mL of binding buffer (50 mM HEPES, 100 mM NaCl, and 10 mM MgCl<sub>2</sub> at pH 7.5) containing 5  $\mu$ g saponin, 0.10 mM GDP, and 0.1 nM [ $\gamma$ -<sup>35</sup>S]GTP (1200 Ci/mmol), and the indicated compound(s). Membranes were collected using GF/C filters using a Brandel Cell Harvester (Gaithersburg, MD). Samples were then analyzed for bound radionucleotide using a TopCount  $\beta$  scintillation counter. Each data point represents the standard error of the mean for triplicate measurements.



**Sphingosine Kinase Assay [34]**

Mouse SK<sub>2</sub> was expressed by the introduction of plasmid DNA into HEK293T cells. SK<sub>2</sub> activity was measured in a solution that consisted of 20 mM Tris-Cl (pH 7.4), 1 mM 2-mercaptoethanol, 2 mM EDTA, 5 mM sodium orthovanadate, 40 mM β-glycerophosphate, 15 mM NaF, 1 mM phenylmethylsulfonyl fluoride, 10 mM MgCl<sub>2</sub>, 0.5 mM 4-deoxypyridoxine, 10% glycerol, and 0.01 mg/mL each of leupeptin, aprotinin, and soybean trypsin inhibitor. To determine the fractional activity of SK<sub>1</sub> versus SK<sub>2</sub>, the buffer was supplemented with either 0.5% Triton X-100 or 1 M KCl, respectively. The buffer was supplemented with substrate (D-erythro-sphingosine, 15 μM; indicated compound 50 μM), [γ-<sup>32</sup>P]ATP (10 mM, specific activity = 8.3 Ci/mmol), and cell extract (0.02-0.03 mg of total protein). After 30 min at 37°C, the reaction mixture was extracted with 2 volumes of chloroform/methanol/HCl (100:200:1), and the components in the organic phase were separated by normal-phase TLC using a 1-butanol/acetic acid/water (3:1:1) solvent system. Radiolabeled enzyme products were detected by autoradiography and identified by migration, relative to authentic standards. For quantification, the silica gel containing radiolabeled lipid was scraped into a scintillation vial and counted.

**Lymphopenia Assay [91]**

Mice were dosed via oral gavage with varying amounts of indicated compound dissolved in 2% hydroxypropyl β-cyclodextrin. Peripheral blood was taken from the orbital sinus of lightly anesthetized mice after 18 hours or 48 hours. The lymphocytes were quantified using a Hemavet 950 blood analyzer (Drew Scientific, Oxford, CT). Each bar represents the standard error of the mean for triplicate measurements.

## 5.2 Computational Methods

All computational work was done using the 2012 Molecular Operating Environment (MOE) software suite [105] and the 2.80 Å resolution crystal structure of the antagonist-bound S1PR<sub>1</sub>-T4L chimeric receptor (PDB: 3V2Y) [14].

### Generation of the Prepared Receptor Structure

First, residues N250-V417 (corresponding to the T4 lysozyme) were deleted from the structure, as this fusion is utilized solely to facilitate the crystallization process and is not present in the natural receptor. The missing loops and termini were not constructed due to limitations in the program's ability to approximate the structure of these large and inherently flexible regions. Instead, the "Structure Preparation" program was used to acetylate any terminal amines (V34 of the truncated N-terminus; N174 of ICL2; A418 of ICL3) and methylamidate any terminal carbonyls (K166 of ICL2; R249 of ICL3; L495 of the truncated C-terminus).

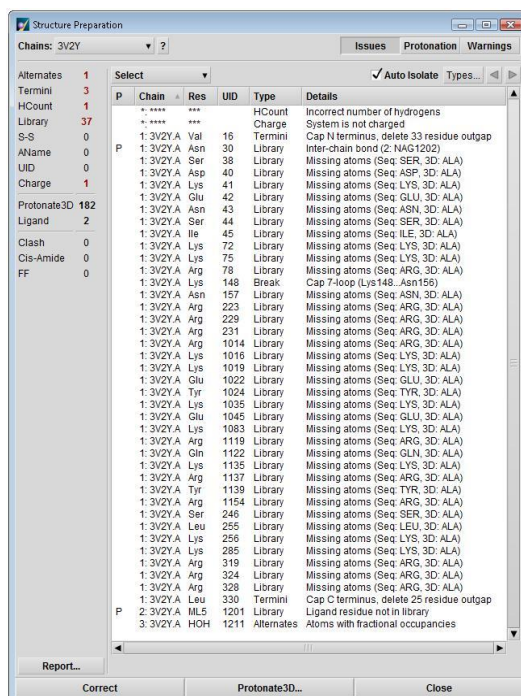
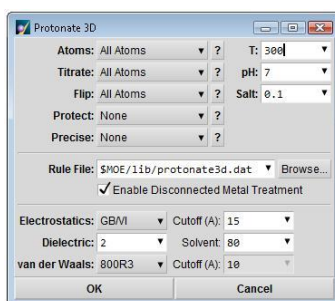


Figure 5.1. Issues fixed using "Structure Preparation" program. The program was used to cap termini and add unresolved heavy atoms to the structure.

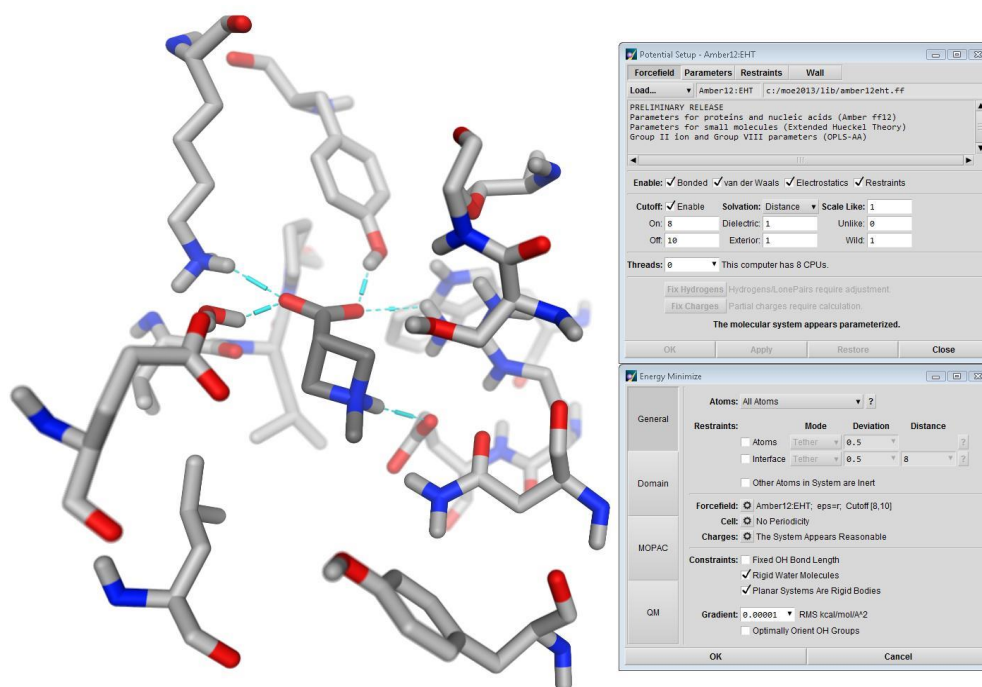
Additionally, the “Structure Preparation” program was used to add heavy atoms to complete the side chains of those residues only partially resolved in the crystal structure (S56, D58, K59, E60, N61, S62, I63, K90, K93, R96, N175, R241, R247, R249, L420, K421, K450, R484, R489, and R493). Finally, the “Protonate 3D” program was employed to add missing hydrogens until the ligand-receptor complex achieved its proper valency and charge (pH = 7, T = 300 K). The structure was checked to ensure that K34, R120<sup>3.28</sup>, and E121<sup>3.29</sup> were charged and that Y29, N101<sup>2.60</sup>, and bound water were neutral and that these groups were properly oriented to optimize interactions with other residues and ligands in the pocket. Overall, these manipulations yielded the prepared receptor, utilized in all subsequent docking studies.



**Figure 5.2. “Protonate 3D” program settings.** The program was used to add hydrogen atoms so that all atoms were of appropriate charge and valency.

## Optimization of Zwitterionic Head Groups

Ligand construction began by first determining the likely conformation of the zwitterionic head groups in the pocket. To achieve this, the ammonium phosphate (2-amino-2-methylpropyl phosphate) and azetidinium carboxylate (both *cis*- and *trans*-*N*-methylazetidinium-3-carboxylate) were each built into the polar portion of the pocket of the prepared receptor in several different conformations. The “Energy Minimize” program was then used to minimize each individual conformer-receptor complex along a gradient of 0.00001 RMS kcal / mol / Å<sup>2</sup> in the Amber 12:EHT forcefield [99]. During minimization each head group was left unconstrained, those residues and solvent molecules possessing atoms within 4.5 Å of the head group were tethered (strength = 50), and all other atoms remained fixed.



**Figure 5.3.** Head group minimization settings for azetidinium carboxylate. The head group (dark gray) and neighboring residues and solvent molecules (light gray) were tethered and then allowed to minimize.

The “pose rescoring” method of the “Dock” program was used to determine the binding energy of each resulting complex using the GBVI/WSA dG scoring methodology [100]. The top conformation for each head group was chosen based upon its calculated binding energy, ensuring that all essential contacts were established.

## Initial Search and Screen of Select Ligands

The most potent ligand within each scaffold was built and underwent a sequential conformational search and virtual screen that crudely determined the pose for bent and linear conformers. For scaffold A, the optimized *trans*-azetidinium carboxylate head group was used as the basis for constructing compound **A-32**. The ligand head group was rigidified, and the “Conformational Search” program was used to perform 30,000 iterations of stochastic stretching / rotation of bonds in the tail region in the MMFF94x forcefield [106]. Conformers that achieved a heavy atom RMSD of  $> 0.35 \text{ \AA}$  were determined to be unique and saved to a corresponding “initial search” library.

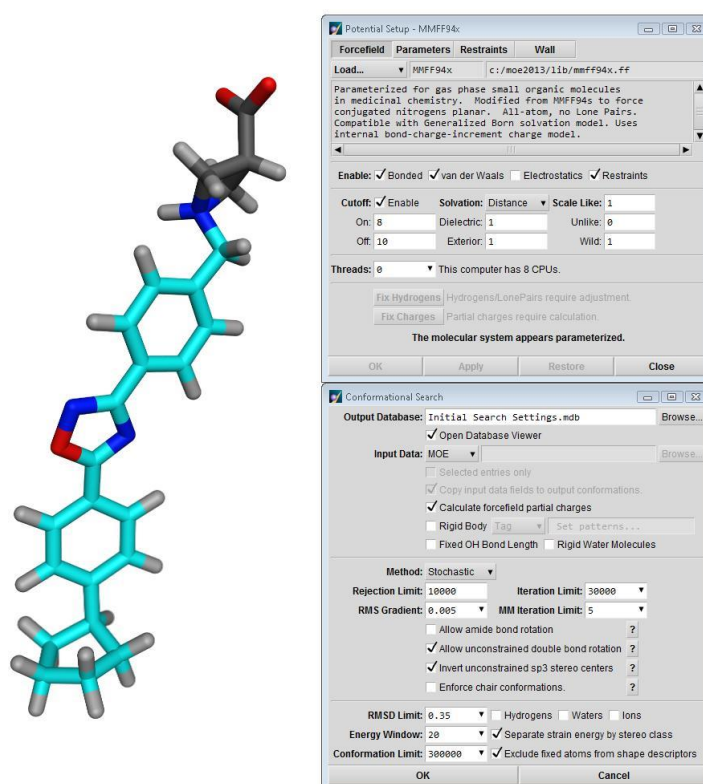


Figure 5.4. Initial conformational search settings for **A-32**. The head group (gray) was rigidified while the tail (cyan) was left free to rotate during search.

The “virtual screen” option of the “Dock” program was utilized to dock this library into the prepared receptor, holding the individual conformers and the receptor rigid throughout the process. This was performed in the AMBER 12:EHT forcefield and the binding energy of each conformer-receptor complex was determined using the GBVI/WSA  $dG$  scoring methodology.

