Structure Based Ligand Design for Monoamine Transporters and Mitogen Activated Kinase 5

Sankar Manepalli

Follow this and additional works at: https://dsc.duq.edu/etd

Recommended Citation
STRUCTURE BASED LIGAND DESIGN FOR MONOAMINE TRANSPORTERS
AND MITOGEN ACTIVATED KINASE 5

A Dissertation
Submitted to the Bayer School of Natural and Environmental Sciences

Duquesne University
In partial fulfillment of the requirements for
the degree of Doctor of Philosophy

By
Sankar Manepalli

May 2012
STRUCTURE BASED LIGAND DESIGN FOR MONOAMINE TRANSPORTERS
AND MITOGEN ACTIVATED KINASE 5

By

Sankar Manepalli

Approved January 31, 2012

Jeffry D. Madura, Ph.D.
Professor of Chemistry and Biochemistry
Duquesne University
(Committee Chair)

Jeffrey D. Evanseck, Ph.D.
Professor of Chemistry and Biochemistry
Duquesne University
(Committee Member)

Rita Mihai, Ph.D.
Associate professor of Chemistry and Biochemistry
Duquesne University
(Committee Member)

Christopher K. Surratt, Ph.D.
Professor of Pharmacology
Division Head, Pharmaceutical Sciences
Mylan School of Pharmacy
Duquesne University
(External Reviewer)

David W. Seybert, Ph.D.
Dean, Bayer School of Natural and Environmental Sciences
Professor of Chemistry and Biochemistry

Ralph Wheeler, Ph.D.
Chair, Department of Chemistry and Biochemistry
Professor of Chemistry and Biochemistry
ABSTRACT

STRUCTURE BASED LIGAND DESIGN FOR MONOAMINE TRANSPORTERS
AND MITOGEN ACTIVATED KINASE 5

By
Sankar Manepalli
May 2012

Dissertation supervised by Jeffry D. Madura

Depression is a major psychological disorder that affects a person's mental and physical abilities. The National Institute of Mental Health (NIMH) classified it as a serious medical illness. It causes huge economic, as well as financial impact on the people, and it is also becoming a major public health issue. Antidepressant drugs are prescribed to mitigate the suffering caused by this disorder. Different generations of antidepressants have been developed with dissimilar mechanisms of action. According to the Center for Disease Control, the usage of antidepressants has skyrocketed by 400 percent increase over 2005-2008 survey period. This dramatic rise in usage indicates that these are the most prescribed drugs in the US. Even with the FDA mandated “black box” warning of increased suicidal thoughts upon use of selected antidepressants, these drugs are still being used at a higher rate.
All classes of antidepressants are plagued by side effects with mainly sexual dysfunction common among them. To avoid the adverse effects, an emphasis is to discover novel structural drug scaffolds that can be further developed as a new generation of antidepressants. The importance of this research is to discover structurally novel antidepressants by performing in silico virtual screening (VS) of chemical databases using the serotonin transporter (SERT). In the absence of a SERT crystal structure, a homology model was developed. The homology model was utilized to develop the first structure-based pharmacophore for the extracellular facing secondary ligand binding pocket. The pharmacophore captured the necessary drug-SERT interaction pattern for SERT inhibitory action. This pharmacophore was employed as one of the filters for VS of candidate ligands. The ten compounds identified were purchased and tested pharmacologically. Out of the ten hits, three structurally novel ligands were identified as lead compounds. Two of these compounds exhibited selectivity towards SERT; the remaining lead compound was selective towards the dopamine transporter and displayed cocaine inhibition. The two SERT selective compounds will provide new opportunities in the development of novel therapeutics to treat depression.

For dopamine transporter (DAT), the study was based on recently developed structurally diverse photo probes. In an effort to better understand the binding profile similarities among these different scaffolds, the photo probes were docked into DAT. The finger print analysis of the interaction pattern of docked poses was performed to identify the inhibitor-binding sites.

For mitogen activated protein kinase 5 (MEK5), given the lack of structural information, a homology model of MEK5 was developed to guide the rational design of
inhibitors. Docking of known MEK5 inhibitors into the homology model was performed to understand the inhibitory interaction profile. Several series of analogues were designed utilizing the generated interaction profile.
DEDICATION

To my family who gave me strength and courage.
ACKNOWLEDGEMENT

First and foremost I would like to thank my advisor, Dr. Jeffry D. Madura, for his constant support and suggestions. He has been an inspiration and I have learned so much more than I expected. Even though sometimes I tested his patience, he was calm and really pushed me towards success. I would also like to thank my committee members Dr. Jeffrey D. Evanseck and Dr. Rita Mihaiescu who have served on my defenses and constantly guided me through my graduate school. I would like to specially thank Dr. Christopher K. Surratt (external reviewer). I am extremely lucky to have him as a pharmacy school colleague who was available for valuable insight and guidance on a professional as well as personal level.

I would like to especially thank Dr. Tammy L. Nolan, postdoctoral research associate in our lab, for our long discussions. I would like to thank Judy Liu, Yurong Huang and especially Laura M. Geffert for the pharmacological testing of serotonin and dopamine transporter inhibitors. I would also like to thank Dr. David Lapinsky and his group members for providing the dopamine transporter inhibitors. I would like to thank Dr. Patrick Flaherty and his group members for providing the mitogen activated kinase 5 inhibitors. I would also like to thank Darlene Monish from Dr. Jane Cavanaugh’s lab for testing kinase inhibitors.

I would like to thank my fellow graduate students on the floor for making my stay memorable. Special thanks to Dr. Josh Plumley and people in Dr. Evanseck’s lab for helping me get accustomed to the new country and its traditions. I would also like to thank Dr. Emilio Esposito for his suggestions throughout my graduate studies. I would also like to especially mention my colleague Kalyan Immadisetty who made me feel at home and
thanks to Sai Pakkala, Sumangala Shetty and Dinesh Nath for our discussions during lunch hours.

There are several people who made my life easier at Duquesne to whom I owe a great deal of gratitude. I would like to especially mention about Sandy Russell and Amy Stroyne who were helpful in many ways. I would also like to thank Sue Patil for her help on a personal as well as professional level.

I am indebted to my wife for her patience and unconditional support.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Dedication</td>
<td>vii</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xiv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xvii</td>
</tr>
<tr>
<td><strong>CHAPTER 1</strong></td>
<td></td>
</tr>
<tr>
<td>1.1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Generation of MAT Homology Models</td>
<td>5</td>
</tr>
<tr>
<td>1.2.1. Pre LeuT</td>
<td>5</td>
</tr>
<tr>
<td>1.2.2. Post LeuT (2005-Present)</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2.1. Dopamine Transporter (DAT)</td>
<td>7</td>
</tr>
<tr>
<td>1.2.2.2. Serotonin Transporter (SERT)</td>
<td>13</td>
</tr>
<tr>
<td>1.2.2.3. Norepinephrine Transporter (NET)</td>
<td>18</td>
</tr>
<tr>
<td>1.3. Molecular Dynamics (MD) on MATs</td>
<td>21</td>
</tr>
<tr>
<td>1.3.1. DAT Dynamics</td>
<td>21</td>
</tr>
<tr>
<td>1.3.2. SERT Dynamics</td>
<td>23</td>
</tr>
<tr>
<td>1.4. Ion (Na⁺, Cl⁻) Binding Sites</td>
<td>24</td>
</tr>
<tr>
<td>1.5. Virtual Screening</td>
<td>28</td>
</tr>
<tr>
<td>1.5.1. Ligand-Based VS</td>
<td>28</td>
</tr>
<tr>
<td>1.5.2. Structure-Based VS</td>
<td>30</td>
</tr>
</tbody>
</table>
1.6. Conclusion ................................................................................................................. 34

CHAPTER 2

2.1. Introduction .................................................................................................................. 44

2.2. Materials and Methods ................................................................................................ 47

2.2.1. Computational Methods ......................................................................................... 48

2.2.1.1. Sequence Retrieval and Integrity Check of Template .......................................... 48

2.2.1.2. Different Sequence Alignments ............................................................................ 49

2.2.1.3. Construction of SERT Homology Models ........................................................... 51

2.2.1.3.1. hSERT Model 1: Manepalli et al. alignment ...................................................... 51

2.2.1.3.2. hSERT Model 2: Beuming et al. alignment ....................................................... 53

2.2.1.3.3. hSERT Model 3: Yamashita et al. alignment ..................................................... 55

2.2.1.3.4. hSERT Model 4: Celik et al. alignment ............................................................ 56

2.2.1.4. Model Refinement and Validation ......................................................................... 57

2.2.1.5. Ligand Binding Site identification .......................................................................... 59

2.2.1.6. Docking of Known Antidepressants ...................................................................... 61

2.2.1.7. Development of 4-point SERT S2 Pharmacophore .............................................. 63

2.2.1.8. Virtual Screening of ZINC Database .................................................................... 64

2.2.2. Experimental Methods ............................................................................................ 67

2.2.2.1. In vitro VS Compound Binding Screen ................................................................. 67

2.2.2.2. hSERT Membrane Binding Assay ......................................................................... 67

2.2.2.3. hDAT and hNET Whole Cell Binding Assays ......................................................... 68

2.2.2.4. Serotonin Uptake Inhibition Screen ..................................................................... 69

2.3. Results and Discussion ............................................................................................... 70
2.3.1. Model Construction and Refinement .......................................................... 70
2.3.2. Docking and Pharmacophore Feature Generation ........................................ 72
2.3.3. Development of 4-point SERT S2 Pharmacophore ....................................... 74
2.3.4. Virtual Screening and Pharmacological Testing ........................................... 76
2.4. Structural Novelty of SM Hit Compounds ....................................................... 85
2.5. Conclusion ........................................................................................................ 85
2.6. Future Directions ............................................................................................. 86

CHAPTER 3
3.1. Introduction ....................................................................................................... 96
3.2. Methods ............................................................................................................. 100
3.3. Results and Discussion ................................................................................... 101
3.4. Conclusion ........................................................................................................ 108

CHAPTER 4
4.1. Introduction ....................................................................................................... 114
4.2. Methods ............................................................................................................. 117
  4.2.1. Retrieval of Sequence .................................................................................. 117
  4.2.2. Search for the Template ............................................................................. 120
  4.2.3. Protein Check of the Template .................................................................. 122
  4.2.4. Sequence Alignment .................................................................................. 124
  4.2.5. MEK5 Homology Model Construction ....................................................... 125
4.2.6. Docking ......................................................................................................... 125
4.3. Results and Discussion ................................................................................... 126
  4.3.1. Homology Modeling ................................................................................... 126

xii
4.3.2. Docking ............................................................................................................. 132

4.4. Conclusion ............................................................................................................. 138
LIST OF TABLES

1.1. Construction of DAT homology models in chronological order ....................... 12
1.2. Construction of SERT homology models in chronological order ...................... 17
1.3. Construction of NET homology models in chronological order ....................... 20
1.4. Novel MAT ligands discovered through MAT virtual screening ....................... 33
2.1. Percent sequence identity and sequence similarity among different alignments
    used to build SERT models .................................................................................... 50
2.2. List of top ten compounds (binders) identified in the evaluation of pharmacophore ....
    .......................................................................................................................... 63
4.1. Various biphenyl analogues of PD0325901 synthesized and tested for ERK1/2 and
    ERK 5 inhibition of phosphorylation ..................................................................... 125
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.</td>
<td>MATs embedded in lipid bilayer</td>
<td>3</td>
</tr>
<tr>
<td>1.2.</td>
<td>The three minimally required conformations of substrate translocation cycle</td>
<td>4</td>
</tr>
<tr>
<td>1.3.</td>
<td>Ligand-accessible regions of MATs (SERT)</td>
<td>9</td>
</tr>
<tr>
<td>1.4.</td>
<td>Sequence alignment of hDAT, hNET, and hSERT amino acid sequences aligned to LeuT following Beuming alignment</td>
<td>11</td>
</tr>
<tr>
<td>1.5.</td>
<td>The top view of proposed ion binding sites in MATs (SERT)</td>
<td>26</td>
</tr>
<tr>
<td>1.6.</td>
<td>The close-up view of the ion-binding sites in MATs (hSERT)</td>
<td>27</td>
</tr>
<tr>
<td>1.7.</td>
<td>Different approaches used in virtual screening to identify novel scaffolds</td>
<td>32</td>
</tr>
<tr>
<td>2.1.</td>
<td>Alignment of LeuT&lt;sub&gt;Aa&lt;/sub&gt; and hSERT amino acid sequences used to build Model 1</td>
<td>52</td>
</tr>
<tr>
<td>2.2.</td>
<td>Alignment of LeuT&lt;sub&gt;Aa&lt;/sub&gt; and hSERT amino acid sequences following Beuming alignment</td>
<td>54</td>
</tr>
<tr>
<td>2.3.</td>
<td>Alignment of LeuT&lt;sub&gt;Aa&lt;/sub&gt; and hSERT amino acid sequences following Yamashita alignment</td>
<td>55</td>
</tr>
<tr>
<td>2.4.</td>
<td>Alignment of LeuT&lt;sub&gt;Aa&lt;/sub&gt; and hSERT amino acid sequences following Celik alignment</td>
<td>56</td>
</tr>
<tr>
<td>2.5.</td>
<td>Polypeptide backbone superposition of the four homology SERT models</td>
<td>58</td>
</tr>
<tr>
<td>2.6.</td>
<td>A sphere of 8 Å created to constrain docking poses in extracellular exposed S2 pocket</td>
<td>60</td>
</tr>
<tr>
<td>2.7.</td>
<td>Chemical structures of SERT inhibitors (antidepressants) used to develop the pharmacophore</td>
<td>62</td>
</tr>
</tbody>
</table>
2.8. Different sequential filters used in structure-based virtual screening of SERT .... 65
2.9. Chemical structures of first batch of hits identified ........................................ 66
2.10. Ramachandran plot analysis of the stereo chemical parameters of model 1 ........................................................................................................................................... 71
2.11. Ligand-accessible regions of SERT substrate and inhibitor binding pockets .................................................................................................................................................. 73
2.12. Four-point pharmacophore for SERT S2 pocket displayed along with excluded volume ................................................................................................................................................... 75
2.13. VS hit compound in vitro MAT binding screen .................................................... 79
2.14. Saturation binding of VS hit compounds .................................................................. 80
2.15. VS hit compound in vitro serotonin uptake inhibition screen .................................. 81
2.16. Structures of hits identified through virtual screening SM-10, SM-11 and SM-14 .............................................................................................................................................................................. 82
2.17. Saturation binding of VS hit compounds SM-14 at hDAT N2A cells ................. 83
2.18. Dopamine uptake inhibition potency of VS hit compound SM-14 at hDAT N2A cells ............................................................................................................................................................................ 84
3.1. Schematic representation of the photoaffinity labeling of a protein ...................... 98
3.2. Structures of tropane- and piperazine-based DAT photoaffinity ligands .............. 99
3.3. Structures of bupropion and its photoactive probe .................................................. 102
3.4. Two-dimensional docking interaction plot of the bupropion photoprobe ............ 103
3.5. PLIF panel of bupropion and photoprobe docked poses ...................................... 104
3.6. DAT binding affinity of pyrovalerone and its analogues ....................................... 105
3.7. Structures and activity profile of methylphenidate analogues ............................ 107
4.1. The signaling cascade of MEK family ................................................................. 116
4.2. Multiple sequence alignment of different isoforms of human MEK5 ................. 118
4.3. Differences between the wild type MEK5 isoform β and its constitutively active mutant .............................................................................................................. 119
4.4. Phylogenetic analysis of the members of MEK family ........................................ 121
4.5. The two different conformations of Phe 223 generated in MEK1 through loop modeling ................................................................................................................. 123
4.6. Sequence alignment of MEK5 β DD and MEK1 used to build homology models.. 124
4.7. Superposition of the fourteen MEK1 crystal structures ..................................... 127
4.8. Sequence alignment of activation loop of MEK1 and MEK5 β DD .................... 130
4.9. Ramachandran plot of the final MEK5 β DD homology model ......................... 131
4.10. Docking interactions of the shortest side chain diphenylamine analogue with MEK5 β DD ................................................................................................................. 135
4.11. Superposition of crystal structure of MEK1 and MEK5 homology model ....... 136
4.12. Different series of inhibitors designed ................................................................ 137
LIST OF ABBREVIATIONS

MAT: Monoamine Transporter
DAT: Dopamine Transporter
NET: Norepinephrine Transporter
SERT: Serotonin Transporter
LeuT: Leucine Transporter
NSS: Neurotransmitter Sodium Symporter
TM: Transmembrane Domain
VS: Virtual Screening
PH4: Pharmacophore
MEK5: Mitogen Activated Protein Kinase
Chapter 1.

