Spring 2015

Modeling the Binding of Inhibitors/Drugs to the Human Serotonin Transporter

Kalyan Immadisetty

MODELING THE BINDING OF INHIBITORS/DRUGS TO THE HUMAN

SEROTONIN TRANSPORTER

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

Kalyan Immadisetty

May 2015

MODELING THE BINDING OF INHIBITORS/DRUGS TO THE SEROTONIN

TRANSPORTER

By

Kalyan Immadisetty

Approved April 8, and 2015

_____   _____
Jeffry D. Madura, Ph.D.   Jeffrey D. Evanseck, Ph.D.
Professor of Chemistry and Biochemistry   Professor of Chemistry and Biochemistry
(Committee Chair)   (Committee Member)

_____   _____
David W. Seybert, Ph.D.   Christopher K. Surratt, Ph.D.
Professor of Chemistry and Biochemistry   Professor of Pharmacology
(Committee Member)   Mylan School of Pharmacy
                                          (External Reviewer)

_____   _____
Philip P. Reeder, Ph.D.   Ralph Wheeler, Ph.D.
Dean and Professor, Bayer School of   Chair, Department of Chemistry and
Natural and Environmental Sciences   Biochemistry
                                          Professor of Chemistry and Biochemistry

ABSTRACT

MODELING THE BINDING OF INHIBITORS/DRUGS TO THE SEROTONIN

TRANSPORTER

By

Kalyan Immadisetty

May 2015

Dissertation supervised by Dr. Jeffry D. Madura

Human serotonin transporter (hSERT), a membrane protein from the neurotransmitter sodium symporter family, is implicated in depression disorder and has been the primary target of antidepressant discovery research for several decades. Since the currently available antidepressants may cause adverse effects and have several limitations, novel drugs are highly desired. However, the efforts to develop better therapeutics are hampered by the lack of a crystal structure of hSERT. Knowledge of the binding site of the drug and its orientation in the protein is crucial in structure-based drug discovery. We employed a novel computational protocol comprised of active site detection, docking, scoring, molecular dynamics simulations, and absolute binding free energy (ABFE) calculations to elucidate the binding site and the binding mode of a dual hSERT/5HT-1A blocker SSA-426 and our in-house hSERT inhibitor DJLDU-3-79 in

hSERT. Through this approach, we propose that both of these inhibitors bind in the S1 pocket of hSERT and in a similar orientation. This disproves the earlier hypothesis that both these inhibitors bind in the S2 site; however, we are in agreement with the earlier hypothesis that both of the ligands orient similarly. Further, we resolved the ambiguity in binding energies and binding trends of the tricyclic antidepressant drugs clomipramine, imipramine, and desipramine with leucine transporter (LeuT) (a bacterial homologue of hSERT) through relative binding free energy (RBFE) calculations. Based on our RBFE results, we proposed that clomipramine should have the highest affinity for LeuT, followed by imipramine and desipramine. Finally, to achieve accuracy in binding energy estimations and to perform all CHARMM simulations, we developed CHARMM general force field parameters (CGenFF) for fifteen monoamine transporter ligands.

DEDICATION

I dedicate this dissertation to my family for supporting and motivating me throughout. I also dedicate this to my friends who helped me in many ways, especially in the last one year.

ACKNOWLEDGEMENT

in every way possible. Lastly I would like to thank Dr. Philip Reeder, dean of BSNES, for offering me tuition scholarship for spring and fall semesters, 2014.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AADS: Automated Active Site Detection, Docking and Scoring Protocol

ABFE: Absolute Binding Free Energy

ADHD: Attention Deficit Hyperactivity Disorder

ALA: Alanine

CGenFF: CHARMM General Force Field

CLO: Clomipramine

DA: Dopamine

DDM: Double Decoupling Method

DES: Desipramine

DRI: Dopamine Reuptake Inhibitor

FEP: Free Energy Perturbation

GAFF: General Amber Force Field

GLY: Glycine hDAT: Plasma Membrane Human Dopamine Transporter Protein

hNET: Plasma Membrane Human Norepinephrine Transporter Protein

hSERT: Plasma Membrane Human Serotonin Transporter Protein

HTS: High-Throughput Screening

5-HT: 5-Hydroxytryptamine

5-HTR: 5-Hydroxytryptamine Receptor

IL: Intracellular Loop

IMI: Imipramine

LeuT: Leucine Transporter

LIE: Linear Interaction Energy

LRA: Linear Response Approximation

MAT: Monoamine Transporter Protein

MAOI: Monoamine Oxidase Inhibitor

MD: Molecular Dynamics

MDMA: Methylenedioxymethamphetamine

MM-PBSA: Molecular Mechanics-Poisson-Boltzmann Surface Area

MM-GBSA: Molecular Mechanics-Generalized Born Surface Area

NE: Norepinephrine

NRI: Norepinephrine Reuptake Inhibitor

NSS: Neurotransmitter Sodium Symporter

OSP: One Step Perturbation

PDLD: Protein Dipoles Langevin Dipoles

PMF: Potential of Mean Force

QM: Quantum Mechanics

RBFE: Relative Binding Free Energy

RESP: Restrained Electrostatic Potential

RE: Replica Exchange

SAR: Structure Activity Relationship

SF: Scoring Function

SRPG: Smooth Reaction Path Generation

SLC6: Solute Carrier 6

SNRI: Serotonin-Norepinephrine Reuptake Inhibitor

SSRI: Selective Serotonin Reuptake Inhibitor

TCA: Tricyclic Antidepressant

TI: Thermodynamic Integration

TM: Transmembrane

VS: Virtual Screening

1. CHAPTER 1

A REVIEW OF MONOAMINE TRANSPORTER-LIGAND INTERACTIONS

1.1. Introduction

Monoamine neurotransmitters serotonin (5-HT), dopamine (DA), and norepinephrine (NE) present in the central nervous system (CNS) control several functions such as mood, appetite, sleep, sexual desire, memory, aggression and learning [1-6]. Functions of the three monoamines are regulated by the respective transporters (MATs): the human serotonin transporter (hSERT), the human dopamine transporter (hDAT), and the human norepinephrine transporter (hNET), by reuptake of excess monoamines from the synapse into the presynaptic nerve terminal, thereby maintaining homeostasis [7-9]. MATs are integral membrane proteins belonging to the neurotransmitter sodium symporter (NSS) family, and uses an ion gradient to transport the monoamines against the concentration gradient [10]. MATs are associated with several mental disorders such as anxiety, depression, drug addiction, Parkinson's disease, and schizophrenia [11-14]. Hence, MATs are established as important drug targets for the treatment of these mental disorders (targets for >30 clinically approved drugs). MATs are also target for drugs of abuse including cocaine, amphetamine, methamphetamine, and methylenedioxymethamphetamine (MDMA, popularly known as "ecstasy"). Particularly, hSERT is the primary target for depression treatment. hDAT is the target for psychostimulants such as cocaine, amphetamine, and methylphenidate [15-18]. Although

to a lesser extent, several antidepressants and cocaine also bind to hDAT and hNET. The three MATs share a common membrane topology characterized by 12 transmembrane (TM) helices and intracellular N- and C-termini [7, 19]. The size of the three MATs is fairly similar (hSERT, hDAT, and hNET have 630, 617, and 620 residues, respectively).

Despite being important drug targets, little definitive information is known about the structure, function and pharmacology of MATs until the recent availability of the Drosophila dopamine transporter (dDAT) crystal structure [20]. Although earlier biochemical and mutagenesis experiments provided valuable insight into the transporter topology and secondary structure of the MATs, it was only after the determination of the leucine transporter (LeuT) crystal structure in 2005 [21] that real advancement in the computational research of MATs has taken place. LeuT is a bacterial protein from *Aquifex aeolicus,* belonging to the NSS family as MATs. LeuT was used as a template to generate the homology models of MATs [21-23]. LeuT shares only 20-25% sequence identity with the MATs; still, it has proved as a useful template in addressing the structural, functional and pharmacological aspects of the MATs [24, 25].

1.1.1. Monoamine transporter inhibitors and their therapeutic applications

Efforts to develop drugs targeting the MATs for the treatment of mood disorders have increased in the last few decades, with the view that MATs regulate the monoamines in the synapse and maintain homeostasis. Drugs acting at MATs are used for treating disorders such as depression, anxiety, attention deficit hyperactivity disorder (ADHD), and obesity (Table 1.1). The first generation of antidepressants developed targeting MATs were tricyclic antidepressants (TCAs) (imipramine, clomipramine, and

desipramine) [26]. However, they have off-target actions at G protein-coupled receptors and cardiac sodium channels, which cause several adverse effects [27]. Drugs acting on specific targets were next developed. These drugs include selective serotonin reuptake inhibitors (SSRIs) (fluoxetine, fluvoxamine, citalopram, sertraline, and paroxetine), selective norepinephrine reuptake inhibitors (NRIs) (reboxetine and atomoxetine), and selective dopamine reuptake inhibitors (bupropion). SSRIs have very minimal or no affinity for other targets and improved side effect profiles compared to the TCAs [28]. Although NRIs can be used to treat depression, they are primarily prescribed to treat ADHD [29]. Later, dual acting inhibitors were developed, which include serotonin and norepinephrine reuptake inhibitors (SNRIs) (duloxetine and desvenlafaxine), hNET and hDAT inhibitors (nomifensine). SNRIs display better antidepressant efficacy and faster onset of action compared to SSRIs [30]. Lately, triple reuptake inhibitors were also identified (tesofensine) [31]. Drugs targeting hDAT such as amphetamine and its analogs methylphenidate, dextroamphetamine, dextromethamphetamine, and modafinil are used to treat ADHD. Inhibitors to treat cocaine addiction (benztropine and analogs) primarily target hDAT as it is the primary target for cocaine [32]. Bupropion, which is a hNET/hDAT blocker, is used for nicotine treatment and obesity [33].

Table 1.1. MAT inhibitors classified according to their application.

Antidepressants							Drugs of Abuse	ADHD
TCA	SSRI	SNRI	NRI	DRI	NRI/DRI	SRI+		
Imipramine	Fluoxetine	Venlafaxine	Reboxetine	Bupropion	Nomifensine	Trazodone	Cocaine	Amphetamine
Desipramine	Fluvoxamine	Duloxetine	Atomoxetine				Amphetamine	Methylphenidate
Clomipramine	Paroxetine	Milnacipran	Mirtazapine				MDMA	Modafinil
Amitriptyline	Sertraline	Desvenlafaxine	Viloxazine					
Nortriptyline	Citalopram	Sibutramine						
Protriptyline	Escitalopram							
Butriptyline								
Opripramol								
Dosulepin								
Doxepin								
Amoxapine								
Loferamine								

Compounds approved for depression treatment are only presented in here. SR1+ refers to compounds with hSERT inhibition and monoamine receptor agonistic or antagonistic effects as well.

4

Despite the availability of several effective drugs to treat mental disorders, attempts are still being made to develop novel treatments targeting MATs, because of the shortcomings associated with the current drugs. For example, currently available antidepressants have several limitations such as late onset of action and a large fraction of non-responding patients, along with several side effects like weight gain, sexual dysfunction, and disturbed sleep. Efforts are launched to develop novel drugs targeted at MATs with lesser limitations than the current drugs. Several experimental and computational techniques are being used to identify and optimize the novel lead compounds targeting the MATs. However, experimental methods, which demand considerable time, effort, and money, can be complemented using various computational tools. The entire drug discovery process takes on average 10-15 years to release the drug into the market. Efficient utilization of computational tools hastens the process of drug discovery, and also significantly reduces the cost. Computational tools were successfully utilized in the drug discovery process of several marketed drugs like zanamivir [34], HIV protease inhibitors [35], and in the development of new inhibitors for hepatitis C protease [36] and beta-secretase [37].

Computing the binding affinity of protein-ligand is a key aspect of rational drug design. Since the experimental determination of binding affinity is expensive and time-consuming, a wide spectrum of computational free energy methods have been developed to simplify the task. Free energy calculations through quantitative estimation of the interaction between ligands and proteins allow comparing computation to the experiment. Select applications of free energy methods include (a) screening of databases to identify a

novel lead compound, (b) lead optimization, (c) identify the impact of mutation on the binding of a ligand to a protein, (d) compare the differential binding of ligands and drugs to the same target, (e) assist in understanding the transport mechanism of ligands through the membrane proteins, (f) mechanism of inhibition of proteins by several drugs, (g) affinity and selectivity of different ions to proteins, (h) folding and unfolding of proteins, and (i) estimation of membrane-water partition coefficient of drugs.

In rational drug design, free energy calculations are highly used in ranking millions of compounds against a target protein to identify the lead compound. However, scoring functions (SF) used in these screening stages had only mixed success in terms of accuracy because of the approximations involved. Fast yet accurate free energy methods that can be applied in the industrial setting are not yet available. Although explicit methods like free energy perturbation (FEP) are available, they need significant computational resources to achieve appropriate sampling; this makes them infeasible to screen the databases. Preparing the system for explicit free energy calculations, like preparing the topology files, is very time-consuming and requires a skillful and proficient modeler to perform the calculations. Another bottleneck for these free energy methods for using in a pharmaceutical setting is the lack of good force fields for drug-like compounds. Because of these reasons, explicit free energy methods can only be used in the lead optimization and not in the screening stages of drug design. Hybrid methods such as MM-PBSA/GBSA and linear interaction energy (LIE) are not only faster than the FEP methods, but they also offer better accuracies than the SFs. However, they are not fast enough to use in the screening stages of drug discovery. Although significant improvements have been seen in the past decade in terms of accuracy and efficiency,

6

free energy methods are not reliable enough to screen large numbers of compounds in a short period of time. Especially, accurately calculating binding affinities of large biomolecules like membrane proteins is still a challenging task. However, considerable progress has been made in sampling aspects, partly due to improved sampling techniques and also due to the availability of advanced computers. In addition, the theory behind various free energy methods is clearly understood, and errors are better characterized than before.

Identifying a lead compound for a target of unknown structure is a challenging task. Homology models of MATs developed using LeuT provide opportunities to apply free energy methods to study various aspects of MATs, apart from the lead identification and optimization. Performing free energy calculations on homology models of membrane proteins is a challenging task. The chances of failure are high, and it is difficult to predict the source of error. Since the sequence homology between LeuT and MATs is low, homology models should be validated before applying free energy calculations as a predictive tool. To achieve good accuracies, a system should be sampled sufficiently and have good force field parameters. These parameters are highly desirable for both the protein and ligand. Since good force fields for drug-like molecules are not yet available, Amber and CHARMM general force fields (GAFF and CGenFF) can be extended to these molecules to achieve accuracy. Inclusion of polarizability in the force field is highly encouraged. In this review, an overview of free energy methods to compute protein-ligand interactions, along with their advantages and drawbacks are presented. Various free energy studies carried on MATs so far are also reviewed.

1.2. Free Energy Methods for Protein-Ligand Interactions

In this section, I reviewed different free energy methods, their application in drug design and their pros and cons. Free energy methods can be broadly classified into explicit, implicit, and hybrid methods (Fig. 1.1).

Free Energy Calculations

Explicit methods   Implicit methods   Hybrid methods

Relative free energy calculations

Absolute free energy calculations

Alchemical methods (Annihilation and Decoupling Approaches)

Non-Alchemical methods (Pulling methods)
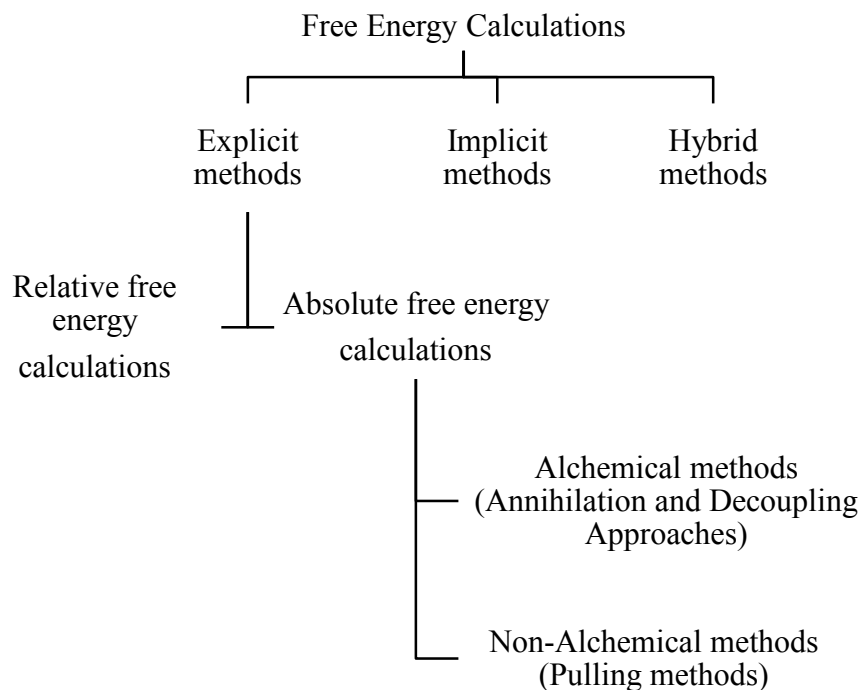
Figure 1.1. Flow chart of free energy methods to compute protein-ligand binding energies.

1.2.1. Explicit methods

Explicit methods, which are capable of quantitatively reproducing the experimental binding affinities, explicitly treat the protein, ligand, and solvent using a molecular force field. These MD based methods extensively sample the system in all degrees of freedom and offer detailed molecular insights [38, 39]. Explicit inclusion of water in these calculations certainly makes the system more realistic, as the role played by water

molecules in the binding of ligands to proteins was illustrated in several crystal structures [40]. However, explicit methods are computationally demanding and slow, making them impractical for computational screening of drugs in rational drug design. Binding energies calculated through these methods are true free energies, as they incorporate both energetic and entropic contributions from all components of the system [41]. Since they simulate experimental conditions (temperature and pressure) in MD, calculated binding energies can be compared to experimental values [42].