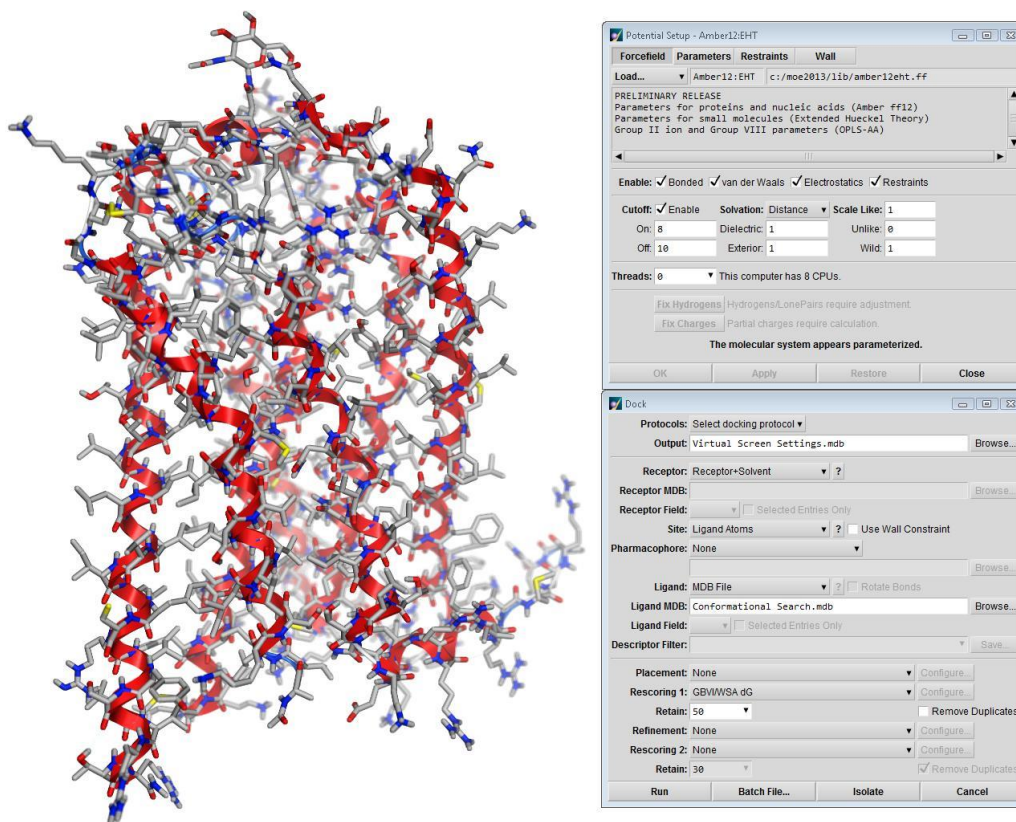
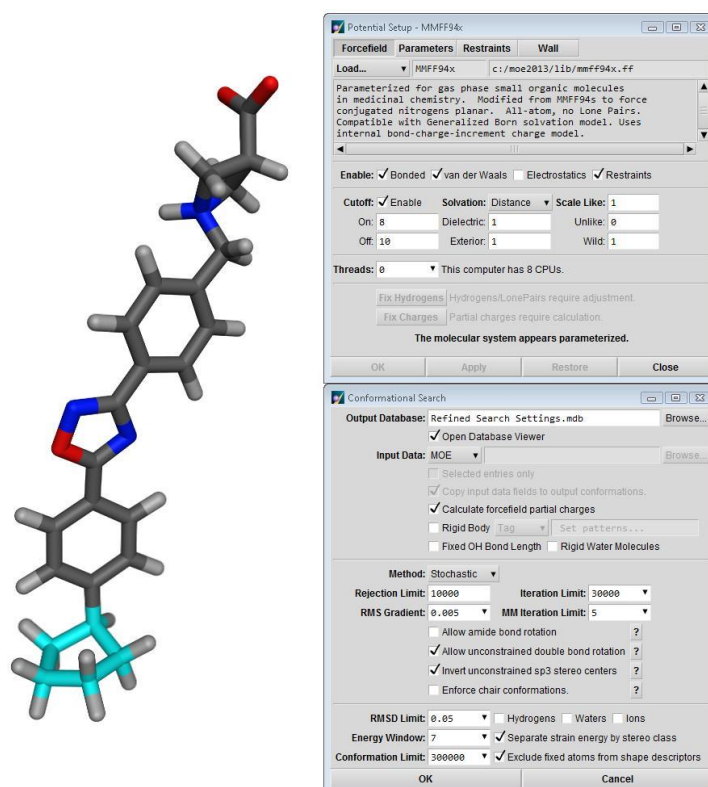


Figure 5.5. Initial virtual screen settings for A-32. The receptor and the ligand are held constant as the library of conformers is docked into the receptor pocket.

The docked conformers and their calculated binding energies were saved to a corresponding “initial screen” library. This process was repeated using **B-18**, **C-24e**, **D-18**, **E-11c**, **F-11t**, **S1P**, and **(S)-fingolimod phosphate** to crudely obtain bent and linear conformers for each compound.

## Refined Search and Screen of All Ligands

Following the initial search and screen, a refined conformational search and virtual screen was used to more precisely determine a bent and linear pose for all compounds. For compound **A-32**, the top bent pose from the “initial screen” library was parameterized so that the entire ligand was rigidified except for the inherently flexible cyclopentyl ring, which was left unconstrained. The “Conformational Search” program was again used to perform 30,000 iterations of stochastic stretching / rotation of the flexible bonds in the MMFF94x forcefield. In this case, conformers with a heavy atom RMSD of  $> 0.05$  Å were considered unique and saved to a “refined search” library.



**Figure 5.6.** Refined conformational search settings for **A-32**. The majority of the compound was rigidified (gray) except for the cyclopentyl ring (blue) during search.

As before, the “virtual screen” option in the “Dock” program was used to dock the conformers of the library into the prepared receptor, holding the individual conformers and the receptor rigid during docking. This was again performed in the AMBER 12:EHT forcefield and the binding energy of each conformer-receptor complex was determined using the GBVI/WSA dG scoring methodology.

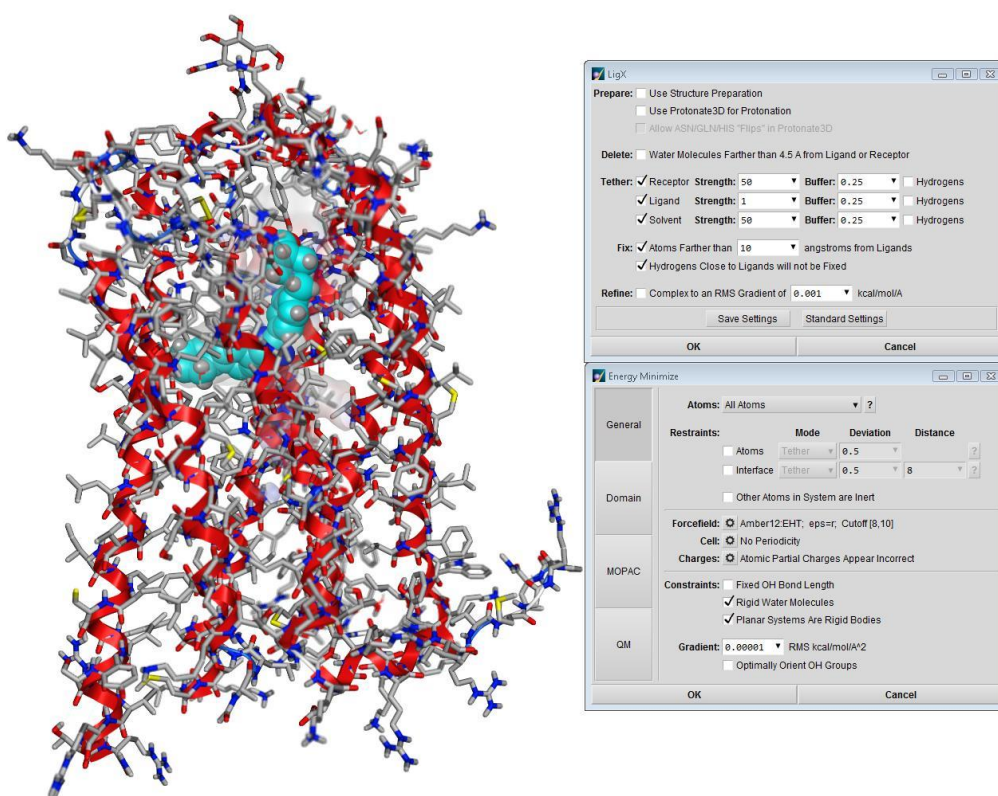
Finally, the docked conformers and their calculated binding energies were saved to a “refined screen” library. The entire process was then repeated using the top linear pose from the “initial screen” library to obtain the corresponding “refined screen” library.

The top bent and linear poses from the “initial screen” library for **A-32** were also used as the basis for constructing corresponding poses for all other members of scaffold A. For these compounds, only the variable / flexible portions of the molecule were given freedom during their “refined searches”. The same virtual screen protocol was then used to obtain the corresponding “refined screen” libraries for each compound in both the bent and linear poses. The “refined search” and “refined screen” libraries for compounds belonging to the remaining scaffolds were generated using top conformers from the “initial screen” libraries for **B-18**, **C-24e**, **D-18**, **E-11c**, and **F-11t** and the above protocol.



## Induced Fit Docking of All Ligands

The top bent pose from the “refined screen” of **A-32** was then added into the prepared receptor. Using the “LigX” program, the ligand (strength = 1) and those receptor and solvent atoms within 10 Å (strength = 50) were tethered to permit modest flexibility during docking; all remaining atoms were held fixed. The oxygen atoms of the head group were then fixed and the ligand-receptor complex was allowed to minimize along a gradient of 0.00001 RMS kcal / mol / Å<sup>2</sup> using the Amber12:EHT forcefield using the “Energy Minimize” program.



**Figure 5.7. Docking settings for A-32.** The ligand-receptor complex is allowed to minimize using the above settings. This process is then repeated with slightly different settings to obtain the final pose and binding energy calculation.

Finally using the “pose rescoring” option in the “Dock” program, the resulting pose was scored to determine the overall binding energy using the GBVI/WSA dG scoring function.

The process was then repeated using the resulting pose and slightly modified parameters. With the ligand added to the pocket, the “LigX” program was then used to tether the ligand (strength = 75),

tether those receptor and solvent atoms within 10 Å (strength = 75), and fix all other atoms. The ligand-receptor complex was allowed to minimize along a gradient of 0.00001 RMS kcal / mol / Å<sup>2</sup> using the Amber12:EHT forcefield using the “Energy Minimize” program.

Finally using the “pose rescoring” option in the “Dock” program, the resulting pose was scored to determine the overall binding energy using the GBVI/WSA dG scoring function. This was used as the final estimate of binding energy for the ligand. The process was then repeated for both bent and linear poses of all the remaining ligands.

### ***In Vitro* Data Preparation**

The final *in silico* result (calculated binding energy) of each pose was then plotted against the corresponding compound's *in vitro* data (pIC<sub>50</sub> or pEC<sub>50</sub>) using Microsoft Excel. Only compounds that have pIC<sub>50</sub> values were included in the regression analysis, though compounds with pEC<sub>50</sub> values are included for comparison. However, before analysis was performed, the biochemical data from each source was standardized by comparing each source's pIC<sub>50</sub> value for **S1P**, used as an internal standard. This adjusted the pIC<sub>50</sub> values for compounds **E-11c** to **E-11p** (+0.1), **E-12a** (-0.1), **F-11q** to **F-11t** (+0.1), and **F-12d** (-0.1).

### 5.3 Synthetic Methods

#### ***Reaction Performance***

All air- or water-sensitive reactions were conducted under an inert nitrogen atmosphere (dried via passage through a tube of Drierite) using a Schlenk line and oven- or flame-dried glassware. Anhydrous dichloromethane ( $\text{CH}_2\text{Cl}_2$ ), *N,N*-dimethylformamide (DMF), methanol (MeOH), tetrahydrofuran (THF), and toluene ( $\text{CH}_3\text{Ph}$ ) were purchased from EMD Millipore, Fisher Scientific, or Sigma-Aldrich. Excluding methanol, the solvents were dried by passage through a column of aluminum oxide and storage over 4 Å molecular sieves. All other building blocks and reagents were purchased from either Acros Chemicals, Fisher Scientific, Matrix Scientific, or Sigma-Aldrich and were used as received.

#### ***Reaction Monitoring***

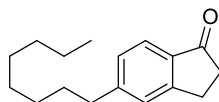
All reactions were monitored by thin-layer chromatography (TLC) using Whatman Partisil K6F 60 Å silica gel plates. Plates were visualized using ultraviolet irradiation or a staining solution: potassium permanganate solution (3 g of  $\text{KMnO}_4$  and 20 g of  $\text{K}_2\text{CO}_3$  in 300 mL water and 5 mL of 5% aqueous NaOH), ninhydrin solution (1.5 g of ninhydrin in 500 mL of 95% ethanol and 5 mL of AcOH), and 2,4-dinitrophenylhydrazine solution (12 g of 2,4-dinitrophenylhydrazine in 60 mL of concentrated  $\text{H}_2\text{SO}_4$ , 80 mL of  $\text{H}_2\text{O}$ , and 200 mL of 95% EtOH).

#### ***Reaction Purification***

Flash chromatography was performed using the indicated solvents and Dynamic Adsorbents silica gel (particle size 0.023 - 0.040 nm). Preparative TLC was performed using the indicated solvents and Analtech Uniplates coated with silica gel to a thickness of 500  $\mu\text{m}$ . The silica was scraped from the glass plate and compounds were desorbed using MeOH.