MONOAMINE TRANSPORTER STRUCTURE, FUNCTION
AND DRUG DISCOVERY: A COMPUTATIONAL REVIEW

1.1 Introduction

Monoamine transporters (MATs) are important proteins that maintain neurotransmitter homeostasis in brain synapses. The monoamine hypothesis shows that dysfunction of these critical gatekeepers result in the imbalance of monoamine neurotransmitters in the synapse, linked to various disease states such as depression, addiction and other major neurological disorders.\textsuperscript{1-6} The control of several body functions such as hunger, sleep and mood is maintained by neurotransmitter homeostasis. A critical understanding of MATs will help design and develop better drugs to treat disorders linked to MATs.\textsuperscript{7}

MATs are membrane proteins that span across the cell membrane lipid bilayer 12 times; the polypeptide N- and C- terminals are in the cytoplasm (Figure 1.1). Experimental studies such as scanning cysteine accessibility mutagenesis (SCAM) were carried out on these proteins to elucidate the mechanistic details of substrate translocation and ligand binding. Computational efforts were performed to understand the atomistic details of these proteins. Briefly, these studies can be divided as “pre-LeuT and post-LeuT” and are described below. Leucine transporter (LeuT), a bacterial transporter homologous to MATs, was crystallized in 2005.\textsuperscript{8} Proteins evolutionarily distant from the neurotransmitter sodium symporter family (NSS) were also used as templates to build the
homology models of MATs, but were later found to be poor tools as the homology was very weak.\textsuperscript{8} Pre-LeuT based MAT efforts were ligand-based and included QSAR studies, ligand-based pharmacophores that were used in virtual screening (VS) efforts to identify novel compounds.

As of now, LeuT is recognized as the closest relative to these MATs, and is shown to have an inverted symmetry between the two domains that are formed by six transmembrane helices each. Only two conformations of the leucine substrate translocation cycle (\textbf{Figure 1.2}), one with open-to out and the other being ligand bound (occluded) were crystallized. The inward-facing conformation is the critical missing piece of the puzzle.

A clear understanding of the MATs conformational transitions during the substrate translocation will help in the design of better drugs. Different MAT homology models were constructed using the two available conformations of LeuT. The homology models were employed in various studies to discern the mechanism of action of various MAT inhibitors that led to opposing conclusions.\textsuperscript{9, 10} There is a dearth of structural diversity in the MAT inhibitor ligands, and no studies were performed to identify novel ligands with increased potency and reduced side effect profile. Recent structure-based efforts have identified diverse novel compounds (hits) that can be optimized further as leads (inhibitors) that bind to MATs.\textsuperscript{11-13}

The following is a review of the computational efforts of MAT homology model construction, their use in understanding MAT dynamics, substrate translocation, inhibitor binding and especially the identification of novel scaffolds through virtual screening.
Figure 1.1. MATs embedded in lipid bilayer. hSERT (green helices) is embedded within a POPE lipid bilayer (Van der Waals spheres) with both polypeptide terminals within the intracellular space. Lipophilic tails (gray) of the phospholipids are oriented toward the center of the cell membrane, arranged tail-to-tail; the polar head groups containing phosphorous (pink), oxygen (red) and nitrogen (blue) atoms are oriented toward the hydrophilic borders of the membrane. The S1 (yellow) and S2 (cyan) substrate binding pockets are displayed as surfaces.
Figure 1.2. The three minimally required conformations of substrate translocation cycle. TMs 1,3,6,8 that form the substrate translocation pore are presented as bars. The outward facing conformation opened towards extracellular space allows the entry of substrate (S) and ions (Na⁺) into the protein to form the occluded conformation. The inward facing conformational transition is achieved with the release of substrate and ions into the cytoplasm.
### 1.2 Generation of MAT Homology Models

#### 1.2.1 Pre LeuT

Despite the lack of a closely related protein crystal structure, MAT homology models were attempted as early as 1994, without the use of a structural template. A DAT model was generated using the combination of 1) multiple sequence alignment (MSA) of 10 neurotransmitter sodium symporter (NSS) transporter cDNAs (hDAT, hNET, hSERT, hGABA, rGABA, rGLY, rTAU, rCHO, rSERT and dog betaine) and 2) hydropathy plots to assign TM domain length.\(^\text{14}\) The DAT homology model was used to discuss the mechanism for substrate/ion binding and cocaine inhibition of the dopamine uptake.\(^\text{14}\) A template-independent SERT homology model also employed MSA and hydropathy plots, using online servers- PHD Predict Protein,\(^\text{15}\) DAS,\(^\text{16}\) TMpred and TopPred2.\(^\text{17}\) A cocaine binding pocket was created in SERT with five TM helices using the data from ligand binding and site-directed MAT mutagenesis.\(^\text{18}\) The remaining seven TM domains were positioned and refined by electrostatic calculations using the electrostatics program GRASP. Follow up docking studies with this SERT model suggested that citalopram and cocaine interact with the charged side chain of D98 (TM1), while the TCA imipramine bind elsewhere.\(^\text{18}\) The TM domain juxtapositions in the generated homology models were largely speculative and not particularly accurate, owing to the lack of a template of 12 TM domain transporters.

The first template-based MAT models were constructed using the 12 TM facilitated diffusion transporter with sequence identity 13% or less. MAT models were derived from the mutational data of two such *E. coli* proteins, the \(\mathrm{Na}^+/\mathrm{H}^+\) antiporter
(NhaA) and the lactose permease (LacY), with the NhaA electron density projection map serving as a template.\textsuperscript{19} The resultant MAT models were used to localize citalopram and cocaine binding as well as the identification of substrate permeation pathway.\textsuperscript{20-22} It was proposed that Y95 (TM1) of SERT to be responsible for the selectivity of the S-enantiomer of citalopram (Celexa\textsuperscript{TM}); the analogous residue was implicated in DAT binding of cocaine and its analogs.\textsuperscript{21} A final round of non LeuT-based MAT homology models employed an inward- (cytoplasm-) facing LacY crystal structure as a template.\textsuperscript{23} The LacY, best characterized of all membrane transporters, is a 12 TM protein with six TM helices in each domain.\textsuperscript{24} Upon creating an inward-facing SERT model, an outward-facing conformation was generated through a rigid body rotation of \( \sim60^\circ \) of both the domains. The SERT conformation thus generated was later used to build the outward-facing DAT and NET models. The MAT models thus developed were utilized to dock cocaine, amphetamine and citalopram to identify the potential inhibitor binding sites.\textsuperscript{25} Unfortunately, the X-ray structures of NhaA and LacY were later found not to appreciably resemble that of a true MAT homolog, the bacterial leucine transporter LeuT.

\subsection*{1.2.2 Post LeuT (2005 - present)}

The publication of a crystal structure (pdb id: 2a65) of the leucine transporter (LeuT) from \textit{Aquifex aeolicus},\textsuperscript{8} a thermophilic bacterium, provided the first reliable MAT template for computational studies as it is the closest prokaryotic orthologue to the NSS family.\textsuperscript{26, 27} The LeuT crystal structure launched a wave of new MAT models that continue to be corroborated by biophysical studies. Several additional LeuT crystal
structures with inhibitors bound were also crystallized.\textsuperscript{28-31} Unless otherwise specified, the models discussed herein were based on the first available LeuT crystal structure (pdb id: 2a65).

\subsection{1.2.2.1 Dopamine Transporter (DAT)}

Several DAT models were constructed using the LeuT as a template (Table 1.1), but the first LeuT-based DAT model was reported by Huang et al. in a study of dopamine interaction with the transporter and the substrate permeation pathway.\textsuperscript{32} Binding free energy calculations on the proposed binding mode of dopamine in DAT were in close agreement to the experimental free energy observed in this study.\textsuperscript{32} Additional molecular dynamic (MD) studies with this DAT model suggest that cocaine initially binds to a site that is distinct from that of the dopamine binding site.\textsuperscript{33} In stark contrast, Beuming et al. used a DAT model and a new multiple sequence alignment (MSA) to suggest that cocaine, benztropine, and amphetamine bind in the primary substrate site also referred to as S1.\textsuperscript{34} Mutagenesis and intramolecular cross-linking studies performed on DAT also supported this proposal.\textsuperscript{35}

Following this, a study by Indarte et al. elaborated the dopamine binding in DAT indicating the likelihood of a primary and a secondary substrate site known as S2. This is the first report that indicated the presence of two distinct binding sites (Figure 1.3)\textsuperscript{10}. The same study docked dopamine and d-amphetamine into three different rDAT models, and from the results it was evident that in addition to the primary binding site, a secondary extracellular pocket exists. It was proposed as the secondary substrate site that acts as a temporary staging area for substrate translocation and a potential site for inhibitors.\textsuperscript{10}
Xhaard et al. proposed an alternate binding mode for substrates by docking dopamine, norepinephrine and serotonin into the S1 site of DAT, NET and SERT homology models respectively. Additionally, the study proposed that the aromatic hydroxyl substituent chelates with the Na1 in the pocket known as the “chelation mode” as opposed to an interaction between the amine nitrogen and the Na1 in the “ionic mode”.

36
Figure 1.3. Ligand-accessible regions of MATs (SERT). Pocket space available to the ligands in the deeper S1 (yellow) and the extracellular S2 (cyan) pockets are displayed as surfaces. Selected residues that define each pocket are displayed as sticks (atomtype color). For clarity, only the main TM contributors to SERT ligand binding pockets are highlighted: TM 1 (pink), TM 3 (green), TM 6 (gold), TM 8 (brown) and EL 4 (blue).
A number of MAT homology models were constructed using Beuming sequence alignment. The alignment was generated using 344 NSS family proteins, and most of the studies utilized models based on this comprehensive alignment (Figure 1.4). Guptaroy et al. used a DAT model to study the local conformational changes associated with a T62D mutation. The highly conserved T62 residue is located in the N-terminal region, and the mutation of T62 favored the inward-facing conformation of DAT suggesting a substantial role in the substrate translocation cycle.

Schmitt et al. generated a DAT model based on the desipramine-bound LeuT crystal structure (pdb id: 2qju) to study the interaction profile of bivalent phenethylamines. The docking studies suggested that the long flexible phenethylamines span the substrate translocation pore of the protein, simultaneously occupying the S1 and S2 pockets. The W84L and D313N DAT mutants also suggested that phenethylamines prefer inward-facing DAT conformation, and further concluded the presence of multiple low-affinity substrate sites throughout the translocation pore.

In a study focused on MAT-inhibitor interactions, the electrostatic potential surface for DAT was calculated, by docking cocaine and clomipramine into S1 and S2 sites respectively. In a recent study by Hong et al. four residues were identified for SCAM studies, which led to the conclusion that DAT prefers the outward-facing conformation in a cholesterol-rich lipid environment.
Figure 1.4. Sequence alignment of hDAT, hNET, and hSERT amino acid sequences aligned to LeuT following Beuming alignment. Residue position in the chain is indicated with numbers (red) before and after each row. Predicted TM regions are indicated above each sequence row. Identical (blue) or similar (yellow) residues across all sequences are highlighted. Gaps between sequences are exhibited with dots. Residues forming the extracellular charged gate (blue spheres), extracellular hydrophobic gate (red stars) and the cytoplasmic charged gate are indicated below the sequence rows.
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Template (PDB ID)</th>
<th>Sequence Alignment</th>
<th>Sequence Identity (%)</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edvardsen</td>
<td>1994</td>
<td>None</td>
<td>Developed</td>
<td>-</td>
<td>MIDAS</td>
</tr>
<tr>
<td>Ravna</td>
<td>2003</td>
<td>NhaA</td>
<td>Developed</td>
<td>7</td>
<td>MIDAS plus</td>
</tr>
<tr>
<td>Ravna</td>
<td>2006</td>
<td>LacY (1pv6)</td>
<td>Developed</td>
<td>5</td>
<td>ICM</td>
</tr>
<tr>
<td>Huang</td>
<td>2007</td>
<td>LeuT</td>
<td>Developed</td>
<td>20.4</td>
<td>Insight II</td>
</tr>
<tr>
<td>Indarte</td>
<td>2008</td>
<td>LeuT</td>
<td>Developed</td>
<td>-</td>
<td>MOE, Robetta 3D-JIGSAW</td>
</tr>
<tr>
<td>Xhaard</td>
<td>2008</td>
<td>LeuT</td>
<td>Developed</td>
<td>-</td>
<td>Malign, MODELLER</td>
</tr>
<tr>
<td>Beuming</td>
<td>2008</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER</td>
</tr>
<tr>
<td>Guptaroy</td>
<td>2008</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER</td>
</tr>
<tr>
<td>Ravna</td>
<td>2009</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>ICM</td>
</tr>
<tr>
<td>Schmitt</td>
<td>2010</td>
<td>LeuT (2qju)</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER, MOE</td>
</tr>
<tr>
<td>Hong</td>
<td>2011</td>
<td>LeuT</td>
<td>Developed</td>
<td>-</td>
<td>Swiss-Pdb Viewer</td>
</tr>
</tbody>
</table>

(-) information not provided

**Table 1.1. Construction of DAT homology models in chronological order.**
1.2.2.2 Serotonin Transporter (SERT)

Of the three MATs, most of the computational focus was on SERT, and so numerous homology models were constructed (Table 1.2). Ravna et al. reported the construction of first SERT model, which was used to identify amino acid residues involved in ligand binding. Using the ICM pocket finder, a pocket formed by TM 1, 3, 6, and 8 that corresponds to the leucine binding site in LeuT was identified. Subsequently, a similar conclusion was reached by a study in which LeuT and SERT substrate binding sites were compared. Additionally, the study suggested that the larger serotonin substrate may be accommodated in a large pocket with smaller side chains, and proposed the importance of D98 (TM1) in SERT for substrate interaction.

To further their understanding of the previously reported SCAM data, both Zhang et al. and Rudnick et al. utilized a SERT model to propose that the residues in TM5 may be involved in substrate permeation pathway. In a similar SCAM study, the role of TM7 was also examined by correlating computational insights with the SCAM data. Through these studies, it was proposed that TM7 was involved in ion dependence and translocation pathway, possibly through the rotation of TM7. With a similar rationale, a SERT homology model was used to support the results from Zn\(^{2+}\) binding experiments, wherein the data suggested that residues in TM1 and TM3 were in close proximity and oriented towards each other, especially V102 and I179. To determine the importance of the N-terminus in amphetamine action, a SERT homology model was developed by Sucic et al. The study concluded that the N-terminus was involved in substrate efflux by acting as a lever that reverses the substrate transport in the presence of amphetamine.
A SERT homology model was generated and used by Jorgensen et al. to identify the potential binding site of escitalopram by docking. The study concluded that the binding mode observed was well in agreement with the already available mutagenesis data. In a study focused on buspirone binding at SERT, two hSERT models were constructed based on the inward-facing LacY (pdb id: 1pv6) and the outward-facing LeuT. Various buspirone analogs were then docked into both the SERT models to ascertain the residues involved in binding. The two distinctive models illustrated the differences in the binding pocket environment created by different conformational states. Specifically, in the LeuT-based SERT model, two binding sites were identified corresponding to high- and low-affinity sites in LeuT.

In an eloquent study, the paired mutant-ligand analogue complementation (PaMLAC) method was used to study the binding of serotonin in the SERT primary substrate pocket. In this study various serotonin analogs were docked into the final SERT model using an induced fit method. The docked poses were experimentally analyzed by coupling with mutational data generated from thirteen different single point mutants, which concluded that serotonin interacts with D98, A173 and T439 residues. A more recent study on the substrate binding by Kaufmann et al. suggested that 5-HT binds in a different orientation, similar to the one proposed by Jorgensen et al.