Explicit methods can be used to compute both relative (RBFE) and absolute (ABFE) binding free energies. RBFE calculations allow us to compute relative binding of two different ligands to a protein, and also to compute the impact of point mutation on the binding of a ligand. RBFE calculations are alchemical in nature and make use of the FEP and thermodynamic integration (TI) formalism [43]. Starting from the 1980s, alchemical methods have been used to compute RBEs [44-47].

FEP methodology utilizes the thermodynamic cycle to compute protein-ligand binding affinities (Fig. 1.2). Since free energy is a state function, it only depends on the initial and final state and not on the path. To compute the RBFE of ligands A and B with a protein P, ligand A is perturbed to ligand B in bound (ΔG_C) and unbound states (ΔG_D), and the RBFE ($\Delta \Delta G_{AB}$) is calculated using equation 1.

$$\Delta \Delta G_{AB} = \Delta G_A - \Delta G_B = \Delta G_D - \Delta G_C \qquad \text{-Equation 1}$$

ΔG_A and ΔG_B are the absolute binding energies of ligands A and B with a protein P. The same methodology can be used to compute the impact of point mutation on ligand binding, except that in both vertical legs of the thermodynamic cycle (Fig. 1.2) ligand remains the same, but protein varies. FEP method can also be used to compute ABEs.

Care should be taken that the configurational ensembles representing the reference and target state sufficiently overlap to get reasonable accuracy [48]. To circumvent this problem, a coupling parameter (λ) is introduced, and the transformation between initial and final stages is a function of λ. λ varies from 0 to 1, and the entire transformation is divided into several windows. Typically, the number of windows should be chosen such that the free energy change between each window is similar.
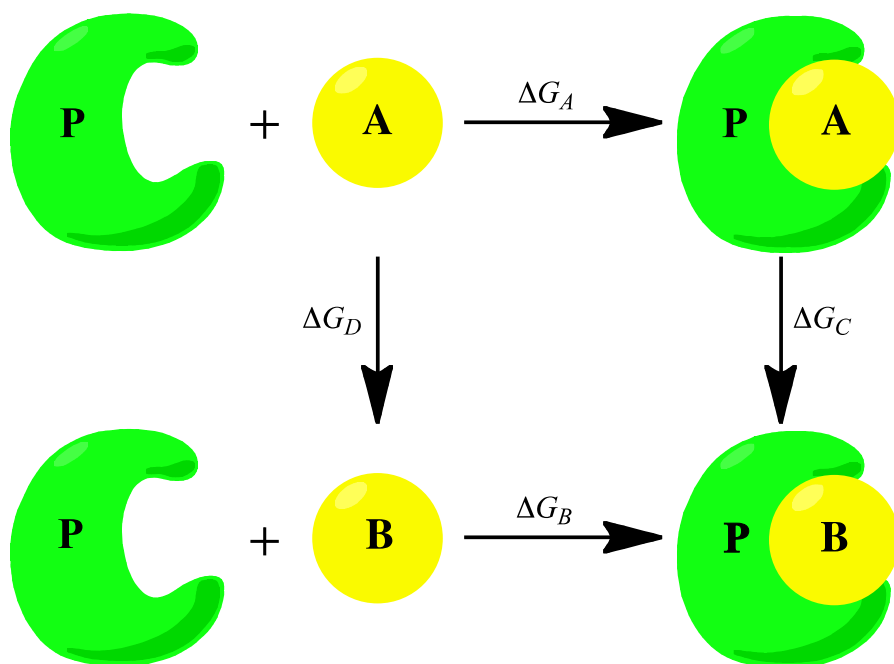
Figure 1.2. Thermodynamic cycle for calculating relative binding energy. ΔG_A and ΔG_B are the absolute binding free energies of ligands (A and B) with the protein (P); ΔG_C and ΔG_D are the free energy changes involved in the transformation of ligand A to B in the bound state and the solvent, respectively.

Several strategies have been designed to handle end-point catastrophes in the dual topology approach. Introduction of a soft-core potential (A modified potential typically used to obtain smoothness in free energy curves) [49] is one among them, and by far the

most elegant method. Limitations of RBFE methods include (a) high computational needs and (b) prior knowledge of at least one complex is required. In addition, it is applicable only to closely related ligands. Although few comparisons were made, generally explicit methods offer better accuracies than SFs [50, 51] and MM-PBSA [52, 53] methods. In most cases, FEP offers accurate RBEs with a mean error lower than 1 kcal/mol. It is much more useful in the lead optimization stages [54] than in the lead identification.

TI methodology is closely related to FEP, and the difference in the free energy between a reference and a target state is an integral over ensemble average of the configurations representative of the respective states governed by λ [55]. Unlike FEP, convergence in TI is represented by the smoothness of the free energy as a function of λ [56]. Both FEP and TI are widely used methods to compute protein-ligand binding affinities.

ABEs can be computed through alchemical and non-alchemical approaches. Alchemical approaches include the widely used annihilation and double decoupling approaches, whereas the non-alchemical approaches are comprised of pulling methods (Fig. 1.1).

Annihilation approaches in principle are FEP [57] and TI [55] based methods and compute binding free energy changes rigorously [43]. These methods make use of the thermodynamic cycle to compute ABEs (Fig. 1.3). In the annihilation approaches, to compute the ABFE (ΔG_{Abs}) of a ligand (L) with a protein (P), two separate MD simulations are performed. The first one has P-L complex in the solvent, and the second one has L alone in the solvent. In simulation one, L is gradually decoupled from the P in the solvent, and the free energy change is ΔG_1. In the second MD simulation, L

11

interactions with the solvent are gradually turned off, and the free energy change is ΔG_2.

ABFE (ΔG_{Abs}) is calculated as

$$\Delta G_{Abs} = \Delta G_2 - \Delta G_1 \qquad \text{- Equation 2}$$

Double annihilation [58] approach had problems incorporating the standard states, and it was later modified in the closely related double decoupling method (DDM) [59] and applied successfully to compute ABEs [60-62].
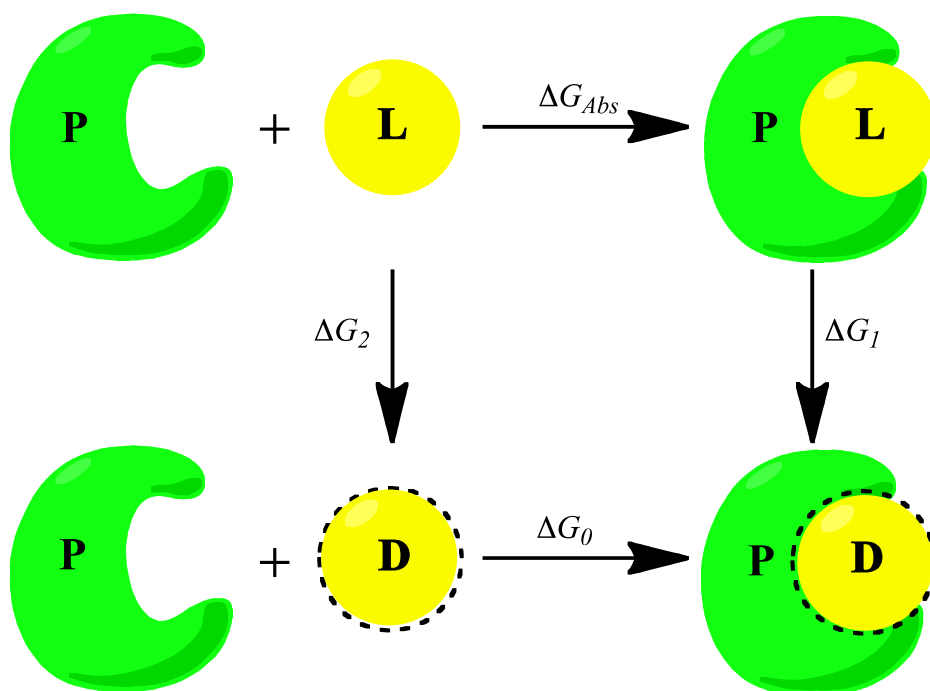
Figure 1.3. Thermodynamic cycle for calculating absolute binding energy. P, L, and D are protein, ligand, and dummy. ΔG_{Abs} is the absolute binding energy of L-P complex; ΔG_1 is the free energy change involved in decoupling L from P and solvent; ΔG_2 is the free energy change involved in turning off the interactions of L with the solvent; ΔG_0 is the absolute binding energy of D-P complex.

The theory behind DDM is shown in Figure 1.4. DDM involves two simulations; the first one is the transfer of L from solvent to the gas phase that yields ΔG_2^o, and the other is

the transfer of L bound to the solvated P into the gas phase, which gives ΔG_1^o. Standard free energy of binding (absolute binding energy) ΔG_{Abs}^o is the difference of ΔG_2^o and ΔG_1^o [59]. As discussed earlier, ΔG_1^o and ΔG_2^o can be computed using either FEP or TI.

$$
\begin{array}{ll}
\left(PL\right)_{Sol} \rightarrow P_{Sol} + L_{Gas} & \Delta G_1^o \\
\underline{L_{Sol} \rightarrow L_{Gas} \qquad\qquad \Delta G_2^o} \\
P_{Sol} + L_{Sol} \rightarrow \left(PL\right)_{Sol} & \Delta G_{Abs}^o = \Delta G_2^o - \Delta G_1^o
\end{array}
$$
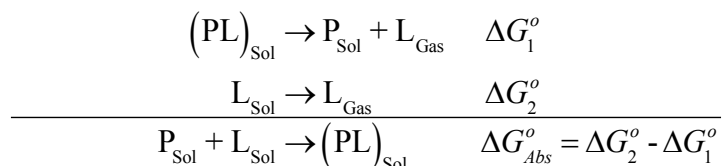
Figure 1.4. Thermodynamics of double decoupling method. (PL)$_{Sol}$ is P-L complex in the solvent; P$_{Sol}$ is P in the solvent; L$_{Sol}$ is L in the solvent; L$_{Gas}$ refers to L in the gas phase.

Recent applications of decoupling approaches include (a) estimating the binding energy contribution of interface water molecules in the HIV-1 protease Kynostatin-272 complex [63], (b) calculation of free energies of water molecules in the hydrophobic pocket of β-lactoglobulin [64], (c) computing the ABEs of nonpolar aromatic ligands with T4 lysozyme L99A mutant [65], (d) estimating the ABEs of FK506-related ligands with FKBP12 (accuracies within 2 kcal/mol were obtained) [66], and (e) calculating the binding energies of charged compounds benzamidine and diazamidine with trypsin. Inclusion of polarization in the force field leads to excellent agreement between the calculated and experimental binding energies (within 0.5 kcal/mol) [67]. This is the most powerful and rigorous approach for computing binding affinities, although computationally expensive. Double decoupling approaches do not require prior experimental knowledge to compute ABEs. Another advantage of DDM is that the path between two endpoints can be defined automatically. Decoupling approaches are

13

impractical for computing binding affinities of charged ligands with proteins, which involve larger binding affinities [68].

As mentioned earlier, non-alchemical methods to compute ABEs are the pulling methods. Pulling methods are potential of mean force (PMF) based and typically used to calculate ABEs. In pulling methods as shown in Figure 1.5, L bound to P is gradually pulled away along a reaction coordinate to a distance representative of the unbound state, while the interactions with the rest of the system (solvent "S") are maintained. Free energy changes are computed as a function of separation distances and summed to give the total free energy change [68, 69]. A major advantage of pulling methods over DDM is that they can be used to compute binding affinities of charged ligands with the receptor [68]. A major drawback is slow convergence compared to alchemical methods, if the protein undergoes large conformational changes during the ligand separation. Pulling methods are impracticable if the pockets are deeply buried in the protein, in which case DDM is highly effective.
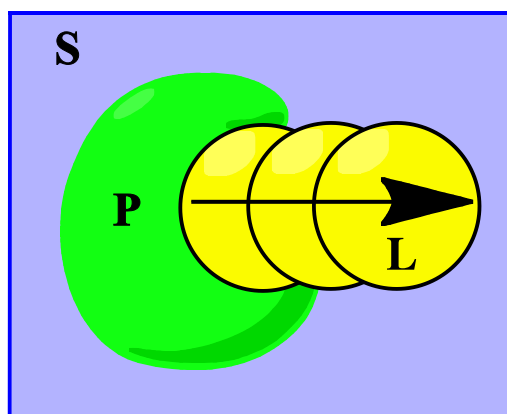
Figure 1.5. Schematic representation of pulling methods. L bound to P is gradually pulled from the pocket into solvent (S), and free energy changes along the path separating the protein and ligand are computed.

Methods based on PMF approach include the smooth reaction path generation method (SRPG) [70] and methods using restraining potentials [68, 71-73]. Recent successful applications of PMF based methods include: (a) estimating the ABFE of phosphotyrosine peptide pYEEI with the Src homology 2 (SH2) domain of human Lck kinase [68], (b) computing ABEs of FK506 and 4-hydroxy-2-butanone with FKBP [69], and (c) calculation of ABFE of pteroic acid with ricin toxin A-chain [74]. Good agreement between the calculated and experimental binding affinities (within 1 kcal/mol) [68, 69, 72] demonstrates the utility of PMF based methods.

In general, the chance of errors in RBFE calculations is smaller compared to ABFE calculations, but both these methods suffer from similar challenges with regards to sampling and accuracies. RBFE calculations may benefit from cancellation of errors if both the ligands share a similar binding mode and conformational changes of the protein show a similar impact on ligand binding. ABFE calculations are computationally expensive, but rigorous compared to RBFE calculations. The major advantages of explicit methods over implicit methods are inclusion of protein flexibility and explicit water in the simulations. The shortcomings of explicit methods include the computational and time expenses. Additionally, binding energies may not always converge due to insufficient sampling of the system. Several efforts have been made to improve convergence and speed [75]. The accuracy of these methods largely depends on the accuracy of the force field parameters. Increase in accuracy of the results by incorporating polarizability in the force fields [76] has been reported [67].

1.2.2. Implicit methods

Implicit methods are typically referred to as scoring functions (SF) and used for ligand docking and ranking of poses [77, 78]. SFs are also used to screen large databases to identify a novel lead compound. They enable us to identify a possible binding mode of a ligand and to approximately estimate its binding affinity with the protein. Implicit methods [79, 80], unlike the accurate methods, are fast enough to screen databases in a reasonable amount of time. However, they compromise on accuracy by treating the protein as a rigid entity and the solvent as a continuum medium; only the ligand is sampled explicitly. SFs offer poor correlation between the predicted and experimental binding affinities. SFs are either force field-based, knowledge-based or empirical [79-81]. Currently, empirical SFs are widely used in docking and scoring activities. Since none of the SFs offer accurate binding energies, consensus-scoring methods are gaining momentum [82-85]. Apart from the above two approximations, there are other limitations for these methods. Ligand entropy [86, 87] and molecular weight bias [88-90] need to be accounted for in these methods. Examples of implicit methods are Poisson-Boltzmann (PB) [91] and Generalized Born (GB)-based [92] models.

The major drawbacks of implicit methods are inflexibility of the protein and implicit treatment of the solvent. Since flexibility is key for the protein function [93], it can be incorporated by including an ensemble of protein conformations in the calculations, [94-96] by explicitly treating only certain degrees of freedom, [94] or both [97]. Keeping in mind the importance of water at the binding interfaces of biomolecules [98], and in the binding of ligands to proteins [99], molecular hydration can be included by explicitly placing water molecules at selected sites in the complexes [100, 101]. For details about

different SFs, please see references [102-104].

1.2.3. Hybrid methods

Hybrid methods are a combination of explicit and implicit methods. They combine MD with the free energy scoring to calculate the binding free energies. Hybrid methods use MD to generate an ensemble of structures and use SFs to calculate the average binding energy. Hybrid methods are not as slow as explicit methods and not as approximate as implicit methods. Although faster than explicit methods, hybrid methods are not fast enough for computational screening of inhibitors. Hybrid methods offer reasonable accuracies, with typical r^2 values in the range of 0.4-0.9 [105-108]. Several hybrid methods are available; notable among them are Molecular Mechanics-Poisson-Boltzmann Surface Area (MM-PBSA), Molecular Mechanics-Generalized Born Surface Area (MM-GBSA), and LIE (Linear Interaction Energy) methods [109].

MM-PBSA/GBSA methods [110, 111] involve calculating the changes in MM, electrostatic and nonpolar solvation energies, and entropic contributions. Entropic contributions are ignored if the ligands are relatively close. However, they are often estimated with normal mode analysis [106], quasi-harmonic analysis [112], or the restrain and release approach [105]. Only the bound and unbound states of a ligand are considered in the calculations. The big drawback of MM-PBSA/GBSA methods is that the ligand is considered in the same configuration in both the unbound and bound state, which is often not true. Other major concerns are the choice of solute dielectric constant, which is crucial for the PBSA method [106], and the overestimation of the entropic contributions [105]. Entropic contributions are often ignored with the assumption that

17

they cancel out when calculating the RBFE of two closely related ligands [113]. These methods performed reasonably well in several cases ($r^2 > 0.8$, and errors < 2 kcal/mol) [110, 111, 114], yet in some cases their performance was inferior to the SFs [52]. Since these methods are based on ensemble averages from MD, they are more appropriate for lead optimization than lead identification. Both absolute and relative binding energies can be estimated through MM-PBSA and GBSA methods.
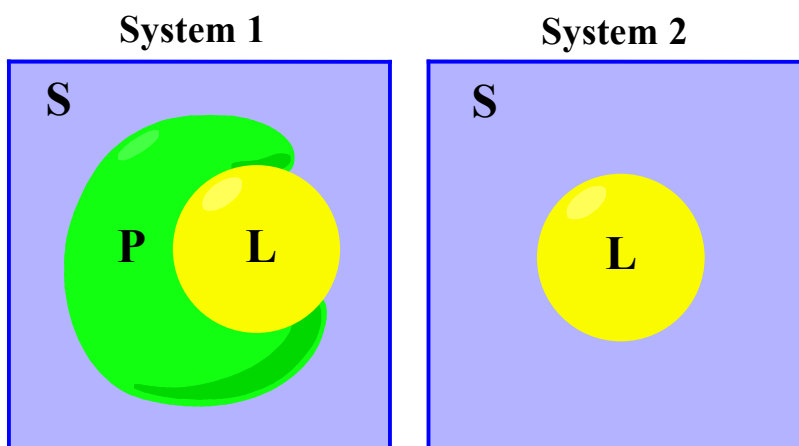
Figure 1.6. Schematic representation of endpoint free energy methods. On the left is system 1 with L-P in the solvent (S), and on the right is system 2 with L itself in the S.