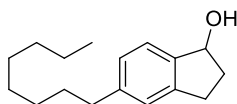
**Product Characterization**

Melting points are uncorrected and were taken using a Barnstead Mel-Temp equipped with a mercury thermometer. Proton ( $^1\text{H}$ ) and carbon ( $^{13}\text{C}$ ) NMR spectra were recorded using a Bruker 600/151 MHz or a Varian 300/54 MHz spectrometer at 298 K. The chemical shifts are reported in ppm ( $\delta$ ) and were standardized using residual solvent peaks:  $\text{CDCl}_3$  ( $\delta$  7.34 for  $^1\text{H}$  NMR and  $\delta$  77.0 for  $^{13}\text{C}$  NMR),  $\text{DMSO-}d_6$  ( $\delta$  2.50 for  $^1\text{H}$  NMR and  $\delta$  39.5 for  $^{13}\text{C}$  NMR),  $\text{CD}_3\text{OD}$  ( $\delta$  3.31 for  $^1\text{H}$  NMR and  $\delta$  47.6 for  $^{13}\text{C}$  NMR), and  $\text{D}_2\text{O}$  ( $\delta$  4.79 for  $^1\text{H}$  NMR). LC-ESI of the final products was performed by the Mass Spectrometry Laboratory in the School of Chemical Sciences at the University of Illinois Urbana-Champaign using a Micromass Q-tof Ultima mass spectrometer. HPLC was carried out using an Agilent 2.1 x 50 mm C-18 column. The eluent was a binary gradient that began with 90:10 -  $\text{H}_2\text{O}$  and 0.01% TFA: MeCN and 0.01% TFA and increased linearly to 0:100 -  $\text{H}_2\text{O}$  and 0.01% TFA: MeCN and 0.01% TFA over the course of 6 minutes followed by isocratic flow of MeCN and 0.01% TFA for an additional 3 minutes. The retention times are abbreviated as  $t_{\text{R}}$ ; mass-to-charge ratios are abbreviated as  $m/z$ .



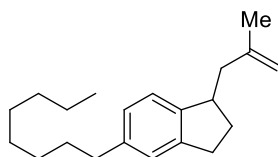
### 5-octyl-1-indanone (2.1)

A round-bottom flask was charged with a 0.5 M solution of 9-BBN in THF (61.4 mL, 30.7 mmol) and then 1-octene (4.82 mL, 30.7 mmol) dropwise via syringe at room temperature. The reaction was stirred overnight at room temperature. A 3 M NaOH solution (11.6 mL, 34.8 mmol) was then added to the octylborane solution and the mixture was stirred for 10 minutes. A two-neck flask equipped with a reflux condenser was charged with Pd(PPh<sub>3</sub>)<sub>4</sub> (950 mg, 0.8 mmol) and 5-bromo-1-indanone (5.00 g, 20.5 mmol) dissolved in THF (23 mL). The octylborane solution was added to the two-neck flask dropwise via syringe at room temperature, and the newly-formed solution was refluxed for 2 hours until the reaction was complete by TLC. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude oil was then dissolved in ethyl acetate and washed three times with water. The organic fraction was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a black tar. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a light brown oil (R<sub>f</sub> = 0.56 in 80:20 - hexanes:ethyl acetate). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.58-7.63 (d, *J* = 7.9 Hz, 1H), 7.20-7.24 (s, 1H), 7.09-7.15 (d, *J* = 7.9 Hz, 1H), 2.99-3.08 (m, 2H), 2.55-2.68 (m, 4H), 1.53-1.65 (m, 2H), 1.15-1.34 (m, 10H), 0.79-0.87 (t, *J* = 7.2 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 206.60, 155.73, 150.82, 135.11, 128.04, 126.40, 123.60, 36.47, 31.92, 31.36, 29.49, 29.36, 29.28, 25.76, 22.72, 14.16.



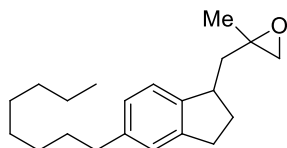
### 5-octyl-1-indanol (2.2)

A round-bottom flask was charged with a solution of **2.1** (5.01 g, 20.5 mmol) in a 1:1 mixture of dry ethanol:THF (60 mL) and cooled to 0 °C. A 2.0 M solution of LiBH<sub>4</sub> in THF (31.8 mL, 63.6 mmol) was added dropwise via syringe. Following addition, the reaction was stirred for 3 hours at room temperature until the reaction was complete by TLC. The solution was cooled to 0°C and quenched with 1 M HCl. The aqueous solution was extracted three times with diethyl ether. The combined organic extracts were dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to afford a yellow oil. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a white, waxy solid (4.32g, 17.5 mmol, 85% yield over two steps, R<sub>f</sub> = 0.37 in 80:20 - hexanes:ethyl acetate). mp = 51-53 °C. <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 7.31-7.34 (d, *J* = 7.6 Hz, 1H), 7.05-7.20 (m, 2H), 5.19-5.23 (dd, *J* = 6.8, 5.0 Hz, 1H), 3.00-3.07 (m, 1H), 2.76-2.83 (m, 1H), 2.57-2.63 (m, 2H), 2.44-2.51 (m, 1H), 1.91-1.99 (m, 1H), 1.78-1.87 (bs, 1H), 1.58-1.65 (m, 2H), 1.23-1.38 (m, 10H), 0.87-0.92 (t, *J* = 7.2 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 143.64, 143.54, 142.50, 127.13, 124.94, 124.06, 76.35, 36.22, 36.09, 32.03, 31.92, 29.84, 29.62, 29.51, 29.39, 22.80, 14.24.



### 5-octyl-1-(2'-methallyl)indane (2.3)

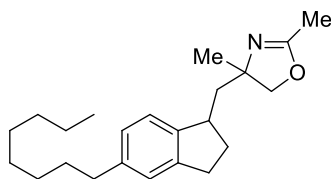
A round-bottom flask was charged with zinc (II) iodide (6.15 g, 19.3 mmol) dissolved in CH<sub>2</sub>Cl<sub>2</sub> (35 mL). A second round-bottom flask was charged with **2.2** (4.32 g, 17.5 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (30 mL). The flasks were cooled to 0 °C and the alcohol solution was added to the zinc iodide solution dropwise. Methallyltrimethylsilane (3.01 mL, 19.3 mmol) was subsequently added to the cooled mixture dropwise via syringe. The resulting solution was allowed to warm to room temperature and stirred for an additional 2 hours until the reaction was complete by TLC. The solution was quenched with saturated aqueous NaHCO<sub>3</sub>. The aqueous solution was extracted three times with CH<sub>2</sub>Cl<sub>2</sub> and the combined organic extracts were dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a yellow oil. The crude product was purified via flash chromatography (hexanes) to afford a clear oil (4.11 g, 14.5 mmol, 83% yield, R<sub>f</sub> = 0.91 in 80:20 - hexanes:ethyl acetate). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.16-7.21 (d, *J* = 7.8 Hz, 1H), 7.08-7.11 (s, 1H), 7.00-7.05 (d, *J* = 7.5 Hz, 1H), 4.79-4.89 (d, *J* = 15.5 Hz, 2H), 3.28-3.41 (m, 1H), 2.78-3.01 (m, 2H), 2.52-2.68 (m, 3H), 2.21-2.35 (m, 1H), 2.10-2.21 (m, 1H), 1.85-1.88 (s, 3H), 1.59-1.79 (m, 3H), 1.24-1.48 (m, 10H), 0.89-0.99 (m, 3H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 144.73, 144.63, 144.21, 141.27, 126.36, 124.55, 123.40, 111.57, 44.03, 42.29, 36.06, 32.24, 32.10, 32.05, 31.30, 29.71, 29.66, 29.48, 22.88, 22.63, 14.28.



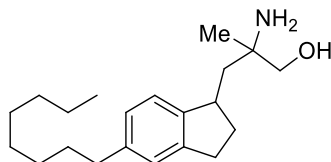
#### 5-octyl-1-(2',3'-epoxy-2'-methylpropane) (2.4)

A round-bottom flask was charged with **2.3** (1.39 g, 4.89 mmol) and NaHCO<sub>3</sub> (1.64 g, 19.5 mmol). The solid mixture was suspended in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and then cooled to -42 °C. A second round-bottom flask was charged with mCPBA (~77%, 2.19 g, ~9.77 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (25 mL), and that solution was added to the initial mixture dropwise via syringe. The resulting solution was allowed to warm to room temperature and stirred for an additional hour until the reaction was complete by TLC. The reaction was quenched with 1 M NaOH and the aqueous solution was extracted with three times with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic extracts were washed three times with 1 M NaOH, one time with saturated aqueous NaHCO<sub>3</sub>, and two times with water. The organic fraction was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to a yellow solid. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a white solid (0.875 g, 2.91 mmol, 60% yield, R<sub>f</sub> = 0.55 in 90:10 - hexanes:ethyl acetate). mp not recorded. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.11-7.18 (m, 1H), 7.06-7.11 (m, 1H), 6.97-7.06 (m, 1H), 3.16-3.37 (m, 1H), 2.53-3.03 (m, 6H), 2.06-2.48 (m, 2H), 1.51-1.94 (m, 4H), 1.43-1.51 (m, 3H), 1.22-1.43 (m, 10H), 0.87-0.98 (m, 3H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 144.04, 143.81, 141.39, 126.47, 126.32, 124.61, 124.49, 123.21, 122.99, 56.40, 54.35, 53.92, 42.27, 42.11, 41.53, 41.35, 35.93, 33.20, 33.06, 31.99, 31.93, 31.55, 31.43, 29.60, 29.53, 29.37, 22.77, 21.44, 21.32, 14.21.



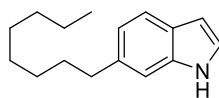
**2,4-dimethyl-4-(5'-octyl-indanylmethyl)-oxazoline (2.5)**

A round-bottom flask was charged with  $\text{AlCl}_3$  (0.39 g, 2.9 mmol) in acetonitrile (4.5 mL, 86 mmol) and cooled to  $-42\text{ }^\circ\text{C}$ . A second round-bottom flask was charged with a solution of **2.4** (0.88 g, 2.9 mmol) in acetonitrile (4.5 mL, 86 mmol) and similarly cooled to  $-42\text{ }^\circ\text{C}$ . The epoxide solution was then added to the aluminum solution dropwise via syringe at  $-42\text{ }^\circ\text{C}$ . The resulting mixture was allowed to warm to room temperature and stirred overnight until complete by TLC. The reaction was quenched with saturated  $\text{K}_2\text{CO}_3$  (9 mL). The liquid was extracted three times with ethyl acetate. The combined organic extracts were then washed two times with saturated  $\text{K}_2\text{CO}_3$  (9 mL) before they were dried over  $\text{MgSO}_4$ , filtered, and concentrated *in vacuo* to a crude yellow solid, which was not further purified due to degradation during chromatography.



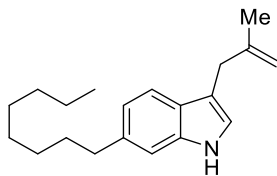
**5-octyl-1-(2'-amino-3'-hydroxy-2'-methylpropane) (2.6 or VPC142072)**

A round-bottom flask was charged with a solution of crude **2.5** (1.00 g) in a 2:1 mixture of THF:1 N HCl (45 mL). The mixture was stirred for 90 minutes at room temperature before solid  $K_2CO_3$  was added until the solution reached a pH of 9.5. The mixture was allowed to stir overnight at room temperature before being extracted three times with ethyl acetate. The combined organic extracts were washed one time with saturated  $K_2CO_3$  and were then dried over  $MgSO_4$ , filtered, and concentrated *in vacuo* to a yellow solid. A two-neck flask equipped with a reflux condenser was charged with a solution of the crude solid in a 2:1:3 mixture of THF:MeOH:1 N LiOH (60 mL). The solution was heated at reflux for 18 hours and was subsequently allowed to cool before it was diluted with ethyl acetate (30 mL) and 3 N NaOH (30 mL). The aqueous layer then was extracted three times with ethyl acetate, and the combined organic extracts were dried over  $Na_2SO_4$ , filtered, and concentrated *in vacuo* to a yellow solid. The crude product was purified via preparatory TLC (80:20 -  $CH_2Cl_2$ :MeOH) to afford a white solid (0.36 g, 1.1 mmol, 39% yield over two steps,  $R_f = 0.23$  in 80:20 -  $CH_2Cl_2$ :MeOH). mp not recorded.  $^1H$  NMR (500 MHz,  $CD_3OD$ )  $\delta$  7.08-7.12 (d,  $J = 7.5$  Hz, 1H), 6.97-6.99 (s, 1H), 6.93-6.97 (d,  $J = 7.7$  Hz, 1H), 3.43-3.56 (m, 2H), 3.34-3.35 (s, 1H), 3.10-3.21 (m, 1H), 2.75-2.93 (m, 2H), 2.52-2.58 (t,  $J = 7.6$  Hz, 2H), 2.38-2.49 (m, 1H), 2.07-2.22 (m, 1H), 1.88-1.89 (s, 2H), 1.65-1.77 (m, 1H), 1.49-1.62 (m, 3H), 1.24-1.36 (m, 10H), 1.21-1.24 (s, 3H), 0.87-0.92 (t,  $J = 6.9$  Hz, 3H).  $^{13}C$  NMR (126 MHz,  $CD_3OD$ )  $\delta$  145.98, 144.67, 142.20, 127.50, 125.21, 124.03, 70.38, 69.85, 55.63, 45.11, 41.28, 36.77, 36.40, 36.28, 33.05, 32.67, 30.60, 30.43, 30.34, 24.00, 23.73, 23.66, 14.45. MS  $m/z$  calculated for  $[2M+1]^+ = 635.54$  and  $m/z$  calculated for  $[M+1]^+ = 318.27$ ; found 635.5 and 318.3.  $t_R = 9.23$  min.