Kaufman et al. built human and drosophila SERT models to identify the differences conferring species selectivity. Different 5-HT analogs were docked; and the most favorable pose oriented 5-HT in a “down” binding mode. The indole nitrogen is located near Y176/Y171 residue pointing toward the interface between TM3 and TM8. This was well in agreement with SCAM and mutagenesis data. A SERT model was
utilized to examine the substrate binding, where several 5-HT analogs were docked into S1 site, and molecular dynamics was performed on the selected protein-ligand complexes. The study concluded that IL1 and TMs 6, and 8 play an important role in the substrate translocation.\textsuperscript{52}

Binding of TCAs to the SERT was studied by using a model constructed using open-to-out conformation of LeuT (pdb id: 3f3a).\textsuperscript{28} Analysis of the docked poses showed protein-ligand interactions that were in agreement with site-directed mutagenesis data, suggesting the presence of a vestibular low-affinity binding site for TCAs. The tricyclic ring of the TCAs occupied the S2 vestibular pocket, while the amine tail probed into S1 site in the generated poses.\textsuperscript{53} Using both the open-to-out LeuT (pdb id: 3f3a) and the occluded conformation (pdb id: 2a65), different SERT models were developed to study inhibitor binding. The high affinity $S$-enantiomer of citalopram (escitalopram) was docked into SERT S1 site, and 64-point mutations were created to experimentally analyze the docked poses. Docking into S2 was not considered in this study because previous mutational studies for nonconserved residues in this region did not produce a significant biological effect.\textsuperscript{54} Andersen et al. proposed two new residues, N177 and F341, in S1 to be critical for the escitalopram binding.\textsuperscript{55}

Most recently, Manepalli et al. utilized a SERT model as a tool for drug discovery.\textsuperscript{13} In this study SERT models were developed based on four different alignments: Yamashita, Beuming, Celik, and the Manepalli alignment, a modified Celik alignment. The Manepalli alignment based model was used later for \textit{in silico} virtual screening (VS) of a subset of ZINC database to identify novel chemotypes. Furthermore the use of this model in optimizing a previously identified hit was also reported.
recently.\textsuperscript{12} Herein, docking and flexible alignment studies were used to guide the chemical transformation of a non-selective MAT inhibitor into a SSRI.\textsuperscript{12}
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Template (PDB ID)</th>
<th>Sequence Alignment</th>
<th>Sequence Identity (%)</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ravna</td>
<td>2001</td>
<td>-</td>
<td>Developed</td>
<td>-</td>
<td>PHD, WHATIF</td>
</tr>
<tr>
<td>Ravna</td>
<td>2003</td>
<td>DAT</td>
<td>Developed</td>
<td>80</td>
<td>ICM</td>
</tr>
<tr>
<td>Ravna</td>
<td>2006</td>
<td>NhaA</td>
<td>Developed</td>
<td>7</td>
<td>ICM</td>
</tr>
<tr>
<td>Ravna</td>
<td>2006</td>
<td>LacY (1pv6)</td>
<td>Developed</td>
<td>5</td>
<td>ICM</td>
</tr>
<tr>
<td>Ravna</td>
<td>2006</td>
<td>LeuT</td>
<td>Developed</td>
<td>20</td>
<td>ICM</td>
</tr>
<tr>
<td>Zhang</td>
<td>2006</td>
<td>LeuT</td>
<td>Developed</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Henry</td>
<td>2006</td>
<td>LeuT</td>
<td>Yamashita</td>
<td>20.5</td>
<td>SCWRL</td>
</tr>
<tr>
<td>White</td>
<td>2006</td>
<td>LeuT</td>
<td>Yamashita</td>
<td>20.5</td>
<td></td>
</tr>
<tr>
<td>Jørgensen</td>
<td>2007</td>
<td>LeuT</td>
<td>Developed</td>
<td>-</td>
<td>MODELER</td>
</tr>
<tr>
<td>Forrest</td>
<td>2007</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>NEST</td>
</tr>
<tr>
<td>Xhaard</td>
<td>2008</td>
<td>LeuT</td>
<td>Developed</td>
<td>-</td>
<td>Malign</td>
</tr>
<tr>
<td>Jarończyk</td>
<td>2008</td>
<td>LacY (1pv6)</td>
<td>Developed</td>
<td>-</td>
<td>ICM, Bioedit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LeuT (2a65)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celik</td>
<td>2008</td>
<td>LeuT</td>
<td>Developed</td>
<td>22.9</td>
<td>MODELER</td>
</tr>
<tr>
<td>Kaufmann</td>
<td>2009</td>
<td>LeuT</td>
<td>Beuming</td>
<td>17</td>
<td>ROSETTA suite</td>
</tr>
<tr>
<td>Ravna</td>
<td>2009</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>ICM</td>
</tr>
<tr>
<td>Sarker</td>
<td>2010</td>
<td>LeuT (3f3a)</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER</td>
</tr>
<tr>
<td>Andersen</td>
<td>2010</td>
<td>LeuT (3f3a)</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER</td>
</tr>
<tr>
<td>Wentthur</td>
<td>2010</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER</td>
</tr>
<tr>
<td>Sucic</td>
<td>2010</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>MOE 2007.08</td>
</tr>
<tr>
<td>Gabrielsen</td>
<td>2011</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>ICM</td>
</tr>
<tr>
<td>Manepalli</td>
<td>2011</td>
<td>LeuT</td>
<td>Developed</td>
<td>23.9</td>
<td>DS 2.5.1</td>
</tr>
</tbody>
</table>

(-) information not provided.

**Table 1.2. Construction of SERT homology models in chronological order.**
1.2.2.3 Norepinephrine Transporter (NET)

In the past, NET was not given enough attention as compared to DAT and SERT, however with the promising results of norepinephrine reuptake inhibitors (NRIs) and dual-acting norepinephrine-dopamine reuptake inhibitors (NDRIs), attention to NET is on the rise (Table 1.3). Aside from the NET models built by Ravna et al. along with SERT and DAT models, few attempts have been reported.\textsuperscript{22, 25, 56-58} A NET homology model was generated to visually examine the effects of mutations on the binding of TCA desipramine and conopeptide $\chi$-Mr1A, a noncompetitive, selective NET inhibitor.\textsuperscript{56} Several residues presumed to be important for $\chi$-Mr1A interactions were located near the extracellular pocket of the transporter. Correlating the mutagenesis data along with docking studies, it was concluded that the binding sites of the conotoxin and TCA overlap while the conotoxin and norepinephrine sites are discrete.\textsuperscript{56}

In a similar unpublished study by Nolan et al. two distinct binding pockets were identified using a NET model based on the open-to-out LeuT crystal structure (pdb id: 3F3A). Adaptive Poisson-Boltzmann Solver (APBS) calculations of docked poses for known NET ligands were compared to experimental data, which concluded that a secondary site exists in the vestibular region of NET where inhibitors bind, particularly the TCAs. In a study on NET ligand binding, Hill et al. used previously built DAT model\textsuperscript{32, 33} as a template to construct a NET model, as there is a high similarity (67%) between DAT and NET.\textsuperscript{57} NE was docked into the S1 site and MD simulations were carried out on the resulting NET-NE complex to examine the influence of cocaine and its analogs, RTI-33 and RTI-113, on substrate binding. The study concluded that cocaine occupies a pocket far removed from the S1 substrate site.\textsuperscript{57} A recent effort to identify the
determinants conferring inhibitor selectivity for SERT and NET was performed.\textsuperscript{58} Herein, the structural details of SERT/NET selectivity were studied for citalopram and talopram, an SSRI and NRI respectively. Mutational data from this study suggested that the binding pocket of SSRIs overlap the S1 pocket whereas talopram binds at a distinct site.\textsuperscript{58}
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Template (PDB ID)</th>
<th>Sequence alignment</th>
<th>Sequence Identity (%)</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ravna</td>
<td>2006</td>
<td>NhaA</td>
<td>Developed</td>
<td>7</td>
<td>ICM</td>
</tr>
<tr>
<td>Ravna</td>
<td>2006</td>
<td>LacY</td>
<td>Developed</td>
<td>5</td>
<td>ICM</td>
</tr>
<tr>
<td>Paczkowski</td>
<td>2007</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>MODELLER</td>
</tr>
<tr>
<td>Hill</td>
<td>2011</td>
<td>DAT</td>
<td>Beuming</td>
<td>66</td>
<td>Insight II</td>
</tr>
<tr>
<td>Andersen</td>
<td>2011</td>
<td>LeuT</td>
<td>Beuming</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>Nolan (unpublished)</td>
<td>2011</td>
<td>LeuT (3f3a)</td>
<td>Beuming</td>
<td>20</td>
<td>DS 2.5.1, APBS</td>
</tr>
</tbody>
</table>

(-) information not provided

Table 1.3. Construction of NET homology models in chronological order.
1.3 Molecular Dynamics (MD) on MATs

1.3.1 DAT Dynamics

Of the three MATs, DAT and SERT were the most studied, which may be due to their involvement in several disease states. Molecular dynamics simulations were performed on DAT models even before LeuT was crystallized. Edvardsen et al. calculated the interaction energies of DAT complexes with the substrate dopamine and two enantiomers of the psychostimulant cocaine. Another study performed by Ravna et al. later concluded that changes in TM helical arrangement constricted the translocation pathway suggesting tropane analogues cannot fit in the putative ligand pocket.

The first full-fledged MD study on DAT models constructed using LeuT was performed by Huang et al. Molecular dynamics on both apo DAT and DAT- dopamine complex were performed to localize the substrate-entry pathway. The calculated strength of dopamine binding to DAT evaluated by molecular mechanic/Poisson-Boltzmann surface area (MM-PBSA) was -6.4 kcal/mol that was similar to the experimental binding energy (-7.4 kcal/mol). Following this study, several DAT complexes were generated by docking ligands like dopamine, amphetamine, cocaine and its analogues, benztropine and its analogues into DAT. Of the different complexes generated, the complexes of DAT-CFT (a cocaine analog) and DAT-DA were studied further using MD simulations. Using mutagenesis studies, intramolecular cross-linking studies, docking and molecular dynamics, it is proposed that cocaine and its higher affinity analog WIN-35,248 can go deeper into DAT and occupy S1 substrate pocket which was not shown previously.

The effect of mutation of a highly conserved T62 residue in the N-terminal of
hDAT in the [RETW] motif on dopamine uptake was studied using free energy perturbation (FEP) and mutational studies. The study concluded that T62 holds DAT in an open to outward conformation towards the extracellular space favoring dopamine uptake. Huang et al. extended their previous MD studies on DAT to explain the mechanism of cocaine inhibition on dopamine uptake. The study concluded that dopamine and cocaine bind at discrete non-overlapping pockets and cocaine decreases the kinetic turnover of DAT by binding in the extracellular pocket. The conclusion of this study directly contradicted the study performed by Beuming et al. which proposed a single overlapping binding site for cocaine and dopamine.

Gedeon et al. performed dynamics on DAT homology model constructed using LeuT. The study concluded that the extracellular salt bridge R85 (TM1) – D476 (TM10) is formed less frequently in the presence of a substrate at S2. Biased steered molecular dynamics was performed by applying a force to dopamine and pulled in both the extracellular and intracellular ways to identify the whole permeation pathway. A stopover site towards extracellular space corresponding to S2 pocket was identified. The study was extended further by placing a second substrate molecule in the S1 pocket, which finally concluded that the presence of a second substrate molecule in the S2 site had an allostERIC effect on the substrate in S1 pocket, causing it to get released into the cytoplasm.
1.3.2 SERT Dynamics

The first full-fledged dynamics on SERT employed complexes of the SERT-serotonin and SERT-escitalopram generated using docking. It was shown that the extracellular hydrophobic gate residue Y176 would embrace different conformations to better accommodate the changes in the binding pocket. Following this study, dynamics was performed to verify whether the SERT model constructed using LacY would convert to SERT built using LeuT. It was noticed that there was no impact on the helical packing of SERT based on LacY, and did not resemble the SERT model constructed using LeuT.

MD studies on the models of wild type SERT, T81A and T81D mutants were performed to focus on how the N-terminus (Thr81) would influence the SERT conformation. It showed the incidence of intracellular structural changes resulting in the opening of intracellular vestibule. The same study also showed that the mutation of T81A would induce an inward facing conformation by increasing the distance between the terminals. This equilibrium shift towards increased SERT inward facing conformation favored the amphetamine-induced efflux of 5-HT. The equilibrium hypothesis was tested experimentally by the removal of first 64 residues at the N-terminus, and tethering of N-terminal completely obliterated the amphetamine induced substrate efflux. The final conclusion of this study was the importance of the N-terminus, which induces conformational changes by acting as a lever.

The trajectory analysis of molecular dynamics performed on SERT showed a substantial movement of TM7 especially, the M370 and S375 residues. SCAM studies were followed up to identify the residues that are exposed towards the hydrophilic
environment. The combined analysis of MD and SCAM suggest the importance of TM 7 in addition to TMs 1, 3, 6 and 8. The same SERT homology model also identified residues, V366 and M370 exposed towards the lipid bilayer, thus reducing their accessibility towards the MTSEA reagents. Finally, they concluded that Na ion exodus rotates TM7 to expose V366 and M370 towards MTSEA reagents.\textsuperscript{45}

An effort to study the substrate translocation in SERT was launched recently.\textsuperscript{52} Different SERT complexes of apo SERT, SERT-serotonin and SERT-escitalopram were generated and MD was performed. Careful analysis of the MD trajectories showed that SERT adopts different conformational changes that assist in the substrate translocation.\textsuperscript{52} Koldsø et al. performed the longest MD simulations on SERT.\textsuperscript{62} The study identified that the outward facing single substrate-bound SERT complex transformed to an inward facing conformation thus strengthening the claim of the absence of allosteric effect. The doubly bound substrate complex in SERT did not show any signs of the substrate translocation; rather, the loosely bound substrate got dissociated towards extracellular pocket. The role of D437 residue in ion transport was shown along with the hydration of substrate and Na2 ion, which is the ion that gets released into the cytoplasm.\textsuperscript{62}

1.4 Ion (Na\textsuperscript{+}, Cl\textsuperscript{−}) Binding Sites

Structure-based computational efforts were directed towards identifying the binding sites for two Na\textsuperscript{+} and one Cl\textsuperscript{−} ion, the essential components for monoamine transporter function (Figure 1.5). While residues interacting with two bound Na\textsuperscript{+} ions in LeuT crystal structures were conserved amongst the NSS proteins (Figure 1.6), little was known about the binding domain of the Cl\textsuperscript{−} ion, as it is not essential for many NSS
members including the LeuT, Tyt1 and TnaT. With this, a SERT homology model was
developed in an effort to elucidate the Cl- ion-binding site in MATs. The negatively
charged residues in LeuT that may replace the role of Cl- ion in the MATs were
identified using the Multi-Conformer Continuum Electrostatics (MCCE) method. It was
identified that the E290 of LeuT was nonconserved among Cl- dependent transporters,
corresponding to a serine in the MATs (S372 in SERT). The placement of a Cl- ion at
this position in the SERT model showed coordination between the hydroxyl and amine
hydrogen’s of Y121, S336, N368, S372 and is in close proximity to Na1 (~5Å)
suggesting that Cl- at this position would stabilize the binding of Na1 ion.
Experimentally, the S372 and N368 mutants displayed decreased function and moreover,
the substitution of these residues with a negatively charged residue produced a Cl- ion
independent transporter with reduced activity. Similar experimental studies support
these findings that likewise suggest that a Cl- ion or a negatively charged residue is
required for transport.
Figure 1.5. The top view of proposed ion binding sites in MATs (SERT). The residues of helices, TM1(yellow), TM2(blue), TM6(pink), TM7(brown) and TM8 (red) that interact with substrate(green sticks), sodium (orange) and chloride (green) ions are displayed as sticks.
Figure 1.6. The close-up view of the ion-binding sites in MAT (hSERT). Substrate (green sticks) Na⁺ ions (orange spheres) and Cl⁻ ions (green spheres) are displayed. The residues interacting with these ions are presented as sticks in atom type color and the corresponding residues in DAT (blue), NET (red) and SERT (green) are shown.
1.5 Virtual Screening

In addition to homology models, ligand-based models were proven useful in the study of transporter interactions and ligand recognition. Prior to the availability of NSS crystal structures, computational approaches relied heavily on known MAT ligands and their associated structure-activity relationships (SAR). A few caveats exist with 3D ligand-based studies including the assumption that all training set compounds bind at the same site of the same protein in a similar conformation.