LIE is also an endpoint method, which is used to estimate the binding free energy from the MD simulations. It is based on the linear response approximation, according to which the protein-ligand binding energy linearly depends on the polar and non-polar energy contributions from the resulting MD averages [115]. Binding energy is the difference of the interaction energy of a ligand with solvent and solvated protein. Only two simulations are required in this method, the ligand itself in water and the protein-ligand complex in water (Fig. 1.6). Recent developments include incorporation of

multiple binding modes of ligands [116], continuum electrostatics [117], quantum mechanical (QM) calculations to treat polarization effects, and new parameterization approaches [118].

Hybrid methods, unlike implicit methods, incorporate protein flexibility by using MD generated conformational ensemble in the free energy calculations. However, they are faster than explicit methods because they skip the intermediate stages between the bound and unbound states, giving them the name 'endpoint' free energy methods [109, 113, 119]. They are approximate because in this method implicit solvent calculations are performed on the structures resulting from MD [109]. Multiple snapshots from a MD trajectory are included for computing free energies and the average is calculated. Overall, they are computationally less demanding and easier to execute than the explicit methods and more reliable and accurate than the implicit methods. Contrary to FEP methods, these methods can handle large structural changes in the ligands. Binding energies calculated through hybrid methods are close to real free energies because of the extensive sampling involved and incorporation of entropic contributions.

The latest advance in free energy methods is to combine various methods to take advantage of the positives in the individual methods. For example, PDLD/S-LRA/β method is a combination of the Protein Dipoles Langevin Dipoles method (PDLD), the Linear Response Approximation (LRA) and LIE. It performed more effectively than the individual endpoint methods [105]. The TI method was combined with the replica exchange (RE) to overcome the sampling problems, and the new method (RETI) was more efficient than the TI method [120-122]. Similarly, the one-step perturbation (OSP) was combined with the LIE to improve accuracy [123].

Good force field parameters are always essential for accuracy in the binding energies. Recent developments include incorporation of polarizability in the force fields [124] (e.g., CHARMM and AMOEBA polarizable force fields), which is critical for accuracy [67]. However, polarizable force fields are not yet extensively validated and are computationally intensive. The choice of method depends on the nature of the problem. The speed and accuracy of different free energy methods are compared in Figure 1.7. For complete technical details of different free energy methods, please see the references [124-126].
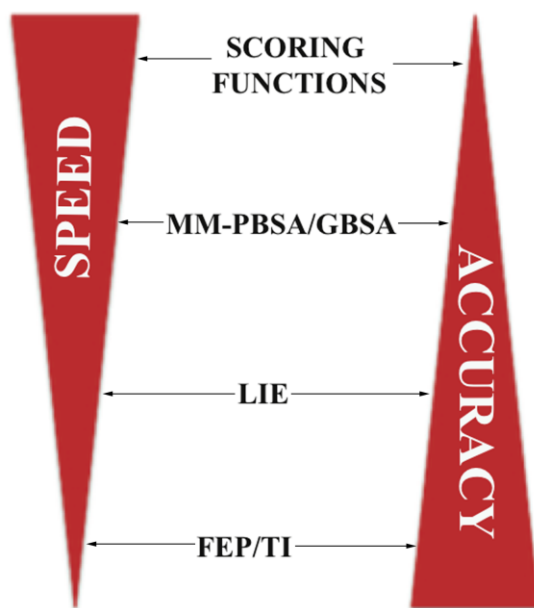
Figure 1.7. Comparison of different free energy methods for accuracy and speed.

1.3. Free Energy Calculations on MATs

In this section, different free energy studies carried on MATs are reviewed. However, the implicit methods used in the docking and virtual screening studies are not reviewed.

Huang et al. [127], through free energy calculations, identified the binding mode of

20

dopamine in hDAT. Free energy calculations were performed on the homology model of hDAT, which was developed using LeuT as a template [21]. An ensemble of hDAT-DA complexes required for the free energy calculations was obtained from MD simulations. Using the MM-PBSA method [110], the calculated binding energy of DA-hDAT was reported as -6.4 kcal/mol, while the experimental binding affinity was -7.4 kcal/mol [128]. A difference of 1 kcal/mol between the calculated and experimental binding energies was attributed to the inherent approximations in the MM-PBSA method. RESP charges for DA were used in these calculations.

Huang et al. [129], through MM-PBSA [110] calculations, demonstrated that cocaine binds stronger to hDAT in the absence of DA in the substrate binding site. The calculated RBFE of cocaine with the hDAT-DA complex and hDAT alone was approximately 4 kcal/mol, which translates to a difference of 1000-fold between the K_d's of cocaine-hDAT-DA and cocaine-hDAT. Binding modes of DA and cocaine in the homology model of hDAT [127] were explored through AutoDock 3.0.5 [130]. The complexes were initially energy minimized and further refined through MD in the sander module of Amber 8 [131]. Atomic charges, determined through RESP fitting, were used for DA and cocaine [132-134].

Hill et al. [135] quantitatively estimated the impact of mutations on the inhibition of hDAT and hNET by RTI-113 through MM-PBSA [110] calculations. The calculated RBFE of RTI-113 with the wild-type hNET-NE and hNET/Y151F-NE was 1.8 kcal/mol, whereas the experimental RBFE was 1.6 kcal/mol [135]. Similarly, the calculated and experimental RBEs of RTI-113 with the wild-type hDAT-DA and hDAT/F155Y-DA were 1.5 kcal/mol and 1.8 kcal/mol, respectively [135]. The homology model of hDAT

[127, 129] was used to develop the hNET homology model [135]. Binding modes of DA, NE and RTI-113 were explored using AutoDock 3.0.5 [130], and the complexes were energy minimized using Amber 9 [136]. RESP charges for the DA, NE and RTI-113 were used in the binding energy calculations.

Henry et al. [137] calculated the shift in pKa's of Cys336 and Cys372 upon deprotonation in hSERT using the FEP approach [138]. In the presence of Na$^+$ in the site-1 (also known as the NA1 site), shift in ΔpKa's of Cys336 and Cys372 were approximately -5.0 and -0.9, respectively. Therefore, the net pKa's of Cys336 and Cys372 were ~3.0 and ~7.1, indicating a higher probability of deprotonation of Cys336 compared to Cys372 in the presence of Na$^+$ in the Na1 site. This was due to the proximity of Cys336 to the Na1 site compared to Cys372. The homology model of hSERT was developed using LeuT as a template [140]. Binding mode of 5-HT was initially modeled identical to the experimentally validated mode of Kaufmann et al. [140] and later refined through a series of energy minimizations and equilibrations. FEP calculations were performed using NAMD 2.7b1 [139] and analyzed using CHARMM c36b2. CHARMM27 [141] force field parameters were used for the protein and lipids; for the deprotonated cysteine, parameters developed by Foloppe and Nillson [142] were used. Bond lengths and bond angles of 5-HT were obtained by QM calculations at the B3LYP/6–31G* level of theory using the RESP fitting approach [143].

For the first time, our group has optimized the novel hSERT inhibitors using the FEP calculations (Not yet published). We developed CGenFF [144] parameters for fifteen ligands of MATs for using in the FEP studies and MD simulations. CGenFF is a force field for drug-like molecules, which is compatible with the CHARMM additive all-atom

force field for the proteins, carbohydrates, lipids and nucleic acids. This allows performing "all-CHARMM" simulations to study protein-drug interactions. Prior to its application to hSERT, FEP method was first validated by calculating the RBEs of three amino acid ligands glycine, alanine, and leucine with LeuT and compared to the experimental binding affinities (Table 1.2). The calculated RBEs were in good agreement with the experimental RBEs.

Table 1.2. Comparison of the calculated and experimental relative binding energies of the three amino acid ligands with LeuT.

Calculation	$\Delta\Delta G_{Calc}$ (kcal/mol)	$\Delta\Delta G_{Expt}$ (kcal/mol)
LeuT$_{Aa}$: Ala $\rightarrow$ Gly	3.26 ± 0.05	2.85 ± 0.22[145]
LeuT$_{Aa}$: Leu $\rightarrow$ Ala	3.39 ± 0.07	3.17 ± 0.30[145]

$\Delta\Delta G_{Calc}$ and $\Delta\Delta G_{Expt}$ are the calculated and experimental relative binding energies. Gly, Ala, and Leu refer to the amino acid ligands glycine, alanine, and leucine. The experimental relative binding affinities ($\Delta\Delta G_{Expt}$) are from the reference 149.

1.4. Conclusions

Here, various free energy studies carried on MATs were reviewed. This review also covers free energy methods available to compute protein-ligand binding affinities, their advantages, pitfalls and applicability in drug design, along with the recent developments. Although computationally expensive and slow, explicit methods offer quantitative agreement with the experiment, making them more practical in the lead optimization stages of drug discovery. On the other end of the spectrum, there are scoring functions that are fast and computationally less demanding than explicit methods. However, these methods are applicable only in the screening stages of drug discovery because of the inherent approximations involved. Hybrid methods, which are a mix of explicit and implicit methods, are faster than accurate methods and more accurate than scoring

functions. However, they are not fast enough for virtual screening since they use MD simulations to sample the endpoints of a free energy simulation. Recent advances include incorporation of polarizability in the force fields, and development of combination methods. Explicit methods largely benefited from the improved force fields and parallel computing architecture. Sampling issues and errors are better understood than before. Despite the recent advancements, free energy methods are not yet ready for the routine use in drug design. Despite several drawbacks, they can still be used to enhance the drug discovery process and to derive valuable information about the structure and function of proteins.

MATs have been targeted for drug discovery for several decades because of their role in the regulation of neurotransmitter signaling in the CNS. Since the crystal structures of MATs were not available until recently, LeuT was used to understand the structure and function of MATs. Homology models of MATs using LeuT were developed, and several computational techniques have been used to explore the molecular details underlying several functional aspects of MATs. Free energy calculations have been applied to MATs for various purposes. However, there was not a single report on the optimization of MAT ligands using free energy methods. This may be partly due to the lack of good force field parameters for drug-like molecules, which are key for the accurate determination of free energies; other reasons include computational costs associated with the explicit methods and sampling issues because of the size of the membrane proteins. Keeping in mind the importance of force field parameters for achieving accuracy, CGenFF parameters were developed for several MAT ligands. RBEs of the three amino acid ligands glycine, alanine, and leucine with LeuT computed through FEP were reported. The calculated

RBEs were in good agreement with the experimental values. Finally, free energy calculations reached a stage where they could be used as a predictive tool rather than simply to support the experimental data.

1.5. References

1.   Carlsson, A., *Perspectives on the discovery of central monoaminergic neurotransmission.* Annual Review of Neuroscience, 1987. **10**(1): p. 19-40.

2.   Clark, L., J. Roiser, R. Cools, D. Rubinsztein, B. Sahakian, and T. Robbins, *Stop signal response inhibition is not modulated by tryptophan depletion or the serotonin transporter polymorphism in healthy volunteers: implications for the 5-HT theory of impulsivity.* Psychopharmacology, 2005. **182**(4): p. 570-578.

3.   Greengard, P., *The neurobiology of slow synaptic transmission.* Science, 2001. **294**(5544): p. 1024-1030.

4.   Olivier, B., *Serotonin and aggression.* Annals of the New York Academy of Sciences, 2004. **1036**(1): p. 382-392.

5.   Romero-Ramos, M., J.A. Rodriguez-Gomez, J.L. Venero, J. Cano, and A. Machado, *Chronic inhibition of the high-affinity dopamine uptake system increases oxidative damage to proteins in the aged rat substantia nigra.* Free Radical Biology and Medicine, 1997. **23**(1): p. 1-7.

6.   Volkow, N.D., J.S. Fowler, G.J. Wang, J. Logan, D. Schlyer, R. MacGregor, R. Hitzemann, and A.P. Wolf, *Decreased dopamine transporters with age in healthy human subjects.* Annals of Neurology, 1994. **36**(2): p. 237-239.

7.   Amara, S.G. and M.J. Kuhar, *Neurotransmitter transporters: recent progress.* Annual Review of Neuroscience, 1993. **16**(1): p. 73-93.

8.   Blakely, R.D., S. Ramamoorthy, S. Schroeter, Y. Qian, S. Apparsundaram, A. Galli, and L.J. DeFelice, *Regulated phosphorylation and trafficking of antidepressant-sensitive serotonin transporter proteins.* Biol Psychiatry, 1998. **44**(3): p. 169-78.

9.   Murphy, D.L. and K.-P. Lesch, *Targeting the murine serotonin transporter: insights into human neurobiology.* Nat Rev Neurosci, 2008. **9**(2): p. 85-96.

10. Chen, N.-H., M.A. Reith, and M. Quick, *Synaptic uptake and beyond: the sodium- and chloride-dependent neurotransmitter transporter family SLC6.* Pflügers Archiv European Journal of Physiology, 2004. **447**(5): p. 519-531.

11. Gainetdinov, R.R., T.D. Sotnikova, and M.G. Caron, *Monoamine transporter pharmacology and mutant mice.* Trends in Pharmacological Sciences, 2002. **23**(8): p. 367-373.

12. Kugaya, A., M. Fujita, and R. Innis, *Applications of SPECT imaging of dopaminergic neurotransmission in neuropsychiatric disorders.* Annals of Nuclear Medicine, 2000. **14**(1): p. 1-9.

13. Rabey, J.M., A. Lerner, M. Sigal, E. Graff, and Z. Oberman, *[3H] Dopamine uptake by platelet storage granules in schizophrenia.* Life Sciences, 1992. **50**(1): p. 65-72.

14. Seeman, P. and H. Niznik, *Dopamine receptors and transporters in Parkinson's disease and schizophrenia.* The FASEB Journal, 1990. **4**(10): p. 2737-2744.

15. Reith, M.E.A., B. de Costa, K.C. Rice, and A.E. Jacobson, *Evidence for mutually exclusive binding of cocaine, BTCP, GBR 12935, and dopamine to the dopamine transporter.* European Journal of Pharmacology: Molecular Pharmacology, 1992. **227**(4): p. 417-425.

16. Ritz, M., R. Lamb, Goldberg, and M. Kuhar, *Cocaine receptors on dopamine transporters are related to self-administration of cocaine.* Science, 1987. **237**(4819): p. 1219-1223.

17. Sandoval, V., E.L. Riddle, Y.V. Ugarte, G.R. Hanson, and A.E. Fleckenstein, *Methamphetamine-induced rapid and reversible changes in dopamine transporter function: an in vitro model.* The Journal of Neuroscience, 2001. **21**(4): p. 1413-1419.

18. Saunders, C., J.V. Ferrer, L. Shi, J. Chen, G. Merrill, M.E. Lamb, L.M.F. Leeb-Lundberg, L. Carvelli, J.A. Javitch, and A. Galli, *Amphetamine-induced loss of human dopamine transporter activity: an internalization-dependent and cocaine-sensitive mechanism.* Proceedings of the National Academy of Sciences, 2000. **97**(12): p. 6850-6855.

19. Nelson, N., *The family of Na+/Cl− neurotransmitter transporters.* Journal of Neurochemistry, 1998. **71**(5): p. 1785-1803.

20. Penmatsa, A., K.H. Wang, and E. Gouaux, *X-ray structure of dopamine transporter elucidates antidepressant mechanism.* Nature, 2013. **503**(7474): p. 85-90.

21. Yamashita, A., S.K. Singh, T. Kawate, Y. Jin, and E. Gouaux, *Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters.* Nature, 2005. **437**(7056): p. 215-223.

22. Beuming, T., L. Shi, J.A. Javitch, and H. Weinstein, *A comprehensive structure-based alignment of prokaryotic and eukaryotic neurotransmitter/Na+ symporters (NSS) aids in the use of the LeuT structure to probe NSS structure and function.* Molecular Pharmacology, 2006. **70**(5): p. 1630-1642.

23. Henry, L.K., J.R. Field, E.M. Adkins, M.L. Parnas, R.A. Vaughan, M.-F. Zou, A.H. Newman, and R.D. Blakely, *Tyr-95 and Ile-172 in transmembrane segments 1 and 3 of human serotonin transporters interact to establish high affinity recognition of antidepressants.* Journal of Biological Chemistry, 2006. **281**(4): p. 2012-2023.

24. Singh, S.K., *LeuT: A prokaryotic stepping stone on the way to a eukaryotic neurotransmitter transporter structure.* Channels, 2008. **2**(5): p. 380-389.

25. Jørgensen, A.M., L. Tagmose, A.M.M. Jørgensen, K.P. Bøgesø, and G.H. Peters, *Molecular dynamics simulations of Na+/Cl−-dependent neurotransmitter transporters in a membrane-aqueous system.* ChemMedChem, 2007. **2**(6): p. 827-840.