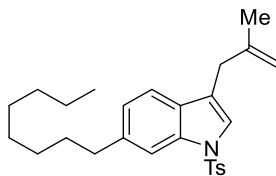
### 6-octylindole (2.7)

A round-bottom flask was charged with a 0.5 M solution of 9-BBN in THF (32.9 mL, 16.5 mmol) and then 1-octene (2.82 mL, 16.5 mmol) dropwise via syringe at room temperature. The reaction was stirred overnight at room temperature. A 3 M NaOH solution (6.22 mL, 18.6 mmol) was then added to the octylborane solution and the mixture was stirred for 10 minutes. A two-neck flask equipped with a reflux condenser was charged with Pd(PPh<sub>3</sub>)<sub>4</sub> (507 mg, 0.4 mmol) and 6-bromoindole (2.15 g, 11.0 mmol) dissolved in THF (12.4 mL). The octylborane solution was added to the two-neck flask dropwise via syringe at room temperature, and the newly-formed solution was refluxed for 4 hours until the reaction was complete by TLC. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude oil was then dissolved in dichloromethane and washed once with a 0.5 M solution of Rochelle's salt and three times with water. The organic fraction was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a black tar. The crude product was purified via flash chromatography (97.5:2.5 - hexanes:ethyl acetate) to afford a viscous purple oil (1.24 g, 5.42 mmol, 49% yield, R<sub>f</sub> = 0.48 in 80:20 - hexanes:ethyl acetate). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.99-8.08 (bs, 1H), 7.54-7.56 (d, *J* = 8.1 Hz, 1H), 7.19-7.21 (s, 1H), 7.12-7.16 (m, 1H), 6.96-6.99 (d, *J* = 8.1 Hz, 1H), 6.50-6.52 (m, 1H), 2.69-2.74 (t, *J* = 8.1 Hz, 2H), 1.63-1.70 (m, 2H), 1.20-1.39 (m, 10H), 0.87-0.91 (t, *J* = 6.8 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 137.24, 136.30, 125.96, 123.65, 121.18, 120.43, 110.45, 102.52, 36.43, 32.27, 32.07, 29.71, 29.56, 29.45, 22.83, 14.27.



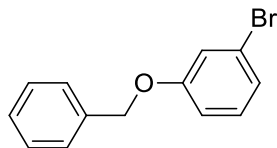
### 3-(2'-methallyl)- 6-octyl-indole (2.8)

A round-bottom flask was charged with **2.7** (0.50 g, 2.18 mmol), zinc triflate (0.48 g, 1.3 mmol), and tetrabutylammonium iodide (0.40 g, 1.1 mmol) in toluene (10 mL). *N,N*-diisopropyl-ethylamine (0.42 mL, 2.2 mmol) was added dropwise via syringe, and the newly formed mixture was allowed to stir at room temperature for 15 minutes. Once complete, 3-bromo-2-methylpropene (0.22 mL, 2.2 mmol) was added via syringe and the resulting mixture was heated to 40 °C and allowed to react at elevated temperature for 5 days. The reaction was quenched with a saturated solution of ammonium chloride and subsequently diluted with water. The mixture was extracted three times with ethyl acetate and the combined organic extracts were dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a reddish-brown oil. The crude product was purified via flash chromatography (97.5:2.5 – hexanes:ethyl acetate) to afford a yellow oil (0.16 g, 0.57 mmol, 26% yield, R<sub>f</sub> = 0.62 in 80:20 - hexanes:ethyl acetate). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.65-7.70 (m, 2H), 7.18-7.20 (s, 1H), 7.12-7.15 (d, *J* = 8.1 Hz, 1H), 6.94-6.96 (s, 1H), 5.01-5.03 (s, 1H), 4.98-5.00 (s, 1H), 3.60-3.63 (s, 2H), 2.85-2.90 (t, *J* = 7.5 Hz, 2H), 1.91-1.93 (s, 3H), 1.80-1.88 (m, 2H), 1.41-1.59 (m, 10H), 1.06-1.09 (t, *J* = 6.8 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 145.12, 136.97, 136.84, 125.89, 121.76, 120.48, 119.00, 113.86, 111.11, 110.50, 36.45, 34.31, 32.30, 32.07, 29.72, 29.63, 29.47, 22.82, 22.34, 14.25.

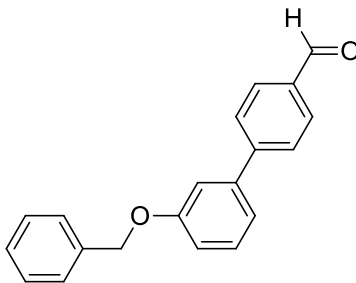


### 3-(2'-methallyl)-6-octyl-1-tosyl-indole (2.9)

A round-bottom flask was charged with a suspension of NaH (0.03 g, 0.078 mmol) in THF (1 mL). A solution of **2.8** (0.10 g, 0.35 mmol) in THF (0.75 mL), cooled to 0°C, was then added to the round-bottom flask. Finally, a solution of TsCl (0.74 g, 0.078 mmol) in THF (0.5 mL) was also cooled and added to the original round-bottom flask. The reaction was allowed to warm to room temperature and stirred for 15 hours. Once complete, the reaction was quenched with water, and the resulting solution was extracted three times with ethyl acetate. The combined organic fractions were washed once with saturated NaHCO<sub>3</sub> and once with brine before they were dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a yellow oil. The crude product was purified via flash chromatography (97.5:2.5 – hexanes:ethyl acetate) to afford a clear oil (0.13 g, 0.29 mmol, 82% yield, R<sub>f</sub> = 0.62 in 80:20 - hexanes:ethyl acetate). <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 7.78-7.79 (s, 1H), 7.71-7.73 (d, *J* = 8.4 Hz, 2H), 7.35-7.37 (d, *J* = 8.0 Hz, 1H), 7.27-7.28 (s, 1H), 7.17-7.19 (d, *J* = 7.9 Hz, 2H), 7.03-7.05 (d, *J* = 8.0 Hz, 1H), 4.81-4.83 (s, 1H), 4.75-4.77 (s, 1H), 3.33-3.35 (s, 2H), 2.67-2.74 (t, *J* = 7.6 Hz, 2H), 2.33-2.34 (s, 3H), 1.68-1.70 (s, 3H), 1.58-1.67 (m, 2H), 1.21-1.36 (m, 10H), 0.87-0.91 (t, *J* = 7.0 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 144.71, 143.19, 140.10, 136.00, 135.41, 129.81, 129.27, 126.85, 124.11, 123.47, 120.99, 119.58, 113.55, 112.34, 36.49, 33.98, 32.16, 32.07, 29.69, 29.47, 29.41, 22.84, 22.24, 21.67, 14.27.

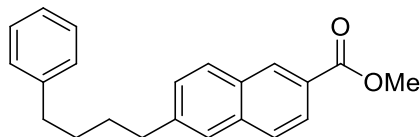
**1-(benzyloxy)-3-bromobenzene (4.1)**

A two-neck flask equipped with a reflux condenser was charged with 3-bromophenol (3.00 g, 17.30 mmol) and anhydrous  $K_2CO_3$  (4.78 g, 34.6 mmol). The solids were suspended in dry acetone (18 mL) before addition of benzyl bromide (2.3 mL, 19 mmol). The mixture was heated at reflux for 2 hours until the reaction was complete by TLC. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude solid was then dissolved using  $CH_2Cl_2$  and water. The aqueous layer was extracted three times with  $CH_2Cl_2$  and the organic extracts were combined, dried over  $MgSO_4$ , filtered, and concentrated *in vacuo* to yield a yellow solid. The crude product was purified via flash chromatography (95:5 - hexanes: $CH_2Cl_2$ ) to afford a white solid (4.38 g, 98% yield,  $R_f = 0.40$  in 95:5 - hexanes: $CH_2Cl_2$ ). mp = 58-60 °C.  $^1H$  NMR (600 MHz,  $CDCl_3$ )  $\delta$  7.36-7.47 (m, 5H), 7.12-7.21 (m, 3H), 6.93-6.95 (m, 1H), 5.06 (s, 2H).  $^{13}C$  NMR (151 MHz,  $CDCl_3$ )  $\delta$  159.63, 136.48, 130.68, 128.75, 128.25, 127.60, 124.16, 122.93, 118.27, 113.92, 70.28.



### 3'-(benzyloxy)-[1,1'-biphenyl]-4-carbaldehyde (4.2)

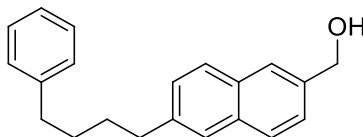
A two-neck flask equipped with a reflux condenser was charged with **4.1** (4.38 g, 16.4 mmol), 4-formylphenylboronic acid (3.68 g, 24.6 mmol), and Pd(PPh<sub>3</sub>)<sub>4</sub> (0.76 g, 0.66 mmol). The solids were suspended in dry THF (13 mL) and a 2 M aqueous solution of Na<sub>2</sub>CO<sub>3</sub> (1.3 mL). The mixture was heated at reflux for 18 hours until the reaction was complete by TLC. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude solid was then dissolved in ethyl acetate and saturated Na<sub>2</sub>CO<sub>3</sub>. The organic fraction was washed two times with more saturated Na<sub>2</sub>CO<sub>3</sub> and one time with saturated NaHCO<sub>3</sub>. The organic layer was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a black tar. The crude product was purified via flash chromatography (gradient 95:5 → 70:30 - hexanes:CH<sub>2</sub>Cl<sub>2</sub>) to afford a clear oil (2.13 g, 45% yield, R<sub>f</sub> = 0.11 in 66:33 - hexanes:CH<sub>2</sub>Cl<sub>2</sub>). <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 9.95-9.97 (s, 1H), 7.83-7.87 (d, *J* = 8.4 Hz, 2H), 7.64-7.67 (d, *J* = 8.4 Hz, 2H), 7.43-7.45 (d, *J* = 7.4 Hz, 2H), 7.28-7.38 (m, 4H), 7.20-7.23 (s, 1H), 7.16-7.19 (d, *J* = 7.2 Hz, 1H), 6.97-7.00 (d, *J* = 7.0 Hz), 5.04-5.07 (s, 2H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 191.69, 159.19, 146.68, 146.66, 141.01, 136.71, 135.21, 130.10, 130.01, 128.55, 127.99, 127.54, 127.45, 119.94, 114.50, 114.08, 69.97.



#### **methyl 6-(4-phenylbutyl)-2-naphthoate (4.4)**

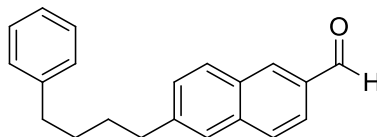
A round-bottom flask was charged with a 0.5 M solution of 9-BBN in THF (33.9 mL, 17.0 mmol) and then 4-phenyl-1-butene (2.6 mL, 17 mmol) dropwise via syringe at room temperature. The reaction was stirred overnight at room temperature. A 3 M NaOH solution (6.5 mL, 19 mmol) was then added to the 4-phenylbutyllborane solution and the mixture was stirred for 10 minutes. A two-neck flask equipped with a reflux condenser was charged with Pd(PPh<sub>3</sub>)<sub>4</sub> (520 mg, 0.45 mmol) and methyl 6-bromo-2-naphthoate (3.00 g, 11.3 mmol) dissolved in THF (9 mL). The 4-phenylbutyllborane solution was added to the two-neck flask dropwise via syringe at room temperature, and the newly-formed solution was refluxed for 4 hours until the reaction was complete by TLC. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude oil was then dissolved in ethyl acetate and washed three times with water. The organic fraction was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a black tar. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a white solid (2.99 g, 83% yield, R<sub>f</sub> = 0.41 in 90:10 - hexanes:ethyl acetate). mp = 70-73 °C. <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 8.58-8.60 (d, *J* = 2.2 Hz, 1H), 8.04-8.07 (dd, *J* = 8.6, 1.7 Hz, 1H), 7.86-7.89 (d, *J* = 8.4 Hz, 1H), 7.79-7.82 (d, *J* = 8.6 Hz, 1H), 7.63-7.65 (d, *J* = 2.2 Hz, 1H), 7.37-7.41 (dd, *J* = 8.4, 1.7 Hz, 1H), 7.27-7.32 (m, 2H), 7.17-7.23 (m, 3H), 3.98-4.00 (s, 3H), 2.81-2.85 (t, *J* = 7.4 Hz, 2H), 2.66-2.70 (t, *J* = 7.5 Hz, 2H), 1.69-1.84 (m, 4H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 167.48, 143.05, 142.50, 135.88, 131.09, 130.94, 129.36, 128.52, 128.41, 128.34, 127.73, 126.73, 126.37, 125.83, 125.41, 52.27, 36.19, 35.92, 31.20, 30.85.