1.5.1 Ligand-Based VS

While most ligand-based models have served to define the necessary features common to known ligands, some of these models were applied in the search for potential new scaffolds through virtual (in silico) screening. An early ligand-based virtual screening study on the DAT featured a pharmacophore model based on just two compounds, cocaine and WIN-35065-2. The model consisted of three features: a nitrogen atom, a carbonyl group and an aromatic ring. Screening of the National Cancer Institute (NCI) database afforded 4094 compounds from which 44 out of the 70 tested experimentally were proven to inhibit DAT at 10 μM. The most potent of these hits exhibited an uptake inhibition potency of 274 nM, 155 nM, and 108 nM at DAT, SERT, and NET respectively. An analogue of the quinuclidine scaffold was identified as a novel DAT inhibitor in a different study. Following the analysis of known DAT ligands lacking a carbonyl atom, the pharmacophore model already proposed was refined to include two additional features- an aromatic group and a tertiary amine. Screening of
the Advanced Chemical Dictionary (ACD) database with the modified pharmacophore model resulted in the identification of eight hits with IC50 values less than 10 μM.68

Applying slight modifications to the distance constraints between the nitrogen atom and two aromatic features, the model afforded indole class of compounds with a NET and SERT affinity of 137 nM and 126 nM respectively.69 In response to the report that a tertiary nitrogen in tropane analogs was unnecessary,70 Enyedy et al. refined existing model by substituting a hydrogen-bond acceptor to tertiary nitrogen. The refined model was used for screening the NCI database that resulted in 1104 hits. Experimental testing of these compounds identified two hits with DAT uptake inhibition potency of 2.3 μM and 0.255 μM respectively.71 A similar pharmacophore was derived from mazindol that consists two lipophilic aromatic rings along with one nitrogen atom. The pharmacophore was used to screen the ACD database containing 225,000 compounds which identified three compounds with Ki values of 128 nM, 48 nM and 24 nM, respectively.72

Using an undisclosed VS method, a potent NET-selective compound with an uptake inhibition potency of 1 nM was identified.73 A pharmacophore model for NET was generated by studying the conformations adopted by milnacipran in solution and used in the design of analogs with better pharmacological profiles at MATs.74 The generation of a pharmacophore model based on the NET-selective χ-conotoxin, a peptide component of marine cone snail venom resulted in the design of a compound which has entered phase II clinical trials.75

In an effort to compare the necessary features for activity amongst MATs, different pharmacophore models were generated separately for each transporter using a
training set containing either tropane analogs or GBR-12909 like compounds. The study highlighted the similarities as well as differences in the potential binding modes by coupling structures with biological activity. It was suggested that minor modifications to a core scaffold can confer selectivity for one transporter over the others.\textsuperscript{76} A study was attempted to differentiate the pharmacophore features necessary for SERT and D\textsubscript{2} receptor binding among a series of dual-acting piperazine ligands. The SERT pharmacophore model consisted of basic nitrogen, a sterically restricted area around the nitrogen, an aromatic region, and a general hydrophobic region.\textsuperscript{77} While these pharmacophore features were the same as those of the D2 receptor, the three dimensional arrangement of features differed, suggesting the flexibility of dual ligands vital for adopting different conformations to bind these two proteins.\textsuperscript{77}

1.5.2 Structure-Based VS

The utility of MAT homology models was not limited to mechanistic and binding studies. Recent reports demonstrate the value of using these models for the identification of novel chemotypes.\textsuperscript{11, 13} To date there have been only two reports describing the discovery of novel MAT ligands through structure-based virtual screening wherein both have focused on the extracellular vestibular region (S2).\textsuperscript{11, 13} The first DAT-based VS approach used a structure-based S2 pharmacophore to screen a chemical database containing over 140,000 compounds. Ten final compounds were selected based on the score (Affinity dG in MOE), predicted pKi, and visual inspection to be pharmacologically tested for activity at the MATs. MI-4, an ifenprodil analog, was discovered to have affinity at all three transporters.\textsuperscript{11}
The first structure-based VS on SERT was approached more recently using a pharmacophore developed from S2 and halogen binding pockets (HBP). Herein, a subset of the ZINC database (~ 1 million compounds) was screened using high-throughput docking in MOE. Visual inspection along with the Affinity dG score aided in the selection of hit compounds for pharmacological analysis. This study resulted in three structurally novel compounds, two selective SERT ligands (SM-10 and SM-11) and one DAT selective compound (SM-14) with no affinity towards the other transporters in both the cases. The SERT model was equally used to refine a hit compound (MI-17) identified through the DAT VS study to create a more selective SSRI (MI-17 hybrid). The first NET-based VS effort employed S1 pocket to screen Kyoto Encyclopedia of Genes and Genomes (KEGG) drug database containing 6436 known drugs. Ten of the final eighteen hits showed micromolar affinity towards NET. The study was more like drug repurposing for already known molecules and not a pure VS approach to identify novel compounds.
Figure 1.7. Different approaches used in virtual screening to identify novel scaffolds. Ligand-based methods use models that employ different pharmacophore features labeled F1-F5 with a volume constraint to screen a small molecule library. Structure-based methods filter through virtual small molecule libraries to identify hits (orange spheres) by docking them into the binding pockets of MATs (green helices). Hybrid approaches use ligand-based VS followed by structure-based VS.
<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>Binding affinity (Ki (nM))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>hDAT</td>
</tr>
<tr>
<td>MI-4</td>
<td><img src="#" alt="MI-4 structure" /></td>
<td>3460±260</td>
</tr>
<tr>
<td>SM-10</td>
<td><img src="#" alt="SM-10 structure" /></td>
<td>ND</td>
</tr>
<tr>
<td>SM-11</td>
<td><img src="#" alt="SM-11 structure" /></td>
<td>ND</td>
</tr>
<tr>
<td>SM-14</td>
<td><img src="#" alt="SM-14 structure" /></td>
<td>15,600±2400</td>
</tr>
<tr>
<td>MI-17</td>
<td><img src="#" alt="MI-17 structure" /></td>
<td>1298±36</td>
</tr>
<tr>
<td>MI-17 hybrid</td>
<td><img src="#" alt="MI-17 hybrid structure" /></td>
<td>2129±177</td>
</tr>
</tbody>
</table>

Table 1.4. Novel MAT ligands discovered through MAT virtual screening.
1.6 Conclusion

Structure-function studies performed on MATs were limited by the experimental techniques; and the recent computational efforts have proven highly valuable in understanding the functioning of these transporters more closely. The computational efforts to understand atomistic aspects of MAT function provided more structural clues in the design and development of new drugs. The design of novel drugs to treat various neurological disorders so far lacked a structural direction. Thus the structure-based VS efforts on MATs were undertaken, which led to the discovery of novel chemo types; and hit-to-lead optimization of these chemotypes helped increase their potency profile.
References


Chapter 2.

DISCOVERY OF NOVEL SELECTIVE SEROTONIN REUPTAKE INHIBITORS THROUGH DEVELOPMENT OF A PROTEIN-BASED PHARMACOPHORE†


The serotonin transporter (SERT), a member of the neurotransmitter sodium symporter (NSS) family, is responsible for the reuptake of serotonin from the synaptic cleft to maintain neurotransmitter homeostasis. SERT is established as an important target in the treatment of anxiety and depression. Because a high-resolution crystal structure is not available, a computational model of SERT was built based upon the x-ray coordinates of the leucine transporter LeuT, a bacterial NSS homolog. The model was used to develop the first SERT structure-based pharmacophore. Virtual screening (VS) of a small molecule structural library using the generated SERT computational model yielded candidate ligands of diverse scaffolds. Pharmacological analysis of the VS hits identified two SERT-selective compounds, potential lead compounds for further SERT-related medication development.

2.1 Introduction

Signaling between cells in the central nervous system is mediated by the controlled release and reuptake of neurotransmitters in the synapse.1 An excess or deficit of the monoamines serotonin, dopamine or norepinephrine in the synapse has been associated with various psychiatric and neurological disorders including depression,
anxiety, compulsivity, attention deficit hyperactivity disorder (ADHD), substance abuse, Parkinson’s disease and schizophrenia. Control of synaptic monoamine levels is affected by the plasma membrane monoamine transporters (MATs), which terminate the action of these biogenic amines via reuptake into the presynaptic cell. The mechanisms of action of drugs related to the above medical conditions typically involve the MAT proteins. Tricyclic antidepressants (TCAs) such as imipramine (Tofranil\textsuperscript{TM}), developed in the 1950s, alleviate depression by blocking serotonin and norepinephrine transporters (SERT and NET, respectively), thereby extending the lifespan of synaptic serotonin and norepinephrine. Unfortunately, the TCAs also block adrenergic, muscarinic acetylcholine and histamine receptors, responsible for a plethora of adverse effects. Selective serotonin reuptake inhibitors (SSRIs), the next generation of antidepressants led by fluoxetine (Prozac\textsuperscript{TM}) in the 1980s, carry far fewer adverse effects compared to the TCAs. The SSRIs block SERT, but because the resultant surge of serotonin can activate any of 14 serotonin receptor types, this drug class is not without its own adverse effects.

The driving force for MAT uptake of monoamine substrate is electrogenic, harnessing the inward Na\textsuperscript{+} gradient across the cell membrane. The SERT, NET and DAT (dopamine transporter) are members of the neurotransmitter:sodium symporter (NSS) family as well as members of a larger group of Na\textsuperscript{+} and Cl\textsuperscript{−} dependent transporters known as the "solute carrier 6" (SLC6) family. The lack of a high-resolution 3-D MAT structure had hindered structure-function and therapeutic development efforts until a breakthrough was achieved in the form of crystallization of the bacterial NSS homolog LeuT, a leucine transporter. The LeuT x-ray structure has provided a template to build
credible MAT computational models.\textsuperscript{16-21} Although model quality increases with sequence identity with the template, structural similarity also plays a significant role.

In the absence of high-resolution 3D structures for SERT, development of a ligand based pharmacophore\textsuperscript{22-25} or QSAR\textsuperscript{26-28} are feasible alternatives to obtain structural information about the binding pocket. Ligand-based approaches analyze a set of ligands and generate possible protein-ligand interaction patterns without knowledge of the protein structure. A limitation of ligand-based approaches is that flexible alignment using dissimilar scaffolds is less reliable; knowledge of the bioactive conformation of at least one active molecule significantly improves alignment accuracy.\textsuperscript{29} These limitations can be overcome by using structure-based approaches, in which diverse scaffolds can be used to capture ligand interactions and the binding pocket environment in general. Docking, a structure-based technique, is capable of reliably predicting the bioactive conformation for a co-crystallized ligand.\textsuperscript{30} In the absence of an experimental 3D structure, however, a reasonable 3D model can be constructed for a receptor using crystallographic or NMR data from genetically and functionally related proteins.\textsuperscript{31, 32}

Comparative modeling correctly predicts the 3D fold of a protein in most cases and has often provided insight into the atomistic details of ligand binding.\textsuperscript{33, 34} Models created by this method have identified potential binding pockets, and have been employed for docking, structure-based pharmacophore generation and virtual screening (VS) to identify new ligands and their interactions with the protein.\textsuperscript{33, 35, 36} Pharmacophore models, defined as a 3D collection of features essential for bioactivity of a ligand, have been widely employed.\textsuperscript{37, 38} The approach of combining a pharmacophore with VS has yielded novel structural scaffolds for ligands of a target receptor.\textsuperscript{39-44}
Application of this technique with the DAT led to the identification of structurally novel VS "hit" compounds which, upon further modification, created a more potent analog through rational drug design.\textsuperscript{33, 43, 45} The present work describes the development of a structure-based pharmacophore and its VS utilization to identify novel ligand structural scaffolds that display SERT selectivity.

### 2.2 Materials and Methods

Molecular modeling studies were performed out using dual-core 3.06 GHz iMac and 2.66 GHz quad-core Intel Xeon macpro. Construction and evaluation of homology models were performed using Discovery Studio 2.5.1.\textsuperscript{46} Docking and pharmacophore studies executed using the modules in Molecular Operating Environment (MOE 2008.09); Chemical Computing Group Inc., S. S. W., Suite #910, Montreal, QC, Canada, H3A 2R7.) $[^{125}]$-RTI-55 and $[^{3}]$H-serotonin were purchased from PerkinElmer (Foster City, CA). The VS hit compounds were purchased from enamine and molport. The N2A-hDAT cell line was a gift from Dr. Margaret Gnegy (University of Michigan). HEK293 cells stably transfected with hSERT or hNET were prepared in collaboration with Dr. Mads Larsen and Dr. Susan Amara (University of Pittsburgh, Pittsburgh, PA).
2.2.1 Computational Methods

2.2.1.1 Sequence Retrieval and Integrity Check of Template

The FASTA sequence of hSERT was downloaded from the Uniprot database (accession number P31645). The X-ray crystal structure of the *Aquifex aeolicus* protein LeuT (2a65) was used as the template and downloaded from the protein data bank (PDB). The template was checked for any missing residues/atoms using the “protein reports and utilities” module in DS 2.5.1. The protein report generated is a compilation of information about the resolution at which a protein was crystallized, the cell and space group of the crystal, atoms/residues missing the coordinates. In addition, it also provides a list of active sites, number of ligands, ions and solvent molecules cocrystallized with the biomolecule. This report can be viewed as concise portrayal about the defects in a crystal structure like nonstandard naming, structural disorder, atom connectivity and incorrect bond order of the amino acid residues. The report indicated that the residues Asn-133 and Ala-134 of the second extracellular loop (EL2) in LeuT (2a65) were missing the coordinates. As EL2 was noticed to play an important role in transport by forming the extracellular lid along with EL4 the coordinates of the missing residues in the loop (EL2) were built into the template in DS. Since the modeled residues were in a loop, the most flexible regions in a protein, EL2 was subjected to the loop refinement protocol in DS to optimize the generated loop region.
2.2.1.2 Different Sequence Alignments

SERT models were created even before LeuT crystallization, which is the closest homologue to MAT family. Even though different alignments were developed using LacY and other proteins as templates, the sequence alignments using LeuT as the template gained more attention due to its close evolutionary relationship with MATs (Table 2.1). Of these alignments, Yamashita et al. proposed the first alignment\textsuperscript{16} using Psi-BLAST\textsuperscript{49} with minor manual adjustments. No evolutionary information about MATs was used in generating this alignment and was generated only for the following human transporters--glycine, gamma-amino butyric acid, dopamine and serotonin leaving out norepinephrine transporter from the alignment. Following this, Beuming et al.\textsuperscript{50} proposed a comprehensive sequence alignment using 177 eukaryotic and 167 prokaryotic protein sequences belonging to NSS family. The alignment considered evolutionary conservation pattern along with wealth of biological data available for NSS proteins. It considerably differed in TM 4,5,9 regions along with differences in EL 2,3,4 loops compared to Yamashita alignment, and was the most used alignment for model generation by different groups. Recently Celik et al.\textsuperscript{51} built SERT models using the proposed Yamashita’s and Beuming’s alignment along with an in-house developed alignment. Even though Celik’s alignment was not as comprehensive as Beuming’s; it considered residues Asp 98, Ala 169 and Ile 172 be pointing into the pocket and not into the lipid-facing region. None of the alignments proposed so far explicitly mentioned about modeling of the missing residues in EL2 that played a substantial role in substrate translocation. This study involved generation of another alignment by using “align sequence to template” protocol in DS. This protocol considers structural restraints into
account in the search for common motifs among the sequences. Though all these alignments looked similar in TM regions, minor differences are present in loop regions that influence studies performed to study dynamic aspects of these proteins. The addition of missing residues to LeuT used in alignment modified the sequence identity by a marginal extent (24%) compared to other alignments, but it is regarded generally as the more the sequence identity with the template used the better the model generated\textsuperscript{52}. The objective of using all four alignments was to identify which alignment produces a model suitable for our study.

<table>
<thead>
<tr>
<th></th>
<th>Manepalli alignment</th>
<th>Celik alignment</th>
<th>Beuming alignment</th>
<th>Yamashita alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Sequence identity</td>
<td>23.9</td>
<td>22.9</td>
<td>20.5</td>
<td>20.0</td>
</tr>
<tr>
<td>% Sequence similarity</td>
<td>46.7</td>
<td>45.6</td>
<td>42.6</td>
<td>42.4</td>
</tr>
</tbody>
</table>

Table 2.1. Percent sequence identity and sequence similarity among different alignments used to build SERT models.
2.2.1.3 Construction of SERT Homology Models

Different hSERT models were then created using the Align2d methodology\textsuperscript{53} and the three previously proposed sequence alignments.\textsuperscript{16, 50, 51}

2.2.1.3.1 hSERT Model 1: Manepalli et al. Alignment

The sequences of hSERT and modified LeuT were aligned using the "align sequence to templates" protocol in DS. An alignment similar to Celik et al.\textsuperscript{51} was generated, with minor manual modifications further increasing the sequence identity with LeuT. N- and C-terminal overhangs were deleted, increasing sequence identity to 24\% and sequence similarity to 47\%. The hSERT model sequence began at Arg-79 in the N-terminal tail and ended with Lys-605 in the C-terminal tail (Figure 2.1). The "build homology model" protocol in DS, which employs MODELLER version 9v4,\textsuperscript{31} was used to construct 20 hSERT models. The models were further refined and ranked by discrete optimized protein energy (DOPE) score and the model with the best DOPE score was designated as “Model 1”.