26. Andersen, J., A.S. Kristensen, B. Bang-Andersen, and K. Stromgaard, *Recent advances in the understanding of the interaction of antidepressant drugs with serotonin and norepinephrine transporters.* Chemical Communications, 2009(25): p. 3677-3692.

27. Gillman, P.K., *Tricyclic antidepressant pharmacology and therapeutic drug interactions updated.* British Journal of Pharmacology, 2007. **151**(6): p. 737-748.

28. Waitekus, A.B. and P. Kirkpatrick, *Duloxetine hydrochloride.* Nat Rev Drug Discov, 2004. **3**(11): p. 907-908.

29. Wong, D.T., P.G. Threlkeld, K.L. Best, and F.P. Bymaster, *A new inhibitor of norepinephrine uptake devoid of affinity for receptors in rat brain.* Journal of Pharmacology and Experimental Therapeutics, 1982. **222**(1): p. 61-65.

30. Tran, P.V., F.P. Bymaster, R.K. McNamara, and W.Z. Potter, *Dual monoamine modulation for improved treatment of major depressive disorder.* Journal of Clinical Psychopharmacology, 2003. **23**(1): p. 78-86.

31. Astrup, A., S. Madsbad, L. Breum, T.J. Jensen, J.P. Kroustrup, and T.M. Larsen, *Effect of tesofensine on bodyweight loss, body composition, and quality of life in obese patients: a randomised, double-blind, placebo-controlled trial.* The Lancet, 2008. **372**(9653): p. 1906-1913.

32. Newman, A.H. and S. Kulkarni, *Probes for the dopamine transporter: new leads toward a cocaine-abuse therapeutic—A focus on analogues of benztropine and rimcazole.* Medicinal Research Reviews, 2002. **22**(5): p. 429-464.

33. Jorenby, D., *Clinical efficacy of bupropion in the management of smoking cessation.* Drugs, 2002. **62**: p. 25-35.

34. von Itzstein, M., W.-Y. Wu, G.B. Kok, M.S. Pegg, J.C. Dyason, B. Jin, T.V. Phan, M.L. Smythe, H.F. White, S.W. Oliver, P.M. Colman, J.N. Varghese, D.M. Ryan, J.M. Woods, R.C. Bethell, V.J. Hotham, J.M. Cameron, and C.R. Penn, *Rational design of potent sialidase-based inhibitors of influenza virus replication.* Nature, 1993. **363**(6428): p. 418-423.

35. Jorgensen, W.L., *The many roles of computation in drug discovery.* Science, 2004. **303**(5665): p. 1813-1818.

36. Liverton, N.J., M.K. Holloway, J.A. McCauley, M.T. Rudd, J.W. Butcher, S.S. Carroll, J. DiMuzio, C. Fandozzi, K.F. Gilbert, S.-S. Mao, C.J. McIntyre, K.T. Nguyen, J.J. Romano, M. Stahlhut, B.-L. Wan, D.B. Olsen, and J.P. Vacca, *Molecular modeling based approach to potent P2-P4 macrocyclic inhibitors of hepatitis C NS3/4A protease.* Journal of the American Chemical Society, 2008. **130**(14): p. 4607-4609.

37. Stauffer, S.R., M.G. Stanton, A.R. Gregro, M.A. Steinbeiser, J.R. Shaffer, P.G. Nantermet, J.C. Barrow, K.E. Rittle, D. Collusi, A.S. Espeseth, M.-T. Lai, B.L. Pietrak, M.K. Holloway, G.B. McGaughey, S.K. Munshi, J.H. Hochman, A.J. Simon, H.G. Selnick, S.L. Graham, and J.P. Vacca, *Discovery and SAR of isonicotinamide BACE-1 inhibitors that bind beta-secretase in a N-terminal 10s-loop down conformation.* Bioorganic & Medicinal Chemistry Letters, 2007. **17**(6): p. 1788-1792.

38. Mackerell, A.D., *Empirical force fields for biological macromolecules: overview and issues.* Journal of Computational Chemistry, 2004. **25**(13): p. 1584-1604.

39. Guvench, O. and A.D. MacKerell, *Comparison of protein force fields for molecular dynamics simulations molecular modeling of proteins*, A. Kukol, Editor. 2008, Humana Press. p. 63-88.

40. Zhou, Z., J. Zhen, N.K. Karpowich, C.J. Law, M.E.A. Reith, and D.-N. Wang, *Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures.* Nat Struct Mol Biol, 2009. **16**(6): p. 652-657.

41. Guvench, O. and A.D. MacKerell Jr, *Computational evaluation of protein-small molecule binding.* Current Opinion in Structural Biology, 2009. **19**(1): p. 56-61.

42.   Tuckerman, M.E. and G.J. Martyna, *Understanding modern molecular dynamics: techniques and applications.* The Journal of Physical Chemistry B, 1999. **104**(2): p. 159-178.

43.   Simonson, T., *Free energy calculations. Computational Biochemistry and Biophysics*, ed. M.A. Becker OM, Roux B, Watanabe M 2001: Marcel Dekker, Inc.

44.   Tembre, B.L. and J.A. Mc Cammon, *Ligand-receptor interactions.* Computers & Chemistry, 1984. **8**(4): p. 281-283.

45.   Warshel, A., F. Sussman, and G. King, *Free energy of charges in solvated proteins: microscopic calculations using a reversible charging process.* Biochemistry, 1986. **25**(26): p. 8368-8372.

46.   Shirts, M.R., D.L. Mobley, and J.D. Chodera, *Chapter 4 Alchemical free energy calculations: ready for prime time?*, in *Annual Reports in Computational Chemistry*, D.C. Spellmeyer and R. Wheeler, Editors. 2007, Elsevier. p. 41-59.

47.   Bash, P., U. Singh, F. Brown, R. Langridge, and P. Kollman, *Calculation of the relative change in binding free energy of a protein-inhibitor complex.* Science, 1987. **235**(4788): p. 574-576.

48.   Lu, N., D.A. Kofke, and T.B. Woolf, *Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods.* Journal of Computational Chemistry, 2004. **25**(1): p. 28-40.

49.   Beutler, T.C., A.E. Mark, R.C. van Schaik, P.R. Gerber, and W.F. van Gunsteren, *Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations.* Chemical Physics Letters, 1994. **222**(6): p. 529-539.

50.   Mobley, D.L., A.P. Graves, J.D. Chodera, A.C. McReynolds, B.K. Shoichet, and K.A. Dill, *Predicting absolute ligand binding free energies to a simple model site.* Journal of Molecular Biology, 2007. **371**(4): p. 1118-1134.

51.   Pearlman, D.A. and P.S. Charifson, *Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system.* Journal of Medicinal Chemistry, 2001. **44**(21): p. 3417-3423.

52.   Pearlman, D.A., *Evaluating the molecular mechanics poisson-boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase.* Journal of Medicinal Chemistry, 2005. **48**(24): p. 7796-7807.

53.   Steinbrecher, T., D.A. Case, and A. Labahn, *A multistep approach to structure-based drug design: studying ligand binding at the human neutrophil elastase.* Journal of Medicinal Chemistry, 2006. **49**(6): p. 1837-1844.

54.   Jorgensen, W.L., *Efficient drug lead discovery and optimization.* Accounts of Chemical Research, 2009. **42**(6): p. 724-733.

55.   Kirkwood, J.G., *Statistical mechanics of fluid mixtures.* The Journal of Chemical Physics, 1935. **3**(5): p. 300-313.

56.   Chipot, C. and K. Schulten, *Understanding structure and function of membrane proteins using free energy calculations*, in *Biophysical Analysis of Membrane Proteins.* 2008, Wiley-VCH Verlag GmbH & Co. KGaA. p. 187-211.

57.   Zwanzig, R.W., *High-temperature equation of state by a perturbation method. I. Nonpolar gases.* The Journal of Chemical Physics, 1954. **22**(8): p. 1420-1426.

58.   Jorgensen, W.L., J.K. Buckner, S. Boudon, and J. Tirado-Rives, *Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water.* The Journal of Chemical Physics, 1988. **89**(6): p. 3742-3746.

59.   Gilson, M.K., J.A. Given, B.L. Bush, and J.A. McCammon, *The statistical-thermodynamic basis for computation of binding affinities: a critical review.* Biophysical journal, 1997. **72**(3): p. 1047-1069.

60.   Boresch, S., F. Tettinger, M. Leitgeb, and M. Karplus, *Absolute binding free energies: a quantitative approach for their calculation.* The Journal of Physical Chemistry B, 2003. **107**(35): p. 9535-9551.

61.   Donnini, S. and A.H. Juffer, *Calculation of affinities of peptides for proteins.* Journal of Computational Chemistry, 2004. **25**(3): p. 393-411.

62.   Roux, B., M. Nina, R. Pomes, and J.C. Smith, *Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study.* Biophysical journal, 1996. **71**(2): p. 670-681.

63.   Lu, Y., C.-Y. Yang, and S. Wang, *Binding free energy contributions of interfacial waters in HIV-1 protease/inhibitor complexes.* Journal of the American Chemical Society, 2006. **128**(36): p. 11830-11839.

64.   Qvist, J., M. Davidovic, D. Hamelberg, and B. Halle, *A dry ligand-binding cavity in a solvated protein.* Proceedings of the National Academy of Sciences, 2008. **105**(17): p. 6296-6301.

65.   Deng, Y. and B. Roux, *Calculation of standard binding free energies: aromatic molecules in the T4 lysozyme L99A mutant.* Journal of Chemical Theory and Computation, 2006. **2**(5): p. 1255-1273.

66.   Wang, J., Y. Deng, and B. Roux, *Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials.* Biophysical journal, 2006. **91**(8): p. 2798-2814.

67.   Jiao, D., P.A. Golubkov, T.A. Darden, and P. Ren, *Calculation of protein-ligand binding free energy by using a polarizable potential.* Proceedings of the National Academy of Sciences, 2008. **105**(17): p. 6290-6295.

68.   Woo, H.-J. and B. Roux, *Calculation of absolute protein-ligand binding free energy from computer simulations.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(19): p. 6825-6830.

69.   Lee, M.S. and M.A. Olson, *Calculation of absolute protein-ligand binding affinity using path and endpoint approaches.* Biophysical journal, 2006. **90**(3): p. 864-877.

70.   Fukunishi, Y., D. Mitomo, and H. Nakamura, *Protein-ligand binding free energy calculation by the smooth reaction path generation (SRPG) method.* Journal of Chemical Information and Modeling, 2009. **49**(8): p. 1944-1951.

71.   Doudou, S., N.A. Burton, and R.H. Henchman, *Standard free energy of binding from a one-dimensional potential of mean force.* Journal of Chemical Theory and Computation, 2009. **5**(4): p. 909-918.

72.   Gan, W. and B. Roux, *Binding specificity of SH2 domains: insight from free energy simulations.* Proteins: Structure, Function, and Bioinformatics, 2009. **74**(4): p. 996-1007.

73.   Ge, X. and B. Roux, *Calculation of the standard binding free energy of sparsomycin to the ribosomal peptidyl-transferase P-site using molecular dynamics simulations with restraining potentials.* Journal of Molecular Recognition, 2010. **23**(2): p. 128-141.

74.   Lee, M.S. and M.A. Olson, *Calculation of absolute ligand binding free energy to a ribosome-targeting protein as a function of solvent model.* The Journal of Physical Chemistry B, 2008. **112**(42): p. 13411-13417.

75.   Rodinger, T. and R. Pomes, *Enhancing the accuracy, the efficiency and the scope of free energy simulations.* Current Opinion in Structural Biology, 2005. **15**(2): p. 164-170.

76. Rick, S.W. and S.J. Stuart, *Potentials and algorithms for incorporating polarizability in computer simulations*, in *Reviews in Computational Chemistry*. 2003, John Wiley & Sons, Inc. p. 89-146.

77. Alonso, H., A.A. Bliznyuk, and J.E. Gready, *Combining docking and molecular dynamic simulations in drug design.* Medicinal Research Reviews, 2006. **26**(5): p. 531-568.

78. Zhong SJ, M.A., MacKerell AD, *Computational identification of inhibitors of protein-protein interactions.* Curr Top Med Chem, 2007. **7**: p. 63-82.

79. Brooijmans, N. and I.D. Kuntz, *Molecular recognition and docking algorithms.* Annual Review of Biophysics and Biomolecular Structure, 2003. **32**(1): p. 335-373.

80. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *Protein–ligand docking: current status and future challenges.* Proteins: Structure, Function, and Bioinformatics, 2006. **65**(1): p. 15-26.

81. Rajamani R, G.A., *Ranking poses in structure-based lead discovery and optimization: current trends in scoring function development.* Curr Opin Drug Discov Dev, 2007. **10**: p. 308-315.

82. Clark, R.D., A. Strizhev, J.M. Leonard, J.F. Blake, and J.B. Matthew, *Consensus scoring for ligand/protein interactions.* Journal of Molecular Graphics and Modelling, 2002. **20**(4): p. 281-295.

83. Feher, M., *Consensus scoring for protein-ligand interactions.* Drug Discovery Today, 2006. **11**(9-10): p. 421-428.

84. Oda, A., K. Tsuchida, T. Takakura, N. Yamaotsu, and S. Hirono, *Comparison of consensus scoring strategies for evaluating computational models of protein-ligand complexes.* Journal of Chemical Information and Modeling, 2005. **46**(1): p. 380-391.

85. Konstantinou-Kirtay, C., J. Mitchell, and J. Lumley, *Scoring functions and enrichment: a case study on Hsp90.* BMC Bioinformatics, 2007. **8**(1): p. 27.

86. Chang, C.-e.A., W. Chen, and M.K. Gilson, *Ligand configurational entropy and protein binding.* Proceedings of the National Academy of Sciences, 2007. **104**(5): p. 1534-1539.

87. Ruvinsky, A.M., *Role of binding entropy in the refinement of protein–ligand docking predictions: analysis based on the use of 11 scoring functions.* Journal of Computational Chemistry, 2007. **28**(8): p. 1364-1372.

88. Pan, Y., N. Huang, S. Cho, and A.D. MacKerell, *Consideration of molecular weight during compound selection in virtual target-based database screening.* Journal of Chemical Information and Computer Sciences, 2002. **43**(1): p. 267-272.

89. Jacobsson, M. and A. Karlen, *Ligand bias of scoring functions in structure-based virtual screening. Journal of Chemical Information and Modeling*, 2006. **46**(3): p. 1334-1343.

90. Carta, G., A.J.S. Knox, and D.G. Lloyd, *Unbiasing scoring functions: a new normalization and rescoring strategy.* Journal of Chemical Information and Modeling, 2007. **47**(4): p. 1564-1571.

91. Honig, B. and A. Nicholls, *Classical electrostatics in biology and chemistry.* Science, 1995. **268**(5214): p. 1144-1149.

92. Still, C., A. Tempczyk, R. Hawley, and T. Hendrickson, *Semianalytical treatment of solvation for molecular mechanics and dynamics.* Journal of the American Chemical Society, 1990. **112**(16): p. 6127-6129.

93. Dodson, G.G., D.P. Lane, and C.S. Verma, *Molecular simulations of protein dynamics: new windows on mechanisms in biology.* EMBO Rep, 2008. **9**(2): p. 144-150.

94. Wong, C.F., *Flexible ligand-flexible protein docking in protein kinase systems.* Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics, 2008. **1784**(1): p. 244-251.

95. Totrov, M. and R. Abagyan, *Flexible ligand docking to multiple receptor conformations: a practical alternative.* Current Opinion in Structural Biology, 2008. **18**(2): p. 178-184.

96. Amaro, R., R. Baron, and J. McCammon, *An improved relaxed complex scheme for receptor flexibility in computer-aided drug design.* Journal of Computer-Aided Molecular Design, 2008. **22**(9): p. 693-705.

97. Nabuurs, S.B., M. Wagener, and J. de Vlieg, *A flexible approach to induced fit docking.* Journal of Medicinal Chemistry, 2007. **50**(26): p. 6507-6518.

98. Li, Z. and T. Lazaridis, *Water at biomolecular binding interfaces.* Physical Chemistry Chemical Physics, 2007. **9**(5): p. 573-581.

99. Young, T., R. Abel, B. Kim, B.J. Berne, and R.A. Friesner, *Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding.* Proceedings of the National Academy of Sciences, 2007. **104**(3): p. 808-813.

100. Rarey, M., B. Kramer, and T. Lengauer, *The particle concept: placing discrete water molecules during protein-ligand docking predictions.* Proteins: Structure, Function, and Bioinformatics, 1999. **34**(1): p. 17-28.

101. Mancera, R., *Molecular modeling of hydration in drug design.* Curr Opin Drug Discov Dev, 2007. **10**: p. 275-280.

102. de Azevedo WFJ, D.R., *Computational methods for calculation of ligand-binding affinity.* Curr Drug Targets, 2008. **9**: p. 1031-9.

103. KontoyianniM, M.P., Suchanek E, SeibelW, *Theoretical and practical considerations in virtual screening: a beaten field?* Curr Med Chem, 2008. **15**: p. 107-16.

104. Moitessier, N., P. Englebienne, D. Lee, J. Lawandi, and C.R. Corbeil, *Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go.* British Journal of Pharmacology, 2008. **153**(S1): p. S7-S26.

105. Singh, N. and A. Warshel, *Absolute binding free energy calculations: on the accuracy of computational scoring of protein–ligand interactions.* Proteins: Structure, Function, and Bioinformatics, 2010. **78**(7): p. 1705-1723.

106. Hou, T., J. Wang, Y. Li, and W. Wang, *Assessing the performance of the MM/PBSA and MM/GBSA Methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations.* Journal of Chemical Information and Modeling, 2010. **51**(1): p. 69-82.