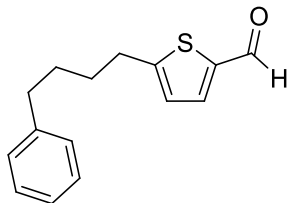
**(6-(4-phenylbutyl)naphthalen-2-yl)methanol (4.5)**

A round-bottom flask was charged with a solution of **4.4** (0.50 g, 1.6 mmol) in toluene (11 mL). The solution was cooled to  $-78\text{ }^{\circ}\text{C}$  before adding a 1.0 M solution of DIBALH in hexanes (3.6 mL, 3.6 mmol) dropwise via syringe. The reaction was stirred for 1 hour at  $-78\text{ }^{\circ}\text{C}$  before it was quenched by addition of methanol (2 mL). The solution was diluted with ethyl acetate washed three times with a saturated solution of Rochelle's salt, one time with water, and one time with brine. The organic layer was dried over  $\text{MgSO}_4$ , filtered, and concentrated *in vacuo* to yield a white solid (0.41 g, 90% yield,  $R_f = 0.38$  in 60:40 - hexanes:ethyl acetate) that required no further purification. mp =  $101\text{-}104\text{ }^{\circ}\text{C}$ .  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.75-7.80 (m, 3H), 7.61-7.63 (s, 1H), 7.45-7.47 (dd,  $J = 8.4, 1.7$  Hz, 1H), 7.34-7.37 (dd,  $J = 8.4, 1.7$  Hz, 1H), 7.29-7.33 (m, 2H), 7.20-7.24 (m, 3H), 4.83-4.84 (d,  $J = 0.9$  Hz, 2H), 2.81-2.85 (t,  $J = 7.4$  Hz, 2H), 2.67-2.72 (t,  $J = 7.5$  Hz, 2H), 2.00-2.05 (bs, 1H), 1.71-1.84 (m, 4H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  142.62, 140.29, 137.58, 133.21, 131.93, 128.52, 128.38, 127.94, 127.88, 127.84, 126.30, 125.78, 125.38, 125.34, 65.58, 36.02, 35.94, 31.20, 31.02.



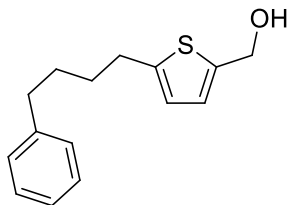
#### 6-(4-phenylbutyl)-2-naphthaldehyde (4.6)

A round-bottom flask was charged with a solution of oxalyl chloride (0.24 mL, 2.8 mmol) in  $\text{CH}_2\text{Cl}_2$  (3 mL). The solution was cooled to  $-78\text{ }^\circ\text{C}$  before addition of DMSO (0.40 mL, 5.7 mmol) dropwise via syringe. The solution was stirred for 10 minutes before addition of **4.5** (0.44 g, 1.4 mmol) in  $\text{CH}_2\text{Cl}_2$  (3 mL) dropwise via syringe at reduced temperature. The reaction was stirred at  $-78\text{ }^\circ\text{C}$  for an additional 40 minutes before addition of triethylamine (1.2 mL, 8.5 mmol) dropwise via syringe at reduced temperature. After 10 minutes of stirring at  $-78\text{ }^\circ\text{C}$ , the reaction was allowed to warm to room temperature and was determined to be complete by TLC. The solution was diluted with ethyl acetate and washed two times with 1 M HCl and one time with brine. The organic layer was then dried over  $\text{MgSO}_4$ , filtered, and concentrated *in vacuo* to yield a light yellow solid (0.39 g, 96% yield,  $R_f = 0.65$  in 60:40 - hexanes:ethyl acetate) that required no further purification. mp =  $54\text{-}57\text{ }^\circ\text{C}$ .  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  10.14-10.15 (s, 1H), 8.27-8.30 (s, 1H), 7.94-7.97 (dd,  $J = 8.5, 1.6$  Hz, 1H), 7.90-7.93 (d,  $J = 8.4$  Hz, 1H), 7.84-7.87 (d,  $J = 8.5$  Hz, 1H), 7.65-7.68 (s, 1H), 7.43-7.46 (dd,  $J = 8.4, 1.7$  Hz, 1H), 7.29-7.34 (m, 2H), 7.19-7.25 (m, 3H), 2.83-2.88 (t,  $J = 7.5$  Hz, 2H), 2.68-2.72 (t,  $J = 7.5$  Hz, 2H), 1.72-1.83 (m, 4H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  192.26, 144.07, 142.38, 136.76, 134.41, 133.57, 131.14, 129.49, 128.69, 128.60, 128.46, 128.37, 126.66, 125.81, 122.89, 36.17, 35.84, 31.11, 30.74.



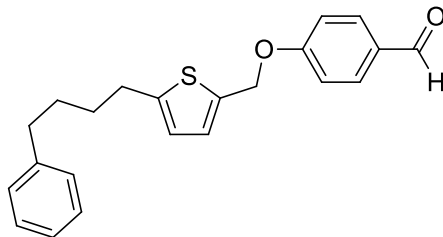
#### 5-(4-phenylbutyl)thiophene-2-carbaldehyde (4.8)

A round-bottom flask was charged with a 0.5 M solution of 9-BBN in THF (33.9 mL, 17.0 mmol) and then 4-phenyl-1-butene (2.6 mL, 17 mmol) dropwise via syringe at room temperature. The reaction was stirred overnight at room temperature. A 3 M NaOH solution (6.5 mL, 19 mmol) was then added to the 4-phenylbutyllborane solution and the mixture was stirred for 10 minutes. A two-neck flask equipped with a reflux condenser was charged with Pd(PPh<sub>3</sub>)<sub>4</sub> (520 mg, 0.45 mmol) and 5-bromothiophene-2-carbaldehyde (2.16 g, 11.3 mmol) dissolved in THF (9 mL). The 4-phenylbutyllborane solution was added to the two-neck flask dropwise via syringe at room temperature, and the newly-formed solution was refluxed for 4 hours until the reaction was complete by TLC. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude oil was then dissolved in ethyl acetate and washed three times with water. The organic fraction was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a black tar. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a light yellow oil (1.75 g, 63% yield, R<sub>f</sub> = 0.76 in 50:50 - hexanes:ethyl acetate). <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 9.81-9.82 (s, 1H), 7.59-7.62 (d, *J* = 3.7 Hz, 1H), 7.27-7.31 (m, 2H), 7.16-7.21 (m, 3H), 6.87-6.89 (d, *J* = 3.8 Hz, 1H), 2.88-2.92 (t, *J* = 6.9 Hz, 2H), 2.64-2.68 (t, *J* = 7.4 Hz, 2H), 1.18-1.80 (m, 4H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 182.76, 157.38, 142.09, 141.79, 137.14, 128.48, 128.47, 126.05, 125.96, 35.65, 30.91, 30.85, 30.79.



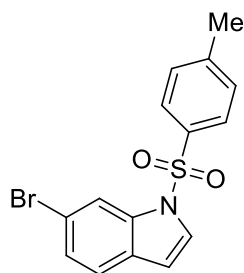
**(5-(4-phenylbutyl)thiophen-2-yl)methanol (4.9)**

A round-bottom flask was charged with a solution of **4.8** (0.80 g, 3.3 mmol) in toluene (23 mL). The solution was cooled to  $-78\text{ }^{\circ}\text{C}$  before adding a 1.0 M solution of DIBALH in hexanes (3.8 mL, 3.8 mmol) dropwise via syringe. The reaction was stirred for 1 hour at  $-78\text{ }^{\circ}\text{C}$  before it was quenched by addition of methanol (2 mL). The solution was diluted with ethyl acetate washed three times with a saturated solution of Rochelle's salt, one time with water, and one time with brine. The organic layer was dried over  $\text{MgSO}_4$ , filtered, and concentrated *in vacuo* to afford a light yellow oil. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a clear oil (0.65 g, 80% yield,  $R_f = 0.64$  in 50:50 - hexanes:ethyl acetate).  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.40-7.45 (m, 2H), 7.30-7.35 (m, 3H), 6.89-6.91 (d,  $J = 3.4$  Hz, 1H), 6.74-6.76 (d,  $J = 3.4$  Hz, 1H), 4.76-4.78 (s, 2H), 3.25-3.37 (bs, 1H), 2.92-2.97 (t,  $J = 6.9$  Hz, 3H), 2.76-2.81 (t,  $J = 7.2$  Hz, 3H), 1.81-1.90 (m, 4H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  145.68, 142.19, 141.35, 128.28, 128.18, 125.61, 125.14, 123.60, 59.62, 35.52, 31.17, 30.72, 29.93.



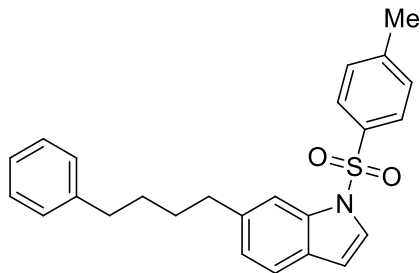
#### 4-((5-(4-phenylbutyl)thiophen-2-yl)methoxy)benzaldehyde (4.10)

A round-bottom flask was charged with a solution of **4.9** (0.65 g, 2.6 mmol), 4-hydroxybenzaldehyde (0.35 g, 2.9 mmol), and triphenylphosphine (1.03 g, 3.93 mmol) in THF (14 mL). The solution was cooled to 0 °C before addition of diisopropyl azodicarboxylate (0.80 g, 3.9 mmol). The flask was allowed to warm to room temperature and stirred overnight until the reaction was complete. The mixture concentrated *in vacuo*. The crude oil was then dissolved in ethyl acetate and washed once with saturated Na<sub>2</sub>CO<sub>3</sub>, once with water, and once with brine. The organic fraction was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a yellow oil. The crude product was purified via flash chromatography (95:5 - hexanes:ethyl acetate) to afford a milky white solid (0.078 g, 8.4% yield, R<sub>f</sub> = 0.71 in 50:50 - hexanes:ethyl acetate). <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 9.89-9.91 (s, 1H), 7.83-7.87 (d, *J* = 8.7 Hz, 2H), 7.27-7.31 (m, 2H), 7.16-7.22 (m, 3H), 7.07-7.10 (d, *J* = 8.7 Hz, 2H), 6.94-6.96 (d, *J* = 3.4 Hz, 1H), 6.67-6.69 (d, *J* = 3.4 Hz, 1H), 5.22-5.23 (s, 2H), 2.82-2.86 (t, *J* = 6.9 Hz, 2H), 2.64-2.68 (t, *J* = 7.1 Hz, 2H), 1.71-1.74 (m, 4H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 190.96, 163.44, 147.51, 142.39, 135.34, 132.11, 130.31, 128.50, 128.42, 127.54, 125.86, 124.01, 115.28, 65.52, 35.73, 31.27, 30.93, 30.17.



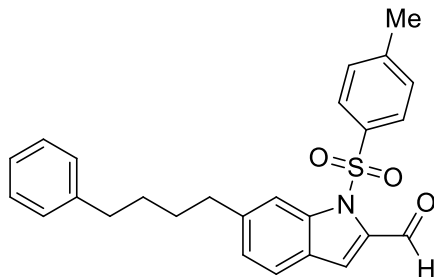
### 6-bromo-1-tosyl-1H-indole (4.12)

A round-bottom flask was charged with sodium hydride dispersed in mineral oil (60% by weight, 1.12 g, 28.1 mmol). The solid was suspended in THF (5 mL) and the mixture was cooled to 0 °C. A second round-bottom flask was charged with 6-bromoindole (5.00 g, 25.5 mmol) in THF (50 mL), and this solution was also cooled to 0 °C. The indole solution was then added to the sodium hydride solution dropwise via syringe and stirred at 0 °C for 30 minutes. A third round-bottom flask was charged with PhSO<sub>4</sub>Cl (5.36 g, 28.1 mmol) in THF (45 mL) and this solution was added to the initial flask dropwise via syringe. The resulting mixture was allowed to come to room temperature and stir over a period of 16 hours. Upon completion, the reaction was quenched with water (10 mL) and then concentrated *in vacuo*. The crude solid was then dissolved using ethyl acetate and water. The aqueous layer was extracted three times with ethyl acetate. The organic extracts were combined, washed one time with brine, dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a red-brown solid. The crude product was purified via flash chromatography (90:10 - hexanes:ethyl acetate) to afford an off-white solid (6.34 g, 18.1 mmol, 61% yield, R<sub>f</sub> = 0.47 in 80:20 - hexanes:ethyl acetate). mp = 128-129 °C. <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 8.16-8.17 (d, J = 1.7 Hz, 1H), 7.75-7.79 (d, J = 8.4 Hz, 2H), 7.52-7.54 (d, J = 3.7 Hz, 1H), 7.37-7.39 (d, J = 8.4 Hz, 1H), 7.32-7.35 (dd, J = 8.3, 1.7 Hz, 1H), 7.24-7.27 (d, J = 8.9 Hz, 2H), 6.60-6.62 (d, J = 3.7 Hz, 1H), 2.35-2.37 (s, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 145.43, 135.60, 135.19, 130.19, 129.68, 126.96, 126.94, 126.80, 122.59, 118.37, 116.73, 108.89, 21.75.