Figure 2.1. Alignment of LeuT<sub>α</sub> and hSERT amino acid sequences used to build Model 1. Identical (blue) and similar (yellow) aligned residues are indicated. The dots and stars correspond to residues lining S1 and S2 pockets respectively. Transmembrane and connecting loop topology is displayed above the sequences. Gaps relative to the other sequence are represented by dots. Position of the residue in its respective chain is indicated before and after each line of sequence.
2.2.1.3.2 hSERT Model 2: Beuming et al. Alignment

The hSERT query sequence and modified LeuT were loaded into DS and aligned as described by Beuming et al. \(\text{(Figure 2.2).}\) Terminal residues not having the matching template coordinates at both ends were omitted, and 20 models were built and assigned DOPE scores as described above. Models based on this alignment resulted in EL2 positioned directly above the substrate permeation pore, which would likely have influenced ligand docking. The dangling EL2 was adjusted using the DS loop refinement protocol. Models were then evaluated using the "verify protein" protocol, which assesses the score of each residue in the known 3D environment. This was followed by superposition of the hSERT model onto LeuT to check for TM domain coincidence. The final version was designated "Model 2".
**Figure 2.2.** Alignment of LeuT<sub>Aa</sub> and hSERT amino acid sequences following Beuming alignment. Identical (blue) and similar (yellow) aligned residues are indicated. Transmembrane and connecting loop topology is displayed above the sequences. Position of the residue in its respective chain is indicated before and after each line of sequence.
2.2.1.3.3 hSERT Model 3: Yamashita et al. Alignment

The hSERT query sequence and modified template were aligned as proposed by Yamashita et al. (Figure 2.3). N- and C-terminal overhangs were deleted, leaving 527 hSERT residues. Twenty models were created and the best model resulting from this alignment was designated as "Model 3".

Figure 2.3. Alignment of LeuT<sub>Aa</sub> and hSERT amino acid sequences following Yamashita alignment. Identical (blue) and similar (yellow) aligned residues are indicated.
2.2.1.3.4 hSERT Model 4: Celik et al. Alignment

Twenty hSERT models were built with DS as described above but following the proposed Celik et al. alignment\(^5^1\) (Figure 2.4), which offered a sequence identity match comparable to the Manepalli alignment and greater than the Beuming and Yamashita alignments. A 527-residue hSERT polypeptide was used to develop “Model 4”, as described for Model 1.

Figure 2.4. Alignment of LeuT\(_{\alpha}\) and hSERT amino acid sequences following Celik alignment. Identical (blue) and similar (yellow) aligned residues are indicated.
2.2.1.4 Model Refinement and Validation

Five models among the twenty homology models created from each alignment were chosen based on Discrete Optimized Protein Energy (DOPE) scores. It is a score generated by MODELLER, which is an atomistic-based statistical potential for model evaluation and structure prediction. This conformational energy measure reflects the stability of a model relative to other models generated. Following this, the stereo chemical quality of the five preferred models was assessed with the validation server that is used to evaluate structures deposited into PDB. This server uses NUCheck, PROCHECK and SFCHECK programs to check various stereo chemical features that symbolize a structure stability. From each group of five, two models were then chosen based on their having the fewest outliers on the phi-psi angle Ramachandran plot. The chosen two models from each alignment were evaluated for 3D fold (tertiary structure), using the “verify protein” protocol in DS. This protocol reduces 3D structures to a simplified 1D representation, compares to the 1D amino acid sequence, and assigns a score. This score measures how compatible a three-dimensional fold is with its one-dimensional sequence. If the score calculated is close to the expected high it implies that generated model has an optimal 3D-fold respective to its sequence and lower scores highlight regions that are poorly modeled and their fold is not optimal to its amino acid sequence. The models selected from the previous evaluation parameters were checked for TM domain assembly and ligand binding pocket volume by superposing the SERT models onto the LeuT crystal structure using MOE (Figure 2.5).
Figure 2.5. Polypeptide backbone superposition of the four homology SERT models. Good overall spatial overlap was observed for Models 1 (Manepalli et al.; red), 2 (Beuming et al.; blue), 3 (Yamashita et al.; yellow) and 4 (Celik et al.; green), with essentially perfect overlap of TM domains. EL2 (arrow) was the most divergent region among the four models.
A single hSERT model was selected from each of the groups resulting in four final models that were generated using different alignments. The all-atom force field AMBER99\textsuperscript{60, 61} was used to add hydrogen atoms and partial charges to the models. Keeping the non-hydrogen atoms fixed in space, hydrogen atoms were minimized with a convergence criterion of 0.05 kcal/mol. The final hSERT models chosen from each alignment were further optimized to minimize the outliers and avoid unnecessary protein contacts. Of the four models, Model 1 had the smallest RMSD on superposition with LeuT and maximal ligand binding pocket volume.

### 2.2.1.5 Ligand Binding Site Identification

The alpha site finder function of MOE 2008.09 was used to identify possible ligand binding sites in Model 1. The site finder is purely geometric; no energy terms are used. Alpha spheres within the hSERT model that were too exposed to solvent were eliminated; spheres that corresponded to locations of tight atomic packing in the model were retained. Spheres were classified as either "hydrophobic" or "hydrophilic" depending on the potential for hydrogen bonding. Following the default settings, alpha spheres were clustered to produce a collection of putative binding sites that are ranked according to the number of hydrophobic atoms within contact distance of the spheres. The sites corresponding to the S1 and S2 pockets were identified and dummy atoms were created in the pockets to specify the algorithm to dock the ligands. As the S2 pocket is a huge cavity exposed towards the extracellular space, a sphere with a radius of 8 Å was created to constrain docking poses in keeping with recent reports (Figure 2.6).\textsuperscript{18, 21}
Figure 2.6. A sphere (red) of 8 Å to constrain docking poses in extracellular exposed S2 pocket. hSERT displayed as blue TM helices.
2.2.1.6 Docking of Known Antidepressants

Ligand coordinates for sixteen antidepressants (Figure 2.7) were obtained from the small molecule crystal database CCSD (Cambridge Crystallographic Structural Database).\textsuperscript{62, 63} Ligands having no coordinates were drawn using MOE. Partial charges and hydrogen atoms were added to all ligands using Merck Molecular Force Field 94X (MMFF94x).\textsuperscript{64-66} Ligand geometry was optimized via energy minimization with a conjugated gradient truncated Newton optimization algorithm with convergence criteria = 0.01 kcal/mol, $\varepsilon = 1$. Leucine or selected TCA drugs were first docked with the LeuT model to test for consistency with the reported cocrystal structures.\textsuperscript{16, 21} Atom placement methods including proxy triangle, alpha triangle and triangle matcher were employed, as were refinement techniques using grid-based and force field-based methods. The combination that produced poses very close to the crystal structure (<1Å) was chosen for hSERT docking. Interaction plots for all docked poses of different antidepressants in hSERT were generated using the MOE ligand interaction module.\textsuperscript{67} A consensus of residues and their predicted ligand interactions was developed and used in creating the pharmacophore features for the vestibular ligand-binding pocket.
Figure 2.7. Chemical structures of SERT inhibitors (antidepressants) used to develop the pharmacophore.
2.2.1.7 Development of 4-point SERT S2 Pharmacophore

A pharmacophore feature was generated at each amino acid side chain – ligand interaction point obtained from initial docking iterations. Pharmacophore points F1, F2, F3 (each with a radius of 1 Å) and F4 (radius = 1.5 Å) were delineated to better encompass the hydrophobic features of the docked ligands. An excluded volume was created, and the algorithm assigned a penalty to poses that encroached on this space. The final pharmacophore product was tested using a database containing known hSERT ligands and decoy ligands. The pharmacophore model successfully discriminated the classic hSERT ligands from the decoys (Table 2.2).

<table>
<thead>
<tr>
<th>Compound</th>
<th>Affinity dG Score (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine</td>
<td>-9.23</td>
</tr>
<tr>
<td>Sertraline</td>
<td>-9.05</td>
</tr>
<tr>
<td>Paroxetine</td>
<td>-8.92</td>
</tr>
<tr>
<td>Citalopram</td>
<td>-8.75</td>
</tr>
<tr>
<td>Fluvoxamine</td>
<td>-8.24</td>
</tr>
<tr>
<td>Zimeldine</td>
<td>-8.07</td>
</tr>
<tr>
<td>Venlafaxine</td>
<td>-7.67</td>
</tr>
<tr>
<td>Trazadone</td>
<td>-7.19</td>
</tr>
<tr>
<td>Imipramine</td>
<td>-6.87</td>
</tr>
</tbody>
</table>

Table 2.2. List of top ten compounds (binders) identified in the evaluation of pharmacophore.
2.2.1.8 Virtual Screening of ZINC Database

An 18 million entry “all-purchasable” subset of the ZINC database containing was downloaded. The subset was “washed” by importing the collection into MOE, which removes unwanted salts, ions and disconnected molecular fragments from the database. The Lipinski’s Rule of Five descriptor available in MOE was applied to the database to remove nondrug-like molecules. The resulting database was then searched for compounds containing toxic groups and eliminated by using toxic moiety descriptor in MOE. From this database a subset containing approximately one million compounds was selected for further use. The tautomers and protomers of these entries were generated at physiological pH. Using the “import conformation” module of MOE, different conformations were generated and minimized using the MMFF94x force field; conformations with strain energy 4 kcal/mol or more above the lowest energy conformation were discarded. This final database was screened using the 4-point SERT S2 pharmacophore. “Hit” compounds acceptable to the pharmacophore filter were further screened to ensure the presence of halogens and a molecular weight < 350 daltons, features common to most of the known antidepressants. These steps yielded a pool of 4,097 hit compounds. Three hSERT docking iterations were performed with this pool of structures. Affinity scores < -6.0 kcal/mol plus visual inspection was used to condense the pool to 68 candidate ligands (Figure 2.8). Of these, 10 of the top-ranked 15 compounds [(coded “SM-1” through “SM-15”), Figure 2.9] were procured from various sources and pharmacologically tested.
Figure 2.8. Different sequential filters used in structure-based virtual screening of SERT.
Figure 2.9. Chemical structures of first batch of hits identified.
2.2.2 Experimental Methods

The experimental (pharmacological assays) were performed by Ms. Laura M. Geffert from Dr. Christopher K. Surratt lab from Mylan School of Pharmacy, Duquesne University.

2.2.2.1 In vitro VS Compound Binding Screen.

Compounds identified from the virtual screen were purchased and dissolved in 100% DMSO to a concentration of 10 mM. Initial one-point competition binding assays were conducted using 10 μM final concentration of VS hit compound and 0.1 nM of the radiolabeled cocaine analog [125I]-RTI-55, a radioligand with high affinity at all three MATs. Stably transfected HEK293-hSERT membranes or N2A-hDAT or HEK293-hNET whole cells were employed for all binding assays. Non-specific binding was determined by inclusion of 10 μM paroxetine, mazindol or desipramine for hSERT, hDAT or hNET assays, respectively. Screening results were analyzed with one-way ANOVA (P < 0.05) with a post-hoc Dunnett's Multiple Comparison Test.

2.2.2.2 hSERT Membrane Binding Assay

hSERT binding affinities were obtained by displacement of [125I]-RTI-55 in membrane binding assays. Membrane was prepared from stable hSERT-HEK cells grown at 37°C in a 5% CO₂ environment on 150 x 25 mm plates. At 95% confluence (3 days of growth), cells were washed twice with 10 mL cold phosphate-buffered saline (DPBS). An additional 10 mL of DPBS was added and cells were harvested by scraping and transferred to cold centrifuge tubes (15 mL), followed by centrifuging for 10 min at low
speed (700 x g). After removal of the supernatant, the cell pellet was resuspended in 500 uL cold TE buffer (50 mM Tris, 1 mM EDTA, pH 7.5). Following centrifuging for 30 min at 100,000 x g at 4°C, the supernatant was discarded and the pellet was frozen for later use or resuspended in ice-cold binding buffer (50 mM Tris, pH 7.5, 100 mM NaCl). Each sample was analyzed for protein content using the Bradford protein assay. For competition binding, membrane was incubated with [\(^{125}\text{I}\)]-RTI-55 (0.1 nM concentration) and increasing concentrations of cold competitor (1 nM to 1 mM), or 10 μM paroxetine to measure non-specific binding. Reactions were carried out in 12 x 75 mm borosilicate glass tubes with gentle shaking at room temperature for 1 hour and terminated by rapid filtration through GF/B filters (Schleicher and Schuell, Keene, NH) presoaked in 0.5% polyethylenimine solution (v/v). Filters were washed twice with 5 mL cold 50 mM Tris buffer and transferred to vials. Radioactivity was determined using a Beckman gamma counter. For saturation binding assays, data were analyzed with GraphPad Prism 5.0 software. IC\(_{50}\) values were generated and converted to K\(_i\) values using the Cheng-Prusoff equation (K\(_i\) = IC\(_{50}\)/ (1 + ([RTI-55]/K\(_d\) RTI-55)).

2.2.2.3 hDAT and hNET Whole Cell Binding Assays.

Whole-cell competition binding assays were performed for hDAT and hNET using stable hDAT-N2A or hNET-HEK293 cell lines, respectively, grown at 37°C in a 5% CO\(_2\) environment. Cell monolayers were grown in 24-well plates to >90% confluence. Cells were washed twice with 1 mL of KRH buffer (25 mM HEPES, pH 7.3, 125 mM NaCl, 4.8 mM KCl, 1.3 mM CaCl\(_2\), 1.2 mM MgSO\(_4\), 1.2 mM KH\(_2\)PO\(_4\), 5.6 mM glucose) supplemented with 50 mM ascorbic acid (KRH/AA). Cells were incubated for 15 minutes
with 1 nM $[^{125}]\text{I}$-RTI-55 supplemented with tropolone (total volume of 500 uL) along with increasing concentrations of drug or 10 M mazindol (hDAT) or 10 M desipramine (hNET). Following incubation, cells were washed twice with 1 mL KRH/AA buffer, then treated with 1 mL 1% SDS with gentle shaking at room temperature for 1 hour. Cell lysates were transferred into vials. Radioactivity was determined using a Beckman gamma counter and analyzed using GraphPad Prism 5.0 as described above.

2.2.2.4 Serotonin Uptake Inhibition Screen.

One-point competition binding assays were performed using stably transfected HEK293-hSERT whole cells grown to >90% confluence on 24-well plates. Cells were washed with KRH buffer, then preincubated with 10 M VS hit compound for 10 min, followed by addition of 10 nM $[^{3}]\text{H}$-serotonin. After an additional 5 min, cells were washed twice, then treated with 1% SDS and shaken for 1 hr. Lysates were transferred to 5 mL scintillation fluid and radioactivity was detected using a Beckman liquid scintillation counter. Non-specific radioligand uptake was determined in the presence of 10 M clomipramine.
2.3 Results and Discussion

2.3.1 Model Construction and Refinement

Four discrete amino acid sequence alignments with LeuT were generated that primarily differed in the loop regions. Manual refinement of the alignments was conducted to remove gaps within the TM regions. Rotamers of the residue side chains lining the hSERT halogen-binding pocket were selected based on their ability to form a sub-pocket, as these residues have been reported to be important for binding of certain SSRIs. A total of 80 SERT homology models were built using Discovery Studio (DS) 2.5.1. Five models from each alignment (20 models total) having the lowest DOPE score were chosen for further evaluation. The models were also deposited into the PDB evaluation server and assessed with respect to stereochemical quality. The plots generated by the server were evaluated for unusual phi-psi angles, bond lengths, bond angles and van der Waals contacts (Figure 2.10). Two models from each alignment were selected for further evaluation, specifically their fitness in a 3D environment assessed by Profiles-3D. This DS module assesses the protein’s tertiary structure compatibility with its sequence. Finally, spatial overlap of the model and template polypeptide backbone atoms was examined to check the integrity of the model's TM domains and the volume of its binding pocket. On this basis, "Model 1" was chosen for further studies.
Figure 2.10. Ramachandran plot analysis of the stereo chemical parameters of model 1. Residues Arg 312 (EL 3) and Val 379 (TM 7) were the outliers and are far from the S2 binding pocket.
2.3.2 Docking and Pharmacophore Feature Generation