107. Guimaraes, C.R.W. and A.M. Mathiowetz, *Addressing limitations with the MM-GB/SA scoring procedure using the watermap method and free energy perturbation calculations.* Journal of Chemical Information and Modeling, 2010. **50**(4): p. 547-559.

108. de Ruiter, A. and C. Oostenbrink, *Free energy calculations of protein-ligand interactions.* Current Opinion in Chemical Biology, 2011. **15**(4): p. 547-552.

109. *Foloppe N, H.R., Towards predictive ligand design with free-energy based computational methods?* Curr Med Chem, 2006. **13**: p. 3583-3608.

110. Kollman, P.A., I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, and T.E. Cheatham, *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models.* Accounts of Chemical Research, 2000. **33**(12): p. 889-897.

111. Wang, J., P. Morin, W. Wang, and P.A. Kollman, *Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and*

predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. Journal of the American Chemical Society, 2001. **123**(22): p. 5221-5230.

112.   Singh, N. and A. Warshel, *A comprehensive examination of the contributions to the binding entropy of protein–ligand complexes.* Proteins: Structure, Function, and Bioinformatics, 2010. **78**(7): p. 1724-1735.

113.   Gilson, M.K. and H.-X. Zhou, *Calculation of protein-ligand binding affinities.* Annual Review of Biophysics and Biomolecular Structure, 2007. **36**(1): p. 21-42.

114.   Suenaga, A., N. Okimoto, Y. Hirano, and K. Fukui, *An efficient computational method for calculating ligand binding affinities.* PLoS ONE, 2012. **7**(8): p. e42846.

115.   Aqvist, J., C. Medina, and J.-E. Samuelsson, *A new method for predicting binding affinity in computer-aided drug design.* Protein Engineering, 1994. **7**(3): p. 385-391.

116.   Stjernschantz, E. and C. Oostenbrink, *Improved ligand-protein binding affinity predictions using multiple binding modes.* Biophysical journal, 2010. **98**(11): p. 2682-2691.

117.   Huang, D. and A. Caflisch, *Efficient evaluation of binding free energy using continuum electrostatics solvation.* Journal of Medicinal Chemistry, 2004. **47**(23): p. 5791-5797.

118.   Valiente, P.A., A. Gil L, P.R. Batista, E.R. Caffarena, T. Pons, and P.G. Pascutti, *New parameterization approaches of the LIE method to improve free energy calculations of PlmII-inhibitors complexes.* Journal of Computational Chemistry, 2010. **31**(15): p. 2723-2734.

119.   Huang N, J.M., *Physics-based methods for studying protein-ligand interactions.* Curr Opin Drug Discov Dev, 2007. **10**: p. 325-331.

120.   Jiang, W., M. Hodoscek, and B. Roux, *Computation of absolute hydration and binding free energy with free energy perturbation distributed replica-exchange molecular dynamics.* Journal of Chemical Theory and Computation, 2009. **5**(10): p. 2583-2588.

121.   Steiner, D., C. Oostenbrink, F. Diederich, M. Zürcher, and W.F. van Gunsteren, *Calculation of binding free energies of inhibitors to plasmepsin II.* Journal of Computational Chemistry, 2011. **32**(9): p. 1801-1812.

122.   Woods, C.J., J.W. Essex, and M.A. King, *The development of replica-exchange-based free-energy methods.* The Journal of Physical Chemistry B, 2003. **107**(49): p. 13703-13710.

123. Oostenbrink, C., *Efficient free energy calculations on small molecule host-guest systems—a combined linear interaction energy/one-step perturbation approach.* Journal of Computational Chemistry, 2009. **30**(2): p. 212-221.

124. Michel, J. and J. Essex, *Prediction of protein–ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations.* Journal of Computer-Aided Molecular Design, 2010. **24**(8): p. 639-658.

125. Truhlar, D., *Chipot, C., Pohorille, A., Eds. Free energy calculations: theory and applications in chemistry and biology.* Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 2008. **121**(1): p. 105-106.

126. Christ, C.D., A.E. Mark, and W.F. van Gunsteren, *Basic ingredients of free energy calculations: a review.* Journal of Computational Chemistry, 2010. **31**(8): p. 1569-1582.

127. Huang, X. and C.-G. Zhan, *How dopamine transporter interacts with dopamine: insights from molecular modeling and simulation.* Biophysical journal, 2007. **93**(10): p. 3627-3639.

128. Dar, D.E., T.G. Metzger, D.J. Vandenbergh, and G.R. Uhl, *Dopamine uptake and cocaine binding mechanisms: the involvement of charged amino acids from the transmembrane domains of the human dopamine transporter.* European Journal of Pharmacology, 2006. **538**(1-3): p. 43-47.

129. Huang, X., H.H. Gu, and C.-G. Zhan, *Mechanism for cocaine blocking the transport of dopamine: insights from molecular modeling and dynamics simulations.* The Journal of Physical Chemistry B, 2009. **113**(45): p. 15057-15066.

130. Morris, G.M., D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, and A.J. Olson, *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* Journal of Computational Chemistry, 1998. **19**(14): p. 1639-1662.

131. Case, D.A.D., T. A.; Cheatham, T. E., III; Simmerling, C. L.;, et al, *AMBER 8.* 2004.

132. Gao, D. and C.-G. Zhan, *Modeling effects of oxyanion hole on the ester hydrolysis catalyzed by human cholinesterases.* The Journal of Physical Chemistry B, 2005. **109**(48): p. 23070-23076.

133. Pan, Y., D. Gao, W. Yang, H. Cho, G. Yang, H.-H. Tai, and C.-G. Zhan, *Computational redesign of human butyrylcholinesterase for anticocaine*

medication. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(46): p. 16656-16661.

134. Zhan, C.-G. and D. Gao, *Catalytic mechanism and energy barriers for butyrylcholinesterase-catalyzed hydrolysis of cocaine.* Biophysical journal, 2005. **89**(6): p. 3863-3872.

135. Hill, E.R., X. Huang, C.-G. Zhan, F. Ivy Carroll, and H.H. Gu, *Interaction of tyrosine 151 in norepinephrine transporter with the 2β group of cocaine analog RTI-113.* Neuropharmacology, 2011. **61**(1-2): p. 112-120.

136. D.A. Case, T.A.D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., et al, *AMBER 9.* 2006.

137. Henry, L.K., H. Iwamoto, J.R. Field, K. Kaufmann, E.S. Dawson, M.T. Jacobs, C. Adams, B. Felts, I. Zdravkovic, V. Armstrong, S. Combs, E. Solis, G. Rudnick, S.Y. Noskov, L.J. DeFelice, J. Meiler, and R.D. Blakely, *A conserved asparagine residue in transmembrane segment 1 (TM1) of serotonin transporter dictates chloride-coupled neurotransmitter transport.* Journal of Biological Chemistry, 2011. **286**(35): p. 30823-30836.

138. Simonson, T., J. Carlsson, and D.A. Case, *Proton binding to proteins: pKa calculations with explicit and implicit solvent models.* Journal of the American Chemical Society, 2004. **126**(13): p. 4167-4180.

139. Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten, *Scalable molecular dynamics with NAMD.* Journal of Computational Chemistry, 2005. **26**(16): p. 1781-1802.

140. Kaufmann, K.W., E.S. Dawson, L.K. Henry, J.R. Field, R.D. Blakely, and J. Meiler, *Structural determinants of species-selective substrate recognition in human and Drosophila serotonin transporters revealed through computational docking studies.* Proteins: Structure, Function, and Bioinformatics, 2009. **74**(3): p. 630-642.

141. Brooks, B.R., C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus, *CHARMM: the biomolecular simulation program.* Journal of Computational Chemistry, 2009. **30**(10): p. 1545-1614.

142. Foloppe, N. and L. Nilsson, *Stabilization of the catalytic thiolate in a mammalian glutaredoxin: structure, dynamics and electrostatics of reduced pig glutaredoxin and its mutants.* Journal of Molecular Biology, 2007. **372**(3): p. 798-816.

143. Anisimov, V.M., G. Lamoureux, I.V. Vorobyov, N. Huang, B. Roux, and A.D. MacKerell, *Determination of electrostatic parameters for a polarizable force field based on the classical drude oscillator.* Journal of Chemical Theory and Computation, 2004. **1**(1): p. 153-168.

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

145. Singh, S.K., C.L. Piscitelli, A. Yamashita, and E. Gouaux, *A competitive inhibitor traps LeuT in an open-to-out conformation.* Science, 2008. **322**(5908): p. 1655-1661.

2. CHAPTER 2

DEVELOPMENT OF CHARMM GENERAL FORCE FIELD PARAMETERS FOR LIGANDS/DRUGS OF MONOAMINE TRANSPORTERS

2.1. Introduction

Monoamine transporters (MATs) are membrane proteins and belong to the NSS family. MATs (hSERT, hDAT and hNET) reuptake respective neurotransmitters from the synapse to the presynaptic neuron thereby maintaining homeostasis in the body. MATs were implicated in several mental disorders including depression, schizophrenia and Parkinson's disease; several psychostimulants exert their function by acting at MATs [1-7]. However, the exact mechanism of action of ligands/drugs at MATs were not clearly understood due to the non-availability of crystal structures of MATs, not at least for hSERT and hNET, which were primarily implicated in the depression disorder. Therefore, over the years several computational techniques including homology modeling, docking, MD simulations and free energy calculations have been utilized to study various aspects of MATs [8-10]. Computational methods were also pursued for drug discovery [1, 11-14]. However, traditionally computational results have been validated through experiments because of the limitations of various computational methods or calculations. For example, free energy calculations are suffering from the non-availability of good force field parameters for the drug-like ligands.

A force field is defined as a functional form and parameter sets that relate chemical structure and conformation to the potential energy. Well-characterized force field parameters are available to treat proteins (e.g., CHARMM, AMBER and MMFF). Especially, accurate prediction of protein-ligand binding energies requires well-characterized ligand force field parameters and is a limitation for performing the binding free energy calculations on a regular basis [8, 15].

The recent availability of general force fields (e.g., CGenFF [16] and GAFF [17]) is a saving grace, and these can be extended to drug-like molecules. CGenFF parameters for 15 MAT ligands including dopamine, serotonin, amphetamine, bupropion, methylphenidate, cocaine, benztropine, clomipramine, imipramine, desipramine, R- and S- fluoxetine, R- and S- citalopram and sertraline (structures are shown in Fig. 2.1) were developed. To our knowledge, this is the first report of extension of the CGenFF to MAT ligands. CGenFF is an organic force field parameterized for drug-like molecules and compatible with the CHARMM force field for biomolecules such as proteins [18] and nucleic acids [19]. This offers the advantage of performing all CHARMM simulations [20, 21] to study protein-ligand interactions. CGenFF is an all-atom force field and explicitly treats all atoms of the system. CGenFF parameterization philosophy was shown in Fig. 2.2. The force field parameters were optimized by comparing against the quantum mechanical (QM) data. CGenFF results were in good agreement with the target data. Development of CGenFF parameters enables the accurate estimation of MAT-ligand binding energies, which helps to resolve the ambiguity of drug binding sites and binding modes in MATs, and also to develop better therapeutics targeted at MATs.

A

HO

HO

NH_3^+

B

H
N

HO

NH_3^+

C

NH_3^+

D

O

O

H_2N^+

E

O

H_2N^+

Cl

F

N

Cl

H
N^+

41

G

H

I

J

K

L
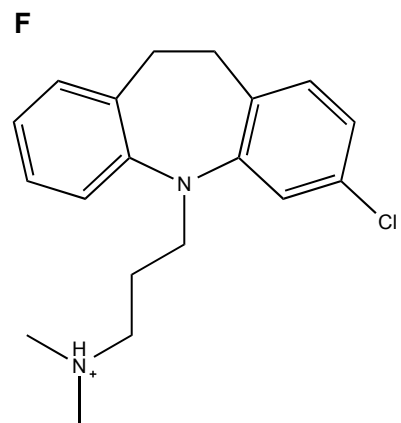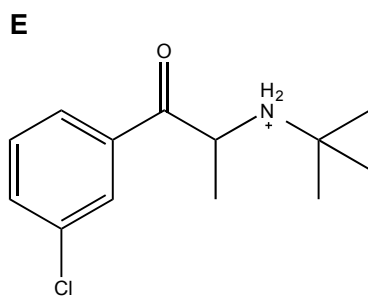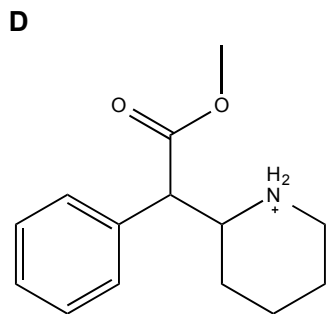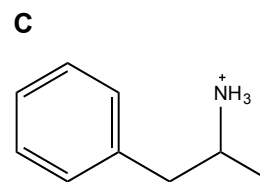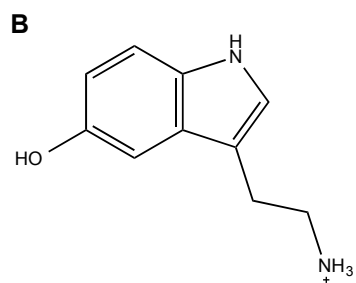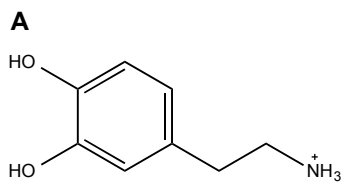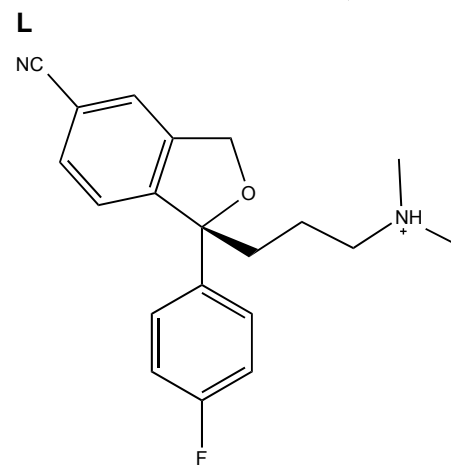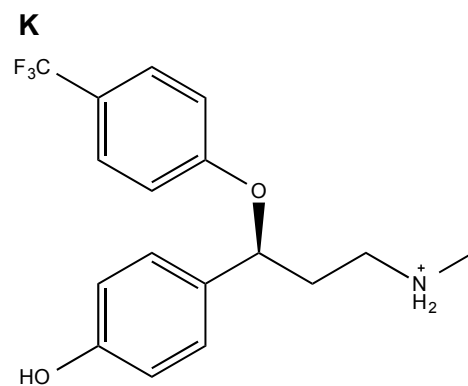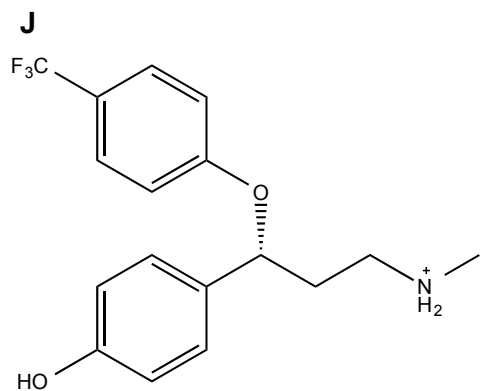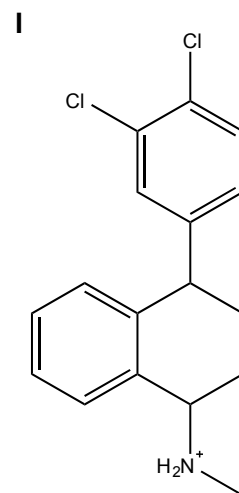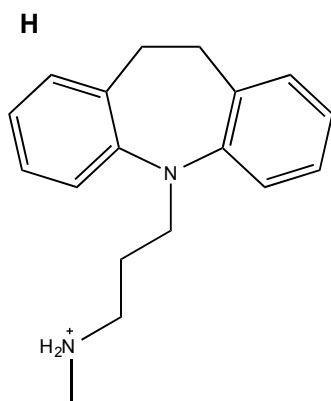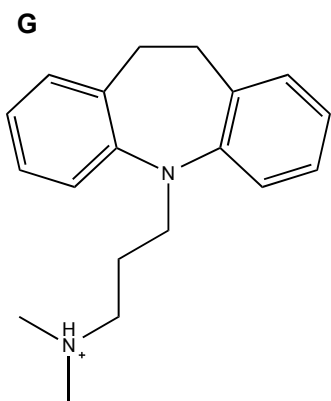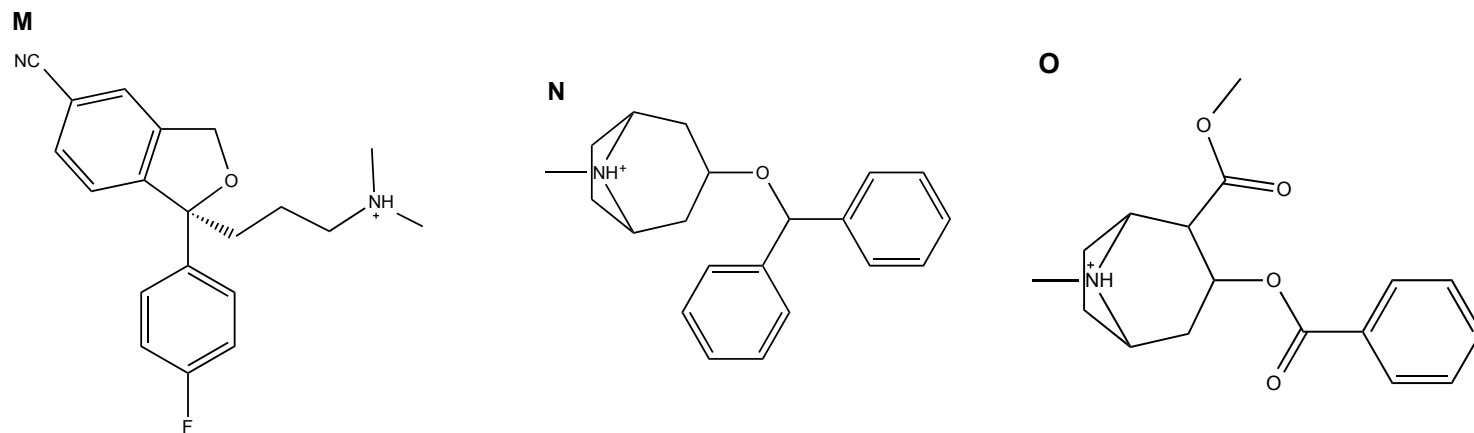
Figure 2.1. Chemical structures of various MAT ligands for which the CGenFF parameters were developed. (A) dopamine, (B) serotonin, (C) amphetamine, (D) methylphenidate, (E) bupropion, (F) clomipramine, (G) imipramine, (H) desipramine, (I) sertraline, (J) R-fluoxetine, (K) S-fluoxetine, (L) R-citalopram, (M) S-citalopram, (N) benztropine, and (O) cocaine.