### 6-(4'-phenylbutyl)-1-tosyl-1H-indole (4.13)

A round-bottom flask was charged with a 0.5 M solution of 9-BBN in THF (54.2 mL, 27.2 mmol). 4-phenyl-1-butene (4.1 mL, 27 mmol) was added to this solution dropwise via syringe at room temperature and allowed to stir overnight. Once complete, a 3 M NaOH solution (10.3 mL) was added to the flask at room temperature and stirred for 10 minutes. A two-neck flask equipped with a reflux condenser was then charged with a solution of **4.12** (6.34 g, 18.1 mmol) and Pd(PPh<sub>3</sub>)<sub>4</sub> (0.84 g, 1.1 mmol) in THF (20 mL). The borane solution was transferred to the two-neck flask dropwise via syringe and the resulting solution was refluxed for a period of 4 hours. The mixture was allowed to cool before it was concentrated *in vacuo*. The crude solid was then dissolved in CH<sub>2</sub>Cl<sub>2</sub> and saturated Na<sub>2</sub>CO<sub>3</sub>. The organic fraction was washed once with a 0.5 M solution of Rochelle's salt and three times with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated *in vacuo* to yield a black tar. The crude product was then purified via flash chromatography (95:5 – hexanes:ethyl acetate) to afford an waxy white solid (6.93 g, 95% yield, R<sub>f</sub> = 0.62 in 80:20 – hexanes:ethyl acetate). mp = 90-95 °C. <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 7.80-7.82 (s, 1H), 7.71-7.74 (d, *J* = 8.3 Hz, 2H), 7.47-7.49 (d, *J* = 3.7 Hz, 1H), 7.39-7.42 (d, *J* = 8.0 Hz, 1H), 7.25-7.31 (m, 2H), 7.16-7.21 (m, 3H), 7.11-7.14 (d, *J* = 8.6 Hz, 2H), 7.03-7.06 (d, *J* = 8.1 Hz, 1H), 6.59-7.60 (d, *J* = 3.7 Hz, 1H), 2.74-2.78 (t, *J* = 7.1 Hz, 2H), 2.63-2.67 (t, *J* = 7.3 Hz, 2H), 2.30-2.31 (s, 3H), 1.63-1.74 (m, 4H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 144.92, 142.68, 139.62, 135.42, 135.33, 129.91, 128.88, 128.55, 128.44, 126.88, 125.96, 125.83, 124.37, 121.11, 113.33, 109.16, 36.31, 36.00, 31.73, 31.05, 21.67.



**6-(4-phenylbutyl)-1-tosyl-1H-indole-2-carbaldehyde (4.14)**

A round-bottom flask was charged with a solution of **4.13** (0.50 g, 1.2 mmol) in THF (10 mL). The mixture was cooled to  $-78\text{ }^{\circ}\text{C}$  and a 2.65 M solution of *n*-BuLi in hexanes (0.47 mL, 1.2 mmol) was added dropwise via syringe and the resulting solution was allowed to stir for 2 hours. Following this, DMF (0.19 mL, 2.5 mmol) was added to the cooled solution dropwise via syringe. The reaction was allowed to come to room temperature and stirred over a period of 24 hours. Once complete, the reaction was quenched with a saturated  $\text{NH}_4\text{Cl}$  solution (14 mL) and then concentrated *in vacuo*. The resulting aqueous solution was extracted three times with  $\text{CH}_2\text{Cl}_2$  and the combined organic extracts were dried over  $\text{MgSO}_4$ , filtered, and concentrated *in vacuo* to yield an orange oil. The crude product was purified via flash chromatography (95:5 – hexanes:ethyl acetate) to afford light orange crystals (0.21 g, 0.48 mmol, 38% yield,  $R_f = 0.54$  in 80:20 – hexanes:ethyl acetate). mp =  $112\text{--}115\text{ }^{\circ}\text{C}$ .  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  10.50-10.51 (s, 1H), 8.02-8.04 (s, 1H), 7.59-7.61 (d,  $J = 8.4\text{ Hz}$ , 2H), 7.49-7.51 (d,  $J = 8.1\text{ Hz}$ , 1H), 7.42-7.43 (s, 1H), 7.27 – 7.31 (m, 2H), 7.17 – 7.21 (m, 3H), 7.12-7.15 (d,  $J = 9.5\text{ Hz}$ , 1H), 7.07-7.10 (d,  $J = 8.5\text{ Hz}$ , 2H), 2.79-2.83 (t,  $J = 7.4\text{ Hz}$ , 2H), 2.65-2.69 (t,  $J = 7.3\text{ Hz}$ , 2H), 2.29-2.30 (s, 3H), 1.64-1.78 (m, 4H).



---

# 6

---

## References

- [1] Baumruker, T.; Billich, A. Sphingolipids in Cell Signaling: Their Function as Receptor Ligands, Second Messengers, and Raft Constituents. *Current Immunology Reviews*. **2006**, *2*, 101-118.
- [2] Gault, C.R.; Obeid, L.M.; Hannun, Y.A. An Overview of Sphingolipid Metabolism: From Synthesis to Breakdown. *Advances in Experimental Medicine and Biology*. **2010**, *688*, 1-23.
- [3] Le Stunff, H.; Milstien, S.; Spiegel, S. Generation and Metabolism of Bioactive Sphingosine-1-Phosphate. *Journal of Cellular Biochemistry*. **2004**, *92*, 882-899.
- [4] Maceyka, M.; Sankala, H.; Hait, N.C.; Le Stunff, H.; Liu, H.; Toman, R.; Collier, C.; Zhang, M.; Satin, L.S.; Merrill, A.H.; Milstien, S.; Spiegel, S. Sphk1 and Sphk2, Sphingosine Kinase Isoenzymes with Opposing Functions in Sphingolipid Metabolism. *Journal of Biological Chemistry*. **2005**, *280*, 37118-37129.
- [5] Hait, N.C.; Allegood, J.; Maceyka, M.; Strub, G.M.; Harikumar, K.B.; Singh, S.K.; Luo, C.; Marmorstein, R.; Kordula, T.; Milstien, S.; Spiegel, S. Regulation of Histone Acetylation in the Nucleus by Sphingosine-1-Phosphate. *Science*. **2009**, *325*, 1254-1257.
- [6] Alvarez, S.E.; Harikumar, K.B.; Hait, N.C.; Allegood, J.; Strub, G.M.; Kim, E.Y.; Maceyka, M.; Jiang, H.L.; Luo, C.; Kordula, T.; Milstien, S.; Spiegel, S. Sphingosine-1-Phosphate Is a Missing Cofactor for the E3 Ubiquitin Ligase Traf2. *Nature*. **2010**, *465*, 1084-1089.
- [7] Maceyka, M.; Payne, S.G.; Milstien, S.; Spiegel, S. Sphingosine Kinase, Sphingosine-1-Phosphate, and Apoptosis. *Biochimica et Biophysica Acta*. **2002**, *1585*, 193-201.
- [8] Kawahara, A.; Nishi, T.; Hisano, Y.; Fukui, H.; Yamaguchi, A.; Mochizuki, N. The Sphingolipid Transporter Spns2 Functions in Migration of Zebrafish Myocardial Precursors. *Science*. **2009**, *323*, 524-527.
- [9] Nijnik, A.; Clare, S.; Hale, C.; Chen, J.; Raisen, C.; Mottram, L.; Lucas, M.; Estabel, J.; Ryder, E.; Adissu, H.; Adams, N.C.; Ramirez-Solis, R.; White, J.K.; Steel, K.P.; Dougan, G.; Hancock, R.E.W. The Role of Sphingosine-1-Phosphate Transporter Spns2 in Immune System Function. *Journal of Immunology*. **2012**, *189*, 102-111.
- [10] Young, N.; Van Brocklyn, J.R. Signal Transduction of Sphingosine-1-Phosphate G Protein-Coupled Receptors. *Scientific World Journal*. **2006**, *6*, 946-966.
- [11] Lagerstrom, M.C.; Schioth, H.B. Structural Diversity of G Protein-Coupled Receptors and Significance for Drug Discovery. *Nature Reviews Drug Discovery*. **2008**, *7*, 339-357.

- [12] Urizar, E.; Claeysen, S.; Deupi, X.; Govaerts, C.; Costagliola, S.; Vassart, G.; Pardo, L. An Activation Switch in the Rhodopsin Family of G Protein-Coupled Receptors: The Thyrotropin Receptor. *Journal of Biological Chemistry*. **2005**, *280*, 17135-17141.
- [13] Brinkmann, V. Sphingosine 1-Phosphate Receptors in Health and Disease: Mechanistic Insights from Gene Deletion Studies and Reverse Pharmacology. *Pharmacology & Therapeutics*. **2007**, *115*, 84-105.
- [14] Hanson, M.A.; Roth, C.B.; Jo, E.; Griffith, M.T.; Scott, F.L.; Reinhart, G.; Desale, H.; Clemons, B.; Cahalan, S.M.; Schuerer, S.C.; Sanna, M.G.; Han, G.W.; Kuhn, P.; Rosen, H.; Stevens, R.C. Crystal Structure of a Lipid G Protein-Coupled Receptor. *Science*. **2012**, *335*, 851-855.
- [15] Chun, E.; Thompson, A.A.; Liu, W.; Roth, C.B.; Griffith, M.T.; Katritch, V.; Kunken, J.; Xu, F.; Cherezov, V.; Hanson, M.A.; Stevens, R.C. Fusion Partner Toolchest for the Stabilization and Crystallization of G Protein-Coupled Receptors. *Structure*. **2012**, *20*, 967-976.
- [16] Parrill, A.L.; Wang, D.; Bautista, D.L.; Van Brocklyn, J.R.; Lorincz, Z.; Fischer, D.J.; Baker, D.L.; Liliom, K.; Spiegel, S.; Tigyi, G. Identification of Edg1 Receptor Residues That Recognize Sphingosine 1-Phosphate. *Journal of Biological Chemistry*. **2000**, *275*, 39379-39384.
- [17] Parrill, A.L.; Lima, S.; Spiegel, S. Structure of the First Sphingosine 1-Phosphate Receptor. *Science Signaling*. **2012**, *5*, 1-4.
- [18] Fujiwara, Y.; Osborne, D.A.; Walker, M.D.; Wang, D.A.; Bautista, D.A.; Liliom, K.; Van Brocklyn, J.R.; Parrill, A.L.; Tigyi, G. Identification of the Hydrophobic Ligand Binding Pocket of the S1p<sub>1</sub> Receptor. *Journal of Biological Chemistry*. **2007**, *282*, 2374-2385.
- [19] van Loenen, P.B.; de Graaf, C.; Verzijl, D.; Leurs, R.; Rognan, D.; Peters, S.L.; Alewijnse, A.E. Agonist-Dependent Effects of Mutations in the Sphingosine-1-Phosphate Type 1 Receptor. *European Journal of Pharmacology*. **2011**, *667*, 105-112.
- [20] Deng, Q.L.; Clemas, J.A.; Chrebet, G.; Fischer, P.; Hale, J.J.; Li, Z.; Mills, S.G.; Bergstrom, J.; Mandala, S.; Mosley, R.; Parent, S.A. Identification of Leu276 of the S1p<sub>1</sub> Receptor and Phe263 of the S1p<sub>3</sub> Receptor in Interaction with Receptor Specific Agonists by Molecular Modeling, Site-Directed Mutagenesis, and Affinity Studies. *Molecular Pharmacology*. **2007**, *71*, 724-735.
- [21] Kluk, M.J.; Hla, T. Signaling of Sphingosine-1-Phosphate Via the S1p/Edg-Family of G-Protein-Coupled Receptors. *Biochimica et Biophysica Acta*. **2002**, *1582*, 72-80.
- [22] Huang, C.L.; Jan, Y.N.; Jan, L.Y. Binding of the G Protein Beta Gamma Subunit to Multiple Regions of G Protein-Gated Inward-Rectifying K<sup>+</sup> Channels. *FEBS Letters*. **1997**, *405*, 291-298.
- [23] Oo, M.L.; Chang, S.H.; Thangada, S.; Wu, M.T.; Rezaul, K.; Blaho, V.; Hwang, S.I.; Han, D.K.; Hla, T. Engagement of S1p<sub>1</sub>-Degradative Mechanisms Leads to Vascular Leak in Mice. *Journal of Clinical Investigation*. **2011**, *121*, 2290-2300.
- [24] Kohout, T.A.; Lefkowitz, R.J. Regulation of G Protein-Coupled Receptor Kinases and Arrestins During Receptor Desensitization. *Molecular Pharmacology*. **2003**, *63*, 9-18.