The MOE alpha site finder was used to identify possible ligand sites in Model 1. A site exposed to the extracellular space was chosen based on evidence for inhibitors binding in this vestibule.\textsuperscript{21} The combination of proxy triangle as the placement method and force field refinement with an affinity scoring function yielded MOE docking poses of leucine into LeuT with an RMSD of 0.31 Å. The same combination was employed to dock the 16 structurally diverse antidepressants that possess a large range of SERT binding affinities (0.1 nM - 9100 nM). It is acknowledged that certain SSRI ligands used in the docking have been reported to access the interior S1 substrate pocket\textsuperscript{71}; however, the VS pocket employed in this study contained elements of the vestibular (S2 substrate) pocket and the halogen-binding pocket (HBP) (Figure 2.11). The latter pocket is defined in part by hSERT residues Leu-99 (TM 1), Trp-103 (TM 1) and Ile-179 (TM 3). The ligand-accessible region of the HBP (green cloud) is adjacent to that of the S2 pocket (orange cloud); the HBP appears to be important for recognition of halogen-containing inhibitors. Poses in which the halogen atom of the ligand was directed toward this pocket were thus favored.
Figure 2.11. **Ligand-accessible regions of SERT substrate and inhibitor binding pockets.** Regions available to ligands in the S1 (blue), S2 (orange) and HBP (green) pockets are displayed as surfaces. The pocket employed for virtual screening combined the S2 and HBP. Residues defining each pocket are displayed as sticks (atomtype color). For clarity, the TM1 (red), TM3 (orange), TM6 (green), TM8 (cyan) and EL4 (pink) helices are highlighted.
2.3.3 Development of 4-point SERT S2 Pharmacophore

Overlaying the highest-ranked poses of different antidepressants a “common interaction pattern” was generated. Residues identified as common interaction points were then chosen to develop a 3D pharmacophore. Pharmacophore feature F1 of radius 1 Å was created midway between Arg-104 (TM 1) and Glu-493 (TM 10) to capture possible ionic or H-bond interactions, particularly with protonated ligand nitrogen atoms. Previous reports of SERT-based pharmacophore models indicated the importance of a positively charged nitrogen atom. Feature F2 (radius 1 Å) was created in the vicinity of Lys-490 (TM 10) to capture possible H-bond interactions. The lysine side chain is among the more flexible of the naturally occurring amino acids; 78 rotamers were possible for Lys-490. The rotamer oriented toward the vestibular pocket, analogous to Asp-401 in LeuT, was chosen to capture possible interactions with ligands. Asp-401 is reported to interact with different antidepressants in LeuT crystal structures; thus, Lys-490 would be expected to play a similar role in hSERT. Feature F3 (radius 1 Å) was created to capture H-bond or cation-π interactions with Tyr-107 (TM 1). Residues 104-106 of hSERT TM 1b are analogous to LeuT residues 30-32, although hSERT Tyr-107 and LeuT Val-33 differ substantially regarding the nature of potential ligand interactions. Feature F4 (radius 1.5 Å) in the HBP was designed for hydrophobic interactions believed to be important in recognizing SSRIs. Residues lining the pocket are conserved between hSERT and LeuT. Substitution of the hSERT TM 3 residue Ile-179 (Ile-111 in LeuT) with alanine, phenylalanine or aspartic acid increased SSRI substrate uptake inhibition potency by 7 – 1080 fold. The TM 1 mutations Ala77Gly in hNET and Ala81Gly in hDAT were observed to improve SSRI affinity, suggesting a negative influence of the methyl side
chain of alanine on SSRI binding. Finally, an exclusion volume was added to make the pharmacophore more stringent and selective (Figure 2.12).

Figure 2.12. Four-point pharmacophore for SERT S2 pocket displayed along with excluded volume. The pharmacophore employed for VS was comprised of features (colored spheres). F1: Donor/acceptor (red). F2: Acceptor (pink). F3: Cation/acceptor (blue). F4: Hydrophobic / π ring (green) and excluded volume (grey).
2.3.4 Virtual Screening and Pharmacological Testing

Approximately one million members of the ZINC small molecule structural library database\(^{68}\) were screened with the 4-point S2 SERT pharmacophore for possible "hit" compounds. A total of 4097 molecules were acceptable to the pharmacophore filter. Docking of this subset into the SERT model followed by monitoring of affinity dG scores with visual inspection of the resulting poses yielded 68 preferred compounds. Due to lack of availability, only ten of the top 15 compounds (labeled SM-1 to SM-15) were initially tested for hSERT, hDAT or hNET affinity in single point in vitro binding assays. Two compounds, SM-10 and SM-11, at 10 μM significantly displaced the cocaine analog \(^{[125]}\text{I}-\text{RTI-55}\) from hSERT membranes, indicating measurable hSERT affinity. At the same final concentration, these two compounds did not displace the radioligand at hDAT or hNET cells. (Figure 2.13) SM-10 and SM-11 were further characterized at hSERT HEK293 membranes in competition assays to obtain \(K_i\) values for inhibition of \(^{[125]}\text{I}-\text{RTI-55}\) binding. The binding affinity values for SM-10 and SM-11 towards hSERT are \(38 \pm 17\) μM and \(17 \pm 7\) μM respectively with SM-11 binding stronger than SM-10 (Figure 2.14). SM-10 and SM-11 were also screened for the ability to inhibit uptake of \(^3\text{H}\)-serotonin by intact hSERT HEK293 cells. Because no inhibition of \(^3\text{H}\)-serotonin uptake was detected at the 10 μM final concentration employed for these two VS hit compounds IC\(_{50}\) values were not obtained (Figure 2.15).

But surprisingly SM-14 (Figure 2.16) was found to be DAT selective by displacing \(^{[125]}\text{I}-\text{RTI-55}\) at hDAT cells alone. A DAT binding affinity \(K_i\) value of \(15.6 \pm 2.4\) μM (mean ± s.e.m.) and a dopamine uptake inhibition potency IC\(_{50}\) value of \(10.5 \pm 4.6\) μM were observed. (Figure 2.17 and Figure 2.18) Given the high sequence similarity (~
50%, as calculated with BLAST) among the three-hMATs, finding a specific inhibitor that targets a single transporter is a challenging task. VS is more precarious with the additional inhibitor binding pocket uncertainty associated with a homology model built using a template sharing < 30% sequence identity. Because of this handicap, and because the goal of the study was to identify SERT ligands of novel structural scaffold, a very high concentration of VS hit compound relative to the SERT ligand radiotracer (100,000-fold) was employed in the pharmacologic screen. The observation that most of the VS hits showed no ability to displace the radioligand at the SERT-HEK cells confirms that the huge molar ratio of VS compounds: radiotracer was itself insufficient for decreasing radiotracer binding (i.e., some SERT affinity is required to register a response in this assay).

The two VS hits pharmacologically verified as SERT ligands displayed micromolar $K_i$ values. These binding affinities are very low compared to the nanomolar or better SERT affinities of certain SSRI therapeutics. Inhibition of serotonin uptake was not observed for any VS hit compound at the 10 μM screening concentration. Substrate uptake inhibition potency is often lower than binding affinity for a given MAT inhibitor, possibly due to differences in MAT conformational or population requirements for the two processes. The 10 μM SM-10 or SM-11 concentration employed may be above the threshold for detecting SERT ligand binding but below the threshold for detecting serotonin uptake inhibition.

Micromolar affinities for VS hit compounds are common when homology models are employed. From this research team, the first reported VS effort based on a DAT homology model yielded a hit compound (MI-4) with low micromolar DAT affinity and
slightly better (high nanomolar) SERT affinity. Without modification, this compound displayed antipsychostimulant and antidepressant properties (Ref. 43 and unpublished data). The same VS effort using a DAT model yielded a hit compound (MI-17) with a SERT $K_i$ value of 284 nM. An MI-17 analog designed to increase SERT selectivity resulted in a compound with a SERT $K_i$ value of 37 nM, and SERT: DAT and SERT: NET selectivity ratios of 50 and over 200 respectively. Thus, VS hit compounds with micromolar affinities appear to be useful as lead compounds for medication development.
Figure 2.13. **VS hit compound in vitro MAT binding screen.** Compounds at 10 μM final concentration were tested for the ability to inhibit [125I]-RTI-55 binding at hSERT HEK293 cells (top panel), hDAT N2A neuroblastoma cells (middle panel), or hNET HEK293 cells (bottom panel). Non-specific binding was assessed by the presence of 10 μM paroxetine (PAR), mazindol (MAZ) or desipramine (DSP) for SERT, DAT and NET, respectively. Data represent n=3 independent experiments performed in duplicate. Data are presented as the mean ± s.e.m. and were analyzed by one-way ANOVA with Dunnett’s multiple comparison post-hoc test. *P < 0.05 vs. total binding for that assay. ****P < 0.0001 vs. total binding for that assay. (Provided by Ms. Laura M. Geffert)
Figure 2.14.  Saturation binding of VS hit compounds. SM-10 (closed circles) and SM-11 (open circles) at hSERT HEK293 cells. Binding affinity Kᵢ values were determined via displacement of $[^{125}\text{I}]-\text{RTI-55}$. n=3 experiments. (Provided by Ms. Laura M. Geffert)
Figure 2.15. VS hit compound in vitro serotonin uptake inhibition screen. Compounds at 10 μM final concentration were tested for the ability to inhibit [3H]-serotonin uptake by hSERT HEK293 cells. Non-specific radioligand uptake was assessed by the presence of 10 μM paroxetine (PAR). Data represent n=3 independent experiments performed in duplicate. Data are presented as the mean ± s.e.m. and were analyzed by one-way ANOVA with Dunnett’s multiple comparison post-hoc test. *P < 0.05 vs. total uptake for that assay. **P < 0.001 vs. total uptake for that assay. (Provided by Ms. Laura M. Geffert)
Figure 2.16. Structures of hits identified through virtual screening SM-10, SM-11 and SM-14.
Figure 2.17. Saturation binding of VS hit compounds SM-14 at hDAT N2A cells. Binding affinity $K_i$ values were determined via displacement of $[^{125}\text{I}]$-RTI-55. (n=3 experiments). (Provided by Ms. Laura M. Geffert)
Figure 2.18. Dopamine uptake inhibition potency of VS hit compound SM-14 at hDAT N2A cells. Uptake inhibition Binding affinity $K_i$ values were determined via displacement of $[^{125}\text{I}]-\text{RTI}-55$. (n=3 experiments). (Provided by Ms. Laura M. Geffert)
2.4 Structural Novelty of SM Hit Compounds

Given their structural flexibility, SM-10, SM-11 and SM-14 greatly differ from the SSRIs and TCAs that were employed to develop the pharmacophore (Figure 2.16). The Tanimoto similarity coefficient was calculated for the three hits to identify the dissimilarity compared to the known inhibitors (SSRI, TCA). A coefficient of less than 0.6 was calculated, which signify the structural novelty of these hits. More specifically, the SSRIs and TCAs currently being used contain either rigid or fused rings when compared to the flexible SM hits. This flexibility would provide the additional advantage of reaching the higher affinity substrate pocket. The similarities between the SSRIs and SM hits are the presence of a charged amine group and halogen atoms, which are required for SERT inhibition. In addition, the flexibility of these molecules with well separated rings presents additional attach points to optimize SERT affinity and can be developed as multi site drugs.

2.4 Conclusion

To our knowledge this is the first study to identify novel SERT-selective reuptake inhibitors through VS performed with an hSERT homology model. Conventional drug development usually involves creation of analogs of established ligands for a given receptor. In contrast, the VS approach affords the possibility of discovering ligands with very different scaffolds, ligands that in all likelihood would never be otherwise identified for that receptor. In the case of the SERT, inhibitors of novel scaffold provide new
opportunities for development of rationally designed multi-site drugs (e.g., SERT/5-HT$_2A$/5-HT$_3$ blockers) that address the therapeutic goal while mitigating the adverse effects associated with the SSRI drug class. It is anticipated that the two SM compounds described herein should serve as lead compounds toward the treatment of depression, anxiety, and other serotonin-related disorders.

2.5 Future Directions

Lead optimization is the next step after lead identification in drug discovery. In this study two novel scaffolds selective to SERT were identified. Of the two leads, SM-10 has a better affinity but it has three stereo centers. Purchasing the different possible stereoisomers and testing them seems to be the next logical step. As the leads were low affinity binders, hit-to-lead optimization has to be performed to identify higher affinity analogues. Even though these leads were identified through S2 pocket virtual screening, the location of the binding pocket and the relative orientation of the ligand in the transporter are speculative. As these novel ligands are flexible an assumption can be made as part of these ligands may reach S1 pocket with rest of the molecule docked into S2. So the studies that identify the binding orientation and modification of the scaffold extend this study.
References


55. [http://deposit.rcsb.org/validate](http://deposit.rcsb.org/validate)


59. Molecular Operating Environment (MOE); Chemical Computing Group Inc., S. S. W., Suite #910, Montreal, QC, Canada, H3A 2R7.


70. Yung-Chi, C.; Prusoff, W. H., Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochemical pharmacology* **1973**, *22*, (23), 3099-3108.


Chapter 3.

ELUCIDATION OF BINDING PROFILE SIMILARITIES AMONG STRUCTURALLY DIVERSE INHIBITORS OF DOPAMINE TRANSPORTER

3.1 Introduction

Dopamine transporter (DAT), norepinephrine transporter (NET) and serotonin transporter (SERT) belong to the family of monoamine neurotransmitters (MATs). They play an important role in maintaining the neurotransmitter levels that control mood, hunger etc.\(^1\) Imbalance in the levels of neurotransmitters is the primary cause for several psychological disorders like addiction, depression etc. Various experimental studies performed imply that DAT is primarily responsible, for the reward/reinforcing properties associated with psychostimulant substances like cocaine and amphetamines.\(^2\)-\(^5\)

Very little is known about how these addictive substances interact with DAT and block dopamine reuptake. Structurally, diverse inhibitors interact with DAT by binding to different non-overlapping binding sites.\(^6\),\(^7\) To identify the exact location of binding pocket, photoaffinity probes of dissimilar scaffolds were developed. In the presence of ultra violet light, they form a covalent bond with the amino acid residues in DAT, which is then identified by mass spectrometric studies (Figure 3.1).

Analogues of the tropane-based\(^8\)-\(^{17}\) with very flexible piperidine\(^{18}\) and piperazine analogues\(^{19}\)-\(^{24}\) were developed and used as photo probes (Figure 3.2) with less focus on non-tropane based analogues.\(^{25}\)-\(^{27}\) Photoaffinity probes based on the tropane analogues
covalently bond to residues far from the binding pocket and may not provide the precise location. So, the probes with photoaffinity groups substituted directly on the inhibitor were developed to provide a much better picture.
Figure 3.1. Schematic representation of the photoaffinity labeling of a protein. The target protein (DAT) is represented by a blue rectangle, photoaffinity ligand (red triangle) and a photoactive group (green star). Upon irradiation by ultra violet light, photoactive group on photoaffinity ligand forms a covalent bond with DAT. The protein is lysed and the polypeptide fragments are analyzed by mass spectrometry to identify the location of covalent linkage.
Tropane analogs

MFZ-2-24

RTI-82

Piperazine-based analogs

DEEP

Figure 3.2. Structures of tropane- and piperazine-based DAT photoaffinity ligands.
3.2 Methods

Numerous analogues synthesized in the scheme towards final photoprobe were also included to better understand the structure activity relationship (interaction pattern) with DAT. The 3D-coordinates of parent compounds and analogues of the three diverse scaffolds (groups) were drawn in molecular operating environment (MOE). Partial charges and hydrogen atoms were added to all ligands using Merck Molecular Force Field 94X (MMFF94x). The geometry of these ligands was then refined using energy minimization module with a convergence criteria of 0.01 kcal/mol in MOE.

A DAT homology model previously generated was employed to dock these compounds. As the binding pocket of these inhibitors was unknown and only extracellular pocket (S2) was accessible to all the analogues, the focus of the study involved docking these ligands into the S2 pocket. Moreover, S2 site was proposed as a staging area for inhibitors before sliding further into the S1 pocket.

The biomolecular force field, assisted model building with energy refinement (AMBER 99) was used to add partial charges and hydrogen atoms to the homology model of DAT. The previously proposed S2 pocket was chosen as the site to dock ligands, and a sphere of 8 Å radius was constructed in this pocket.

Three different docking runs were performed for each group, which included the parent compound and their analogues. The separate databases generated for each group were then merged together into a single database. The compound database was then provided as an input for the protein ligand interaction fingerprint (PLIF) panel in MOE. The PLIF tool generates a population histogram that represents the protein-ligand interaction pattern. PLIF is a computational tool to encapsulate the interaction pattern by
generating fingerprints of docked poses. A fingerprint is a sequence of binary bit strings that associates different structural features of a molecule, and is a commonly employed method in various ligand-based approaches.\textsuperscript{35-37}

### 3.3 Results and Discussion

All analogs towards the synthesis of final photoactive probes for dissimilar scaffolds were also included in the study to understand the structure-activity relationship (SAR).\textsuperscript{25-27} An assumption was made that all these diverse inhibitors bind in the same pocket, and were docked into the S2 pocket of DAT.