2.2. Methods

CGenFF parameterization philosophy

As described earlier, CGenFF is the extension of CHARMM force field for organic molecules. The potential energy function of CGenFF is shown in equation 1.

$$E\left(\vec{R}\right) = \sum_{bonds} K_l \left(l - l_o\right)^2 + \sum_{angle} K_\theta \left(\theta - \theta_o\right)^2 + \sum_{UB} K_s \left(S - S_o\right)^2 + \sum_{dihedrals} K_\chi \left(1 + \cos\left(n\chi - \delta\right)\right) +$$

$$\sum_{impropers} K_\omega \left(\omega - \omega_o\right)^2 + \sum_{electrostatic} \frac{q_i q_j}{4\pi D r_{ij}} + \sum_{vanderwaals} \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{min,ij}}{r_{ij}}\right)^6 \right]$$

- Equation 1

$E\left(\vec{R}\right)$ is the potential energy of a system as a function of Cartesian coordinates and is a sum of intermolecular and intramolecular energy terms. The intermolecular component is comprised of electrostatic and van der Waals terms; bonds, angles, Urey-Bradley, dihedrals and impropers constitute the intramolecular part of the potential energy. q_i and q_j are the partial atomic charges of atoms i and j, respectively, and r_{ij} is the distance separating the atoms i and j. $R_{min,ij}$ is the distance at which Lennard-Jones (LJ) potential reaches minimum and ε_{ij} is the well depth. l_o, θ_o, S_o and ω_o are the equilibrium bonds, angles, Urey-Bradley and impropers, respectively; χ, n and δ are the dihedral angle, multiplicity and phase; K_l, K_θ, K_s, K_χ and K_ω are the force constants of respective intramolecular terms.

The CGenFF parameterization procedure explained elsewhere is followed [16] and is shown in Figure 2.2. Briefly, in the first-step initial parameters were assigned to the novel ligands. The available parameters were taken directly from the CGenFF

parameter file, and the missing parameters were assigned based on analogy. Bond lengths were chosen such that at least one atom type was similar, whereas for the bond angles a middle atom type was similar and for the dihedrals two middle atom types were similar. Lately, the development of automated procedures for assigning the initial set of parameters such as partial charges, atom types and bonded parameters [22, 23] simplifies the parameterization procedure.
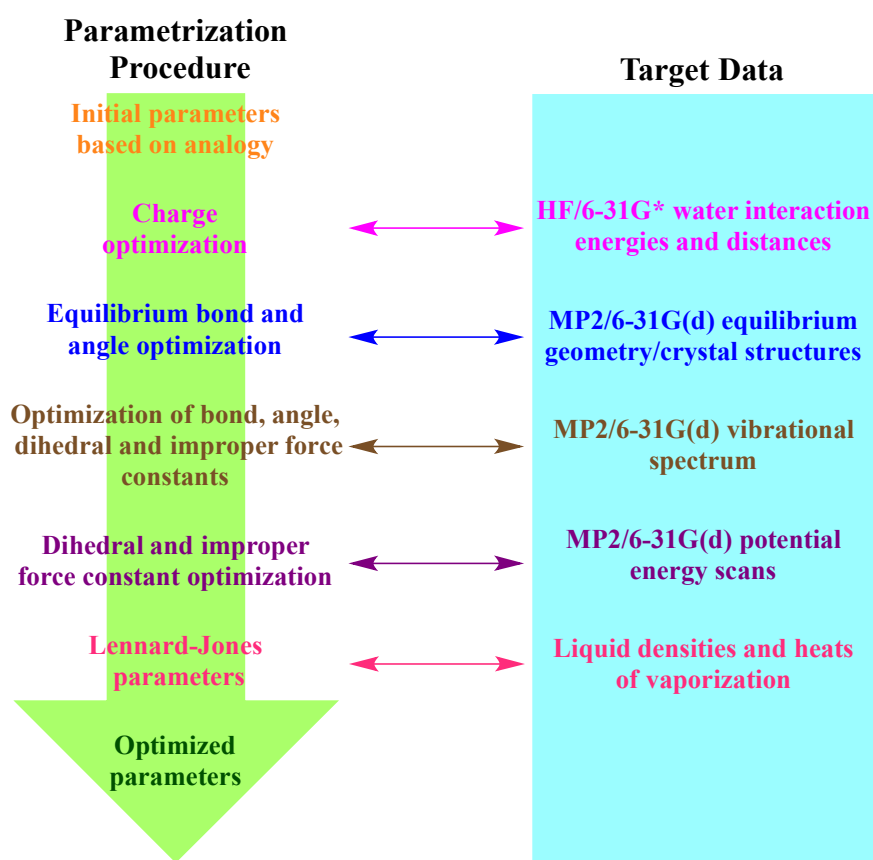
Figure 2.2. CGenFF parameterization philosophy and target data.

In the next step, partial atomic charges were optimized with emphasis on water interaction energies. Prior to the calculation of water interaction energies, equilibrium geometries of the ligands were generated at MP2/6-31G(d) level of theory [24, 25]. For

the purpose of charge optimization, TIP3P waters [26] were placed around the MP2/6-31G(d) optimized geometry of the ligands in various orientations and interaction energies of the individual hydrates were calculated at HF/6-31G* level of theory [27, 28]. Similarly, interaction energies of the monohydrates were calculated using the CGenFF charges and compared against the HF/6-31G* interaction energies. CGenFF charges on the individual atoms were adjusted and energies recalculated until they correlated well with the HF/6-31G* interaction energies. QM interaction energies were not scaled since all our compounds were protonated and have a net charge. Further, the equilibrium bonds, angles, dihedrals and impropers were optimized using the equilibrium structure generated at the MP2/6-31G(d) level of theory. Next, is the optimization of bond and angle force constants (target data is MP2/6-31G(d) vibration spectra), followed by the dihedral and improper force constants (target data is MP2/6-31G(d) potential energy scans); however, we did not pursue this step. The last step in the parameterization is the optimization of LJ parameters, however, in the CGenFF parameterization philosophy this step is not typically needed and we did not optimize LJ parameters.. The entire parameterization was performed in an iterative manner until the required quality was attained. All QM calculations were carried with Gaussian 09 [29] and empirical calculations with the CHARMM program [30].

2.3. Results and Discussion

All the ligands from Figure 2.1 were parameterized via the CGenFF procedure, and the results are shown below.

2.3.1. Charge optimization

a) Amphetamine

As described earlier in the methods section, charges were optimized using the water interaction energies.
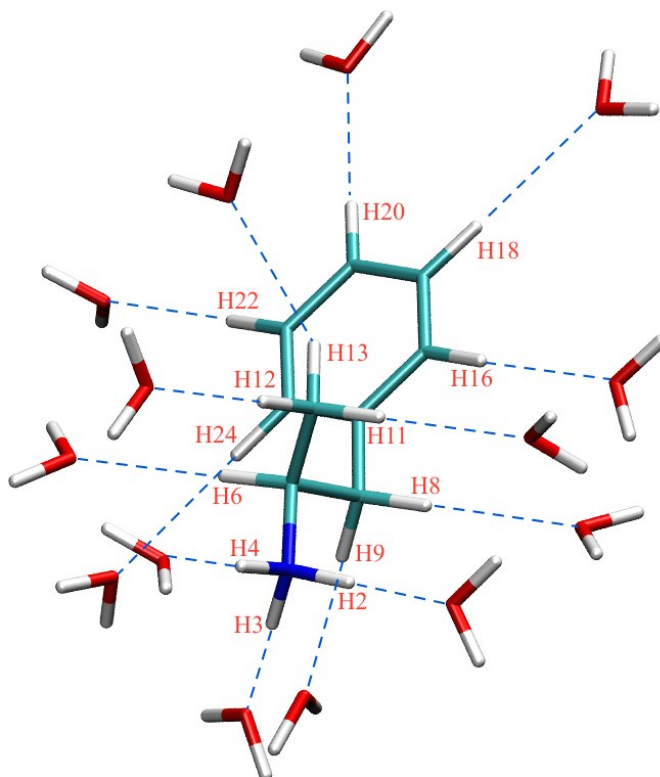
Figure 2.3. Orientation of water molecules around amphetamine used for charge optimization.

Fourteen monohydrates of amphetamine in different configurations were generated using the MP2/6-31G(d) optimized geometry, and HF/6-31G* water interaction energies of these individual hydrates were compared against the CGenFF interaction energies. The CGenFF charges on the individual atoms were slightly adjusted, and the CGenFF interaction energies were recalculated to match the HF/6-31G* interaction energies.

According to the CGenFF parameterization philosophy, ideally the CGenFF and HF/6-31G* interaction energy differences should be within 0.2 kcal/mol, and the CGenFF interaction distances should be 0.2Å less than the QM distances since the interaction distances in bulk are typically shorter than in vacuum [16].

Table 2.1. Comparison of CGenFF and HF/6-31G* water interaction energies and distances of amphetamine.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H24…OHH	-5.45	-5.29	0.15	2.55	2.65	0.1
H22…OHH	-4.28	-3.93	0.34	2.41	2.60	0.19
H20…OHH	-3.96	-3.54	0.41	2.43	2.62	0.19
ZH18…OHH	-3.97	-3.63	0.34	2.45	2.62	0.17
H16…OHH	-4.27	-4.32	-0.05	2.61	2.68	0.07
H8.....OHH	-7.18	-6.91	0.26	2.42	2.58	0.16
H9.....OHH	-7.01	-6.85	0.15	2.45	2.58	0.16
H6.....OHH	-7.93	-7.91	0.02	2.31	2.54	0.23
H11…OHH	-4.97	-5.17	-0.20	2.61	2.60	-0.01
H12…OHH	-6.80	-6.51	0.28	2.43	2.56	0.13
H13…OHH	-6.99	-7.00	-0.01	2.43	2.54	0.11
H2.....OHH	-16.87	-17.06	-0.19	1.84	1.77	-0.07
H3.....OHH	-17.19	-17.14	0.05	1.83	1.76	-0.07
H4.....OHH	-17.08	-17.16	-0.08	1.84	1.76	-0.08
AD	0.105					
RMSD	0.197					
AAE	0.18					

All energies are in kcal/mol and distances in Å. ΔE is the water interaction energy, ΔΔE is the difference of HF/6-31G* and CGenFF water interaction energies, r is the interaction distance, and Δr is the difference of HF and CGenFF interaction distances.

As shown in Table 2.1 and Figure 2.3, the CGenFF interaction energies were in good agreement with the HF/6-31G* data. The interaction energy differences of nine hydrates were in the ideal range, and the differences of the remaining six hydrates were less than 0.4 kcal/mol. The interaction distances of three ammonium hydrogens (H2, H3 and H4) with water were ideally lower than HF/6-31G*, which might be because they are regular

hydrogen bonds; whereas the interaction distances of the remaining hydrates were greater than or equal to HF/6-31G* (but the deviation was lesser than 0.2Å) due to the weaker hydrogen bonding nature of the interactions.

b) Dopamine

Table 2.2. Interaction energies and distances of dopamine - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H7…OHH	-6.67	-6.49	0.18	2.45	2.60	0.15
H8…OHH	-4.65	-4.46	0.19	2.33	2.56	0.23
H5…OHH	-7.55	-7.35	0.20	2.32	2.52	0.20
H10…OHH	-7.03	-6.38	0.65	2.41	2.59	0.18
H11…OHH	-6.68	-5.49	1.19	2.33	2.61	0.28
H13…OHH	-8.72	-7.96	0.77	2.27	2.55	0.28
H14…OHH	-8.70	-7.59	1.10	2.23	2.53	0.30
H15…OHH	-2.78	-1.97	0.80	2.54	2.50	-0.04
H16…OHH	-17.06	-16.87	0.18	1.82	1.77	-0.05
H17…OHH	-16.94	-17.01	-0.07	1.84	1.77	-0.07
O4…HOH	-3.05	-3.23	-0.18	2.35	2.22	-0.13
H4…OHH	-12.06	-12.16	-0.10	1.87	1.81	-0.05
O1…HOH	0.30	0.09	0.21	2.21	2.03	-0.18
H1…OHH	-2.81	-2.98	-0.16	3.3	3.44	0.14
AD	0.32					
RMSD	0.47					
AAE	0.43					

Fourteen monohydrates of dopamine were generated (Fig. 2.4) in various configurations and the water interaction energies and distances were calculated to optimize the charges. The interaction energies of the real hydrates H16…OHH, H17…OHH, O4…HOH, H4…OHH, O1…HOH, and H1…OHH were in the ideal range (i.e., deviation within 0.2 kcal/mol), except for H15…OHH, which was 0.8 kcal/mol (Table 2.2). The interaction distances of the real hydrogen bonds were lower than the QM distances, except for H1…OHH. The interaction energy differences of aliphatic hydrogens (H10, H11, H13,

and H14) were not in the ideal range as expected; however, the deviations were less than 0.2 kcal/mol for the three aromatic hydrogens. The interaction distances of the three aromatic and aliphatic hydrogens with water were greater than QM distances, probably due to the feeble hydrogen bond nature of the interactions.
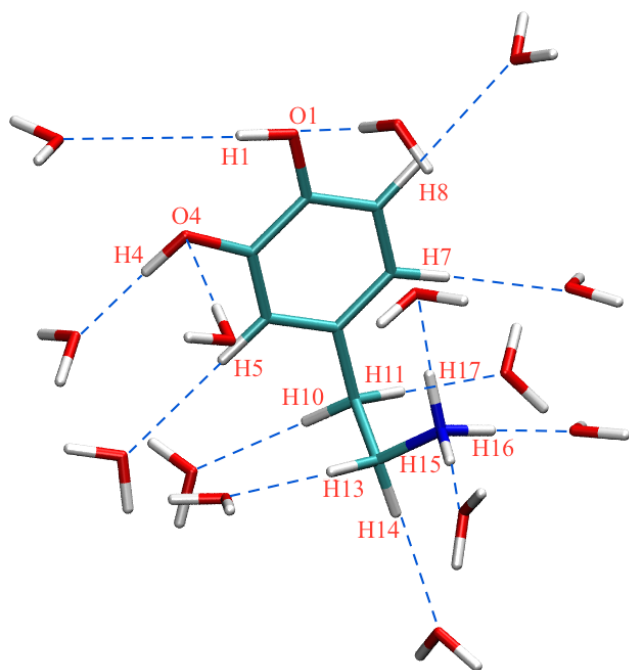
Figure 2.4. Orientation of water molecules around dopamine used for charge optimization.

c) Serotonin

Fifteen monohydrates of serotonin were generated in multiple configurations (Fig. 2.5). The interaction energy differences of the real hydrogen bonds (N9…HOH, H9…OHH, 01…HOH, H1…OHH, H17…OHH, and H18…OHH) were below 0.35 kcal/mol, except for the H19…OHH because of the bad atom position (Table 2.3). The interaction energy differences of the remaining hydrates were below 0.3 kcal/mol, except for one hydrate (H11…OHH), which was -0.62 kcal/mol. The interaction distances of all the real

hydrates were lower than the QM distances except for the hydrate H19…OHH. The interaction distances of remaining hydrates (except H8…OHH) were greater than QM distances.

Table 2.3. Interaction energies and distances of serotonin - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H14…OHH	-7.88	-7.62	0.25	2.27	2.53	0.26
H15…OHH	-8.15	-7.99	0.15	2.25	2.51	0.26
H11…OHH	-6.36	-6.98	-0.62	2.46	2.54	0.08
H12…OHH	-6.32	-6.01	0.29	2.35	2.57	0.22
H8…OHH	-6.77	-6.91	-0.14	2.26	2.22	-0.04
H5…OHH	-5.59	-5.64	-0.05	2.33	2.52	0.19
H6…OHH	-5.96	-5.82	0.13	2.31	2.51	0.2
H2…OHH	-6.20	-6.01	0.18	2.58	2.64	0.06
N9…HOH	-3.34	-3.53	-0.19	3.12	2.94	-0.18
H9…OHH	-10.46	-10.53	-0.07	1.95	1.84	-0.11
01…HOH	-3.93	-3.83	0.09	2.07	1.95	-0.12
H1…OHH	-10.80	-10.97	-0.17	1.92	1.83	-0.09
H17…OHH	-16.20	-15.92	0.27	1.84	1.77	-0.07
H18…OHH	-16.18	-16.53	-0.35	1.86	1.78	-0.08
H19…OHH	-11.05	-2.96	8.09	2.49	2.56	0.54
AD	-0.01					
RMSD	0.25					
AAE	0.21					

Figure 2.5. Orientation of water molecules around serotonin used for charge optimization.

d) Methylphenidate

Twenty-two monohydrates of methylphenidate were generated to optimize the CGenFF charges (Fig. 2.6). Although the interaction energy differences of half of the hydrates were in the ideal range, differences of the remaining half deviated significantly (Table 2.4). The interaction energy differences indicate that the charges require better fitting to minimize the energy differences. The interaction distances of real hydrates were lower than the QM distances as expected, except for the hydrate H14…OHH; whereas the interaction distances of the rest of the hydrates (aliphatic and aromatic) were greater than QM distances.