- [25] Mullershausen, F.; Zecri, F.; Cetin, C.; Billich, A.; Guerini, D.; Seuwen, K. Persistent Signaling Induced by Fty720-Phosphate Is Mediated by Internalized S1p1 Receptors. *Nature Chemical Biology*. **2009**, *5*, 428-434.
- [26] Oo, M.L.; Thangada, S.; Wu, M.T.; Liu, C.H.; Macdonald, T.L.; Lynch, K.R.; Lin, C.Y.; Hla, T. Immunosuppressive and Anti-Angiogenic Sphingosine 1-Phosphate Receptor-1 Agonists Induce Ubiquitinylation and Proteasomal Degradation of the Receptor. *Journal of Biological Chemistry*. **2007**, *282*, 9082-9089.
- [27] Graler, M.H.; Goetzl, E.J. The Immunosuppressant Fty720 Down-Regulates Sphingosine 1-Phosphate G Protein-Coupled Receptors. *FASEB Journal*. **2004**, *18*, 551-553.
- [28] Jongsma, M.; van Unen, J.; van Loenen, P.B.; Michel, M.C.; Peters, S.L.; Alewijnse, A.E. Different Response Patterns of Several Ligands at the Sphingosine-1-Phosphate Receptor Subtype 3 (S1p<sub>3</sub>). *British Journal of Pharmacology*. **2009**, *156*, 1305-1311.
- [29] Young, N.; Van Brocklyn, J.R. Signal Transduction of Sphingosine-1-Phosphate G Protein-Coupled Receptors. *Scientific World Journal*. **2006**, *6*, 946-966.
- [30] Ishii, I.; Fukushima, N.; Ye, X.; Chun, J. Lysophospholipid Receptors: Signaling and Biology. *Annual Review of Biochemistry*. **2004**, *73*, 321-354.
- [31] Adachi, K.; Kohara, T.; Nakao, N.; Arita, M.; Chiba, K.; Mishina, T.; Sasaki, S.; Fujita, T. Design, Synthesis, and Structure-Activity-Relationships of 2-Substituted-2-Amino-1,3-Propanediols: Discovery of a Novel Immunosuppressant, Fty720. *Bioorganic & Medicinal Chemistry Letters*. **1995**, *5*, 853-856.
- [32] Brinkmann, V.; Billich, A.; Baumruker, T.; Heining, P.; Schmouder, R.; Francis, G.; Aradhye, S.; Burtin, P. Fingolimod (Fty720): Discovery and Development of an Oral Drug to Treat Multiple Sclerosis. *Nature Reviews Drug Discovery*. **2010**, *9*, 883-897.
- [33] Paugh, S.W.; Payne, S.G.; Barbour, S.E.; Milstien, S.; Spiegel, S. The Immunosuppressant Fty720 Is Phosphorylated by Sphingosine Kinase Type 2. *FEBS Journal*. **2003**, *554*, 189-193.
- [34] Kharel, Y.; Lee, S.; Snyder, A.H.; Sheasley-O'Neill, S.L.; Morris, M.A.; Setiady, Y.; Zhu, R.; Zigler, M.A.; Burcin, T.L.; Ley, K.; Tung, K.S.K.; Engelhard, V.H.; Macdonald, T.L.; Pearson-White, S.; Lynch, K.R. Sphingosine Kinase 2 Is Required for Modulation of Lymphocyte Traffic by Fty720. *Journal of Biological Chemistry*. **2005**, *280*, 36865-36872.
- [35] Mandala, S.; Hajdu, R.; Bergstrom, J.; Quackenbush, E.; Xie, J.; Milligan, J.; Thornton, R.; Shei, G.J.; Card, D.; Keohane, C.; Rosenbach, M.; Hale, J.; Lynch, C.L.; Rupprecht, K.; Parsons, W.; Rosen, H. Alteration of Lymphocyte Trafficking by Sphingosine-1-Phosphate Receptor Agonists. *Science*. **2002**, *296*, 346-349.
- [36] Clemens, J.J.; Davis, M.D.; Lynch, K.R.; Macdonald, T.L. Synthesis of Para-Alkyl Aryl Amide Analogues of Sphingosine-1-Phosphate: Discovery of Potent S1p Receptor Agonists. *Bioorganic & Medicinal Chemistry Letters*. **2003**, *13*, 3401-3404.
- [37] Davis, M.D.; Clemens, J.J.; Macdonald, T.L.; Lynch, K.R. Sphingosine 1-Phosphate Analogs as Receptor Antagonists. *Journal of Biological Chemistry*. **2005**, *280*, 9833-9841.

- [38] Foss, F.W., Jr.; Clemens, J.J.; Davis, M.D.; Snyder, A.H.; Zigler, M.A.; Lynch, K.R.; Macdonald, T.L. Synthesis, Stability, and Implications of Phosphothioate Agonists of Sphingosine-1-Phosphate Receptors. *Bioorganic & Medicinal Chemistry Letters*. **2005**, *15*, 4470-4474.
- [39] Foss, F.W., Jr.; Snyder, A.H.; Davis, M.D.; Rouse, M.; Okusa, M.D.; Lynch, K.R.; Macdonald, T.L. Synthesis and Biological Evaluation of G-Aminophosphonates as Potent, Subtype-Selective Sphingosine 1-Phosphate Receptor Agonists and Antagonists. *Bioorganic & Medicinal Chemistry*. **2007**, *15*, 663-677.
- [40] Aguilar, N.; Mir, M.; Grima, P.M.; Lopez, M.; Segarra, V.; Esteban, L.; Moreno, I.; Godessart, N.; Tarrason, G.; Domenech, T.; Vilella, D.; Armengol, C.; Cordoba, M.; Sabate, M.; Casals, D.; Dominguez, M. Discovery of a Novel Class of Zwitterionic, Potent, Selective and Orally Active S1p<sub>1</sub> Direct Agonists. *Bioorganic & Medicinal Chemistry Letters*. **2012**, *22*, 7672-7676.
- [41] Asano, M.; Nakamura, T.; Sekiguchi, Y.; Mizuno, Y.; Yamaguchi, T.; Tamaki, K.; Shimozato, T.; Doi-Komuro, H.; Kagari, T.; Tomisato, W.; Inoue, R.; Yuita, H.; Oguchi-Oshima, K.; Kaneko, R.; Nara, F.; Kawase, Y.; Masubuchi, N.; Nakayama, S.; Koga, T.; Namba, E.; Nasu, H.; Nishi, T. Synthesis and Sar of 1,3-Thiazolyl Thiophene and Pyridine Derivatives as Potent, Orally Active and S1p<sub>3</sub>-Sparing S1p<sub>1</sub> Agonists. *Bioorganic & Medicinal Chemistry Letters*. **2012**, *22*, 3083-3088.
- [42] Hale, J.J.; Lynch, C.L.; Neway, W.; Mills, S.G.; Hajdu, R.; Keohane, C.A.; Rosenbach, M.J.; Milligan, J.A.; Shei, G.J.; Parent, S.A.; Chrebet, G.; Bergstrom, J.; Card, D.; Ferrer, M.; Hodder, P.; Strulovici, B.; Rosen, H.; Mandala, S. A Rational Utilization of High-Throughput Screening Affords Selective, Orally Bioavailable 1-Benzyl-3-Carboxyazetidine Sphingosine-1-Phosphate-1 Receptor Agonists. *Journal of Medicinal Chemistry*. **2004**, *47*, 6662-6665.
- [43] Li, Z.; Chen, W.; Hale, J.J.; Lynch, C.L.; Mills, S.G.; Hajdu, R.; Keohane, C.A.; Rosenbach, M.J.; Milligan, J.A.; Shei, G.J.; Chrebet, G.; Parent, S.A.; Bergstrom, J.; Card, D.; Forrest, M.; Quackenbush, E.J.; Wickham, L.A.; Vargas, H.; Evans, R.M.; Rosen, H.; Mandala, S. Discovery of Potent 3,5-Diphenyl-1,2,4-Oxadiazole Sphingosine-1-Phosphate-1 (S1p<sub>1</sub>) Receptor Agonists with Exceptional Selectivity against S1p<sub>2</sub> and S1p<sub>3</sub>. *Journal of Medicinal Chemistry*. **2005**, *48*, 6169-6173.
- [44] Dyckman, A.J. Recent Advances in the Discovery and Development of Sphingosine-1-Phosphate-1 Receptor Agonists. *Annual Reports in Medicinal Chemistry*. **2012**, *47*, 195-207.
- [45] Parrill, A.L.; Tigyi, G. Integrating the Puzzle Pieces: The Current Atomistic Picture of Phospholipid-G Protein Coupled Receptor Interactions. *Biochimica et Biophysica Acta*. **2013**, *1831*, 2-12.
- [46] Archbold, J.K.; Martin, J.L.; Sweet, M.J. Towards Selective Lysophospholipid GPCR Modulators. *Trends in Pharmacological Sciences*. **2014**, *35*, 219-226.
- [47] Rivera, J.; Proia, R.L.; Olivera, A. The Alliance of Sphingosine-1-Phosphate and Its Receptors in Immunity. *Nature Reviews Immunology*. **2008**, *8*, 753-763.
- [48] LeBien, T.W.; Tedder, T.F. B Lymphocytes: How They Develop and Function. *Blood*. **2008**, *112*, 1570-1580.

- [49] Reinherz, E.L.; Schlossman, S.F. The Differentiation and Function of Human T Lymphocytes. *Cell*. **1980**, *19*, 821-827.
- [50] Starr, T.K.; Jameson, S.C.; Hogquist, K.A. Positive and Negative Selection of T Cells. *Annual Review of Immunology*. **2003**, *21*, 139-176.
- [51] Rajewsky, K. Clonal Selection and Learning in the Antibody System. *Nature*. **1996**, *381*, 751-758.
- [52] Springer, T.A. Traffic Signals for Lymphocyte Recirculation and Leukocyte Emigration - the Multistep Paradigm. *Cell*. **1994**, *76*, 301-314.
- [53] Guermonprez, P.; Valladeau, J.; Zitvogel, L.; Thery, C.; Amigorena, S. Antigen Presentation and T Cell Stimulation by Dendritic Cells. *Annual Review of Immunology*. **2002**, *20*, 621-667.
- [54] Venkataraman, K.; Thangada, S.; Michaud, J.; Oo, M.L.; Ai, Y.; Lee, Y.M.; Wu, M.; Parikh, N.S.; Khan, F.; Proia, R.L.; Hla, T. Extracellular Export of Sphingosine Kinase-1a Contributes to the Vascular S1p Gradient. *Biochemical Journal*. **2006**, *397*, 461-471.
- [55] Hla, T.; Venkataraman, K.; Michaud, J. The Vascular S1p Gradient-Cellular Sources and Biological Significance. *Biochimica et Biophysica Acta*. **2008**, *1781*, 477-482.
- [56] Matloubian, M.; Lo, C.G.; Cinamon, G.; Lesneski, M.J.; Xu, Y.; Brinkmann, V.; Allende, M.L.; Proia, R.L.; Cyster, J.G. Lymphocyte Egress from Thymus and Peripheral Lymphoid Organs Is Dependent on S1p Receptor 1. *Nature*. **2004**, *427*, 355-360.
- [57] [Http://www.nationalmssociety.org/](http://www.nationalmssociety.org/).
- [58] Lublin, F.D.; Reingold, S.C. Defining the Clinical Course of Multiple Sclerosis: Results of an International Survey. *Neurology*. **1996**, *46*, 907-911.
- [59] Disanto, G.; Morahan, J.M.; Barnett, M.H.; Giovannoni, G.; Ramagopalan, S.V. The Evidence for a Role of B Cells in Multiple Sclerosis. *Neurology*. **2012**, *78*, 823-832.
- [60] Ballabh, P.; Braun, A.; Nedergaard, M. The Blood-Brain Barrier: An Overview: Structure, Regulation, and Clinical Implications. *Neurobiology of Disease*. **2004**, *16*, 1-13.
- [61] Larochelle, C.; Alvarez, J.I.; Prat, A. How Do Immune Cells Overcome the Blood-Brain Barrier in Multiple Sclerosis? *FEBS Letters*. **2011**, *585*, 3770-3780.
- [62] Bruck, W. The Pathology of Multiple Sclerosis Is the Result of Focal Inflammatory Demyelination with Axonal Damage. *Journal of Neurology*. **2005**, *252*, v3-v9.
- [63] Agius, M.; Meng, X.Y.; Chin, P.; Grinspan, A.; Hashmonay, R. Fingolimod Therapy in Early Multiple Sclerosis: An Efficacy Analysis of the Transforms and Freedoms Studies by Time since First Symptom. *CNS Neuroscience & Therapeutics*. **2014**, *20*, 446-451.