The homology model of DAT used in this study was constructed using the occluded conformation of LeuT (pdb id:2a65) with the primary substrate pocket (S1) completely blocked by the extracellular gates.\textsuperscript{38} As the S2 pocket is large and exposed towards the extracellular space, a sphere was created to constraint the docking poses in DAT. Docking runs were performed to generate various docked poses in MOE.

The fingerprints were then generated for each docked pose using PLIF analysis. It considers six different non-bonded interactions [side chain hydrogen bonds (donor/acceptor), back bone hydrogen bonds (donor/acceptor), ionic interactions and surface contacts] into account to generate a fingerprint for each pose. Clustering of these fingerprints was then performed to generate a histogram that showed the interaction pattern of each docked pose.
**Bupropion:** A number of bupropion analogues with varying substitutions on aromatic ring were synthesized (see Figure 3.3). It was perceived that halogen substituted analogues had higher DAT affinity, which implies that bulky substitutions were well tolerated. It was also noticed that the parent compound bupropion (441 nM) is seven fold stronger than its photoprobe analog, SADU-3-72 (3071 nM) in DAT N2A neuroblastoma cells. The fingerprint analysis (Figure 3.5) identified that the terminal amino group interacts with Asp 312 (TM6) and the azide group forms an ionic bond with Asp 475 (TM10), which wasn’t detected for bupropion (Figure 3.4). A clear trend of increased interaction profile with Asp 475 was identified (Figure 3.5) for the photoprobe. The trend suggested that the highly reactive nitrene generated upon UV light would covalently link to Asp 475.

![Figure 3.3. Structures of bupropion and its photoactive probe.](image)

(±) R¹ = -Cl, R² = -H, Bupropion

(±) R¹ = -I, R² = -N₃, SADU-3-72

**Figure 3.3. Structures of bupropion and its photoactive probe.**
Figure 3.4. Two-dimensional docking interaction plot of the bupropion photoprobe. The hydrophobic residues (green spheres) and hydrophilic residues (pink spheres) in the pocket are displayed. The azide group (N₃) interacts with Asp 475 demonstrating the possibility of covalent bond formation with the residue.
Figure 3.5. **PLIF panel of bupropion and photoprobe docked poses.** Each docked pose is indicated by a row, black rectangles in each row represent the interacting residues. A residue may interact with a docked pose in six different ways, with more than one rectangle observed for a single residue. The red bracket represents bupropion poses and the blue bracket indicates the photoprobe docked poses. The panel clearly shows the increased interaction profile of photoprobe with Asp 475.
**Pyrovalerone:** The differences between the bupropion and pyrovalerone were very minimal with a great difference in affinity. Bupropion consists a secondary nitrogen, where as the pyrovalerone contains a tertiary nitrogen. The other minor difference was the pyrovalerone had a longer alkyl chain (n-propyl group) compared to the shorter methyl chain in bupropion. The larger pyrovalerone is fifty fold stronger than bupropion in DAT.

Substitution on the aromatic ring of pyrovalerone was also better tolerated as in bupropion without any appreciable loss of affinity (Figure 3.6). The pyrovalerone photoprobe was designed by substituting the 4’-methyl group with photoreactive azide (-N₃) and iodine (I) at the 3’ position. The bulky substitutions on the final photoprobe result in ten-fold less affinity compared to pyrovalerone.

![Pyrovalerone structure](image)

<table>
<thead>
<tr>
<th>Aromatic substituent</th>
<th>Compound</th>
<th>^3H-WIN binding inhibition Kᵢ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R¹ (3’) R² (4’)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H Me</td>
<td>Pyrovalerone</td>
<td>8 ± 2</td>
</tr>
<tr>
<td>I N₃</td>
<td>Photoprobe</td>
<td>78 ± 18</td>
</tr>
<tr>
<td>H NH₂</td>
<td>Amino pyrovalerone</td>
<td>5 ± 1</td>
</tr>
<tr>
<td>I NH₂</td>
<td>Amino, iodo-pyrovalerone</td>
<td>28 ± 8</td>
</tr>
</tbody>
</table>

**Figure 3.6. DAT binding affinity of pyrovalerone and its analogues.** (Provided by Dr. David J. Lapinsky)
The five-fold affinity difference between the amino substituted pyrovalerone and the dual substituted pyrovalerone suggest the detrimental effect of bulky iodine group. Similar to the bupropion photoprobe, the pyrovalerone photoprobe also showed an increased interaction pattern with Asp 475. As the scaffolds of bupropion and pyrovalerone were similar, the ionic interaction with Asp 475 is no strange observation. It is proposed that the photoprobes based on these scaffolds would form a covalent bond with the charged residue.

**Methylphenidate:** It is structurally larger and diverse compared to bupropion and pyrovalerone. The iodo and azido-substitution on the same ring motif, may lead to a low efficient photoprobe for this scaffold, and so methylphenidate was hybridized to better accommodate the iodo and azido groups on separate rings without a loss in affinity (Figure 3.7). The fingerprint analysis of methylphenidate showed no H-bonding pattern due to the lack of H-bond groups. The PLIF analysis identified that the methylphenidate and iodo-substituted analogs interact with Thr 472 and Asp 475 residues.

Much like the other two scaffolds, the halogen substitution on the aromatic ring was well tolerated on this scaffold. The para-azido substitution on the hybrid analogues had better affinity compared to the ortho and meta-substituted analogs. The azido group substituted on the hybrid analogues formed a H-bond with Ser 308 (TM6) and an ionic interaction with Asp 312 (TM6). Based on the PLIF pattern identified, it is speculated that the nitrene generated will covalently link to either one of these residues.
(±) X = -H, Methylphenidate
(±) X = 4-I, 4-I-methylphenidate
(±) X = 3-I, 3-I-methylphenidate

(±) p-N$_3$-4-I = p-N$_3$-N-Bn-4-I-methylphenidate
(±) m-N$_3$-4-I = m-N$_3$-N-Bn-4-I-methylphenidate
(±) o-N$_3$-4-I = o-N$_3$-N-Bn-4-I-methylphenidate
(±) p-N$_3$-3-I = p-N$_3$-N-Bn-3-I-methylphenidate
(±) m-N$_3$-3-I = m-N$_3$-N-Bn-3-I-methylphenidate
(±) o-N$_3$-3-I = o-N$_3$-N-Bn-3-I-methylphenidate

<table>
<thead>
<tr>
<th>Substitution</th>
<th>Compound</th>
<th>$^3$H-WIN binding inhibition $K_i$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Methylphenidate</td>
<td>25 ± 1</td>
</tr>
<tr>
<td>4-I</td>
<td>4-I-methylphenidate</td>
<td>14 ± 3</td>
</tr>
<tr>
<td>3-I</td>
<td>3-I-methylphenidate</td>
<td>4.5 ± 1</td>
</tr>
<tr>
<td>p-N$_3$-4-I</td>
<td>p-N$_3$-N-Bn-4-I-methylphenidate</td>
<td>363 ± 28</td>
</tr>
<tr>
<td>m-N$_3$-4-I</td>
<td>m-N$_3$-N-Bn-4-I-methylphenidate</td>
<td>2754 ± 169</td>
</tr>
<tr>
<td>o-N$_3$-4-I</td>
<td>o-N$_3$-N-Bn-4-I-methylphenidate</td>
<td>517 ± 65</td>
</tr>
<tr>
<td>p-N$_3$-3-I</td>
<td>p-N$_3$-N-Bn-3-I-methylphenidate</td>
<td>658 ± 70</td>
</tr>
<tr>
<td>m-N$_3$-3-I</td>
<td>m-N$_3$-N-Bn-3-I-methylphenidate</td>
<td>2056 ± 73</td>
</tr>
<tr>
<td>o-N$_3$-3-I</td>
<td>o-N$_3$-N-Bn-3-I-methylphenidate</td>
<td>1112 ± 163</td>
</tr>
</tbody>
</table>

Figure 3.7. Structures and activity profile of methylphenidate analogues. (Provided by Dr. David J. Lapinsky)
3.4 Conclusion

Numerous DAT photoaffinity probes with diverse scaffolds were developed for identifying the inhibitor-binding pocket in DAT. They include tropane-based and piperazine-based analogues with a very long flexible linker, separating the azide group and the pharmacophore. These scaffolds have an inherent problem with the covalent bond forming far from the binding pocket. This defeats the purpose of the scaffolds, and so the search for probes with the photoactive groups directly on the scaffold or with a short tether would help in better identification of the binding pocket of DAT inhibitors.

The PLIF analysis of bupropion photoprobe showed the increased possibility of covalent bond formation with the Asp 475. It was noticed that the pyrovalerone photoprobe also showed a similar interaction pattern as the bupropion photoprobe. This is no strange observation being the two scaffolds were very similar. The methylphenidate and its hybrid analogs were bigger compared to the probes of other scaffolds and so a difference in interaction pattern was observed. It was identified that Ser 308 and Asp 312 residues were closer to the azido group, and the methylphenidate probe may form a covalent bond with either one of these residues.
References


Chapter 4.

RATIONAL DESIGN OF NOVEL ORTHO-AMIDO DIPHENYLAMINES AS SELECTIVE MEK5 INHIBITORS THROUGH DEVELOPMENT OF A MEK5 HOMOLOGY MODEL

4.1 Introduction

Kinases are important cellular proteins that catalyze the transfer of a terminal phosphoryl group from adenosine triphosphate (ATP) to specific hydroxyl groups of serine, threonine or tyrosine residues.\(^1\) Mutations of these residues to acidic counterparts (often called DD, EE, or DE mutants) render kinases with significantly elevated activity, commonly observed in several cancers. Mitogen activated protein kinase [(MAPKs), also called as extracellular signal regulated kinases (ERKs)] belong to the family of Ser/Thr kinases. The MAPK pathway consists of a series of numerous interconnected signal transduction pathways that transfer extracellular responses into the cell nucleus, to control several cellular responses such as embryogenesis, proliferation and cell death (Figure 4.1).\(^2\)\(^-\)\(^7\)

The human genome encodes seven MEK enzymes that regulate the activity of MAP kinase pathways. Aside from the highly similar MEK1/MEK2, other members of the family consist of MEK3, MEK4, MEK5, MEK6 and MEK7. Of all these proteins, MEK5 present as an attractive target for therapeutic intervention, because of the detection of elevated levels in squamous cell carcinoma\(^8\), prostrate\(^9\) and breast cancers.\(^10\)
Drug resistance to the approved anticancer agents like imatinib, gefitinib and erlotinib has become a growing concern.\textsuperscript{11} To address this, another interesting concept is in the developmental stages, where inhibitors with more specificity and minimal side effects are being developed. The inhibitors attack the targets downstream ensuring that only certain functions are eliminated in the signaling cascade. These inhibitors are called non-classical kinase inhibitors, where they occupy a pocket allosteric to the substrate ATP pocket, there by becoming effective against only a selected panel of kinases.\textsuperscript{12} Two studies of MEK5 inhibitors were reported. One study involves the development of oxindole derivatives\textsuperscript{13} as classical kinase inhibitors (ATP-competitive); where as the other study\textsuperscript{14} reported the development of non-classical (allosteric) inhibitors. It is assumed that the allosteric inhibitor binding induces a conformational change in the unphosphorylated kinase (MEK) by locking it into a catalytically inactive conformation.\textsuperscript{3}
Figure 4.1. The signaling cascade of MEK family.\textsuperscript{14} (Adapted with permission from Flaherty et al. BMCL 2010,20, 2892-2896.)
4.2 Methods

4.2.1 Retrieval of sequence

Three different isoforms- A, B and C of human MEK5 were identified (Figure 4.2).\textsuperscript{15-17} The splice variant B (referred to as MEK5 β further) is the canonical sequence; and longer (448 residues) than the remaining isoforms, expressed primarily in cytosol. The isoforms, “A” and “C” were comparatively shorter, with isoform A being the shortest. The fasta sequence of human MEK5 β (accession number Q13163) was downloaded from the Uniprot database\textsuperscript{18}. MEK5 β has a conserved motif “S\textsuperscript{311}xxxT\textsuperscript{315}” which is the phosphorylation, and activation motif. Mutagenesis of Ser 311 and Thr 315 to Asp residue constitutively activates MEK5, whose levels are observed to be significantly elevated in different types of cancer. By mutating these two residues to aspartate in the downloaded primary sequence a constitutively active mutant was generated (Figure 4.3).\textsuperscript{16}
<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>MNMLALGPFFPAMEQVVLVIRIKIPNSGAVDVTSHGPQQLLFDVLVIGQVPPEATITAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>MNMLALGPFFPAMEQVVLVIRIKIPNSGAVDVTSHGPQQLLFDVLVIGQVPPEATITAF</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>MNMLALGPFFPAMEQVVLVIRIKIPNSGAVDVTSHGPQQLLFDVLVIGQVPPEATITAF</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>EYDEEDGRDITVSSDEEMKAMLQYYSTVMEQQVNQVPLQIFPRACKPFGERTNIHGL</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>EYDEEDGRDITVSSDEEMKAMLQYYSTVMEQQVNQVPLQIFPRACKPFGERTNIHGL</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>EYDEEDGRDITVSSDEEMKAMLQYYSTVMEQQVNQVPLQIFPRACKPFGERTNIHGL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>KVNTAGPSQHSSFAPVSDLPSNLSKQLSAELKILANGQREQDIRYRDGLTGHHNGGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>KVNTAGPSQHSSFAPVSDLPSNLSKQLSAELKILANGQREQDIRYRDGLTGHHNGGT</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>KVNTAGPSQHSSFAPVSDLPSNLSKQLSAELKILANGQREQDIRYRDGLTGHHNGGT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>YKAYVPSGKIALAVVIIIDITIEQIMSELEIIYKCDSYYIGFYGAFVVENRISIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>YKAYVPSGKIALAVVIIIDITIEQIMSELEIIYKCDSYYIGFYGAFVVENRISIC</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>YKAYVPSGKIALAVVIIIDITIEQIMSELEIIYKCDSYYIGFYGAFVVENRISIC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>TEFMGGSLDVYRKMPEHVGRIAVAVVGLTLYLWNLILKRDVFSPMLVNTKQVKLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>TEFMGGSLDVYRKMPEHVGRIAVAVVGLTLYLWNLILKRDVFSPMLVNTKQVKLC</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>TEFMGGSLDVYRKMPEHVGRIAVAVVGLTLYLWNLILKRDVFSPMLVNTKQVKLC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>DFVSQTLVNSIAKTVGTNAAYMAPERISEQVYHSVDWLSLGISEFELALGFPYIPIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>DFVSQTLVNSIAKTVGTNAAYMAPERISEQVYHSVDWLSLGISEFELALGFPYIPIQ</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>DFVSQTLVNSIAKTVGTNAAYMAPERISEQVYHSVDWLSLGISEFELALGFPYIPIQ</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>KNQGSLMPQQLQVCIQEDAVEFVPLVGEFGEFVFHSVITTCMQKPYERPAPREELMGPFIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>KNQGSLMPQQLQVCIQEDAVEFVPLVGEFGEFVFHSVITTCMQKPYERPAPREELMGPFIV</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>KNQGSLMPQQLQVCIQEDAVEFVPLVGEFGEFVFHSVITTCMQKPYERPAPREELMGPFIV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>hMEK5-Isoform-B</th>
<th>QFNDGAAVSMWVCRAEERSSQGiFPP----------</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMEK5-Isoform-A</td>
<td>QFNDGAAVSMWVCRAEERSSQGiFPP----------</td>
</tr>
<tr>
<td>hMEK5-Isoform-C</td>
<td>QFNDGAAVSMWVCRAEERSSQGiFPP----------</td>
</tr>
</tbody>
</table>

**Figure 4.2.** Multiple sequence alignment of different isoforms of human MEK5. Consensus of residues among the three isoforms displayed, stars represent identical residues, dots represent dissimilar residues, and gaps shown by dashes.
Isoform B (wild type)

MLWLALGPFPAMENQVLVIRIKIPNSGAVDWTVHSGPQLLFRDVLDVIGQVLPEA
TTTAFYEDEDEDGDRITVRSDEEMKAMLSYYYSTVMEQQVNGQLIEPLQIFPRACK
PPGERNIHGLKVNRAGPSQHSSPAVSDLPSNSLKSSAELKKILANGQMNEQDI
RYRDTLGHLNGGTGYKAYHVPSGKILAVKVILLDLITLQKQIMSELELYKCDSS
YIIGFYGAAFFVENRISICTEFMDDGSLDVYRKMPEHVLGRIAFAVVKGLTYLWSL
KILHRDVKPSNMLVNTRGQVKLCDFGVSTQLVN