Table 2.4. Interaction energies and distances of methylphenidate - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	$\Delta\Delta E$ (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H7…OHH	-7.75	-6.48	1.27	2.39	2.62	0.24
H4…OHH	-4.18	-3.62	0.56	2.73	2.83	0.10
H3…OHH	-3.75	-2.39	1.36	2.46	2.71	0.25
H2…OHH	-3.77	-2.17	1.60	2.44	2.7	0.26
H1…OHH	-3.67	-2.18	1.49	2.46	2.71	0.25
H6…OHH	-2.5	-1.84	0.66	2.89	4.49	1.60
H11…OHH	-7.45	-8.69	-1.24	2.22	2.50	0.28
H16…OHH	-8.08	-9.44	-1.36	2.27	2.49	0.22
H17…OHH	-7.85	-8.67	-0.82	2.26	2.50	0.24
H19…OHH	-5.90	-6.63	-0.73	2.47	2.56	0.09
H20…OHH	-5.90	-6.51	-0.61	2.37	2.55	0.18
H22…OHH	-5.12	-5.49	-0.37	2.42	2.58	0.16
H23…OHH	-5.41	-6.80	-1.39	2.48	2.56	0.08
H25...OHH	-5.14	-5.78	-0.64	2.48	2.57	0.09
H26...OHH	-6.42	-7.06	-0.64	2.69	2.6	-0.09
H13…OHH	-15.15	-17.42	-2.27	1.90	1.78	-0.12
H14…OHH	-11.04	-7.80	3.24	2.40	2.59	0.19
O8…HOH	-1.49	-4.53	-3.04	2.23	1.79	-0.44
O9….OHH	1.70	-0.18	-1.88	2.84	2.08	-0.76
H101…OHH	-5.44	-5.61	-0.17	2.44	2.57	0.13
H102..OHH	-4.91	-4.43	0.48	2.32	2.53	0.21
H103..OHH	-4.91	-4.92	-0.01	2.38	2.54	0.16
AD	-0.20					
RMSD	1.46					
AAE	1.14					

Figure 2.6. Orientation of water molecules around methylphenidate used for charge optimization.

e) Bupropion

Fifteen monohydrates of bupropion were generated to optimize the charges (Fig. 2.7). The interaction energy differences of most of the hydrates were deviating significantly from the ideal range (Table 2.5). The interaction distances of the real hydrates (H14…OHH, H15…OHH, and O7…HOH) and also CL2…HOH were smaller than the QM distances; however, they were larger for the remaining hydrates.

Table 2.5. Interaction energies and distances of bupropion - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
CL2…HOH	-0.27	-1.13	-0.86	3.02	2.39	-0.63
H3…OHH	-4.35	-2.01	2.34	2.26	2.7	0.44
H1…OHH	-5.63	-2.46	3.17	2.26	2.66	0.40
H6…OHH	-5.79	-3.33	2.46	2.32	2.66	0.34
H5…OHH	-8.45	-7.55	0.90	2.54	2.76	0.22
H8…OHH	-9.45	-10.68	-1.23	2.33	2.51	0.18
H10…OHH	-7.75	-7.49	0.26	2.34	2.53	0.19
H11…OHH	-5.08	-6.40	-1.32	2.56	2.59	0.03
H12…OHH	-7.25	-7.28	-0.03	2.31	2.51	0.20
H17…OHH	-8.27	-9.85	-1.58	2.21	2.46	0.25
H19…OHH	-8.45	-9.41	-0.96	2.23	2.46	0.23
H18…OHH	-8.51	-10.0	-1.49	2.26	2.46	0.20
H14…OHH	-4.16	-6.19	-2.03	3.19	2.89	-0.30
H15…OHH	-15.70	-19.21	-3.51	1.88	1.77	-0.11
O7…HOH	-0.39	-2.45	-2.06	2.22	1.71	-0.51
AD	-0.39					
RMSD	1.89					
AAE	1.53					

Figure 2.7. Orientation of water molecules around bupropion used for charge optimization.

f) Benztropine

Twenty-seven monohydrates of benztropine have been generated (Fig. 2.8) to optimize the partial atomic charges. Except for the eleven hydrates (including the real hydrates H5...OHH and O24…HOH), the interaction energy differences of rest of the hydrates were below 1 kcal/mol (Table 2.6). The interaction distances of the real hydrates H5...OHH and O24…HOH were lower than the QM distances, but distances of rest of the aliphatic and aromatic hydrates were greater than the CGenFF distances.

Table 2.6. Interaction energies and distances of benztropine - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H3...OHH	-7.75	-8.78	-1.03	2.60	2.64	0.04
H4...OHH	-7.75	-8.82	-1.07	2.26	2.48	0.22
H2...OHH	-7.95	-8.70	-0.75	2.26	2.47	0.21
H9...OHH	-3.97	-4.19	-0.22	4.12	2.55	-1.57
H10...OHH	-6.45	-7.14	-0.69	2.39	2.56	0.17
H12...OHH	-6.39	-7.25	-0.86	2.42	2.56	0.14
H13...OHH	-6.30	-7.35	-1.05	2.41	2.55	0.14
H17...OHH	-5.62	-5.51	0.11	2.38	2.58	0.20
H18...OHH	-5.74	-6.19	-0.45	2.44	2.59	0.15
H22...OHH	-3.26	-4.04	-0.78	3.23	3.14	-0.09
H23...OHH	-5.28	-6.38	-1.10	2.55	2.57	0.02
H7...OHH	-7.84	-7.54	0.30	2.29	2.55	0.26
H15...OHH	-7.44	-8.32	-0.88	2.33	2.52	0.19
H20...OHH	-5.38	-5.65	-0.27	3.54	3.65	0.11
H37...OHH	-2.63	-1.79	0.84	2.81	3.05	0.24
H35...OHH	-2.71	-1.65	1.06	2.55	2.75	0.20
H33...OHH	-2.88	-1.70	1.18	2.49	2.72	0.23
H31...OHH	-2.91	-1.71	1.20	2.49	2.72	0.23
H29...OHH	-3.18	-2.20	0.98	2.52	2.72	0.20
H26...OHH	-4.31	-4.28	0.03	2.58	2.65	0.07
H48...OHH	-4.05	-3.90	0.15	2.55	2.68	0.13
H46...OHH	-3.64	-2.81	0.83	2.49	2.70	0.21
H44...OHH	-3.61	-2.46	1.15	2.46	2.70	0.24
H42…OHH	-3.50	-2.24	1.26	2.46	2.71	0.25
H40…OHH	-2.46	-1.57	0.89	2.75	2.86	0.11
H5...OHH	-13.73	-18.17	-4.44	1.95	1.79	-0.16
O24…HOH	-1.15	-3.26	-2.11	2.84	2.62	-0.22
AD	-0.21					
RMSD	1.26					
AAE	0.94					

Figure 2.8. Orientation of water molecules around benztropine used for charge optimization.

g) Cocaine

Twenty-six monohydrates were generated, and the interaction energies to optimize the CGenFF charges were calculated (Fig. 2.9). The interaction energy differences of the real hydrogen bonds (H13…OHH, O3...HOH, O20…HOH, O1…HOH and O21…HOH) were lower than 0.28 kcal/mol and close to the ideal range. The interaction energy differences of rest of the hydrates were below 1.26 kcal/mol, except for the H161...OHH, which was -2.91 kcal/mol because of the bad atom position. The interaction distances of the real hydrogen bonds were lower than the QM distances, except for the hydrate O1…HOH. The interaction distances of the remaining hydrates (except H161...OHH) were greater than QM distances (Table 2.7).

Table 2.7. Interaction energies and distances of cocaine - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H181…OHH	-7.73	-7.47	0.26	2.27	2.51	0.24
H182…OHH	-7.93	7.81	0.11	2.27	2.50	0.23
H183…OHH	-8.98	-8.39	1.16	2.23	2.42	0.19
H221…OHH	-3.96	-2.69	1.26	2.45	2.69	0.24
H222…OHH	-3.03	-2.96	0.06	2.55	2.67	0.12
H223…OHH	-3.68	-2.94	0.73	2.44	2.62	0.18
H5…OHH	-2.10	-1.88	0.21	2.36	2.65	0.29
H6…OHH	-3.26	-3.02	0.24	2.47	2.63	0.16
H7…OHH	-3.49	-3.34	0.14	2.44	2.6	0.16
H8…OHH	-3.01	-2.80	0.21	2.53	2.65	0.12
H9…OHH	-2.35	-2.42	-0.07	3.00	4.49	1.49
H10…OHH	-3.86	-3.56	0.30	2.53	2.73	0.2
H111…OHH	-3.90	-3.96	-0.06	2.44	2.66	0.22
H112…OHH	-6.41	-6.06	0.35	2.55	2.64	0.09
H12…OHH	-8.17	-7.89	0.28	2.26	2.53	0.27
H14…OHH	-7.38	-7.14	0.24	2.29	2.56	0.27
H15...OHH	-5.92	-5.60	0.32	2.58	2.72	0.14
H161...OHH	-0.93	-3.85	-2.91	6.76	2.77	-3.99
H162...OHH	-6.21	-5.94	0.27	2.41	2.58	0.17
H171…OHH	-6.33	-6.45	-0.12	2.39	2.56	0.17
H172...OHH	-7.05	-6.57	0.48	2.21	2.49	0.28
H13…OHH	-15.10	-14.82	0.28	1.89	1.85	-0.04
O3...HOH	-3.40	-3.59	-0.19	2.06	1.85	-0.21
O20…HOH	-2.21	-2.44	-0.23	2.16	1.88	-0.28
O1…HOH	-3.19	-3.36	-0.17	2.20	2.26	0.06
O21…HOH	-2.41	-2.56	-0.15	2.11	1.93	-0.18
AD	-0.09					
RMSD	0.68					
AAE	0.39					

Figure 2.9. Orientation of water molecules around cocaine used for charge optimization.

h) Clomipramine

Twenty-four hydrates (Fig. 2.10) of clomipramine were generated to optimize the charges. As shown in Table 2.8, most of the interaction energy differences were far away from the ideal situation and require further optimization. The positive QM energies of the hydrates NZ…HOH and HB23...OHH indicate that these hydrates have to be carefully analyzed, and energies are to be recalculated. Interestingly, the interaction distances of CL...HOH and NZ…HOH were lower than the QM distances, although energies deviated significantly. The interaction distances of rest of the hydrates were greater than the QM distances as expected since they were not real hydrogen bonds.

Table 2.8. Interaction energies and distances of clomipramine - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
HA11...OHH	-6.34	-10.40	-4.06	2.33	2.46	0.13
HA12...OHH	-7.72	-10.40	-2.68	2.25	2.45	0.2
HA13...OHH	-7.39	-10.01	-2.62	2.28	2.46	0.18
HB21...OHH	-6.34	-10.21	-3.87	2.52	2.46	-0.06
HB22...OHH	-7.72	-9.89	-2.17	2.29	2.46	0.17
HG1...OHH	-8.06	-9.55	-1.49	2.32	2.49	0.17
HG2...OHH	-7.95	-9.18	-1.23	2.30	2.51	0.21
HD1...OHH	-5.60	-5.96	-0.36	2.65	2.67	0.02
HD2...OHH	-6.05	-7.74	-1.69	2.36	2.58	0.22
HE1...OHH	-7.15	-6.14	1.01	2.36	2.56	0.2
HE2...OHH	-5.60	-4.16	1.44	2.48	2.95	0.47
HD4...OHH	-5.88	-2.31	3.57	2.54	4.31	1.77
HE4...OHH	-3.92	-1.84	2.08	2.43	2.74	0.31
HZ2...OHH	-4.52	-1.89	2.63	2.32	2.70	0.38
HE5...OHH	-5.27	-1.86	3.41	2.37	2.73	0.36
HE11...OHH	-5.00	-1.58	3.42	2.54	2.80	0.26
HE12...OHH	-5.46	-1.17	4.29	3.35	3.18	-0.17
HE21...OHH	-3.70	-1.06	2.64	2.60	2.85	0.25
HE22...OHH	-3.64	-0.46	3.18	2.51	2.88	0.37
HE3...OHH	-4.09	-1.44	2.65	2.44	2.74	0.3
HZ1...OHH	-3.82	-1.71	2.11	2.42	2.72	0.3
HD3...OHH	-5.21	-1.92	3.29	2.86	2.96	0.1
CL...HOH	6.04	-1.71	-7.75	2.75	2.34	-0.41
NZ…HOH	7.45	0.95	-6.5	3.58	3.39	-0.19
AD	0.05					
RMSD	3.40					
AAE	2.91					

Figure 2.10. Orientation of water molecules around clomipramine used for charge optimization.

i) Imipramine

Twenty-six monohydrates (Fig. 2.11) of imipramine were generated, and charges were adjusted to match the interaction energies. The interaction energy differences of the real hydrates (HB1…OHH and NZ…HOH) were below 1.29 kcal/mol. Half of the remaining hydrates were close to the ideal range and the rest deviated significantly (Table 2.9). The interaction distances of the real hydrogen bonds were smaller than the QM distances; whereas the distances of remaining hydrates were greater than QM, except for a few.

Table 2.9. Interaction energies and distances of imipramine - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
HA11…OHH	-9.78	-9.75	0.03	2.22	2.46	0.24
HA12…OHH	-9.20	-9.71	-0.51	2.25	2.46	0.21
HA13…OHH	-9.74	-9.93	-0.19	2.23	2.46	0.23
HB21…OHH	-9.21	-9.59	-0.38	2.27	2.47	0.2
HB22…OHH	-9.21	-9.83	-0.62	2.27	2.47	0.2
HB23…OHH	-6.62	-8.95	-2.33	2.18	2.48	0.3
HG1…OHH	-9.59	-8.91	0.68	2.26	2.50	0.24
HG2…OHH	-8.80	-9.19	-0.39	2.32	2.50	0.18
HD1…OHH	-7.11	-7.35	-0.24	2.53	2.60	0.07
HD2…OHH	-3.92	-4.54	-0.62	4.30	3.26	-1.04
HE1…OHH	-4.69	-3.81	0.88	2.61	3.01	0.4
HE2…OHH	-5.14	-3.71	1.43	2.95	2.70	-0.25
HD3...OHH	-4.16	-1.89	2.27	2.81	2.95	0.14
HD4…OHH	-3.74	-3.19	0.55	2.56	2.86	0.3
HE6…OHH	-3.64	-2.48	1.16	2.53	2.72	0.19
HZ2…OHH	-3.44	-2.11	1.33	2.59	2.71	0.12
HE5…OHH	-3.03	-1.78	1.25	2.94	2.74	-0.2
HE11...OHH	-3.05	-1.48	1.57	2.62	2.81	0.19
HE12...OHH	-3.71	-1.31	2.40	2.62	3.10	0.47
HE21...OHH	-4.41	-0.31	4.10	2.57	2.90	0.33
HE22...OHH	-4.09	-1.07	3.02	2.70	2.83	0.13
HZ1...OHH	-4.06	-1.69	2.37	2.52	2.72	0.20
HE3…OHH	-4.19	-1.41	2.78	2.55	2.74	0.19
HE4…OHH	-3.97	-1.85	2.12	2.58	2.74	0.16
HB1…OHH	-17.17	-18.46	-1.29	1.86	1.80	-0.06
NZ…HOH	-0.79	-0.10	0.69	3.29	2.73	-0.56
AD	0.84					
RMSD	1.49					
AAE	1.20					
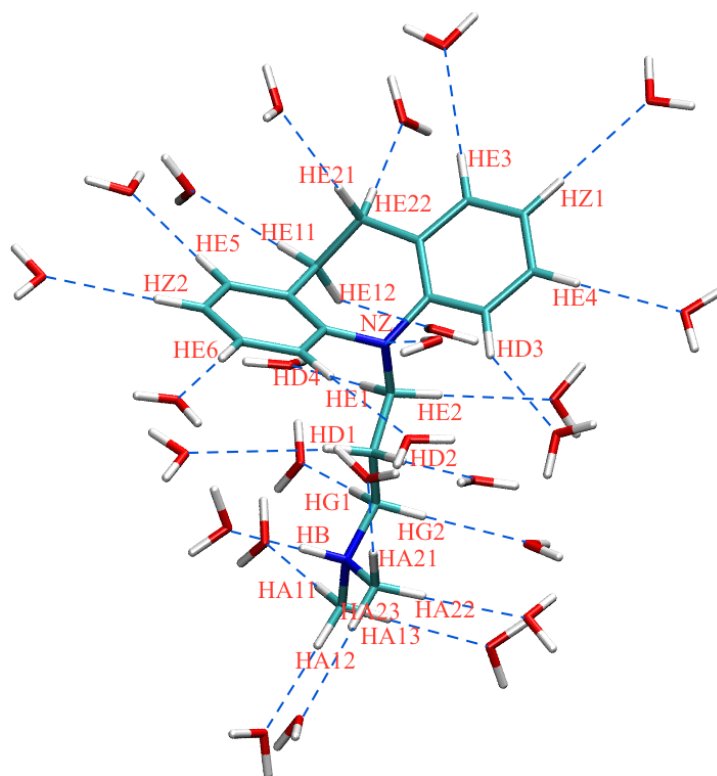
Figure 2.11. Orientation of water molecules around imipramine used for charge optimization.

j) Desipramine

Twenty-four monohydrates of desipramine were generated to optimize the charges (Fig. 2.12). The interaction energy differences of more than half of the hydrates deviated significantly from the ideal situation (Table 2.10), including the real hydrogen bonds. The interaction distances of the hydrates NZ…HOH, HB1…OHH, and HB2…OHH were lower than the QM distances, whereas for the rest they were greater than QM distances. The QM interaction distance of the hydrate HE22...OHH was huge (6.68Å), probably due to the bad atom position.