- [64] Ruggieri, S.; Tortorella, C.; Gasperini, C. Pharmacology and Clinical Efficacy of Dimethyl Fumarate (Bg-12) for Treatment of Relapsing-Remitting Multiple Sclerosis. *Therapeutics and Clinical Risk Management*. **2014**, *10*, 229-239.
- [65] Chiba, K.; Kataoka, H.; Seki, N.; Maeda, Y.; Sugahara, K. Fingolimod (Fty720), the Sphingosine 1-Phosphate Receptor Modulator, as a New Therapeutic Drug in Multiple Sclerosis. *Inflammation and Regeneration*. **2011**, *31*, 167-174.
- [66] Goverman, J. Autoimmune T Cell Responses in the Central Nervous System. *Nature Reviews Immunology*. **2009**, *9*, 393-407.
- [67] Garris, C.S.; Wu, L.F.; Acharya, S.; Arac, A.; Blaho, V.A.; Huang, Y.X.; Moon, B.S.; Axtell, R.C.; Ho, P.P.; Steinberg, G.K.; Lewis, D.B.; Sobel, R.A.; Han, D.K.; Steinman, L.; Snyder, M.P.; Hla, T.; Han, M.H. Defective Sphingosine 1-Phosphate Receptor 1 (S1p<sub>1</sub>) Phosphorylation Exacerbates T<sub>H</sub>17-Mediated Autoimmune Neuroinflammation. *Nature Immunology*. **2013**, *14*, 1166-1172.
- [68] McVerry, B.J.; Garcia, J.G. In Vitro and in Vivo Modulation of Vascular Barrier Integrity by Sphingosine 1-Phosphate: Mechanistic Insights. *Cellular Signalling*. **2005**, *17*, 131-139.
- [69] Quancard, J.; Bollbuck, B.; Janser, P.; Angst, D.; Berst, F.; Buehlmayer, P.; Streiff, M.; Beerli, C.; Brinkmann, V.; Guerini, D.; Smith, P.A.; Seabrook, T.J.; Traebert, M.; Seuwen, K.; Hersperger, R.; Bruns, C.; Bassilana, F.; Bigaud, M. A Potent and Selective S1p<sub>1</sub> Antagonist with Efficacy in Experimental Autoimmune Encephalomyelitis. *Chemistry & Biology*. **2012**, *19*, 1142-1151.
- [70] Ontaneda, D.; Cohen, J.A. Potential Mechanisms of Efficacy and Adverse Effects in the Use of Fingolimod (Fty720). *Expert Review of Clinical Pharmacology*. **2011**, *4*, 567-570.
- [71] Koyrakh, L.; Roman, M.I.; Brinkmann, V.; Wickman, K. The Heart Rate Decrease Caused by Acute Fty720 Administration Is Mediated by the G Protein-Gated Potassium Channel I<sub>kach</sub>. *American Journal of Transplantation*. **2005**, *5*, 529-536.
- [72] [Http://Www.Pharma.Us.Novartis.Com/Product/Pi/Pdf/Gilenya.Pdf](http://www.pharma.us.novartis.com/product/pi/pdf/gilenya.pdf).
- [73] Legangneux, E.; Gardin, A.; Johns, D. Dose Titration of Baf312 Attenuates the Initial Heart Rate Reducing Effect in Healthy Subjects. *British Journal of Clinical Pharmacology*. **2013**, *75*, 831-841.
- [74] Forrest, M.; Sun, S.Y.; Hajdu, R.; Bergstrom, J.; Card, D.; Doherty, G.; Hale, J.; Keohane, C.; Meyers, C.; Milligan, J.; Mills, S.; Nomura, N.; Rosen, H.; Rosenbach, M.; Shei, G.J.; Singer, I.I.; Tian, M.; West, S.; White, V.; Xie, J.; Proia, R.L.; Mandala, S. Immune Cell Regulation and Cardiovascular Effects of Sphingosine 1-Phosphate Receptor Agonists in Rodents Are Mediated Via Distinct Receptor Subtypes. *Journal of Pharmacology and Experimental Therapeutics*. **2004**, *309*, 758-768.
- [75] Gergely, P.; Nuesslein-Hildesheim, B.; Guerini, D.; Brinkmann, V.; Traebert, M.; Bruns, C.; Pan, S.; Gray, N.S.; Hinterding, K.; Cooke, N.G.; Groenewegen, A.; Vitaliti, A.; Sing, T.; Luttringer, O.; Yang, J.; Gardin, A.; Wang, N.; Crumb, W.J.; Saltzman, M.; Rosenberg, M.; Wallstrom, E. The Selective Sphingosine 1-Phosphate Receptor Modulator Baf312 Redirects Lymphocyte Distribution and Has Species-Specific Effects on Heart Rate. *Br J Pharmacol*. **2012**, *167*, 1035-1047.

[76] Moberly, J.B.; Ford, D.M.; Zahir, H.; Chen, S.Q.; Mochizuki, T.; Truitt, K.E.; Vollmer, T.L. Pharmacological Effects of Cs-0777, a Selective Sphingosine 1-Phosphate Receptor-1 Modulator: Results from a 12-Week, Open-Label Pilot Study in Multiple Sclerosis Patients. *Journal of Neuroimmunology*. **2012**, *246*, 100-107.

[77] Jaillard, C.; Harrison, S.; Stankoff, B.; Aigrot, M.S.; Calver, A.R.; Duddy, G.; Walsh, F.S.; Pangalos, M.N.; Arimura, N.; Kaibuchi, K.; Zalc, B.; Lubetzki, C. Edg8/S1p5: An Oligodendroglial Receptor with Dual Function on Process Retraction and Cell Survival. *Journal of Neuroscience*. **2005**, *25*, 1459-1469.

[78] van Doorn, R.; Lopes Pinheiro, M.A.; Kooij, G.; Lakeman, K.; van het Hof, B.; van der Pol, S.M.; Geerts, D.; van Horssen, J.; van der Valk, P.; van der Kam, E.; Ronken, E.; Reijerkerk, A.; de Vries, H.E. Sphingosine 1-Phosphate Receptor 5 Mediates the Immune Quiescence of the Human Brain Endothelial Barrier. *Journal of Neuroinflammation*. **2012**, *9*, 133-147.

[79] Selmaj, K.; Li, D.K.; Hartung, H.P.; Hemmer, B.; Kappos, L.; Freedman, M.S.; Stuve, O.; Rieckmann, P.; Montalban, X.; Ziemssen, T.; Auberson, L.Z.; Pohlmann, H.; Mercier, F.; Dahlke, F.; Wallstrom, E. Siponimod for Patients with Relapsing-Remitting Multiple Sclerosis (Bold): An Adaptive, Dose-Ranging, Randomised, Phase 2 Study. *Lancet Neurology*. **2013**, *12*, 756-767.

[80] Pan, S.; Gray, N.S.; Gao, W.; Mi, Y.; Fan, Y.; Wang, X.; Tuntland, T.; Che, J.; Lefebvre, S.; Chen, Y.; Chu, A.; Hinterding, K.; Gardin, A.; End, P.; Heining, P.; Bruns, C.; Cooke, N.G.; Nuesslein-Hildesheim, B. Discovery of Baf312 (Siponimod), a Potent and Selective S1p Receptor Modulator. *ACS Medicinal Chemistry Letters*. **2013**, *4*, 333-337.

[81] Violin, J.D.; Lefkowitz, R.J. Beta-Arrestin-Biased Ligands at Seven-Transmembrane Receptors. *Trends in Pharmacological Sciences*. **2007**, *28*, 416-422.

[82] David, O.J.; Kovarik, J.M.; Schmouder, R.L. Clinical Pharmacokinetics of Fingolimod. *Clinical Pharmacokinetics*. **2012**, *51*, 15-28.

[83] Evindar, G.; Deng, H.F.; Bernier, S.G.; Doyle, E.; Lorusso, J.; Morgan, B.A.; Westlin, W.F. Exploring Amino Acids Derivatives as Potent, Selective, and Direct Agonists of Sphingosine-1-Phosphate Receptor Subtype-1. *Bioorganic & Medicinal Chemistry Letters*. **2013**, *23*, 472-475.

[84] Pan, S.F.; Mi, Y.A.; Pally, C.; Beerli, C.; Chen, A.; Guerini, D.; Hinterding, K.; Nuesslein-Hildesheim, B.; Tuntland, T.; Lefebvre, S.; Liu, Y.; Gao, W.Q.; Chu, A.A.; Brinkmann, V.; Bruns, C.; Streiff, M.; Cannet, C.; Cooke, N.; Gray, N. A Monoselective Sphingosine-1-Phosphate Receptor-1 Agonist Prevents Allograft Rejection in a Stringent Rat Heart Transplantation Model. *Chemistry & Biology*. **2006**, *13*, 1227-1234.

[85] Fazekas, F.; Berger, T.; Fabjan, T.H.; Ledinek, A.H.; Jakab, G.; Komoly, S.; Kraus, J.; Kurca, E.; Kyriakides, T.; Lisy, L.; Milanov, I.; Panayiotou, P.; Jazbec, S.S.; Talab, R.; Traykov, L.; Turcani, P.; Vass, K.; Vella, N.; Havrdova, E. Fingolimod in the Treatment Algorithm of Relapsing Remitting Multiple Sclerosis: A Statement of the Central and East European (Cee) Ms Expert Group. *Wiener Medizinische Wochenschrift*. **2012**, *162*, 354-366.

[86] Eaton, T.A.; Houk, K.N.; Watkins, S.F.; Fronczek, F.R. Geometries and Conformational Processes in Phencyclidine and a Rigid Adamantyl Analog - Variable-Temperature Nmr, X-Ray Crystallographic, and Molecular Mechanics Studies. *Journal of Medicinal Chemistry*. **1983**, *26*, 479-486.

- [87] Harrold, M.W. The Influence of Conformational Isomerism on Drug Action and Design. *American Journal of Pharmaceutical Education*. **1996**, *60*, 192-197.
- [88] Jary, E.; Bee, T.; Walker, S.R.; Chung, S.K.; Seo, K.C.; Morris, J.C.; Don, A.S. Elimination of a Hydroxyl Group in Fty720 Dramatically Improves the Phosphorylation Rate. *Molecular Pharmacology*. **2010**, *78*, 685-692.
- [89] Huang, T. Synthesis of Conformation-Restricted Analogs of Fty720 as Sphingosine 1-Phosphate Receptor Pro-Drugs. *Dissertation - University of Virginia*. **2012**.
- [90] Zhu, R. Synthesis and Biological Evaluation of Conformationally Constrained Sphingosine 1-Phosphate Analogs as Subtype Selective S1p Receptor Agonist/Antagonists. *Dissertation - University of Virginia*. **2008**.
- [91] Zhu, R.; Snyder, A.H.; Kharel, Y.; Schaffter, L.; Sun, Q.; Kennedy, P.C.; Lynch, K.R.; Macdonald, T.L. Asymmetric Synthesis of Conformationally Constrained Fingolimod Analogues - Discovery of an Orally Active Sphingosine 1-Phosphate Receptor Type-1 Agonist and Receptor Type-3 Antagonist. *Journal of Medicinal Chemistry*. **2007**, *50*, 6428-6435.
- [92] Bin, M.; Guckian, K.M.; Lin, E.Y.S.; Lee, W.C.; Scott, D.; Kumaravel, G.; Macdonald, T.L.; Lynch, K.R.; Black, C.; Chollate, S.; Hahm, K.; Hetu, G.; Jin, P.; Luo, Y.; Rohde, E.; Rossomando, A.; Scannevin, R.; Wang, J.; Yang, C.H. Stereochemistry-Activity Relationship of Orally Active Tetralin S1p Agonist Prodrugs. *Bioorganic & Medicinal Chemistry Letters*. **2010**, *20*, 2264-2269.
- [93] Yang, S.Y. Pharmacophore Modeling and Applications in Drug Discovery: Challenges and Recent Advances. *Drug Discovery Today*. **2010**, *15*, 444-450.
- [94] Jorgensen, W.L. The Many Roles of Computation in Drug Discovery. *Science*. **2004**, *303*, 1813-1818.
- [95] Katritch, V.; Cherezov, V.; Stevens, R.C. Structure-Function of the G Protein-Coupled Receptor Superfamily. *Annual Review of Pharmacology and Toxicology*. **2013**, *53*, 531-556.
- [96] Kennedy, P.C.; Zhu, R.; Huang, T.; Tomsig, J.L.; Mathews, T.P.; David, M.; Peyruchaud, O.; Macdonald, T.L.; Lynch, K.R. Characterization of a Sphingosine 1-Phosphate Receptor Antagonist Prodrug. *Journal of Pharmacology and Experimental Therapeutics*. **2011**, *338*, 879-889.
- [97] Evindar, G.; Bernier, S.G.; Kavarana, M.J.; Doyle, E.; Lorusso, J.; Kelley, M.S.; Halley, K.; Hutchings, A.; Wright, A.D.; Saha, A.K.; Hannig, G.; Morgan, B.A.; Westlin, W.F. Synthesis and Evaluation of Alkoxy-Phenylamides and Alkoxy-Phenylimidazoles as Potent Sphingosine-1-Phosphate Receptor Subtype-1 Agonists. *Bioorganic & Medicinal Chemistry Letters*. **2009**, *19*, 369-372.
- [98] Evindar, G.; Satz, A.L.; Bernier, S.G.; Kavarana, M.J.; Doyle, E.; Lorusso, J.; Taghizadeh, N.; Halley, K.; Hutchings, A.; Kelley, M.S.; Wright, A.D.; Saha, A.K.; Hannig, G.; Morgan, B.A.; Westlin, W.F. Synthesis and Evaluation of Arylalkoxy- and Biarylalkoxy-Phenylamide and Phenylimidazoles as Potent and Selective Sphingosine-1-Phosphate Receptor Subtype-1 Agonists. *Bioorganic & Medicinal Chemistry Letters*. **2009**, *19*, 2315-2319.



[99] Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force-Field for Molecular Mechanical Simulation of Nucleic-Acids and Proteins. *Journal of the American Chemical Society*. **1984**, *106*, 765-784.

[100] Corbeil, C.R.; Williams, C.I.; Labute, P. Variability in Docking Success Rates Due to Dataset Preparation. *Journal of Computer Aided Molecular Design*. **2012**, *26*, 775-786.

[101] Portaccio, E. Evidence-Based Assessment of Potential Use of Fingolimod in Treatment of Relapsing Multiple Sclerosis. *Core Evidence*. **2011**, *6*, 13-21.

[102] 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 Science*. **2014**, *39*, 233-244.

[103] Yuan, S.; Wu, R.; Latek, D.; Trzaskowski, B.; Filipek, S. Lipid Receptor S1p<sub>1</sub> Activation Scheme Concluded from Microsecond All-Atom Molecular Dynamics Simulations. *PLoS Computational Biology*. **2013**, *9*, 1-8.

[104] Obinata, H.; Hla, T. Fine-Tuning S1p Therapeutics. *Chemistry & Biology*. **2012**, *19*, 1080-1082.

[105] Molecular Operating Environment (MOE), C.C.G.I., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.

[106] Halgren, T.A. Merck Molecular Force Field . I. Basis, Form, Scope, Parameterization, and Performance of Mmff94. *Journal of Computational Chemistry*. **1996**, *17*, 490-519.