Isoform B DD (constitutively active mutant)

MLWLALGPFPAMENQVLVIRIKIPNSGAVDWTVHSGPQLLFRDVLDVIGQVLPEA
TTTAFYEDEDEDGDRITVRSDEEMKAMLSYYYSTVMEQQVNGQLIEPLQIFPRACK
PPGERNIHGLKVNRAGPSQHSSPAVSDLPSNSLKSSAELKKILANGQMNEQDI
RYRDTLGHLNGGTGYKAYHVPSGKILAVKVILLDLITLQKQIMSELELYKCDSS
YIIGFYGAAFFVENRISICTEFMDDGSLDVYRKMPEHVLGRIAFAVVKGLTYLWSL
KILHRDVKPSNMLVNTRGQVKLCDFGVSTQLVN

Figure 4.3. Differences between the wild type MEK5 isoform β and its constitutively active mutant. The dissimilarities are indicated with pink colored residues.
4.2.2 Search for the template

Seven members have been identified that belong to the MEK family. The phylogenetic analysis reveals that MEK5 is more closely related to MEK1 with 48% sequence identity (Figure 4.4). A search for the MEK1 crystal structures identified 14 different crystal structures in PDB (accessed Dec 2010). The protein with pdb id (3eqc) was selected as the template and the coordinates were downloaded and loaded into Discovery studio 2.5.1. (DS)
Figure 4.4. Phylogenetic analysis of the members of MEK family. MEK1, MEK2 and MEK5 belong to a group, where as the rest of the family members are more evolutionarily related to each other.
4.2.3 Protein check of the template

The template was checked for missing residues/atoms (protein quality) using protein reports and utilities module in DS. The report identified missing residues at both the polypeptide terminals, gaps between residues Ala 220 and Val 224, gaps between Cys 277 and Pro 307 in the template. As the missing residues at both the terminals and the residues between Cys 277 and Pro 307 in the alpha helix domain were far from the binding pocket, no effort was taken to build the 3D coordinates. The residues between Ala 220 and Val 224 in the activation loop were built and refined by the loop refinement protocol in DS. Of the five residues modeled, only Phe 223 had two different plausible conformations generated (Figure 4.5). The conformation pointing into the allosteric pocket is labeled as Phe 223_in and the one pointing towards p-loop is labeled as Phe 223_out.
Figure 4.5. The two different conformations of Phe 223 generated in MEK1 through loop modeling. The amino acid residue is displayed as ball and sticks. The conformation pointing into the allosteric pocket is labeled as Phe 223_in (orange) and the conformation pointing up towards p-loop is labeled as Phe 223_out (green).
4.2.4 Sequence Alignment

Using default settings on the clustalW2 server, the amino acid sequences of active mutant MEK5 βDD and modified 3eqc were aligned (Figure 4.6). The two important motifs DFG and HRD were made sure to align between the sequence and template.

Figure 4.6. Sequence alignment of MEK5 β DD and MEK1 used to build homology models. Conserved residues between MEK1 (top) and MEK5 β DD (bottom) are highlighted in yellow; similar residues highlighted in red. Secondary structure elements are displayed on top of the sequences with helices as coils, beta-strands as blue arrows and loops as black lines. Gaps relative to the other sequence are shown as dots. Residues are numbered and indicated at the beginning and end of each line. Stars at the bottom identify the phosphorylation residues: Ser 311 and Thr 315 mutated to Asp and black boxes highlight the conserved domains HRD and DFG.
4.2.5 MEK5 model construction

The two different templates that varied in the Phe 223 conformation were used to construct homology models. The model building resulted in the generation of two different MEK5 β DD models, which differed in the orientation of Tyr 316, an analogous residue to Phe 223 in MEK1. Both the templates were used to construct homology models, as there was no knowledge of the preferred conformation among the MEK family and with the loop being very flexible. Ten models of MEK5 β DD were built using each template resulting in the creation of a total of twenty models. A model was picked from each set based on the discrete optimized protein score (DOPE) and 3D profile scores for further docking calculations.

4.2.6 Docking

Various diphenyl amino analogues tested at MEK1 and MEK5 were used to comprehend the interaction pattern of allosteric inhibitors with MEK5 (unreported). The coordinates of these ligands were drawn using builder in molecular operating environment (MOE) and partial charges, hydrogen atoms were added using Merck molecular force field 94X (MMFF94x). The geometry of ligands was optimized using energy minimization module with a convergence criteria of 0.01 kcal/mol in MOE. The coordinates of Mg$^+$ ion were copied from the template MEK1 to MEK5 models to mimic the natural cellular environment. Three independent docking runs of the diphenylamine analogues were performed at both the homology models of MEK5 β DD.
4.3 Results and Discussion

4.3.1 Homology Modeling

MEK5 belongs to the family of Ser/Thr kinases and so the phosphorylation motif residues, Ser 311 and Thr 315 present in the activation loop were both mutated to Asp, to mimic the constitutively active mutant MEK5 β DD. The rotamers that mirrored the coordinates of Ser and Thr were selected; and keeping the rest of residues fixed, the new coordinates were energy minimized in MOE. As MEK5 is more evolutionarily related to MEK1, the search for MEK1 in PDB resulted in identification of 14 X-ray structures.

The pdb id (3eqc) was chosen as the crystal structure was obtained at a higher crystallographic resolution. The protein report identified missing residues in the template with a gap between the residues Ala 220 and Val 224. The missing residues Asn 221, Ser 222 and Phe 223 present in the activation loop may play a significant role in the allosteric inhibition and ERK phosphorylation, as they were closer to the allosteric pocket. The other reason to construct the 3D coordinates is the uncertainty in the different conformations adopted by Phe 223 revealed by fourteen different MEK crystal structures (Figure 4.7).
Figure 4.7. Superposition of the fourteen MEK1 crystal structures. The protein, MEK1 is displayed as wires with ATP displayed as sticks (atom type color), allostERIC inhibitor is exhibited as ball and sticks (atom type color) and Mg$^{2+}$ ion as brown sphere. Different conformations adopted by Phe 223 residue present in the activation loop show the substantial loop movement.
The ambiguity in the Phe 223 conformation and missing residues next to the allostERIC pocket in the template led the effort to construct the 3D coordinates. The coordinates of these residues were built using the loop refinement protocol available in DS. Of the three residues modeled, the protocol generated two possible different conformations with equal probability for Phe 223, and a single optimal conformation for rest of the residues.

The templates with different conformations of Phe 223 were selected based on a low DOPE score. DOPE is an atomistic based statistical potential for the model evaluation and structure prediction. This score is viewed as a conformational energy that measures the relative stability of one conformation with respect to others. Therefore, the two templates with different rotamers of Phe 223 were used to construct the models of MEK5 β DD.

By aligning the conserved motifs (HRD and DFG) followed by nonconserved domains between MEK1 and MEK5, a sequence alignment was generated. Even though the overall sequence identity (~40%) is less, there is a higher homology (~80%) in the binding pockets. The sequence alignment identified Tyr 316, to be an analogous residue to Phe 223 present in the activation loop (Figure 4.7). Even though both residues were aromatic, Tyr 316 may form additional H-bond and cation-π interactions with allostERIC inhibitors.

Using the sequence alignment and two MEK1 templates, ten models of MEK5 active mutant were constructed for each template. So a total of ten MEK5 models with Tyr 316 pointing into the allostERIC pocket and ten models with Tyr 316 pointing up towards the p-loop were generated. Two models with favorable DOPE scores were selected from
each group. The Ramachandran plots (Figure 4.8) of these two models in both the groups were analyzed to pick a final model with a least number of outliers, resulting in two final MEK5 models with different conformations of Tyr 316. The all-atom force field, assisted model building with energy refinement (AMBER 99) was used to add partial charges and hydrogen atoms to MEK5 homology models.26,27
Figure 4.8. Sequence alignment of activation loop of MEK1 (top) and MEK5 β DD (bottom). Identical residues are highlighted by yellow and similar residues by red. Black boxes represent the conserved motifs. The analogous residue to Phe 223 (MEK1) is Tyr 316 (MEK5) is represented by a blue sphere at the bottom.
Figure 4.9. Ramachandran plot of the final MEK5 β DD homology model. The outliers in the model that are away from the binding pockets are represented by plus signs with non-outliers as green and yellow spheres.
4.3.2 Docking

A series of diphenylamines with amine side chain modification were synthesized previously, and tested for inhibition of relative phosphorylation at ERK1/2 and ERK5 (unreported) (Table 4.1). The ability of diphenylamine analogues to block *in vivo* EGF-induced ERK1/2 and EKR5 phosphorylation was measured using this assay. This assay is purely a qualitative measurement of inhibition of relative phosphorylation of ERK5 induced by mutagens.

The 3D structures of diphenylamine analogues were built and refined using MOE. The homology models of MEK5 β DD were then loaded in MOE. The magnesium ion present in the MEK1 crystal structure was added to both MEK5 final models to mimic the natural cellular environment. As there is a higher homology in the allosteric pockets between MEK1 and MEK5; and PD0325901 being a diphenylamine analogue, a reasonable assumption was made that the newly designed analogues will also bind in a similar fashion to MEK5. So the co-crystallized MEK1 inhibitor, PD0325901 was also added to both the models to guide the docking studies. As the activation and p-loops were very flexible, both models of MEK5 β DD were used further to dock diphenylamine analogues.

Three independent docking runs were performed keeping the side chains of pocket residues in 6Å relaxed and backbone atoms fixed. As the allosteric pocket is bound and formed by mostly flexible loops, docking was modeled by allowing the side chain flexibility. Of the seven analogues docked, iodine atom on the ligand with the shortest side chain was found to interact electrostatically with the backbone carbonyl of Ile 225.
The amide side chain was observed to form H-bond with Lys 195, Asp 301 and with hydroxyl group of Tyr 316 (Figure 4.9).

The piperazine, morpholine and piperidine analogues oriented in a similar fashion with heterocyclic rings forming hydrophobic stacking with Tyr 316. The rest of the analogues with beta-hydroxyl side chain, and amino acid substitutions were observed to form favorable interactions in the pocket. Even though the last three analogues were computationally favorable, no inhibition of relative phosphorylation was observed experimentally at both ERK1/2 and ERK 5. The reason may be due to the pocket being hydrophobic and the analogues being hydrophilic. Careful visual inspection of all the docked poses of analogues and scores showed that Tyr 316_in conformation was favored over the Tyr 316_out conformation. Therefore the MEK5 homology model with Tyr 316_in conformation was chosen as the final model for further studies.

As the diphenylamine analogues were active at both MEK1 and MEK5, and some compounds showed selectivity towards MEK5 (Table 4.1), the structures of both proteins were superposed to identify the differences in the pockets (Figure 4.10). Using the interaction pattern identified by diphenylamine analogues and the structural differences between these two proteins, three different series of inhibitors were designed (Figure 4.11). The first series was to simulate the in vivo Mg ion, the Tyr 316 series was designed to interact with Tyr 316 residue in the activation loop and lock the protein in an inactive conformation, where as the central ring modifications were to find the minimal halogen substitution on rings.
Table 4.1. Various biphenyl analogues of PD0325901 synthesized and tested for inhibition of phosphorylation of the substrate. (Provided by Dr. Patrick T. Flaherty)

<table>
<thead>
<tr>
<th>Analogue (R)</th>
<th>Amide variation</th>
<th>pERK1/2 relative % phosphorylation</th>
<th>pERK5 relative % phosphorylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-NH₂</td>
<td>NH₂</td>
<td>3.19</td>
<td>41.0</td>
</tr>
<tr>
<td>1-methyl piperazine</td>
<td></td>
<td>254.0</td>
<td>63.0</td>
</tr>
<tr>
<td>morpholine</td>
<td></td>
<td>251.0</td>
<td>72.0</td>
</tr>
<tr>
<td>piperidine</td>
<td></td>
<td>233.0</td>
<td>82.0</td>
</tr>
<tr>
<td>β-hydroxy ethyl</td>
<td></td>
<td>1.65</td>
<td>45.0</td>
</tr>
<tr>
<td>(S) serine</td>
<td></td>
<td>218.0</td>
<td>100.0</td>
</tr>
<tr>
<td>(R) serine</td>
<td></td>
<td>174.0</td>
<td>121.0</td>
</tr>
<tr>
<td>blank</td>
<td></td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Figure 4.10. Docking interactions of the shortest side chain diphenylamine analogue with MEK5 β DD. Residues interacting with the ligand (sticks in atomtype color) displayed as lines (atomtype color) and MEK5 is represented as yellow cartoon. The 2D interaction plot shows the ligand interactions in the allosteric pocket.
Figure 4.11. Superposition of crystal structure of 3EQC (grey) over model of MEK5 β DD (red). The allostERIC ligand of 3EQC is shown as sticks. Residues differing in both the catalytic site as well as the allostERIC pocket are highlighted and colored according to the protein chain.
Figure 4.12. Different series of MEK5 inhibitors designed. Mg mimetic is to simulate the Mg ion role, Tyr 316 series is supposed to interact and lock Tyr 316 residue present in the activation loop; central ring modifications are to identify minimal substitution on the ring without substantial loss of activity.
4.4 Conclusion

MEK5 is significantly upregulated in breast and prostate cancers. A series of ATP competitive ligands were developed for kinases, but the mutational strategy played by the cancer downplays these inhibitors. Moreover, ATP competitive ligands were not selective and had failed before reaching preclinical trials. The new focus of developing small molecules for kinases is through allosteric inhibitors thereby stopping the signaling cascade necessary for cancer cell survival. In this study, a homology model of MEK5 β DD was developed and refined. The model was utilized to guide the rational design of allosteric inhibitors through docking and different series of analogues were proposed to inhibit ERK5 phosphorylation.
References


Conclusions and Future Directions

Based on the results presented in this dissertation, several new research objectives are recommended.

The usage of antidepressants in the market is limited by their side effects. New antidepressants with reduced side activity profile are sought. Among the newer generation is the recently approved antidepressant “Vilazodone”. It is a dual acting antidepressant with SERT inhibitory action, and partial agonistic activity at the 5-HT$_{1A}$ receptor. The physiological role of 5-HT$_{1A}$ receptor is to inhibit feedback mechanism that decreases serotonin concentration in the synapse. The blockade of 5-HT$_{1A}$ receptor in addition to SERT results in rapid onset of antidepressant action with less side effect profile.

From a SERT perspective, two selective novel scaffolds (SM-10, SM-11) were identified from virtual screening. Additionally, a DAT selective compound (SM-14) was identified, which could be developed towards treating cocaine addiction. The two SERT selective compounds can serve as promising drug candidates for further drug development, and rational optimization of these two hits could be performed. The substitution of non-halogenated heterocyclic rings with serotonin-like indole moiety could be performed to increase the SERT affinity. Like vilazodone, the two SERT-selective hits could also be rationally optimized as multi-site drugs.

More specifically, SM-10 having the lowest affinity towards SERT has three stereo centers present. The isolation and affinity measurement of different enantiomers of SM-10 could also be performed to identify the higher affinity analogue. Moreover, different combinations of halogen substitution on the phenyl ring could also be identified for optimal SERT inhibition.
Computationally, the hits identified showed an array of interactions with residues in the S2 pocket. Mutagenesis studies on these residues could be carried out in an effort to confirm that these ligands are indeed binding in the S2 pocket as predicted. Specifically, the residues Arg104 (TM1), Tyr107 (TM1), Lys490 (TM10), and Glu493 (TM10) could be mutated to alanine to disrupt the H-bonding potential. Likewise, the halogen binding pocket residues Leu 99 (TM1), Trp103 (TM1) and Ile179 (TM3) could also be mutated to ascertain the importance of this pocket.

Inhibitor binding pocket identification in MATs is a challenging task. The DAT study developed and analyzed protein-ligand interaction fingerprints (PLIF) for three structurally different photoprobés. Preliminary results suggest the role of Ser308, Asp312 and Asp475 residues forming a covalent bond with the photoprobés. Mutational studies of these residues could be performed to ascertain the location of covalent bond formation. Additionally, mass spectrometry studies could be performed to better corroborate with the PLIF and mutational data.

The MEK5 kinase study led to the rational design of several series of inhibitors. The MEK5 inhibitory cellular assay could be performed on the inhibitors to better understand the mechanics of inhibition. Moreover, the biological results could help in the validation and refinement of the MEK5 homology model. The biological activity of the central ring modification series could provide minimal substitution required for MEK5 inhibition. In addition, this series will help in improving the solubility of these analogs by eliminating the hydrophobic halogen groups on the rings.