Table 2.10. Interaction energies and distances of desipramine - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
HA11…OHH	-10.02	-9.78	0.24	2.20	2.46	0.26
HA12…OHH	-2.74	-4.30	-1.56	1.94	2.27	0.33
HA13…OHH	-9.54	-9.94	-0.4	2.24	2.45	0.21
HG1…OHH	-9.69	-9.52	0.17	2.24	2.48	0.24
HG2…OHH	-9.99	-9.24	0.75	2.23	2.48	0.25
HD1…OHH	-4.11	-5.03	-0.92	5.74	4.08	-1.66
HD2…OHH	-7.39	-7.21	0.18	2.5	2.59	0.09
HE1…OHH	-5.03	-3.98	1.05	2.59	2.70	0.11
HE2…OHH	-4.81	-3.56	1.25	3.04	3.08	0.04
HE3…OHH	-4.49	-1.91	2.58	2.54	2.73	0.19
HE4...OHH	-4.31	-2.71	1.6	2.57	2.72	0.15
HE5…OHH	-4.02	-1.44	2.58	2.56	2.74	0.18
HE6…OHH	-3.73	-1.89	1.84	2.59	2.74	0.15
HZ1...OHH	-4.36	-2.25	2.11	2.51	2.71	0.2
HZ2…OHH	-3.92	-1.73	2.19	2.53	2.72	0.19
HE11...OHH	-3.49	-0.37	3.12	2.56	2.89	0.33
HE12...OHH	-4.00	-1.12	2.88	2.61	2.83	0.22
HD3...OHH	-4.32	-2.96	1.36	2.79	2.88	0.09
HD4…OHH	-4.08	-2.45	1.63	2.75	2.9	0.15
HE21...OHH	-4.40	-1.58	2.82	2.64	2.81	0.17
HE22...OHH	-3.12	-1.51	1.61	6.68	3.06	-3.62
NZ…HOH	-1.56	-0.20	1.36	3.27	2.89	-0.38
HB1…OHH	-17.67	-20.41	-2.74	1.84	1.75	-0.09
HB2…OHH	-17.54	-20.30	-2.76	1.85	1.75	-0.1
AD	0.95					
RMSD	1.66					
AAE	1.30					

Figure 2.12. Orientation of water molecules around desipramine used for charge optimization.

k) R-Fluoxetine

Twenty-one hydrates of R-fluoxetine were generated to optimize the charges (Fig. 2.13 and Table 2.11). The interaction energy difference of the hydrate H33…OHH was huge, probably due to the bad atom position. The interaction distances of real hydrogen bonds (O15…HOH, H36…OHH, H37…OHH, and H38…OHH) and also F12...HOH, F13...HOH, and F14...HOH were below QM distances; whereas, for most of the hydrates, they were greater than QM distances.

Table 2.11. Interaction energies and distances of fluoxetine (R) - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H33…OHH	11.70	-8.23	-19.93	2.23	2.75	0.52
H34…OHH	-5.80	-9.77	-3.97	3.14	2.46	-0.68
H30…OHH	-7.72	-5.49	2.23	2.46	3.39	0.93
H31…OHH	-6.72	-7.76	-1.04	2.66	2.56	-0.1
H17…OHH	-6.98	-3.59	3.39	2.42	2.73	0.31
H20…OHH	-4.45	-2.93	1.52	2.68	3.53	0.85
H22…OHH	-3.99	-2.06	1.93	2.44	2.74	0.3
H24…OHH	-4.02	-1.83	2.19	2.41	2.72	0.31
H26…OHH	-4.12	-1.55	2.57	2.43	2.73	0.3
H28…OHH	-4.99	-0.35	4.64	2.48	3.01	0.53
H12…OHH	-6.27	-2.93	3.34	2.48	2.75	0.27
H14…OHH	-4.36	-4.24	0.12	2.33	2.58	0.25
H9…OHH	-6.27	-1.76	4.51	2.50	3.5	1
H7…OHH	-4.41	-4.18	0.23	2.33	2.58	0.25
F12...HOH	3.64	-1.82	-5.46	2.18	2.06	-0.12
F13...HOH	2.96	-1.33	-4.29	2.16	1.96	-0.2
F14...HOH	4.80	-1.31	-6.11	2.21	1.96	-0.25
O15...HOH	2.66	-5.28	-7.94	2.33	1.74	-0.59
H36…OHH	-17.75	-22.53	-4.78	1.81	1.71	-0.1
H37…OHH	-17.87	-22.58	-4.71	1.82	1.72	-0.1
H38…OHH	-17.48	-22.38	-4.9	1.82	1.71	-0.11
AD	-1.73					
RMSD	4.42					
AAE	5.72					

Figure 2.13. Orientation of water molecules around R-fluoxetine used for charge optimization.

l) S-Fluoxetine

Twenty-one hydrates of R-fluoxetine were generated to optimize the charges (Fig. 2.14 and Table 2.12). The interaction energy difference of the hydrate H12…OHH was huge due to the bad atom position. The interaction distances of the real hydrogen bonds (H36…OHH, H37…OHH, and H38…OHH) and also F12...HOH, F13...HOH, and F14...HOH were below QM distances (except O15…HOH), whereas the distances of the rest of the hydrates were greater than QM distances.

Table 2.12. Interaction energies and distances of fluoxetine (S) - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H33...OHH	-6.71	-5.49	1.22	2.36	3.83	1.47
H34...OHH	-7.57	-5.76	1.81	2.03	2.58	0.55
H30...OHH	-7.78	-7.20	0.58	2.7	2.61	-0.09
H31...OHH	-7.27	-4.92	2.35	2.34	4.27	1.93
H17...OHH	-5.87	-3.30	2.57	2.36	2.76	0.4
H20...OHH	-4.35	-3.36	0.99	2.36	2.72	0.36
H22...OHH	-4.17	-2.49	1.68	2.4	2.71	0.31
H24...OHH	-4.15	-2.33	1.82	2.45	2.70	0.25
H26...OHH	-4.25	-2.38	1.87	2.29	2.71	0.42
H28...OHH	-5.08	-2.12	2.96	2.51	3.6	1.09
H12...OHH	30.09	-3.20	-33.29	2.42	2.77	0.35
H14...OHH	-3.66	-4.22	-0.56	2.4	2.58	0.18
H9...OHH	-3.31	-3.65	-0.34	2.41	2.77	0.36
H7...OHH	-3.75	-4.21	-0.46	2.45	2.58	0.13
F12...HOH	3.92	-1.56	-5.48	2.19	1.97	-0.22
F13...HOH	4.45	-0.58	-5.03	2.19	2.0	-0.19
F14...HOH	2.69	0.45	-2.24	2.15	2.05	-0.1
O15...HOH	-5.72	-1.32	4.4	2.33	2.69	0.36
H36...OHH	-17.75	-22.32	-4.57	1.81	1.71	-0.1
H37...OHH	-17.68	-22.49	-4.81	1.82	1.71	-0.11
H38...OHH	-17.56	-22.17	-4.61	1.82	1.72	-0.1
AD	-0.29					
RMSD	3.08					
AAE	2.54					

Figure 2.14. Orientation of water molecules around
S-fluoxetine used for charge optimization.

m) R-Citalopram

A total of twenty-five monohydrates (Fig. 2.15) of R-citalopram were generated to optimize the partial charges. The interaction energy differences of the three real hydrates (N1…HOH, H15…OHH and O11…HOH) and F27...HOH were below 1.4 kcal/mol (Table 2.13). The interaction energy differences of half of the remaining twenty-one hydrates were in the ideal range and rest of them deviated significantly. The interaction distances of the three real hydrogen bonds and F27...HOH were lower than the QM distances. The distances of half of the remaining hydrates were also lower than the QM distances, whereas the distances of the other half were greater than QM.

Table 2.13. Interaction energies and distances of citalopram (R) - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H17…OHH	-8.79	-11.56	-2.77	2.23	1.99	-0.24
H18…OHH	-2.08	-8.60	-6.52	2.03	1.90	-0.13
H19…OHH	-8.05	-10.72	-2.67	2.28	2.01	-0.27
H21…OHH	-8.88	-11.11	-2.23	1.87	2.00	0.13
H22…OHH	-8.58	-10.66	-2.08	2.23	2.00	-0.23
H23…OHH	-7.62	-9.56	-1.94	2.24	2.23	-0.01
H36…OHH	-8.24	-6.78	1.46	2.45	2.66	0.21
H37…OHH	-0.91	-8.18	-7.27	2.36	3.53	1.17
H34…OHH	-5.12	-4.95	0.17	4.68	2.70	-1.98
H35…OHH	-6.30	-6.22	0.08	2.54	2.67	0.13
H32…OHH	-6.16	-5.26	0.9	2.58	2.74	0.16
H33…OHH	-4.29	-4.46	-0.17	2.67	2.61	-0.06
H25…OHH	-1.94	-1.79	0.15	2.528	2.95	0.422
H26…OHH	-3.64	-1.62	2.02	3.04	2.73	-0.31
H29…OHH	-4.89	-3.41	1.48	2.37	2.74	0.37
H28…OHH	-3.77	-1.90	1.87	2.57	2.53	-0.04
H30...OHH	-4.15	-4.40	-0.25	2.37	2.66	0.29
H31...OHH	-3.92	-5.49	-1.57	2.62	2.55	-0.07
H8.....OHH	-5.76	-5.46	0.3	2.46	2.40	-0.06
H4….OHH	-5.52	-4.99	0.53	2.27	2.40	0.13
H5….OHH	-6.12	-5.08	1.04	2.28	2.60	0.32
F27...HOH	-1.02	-2.42	-1.4	2.38	1.95	-0.43
N1….HOH	-2.87	-3.16	-0.29	2.27	2.00	-0.27
H15…OHH	-15.33	-13.95	1.38	2.29	1.85	-0.44
O11…HOH	-4.05	-3.56	0.49	1.99	1.81	-0.18
AD	-0.69					
RMSD	2.35					
AAE	1.78					

Figure 2.15. Orientation of water molecules around
R-citalopram used for charge optimization.

n) S-Citalopram

Twenty-five monohydrates (Fig. 2.16) of S-citalopram were generated and optimized the
charges. Although the interaction energy differences of half of the hydrates are in the
ideal range, the remaining half deviated significantly (Table 2.14). The interaction
distances of the real hydrates (H15…OHH, N1…HOH, and O11…HOH) and F27...HOH
were smaller than the QM interaction distances, whereas the distances of the rest were
greater, except for six hydrates (Table 2.14).

72

Table 2.14. Interaction energies and distances of citalopram (S) - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H17…OHH	-8.29	-11.41	-3.12	2.22	1.99	-0.23
H18…OHH	-8.06	-10.42	-2.36	2.23	1.99	-0.24
H19…OHH	-8.19	-10.37	-2.18	2.23	2.01	-0.22
H21…OHH	-8.28	-11.02	-2.74	2.22	2.00	-0.22
H22…OHH	-8.17	-10.54	-2.37	2.23	2.00	-0.23
H23…OHH	-6.32	-8.60	-2.28	2.24	2.16	-0.08
H36…OHH	-8.15	-6.18	1.97	2.25	2.64	0.39
H37…OHH	-8.15	-6.19	1.96	2.28	2.63	0.35
H34…OHH	1.41	-4.16	-5.57	2.38	4.02	1.64
H35…OHH	-3.99	-4.76	-0.77	2.77	2.82	0.05
H32…OHH	-5.56	-5.32	0.24	2.59	2.66	0.07
H33…OHH	-3.68	-4.10	-0.42	2.57	2.67	0.10
H25…OHH	-3.72	-3.63	0.09	3.75	3.89	0.14
H26…OHH	-3.19	-1.99	1.20	2.39	2.72	0.33
H29…OHH	-2.70	-3.35	-0.65	2.70	2.90	0.20
H28…OHH	-3.07	-2.28	0.79	2.41	2.53	0.12
H30...OHH	-3.11	-4.09	-0.98	2.49	2.61	0.12
H31...OHH	-3.01	-4.97	-1.96	2.48	2.56	0.08
H8.....OHH	-4.07	-4.45	-0.38	2.29	2.41	0.12
H4….OHH	-3.79	-4.07	-0.28	2.30	2.41	0.11
H5….OHH	-3.49	-3.97	-0.48	2.52	2.64	0.12
H15…OHH	-15.04	-14.72	0.32	1.87	1.85	-0.02
F27...HOH	-0.60	-2.68	-2.08	2.25	1.95	-0.30
N1….HOH	-2.31	-3.09	-0.78	2.28	2.01	-0.27
O11…HOH	-6.41	-7.45	-1.04	2.02	1.91	-0.11
AD	-0.95					
RMSD	1.69					
AAE	1.29					

Figure 2.16. Orientation of water molecules around S-citalopram used for charge optimization.

o) **Sertraline**

Twenty monohydrates of sertraline were generated to optimize the charges (Fig. 2.17). Except for a few, most of the interaction energy differences deviated significantly from the ideal range (Table 2.15). The interaction distances of real hydrogen bonds (H6…OHH and H7…OHH), CL34...HOH, and CL36...HOH are lower than the QM distances, and the rest of them are greater than QM distances.

After optimizing the partial atomic charges; bonds, angles, dihedrals and impropers were optimized.

Table 2.15. Interaction energies and distances of sertraline - water complexes in different geometries.

Interaction geometry	ΔE (HF)	ΔE (CGenFF)	ΔΔE (HF-CGenFF)	r (HF)	r (CGenFF)	Δr (HF-CGenFF)
H2…OHH	-7.28	-10.05	-2.77	2.28	2.46	0.18
H3…OHH	-7.63	-10.05	-2.42	2.23	2.45	0.22
H4…OHH	-6.25	-9.19	-2.94	2.58	2.58	0
H9…OHH	-7.58	-10.10	-2.52	2.31	2.48	0.17
H17…OHH	-5.49	-3.61	1.88	2.35	2.65	0.30
H14…OHH	-4.60	-7.09	-2.49	3.08	2.92	-0.16
H15…OHH	-4.78	-4.05	0.73	2.43	2.64	0.21
H11…OHH	-6.06	-7.62	-1.56	2.50	2.61	0.11
H12…OHH	-4.92	-4.98	-0.06	2.37	2.57	0.20
H21…OHH	-5.00	-5.55	-0.55	2.47	2.64	0.17
H23…OHH	-4.18	-3.10	1.08	2.39	2.67	0.28
H25…OHH	-4.22	-2.70	1.52	2.37	2.68	0.31
H27…OHH	-4.31	-2.84	1.47	2.42	2.68	0.26
H30…OHH	-4.93	-4.82	0.11	2.42	2.64	0.22
H32…OHH	-4.63	-4.66	-0.03	2.29	2.56	0.27
H38...HOH	-1.77	-3.25	-1.48	4.90	2.67	-2.23
CL34...HOH	1.27	-0.59	-1.86	3.59	2.44	-1.15
CL36...HOH	1.41	-0.76	-2.17	3.12	2.42	-0.70
H6…OHH	-14.04	-20.07	-6.03	1.86	1.75	-0.11
H7…OHH	-3.45	-7.08	-3.63	4.45	4.01	-0.44
AD	-1.18					
RMSD	2.03					
AAE	1.68					

Figure 2.17. Orientation of water molecules around sertraline used for charge optimization.

2.3.2. Geometry optimization

After the charges were optimized, the geometries were optimized using MP2/6-31G(d) optimized equilibrium geometry as the target data. Bond lengths followed by bond angles, dihedrals, and impropers were optimized.

CGenFF optimized equilibrium geometry for almost all the compounds were reasonably in good agreement with the MP2 optimized equilibrium geometry.

a) Amphetamine

Amphetamine geometry was optimized using the MP2/6-31G(d) equilibrium geometry as the target. The CGenFF geometry fits well with the QM geometry. All the bond differences were below 0.03Å, and all the angle differences were below 1.8Å, except for

angles C5-N1-H3 and N1-C5-H6, which were 2.1º and 2.6º, respectively. All the dihedral

and improper angle differences were below 3º, except for the C5-C7-C14-C23 and C7-

N1-C5-C10, which were 4.2º and 4.8º, respectively (Table 2.16). These results

demonstrate that CGenFF geometry matches well with the QM optimized geometry.

Table 2.16. CGenFF equilibrium geometry of amphetamine compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/º			
N1-H2	1.03	1.04	-0.01	H2-C5-N1	110.4	109	1.4
N1-H3	1.03	1.04	-0.01	C5-N1-H3	111.9	109.8	2.1
N1-H4	1.03	1.04	-0.01	C5-N1-H4	111.6	109.8	1.8
N1-C5	1.53	1.51	0.02	N1-C5-C7	108.1	108.3	-0.2
C5-H6	1.09	1.12	-0.03	N1-C5-C10	108.0	108.3	-0.3
C5-C10	1.52	1.53	-0.01	N1-C5-H6	104.7	107.3	-2.6
C10-H11	1.09	1.11	-0.02	C5-C7-C14	110.1	110.3	-0.2
C10-H12	1.09	1.11	-0.02	C5-C7-H8	109.8	111.2	-1.4
C10-H13	1.10	1.11	-0.01	C5-C7-H9	109.5	111.2	-1.7
C5-C7	1.53	1.54	-0.01	N1-C5-C10	108.0	108.5	-0.53
C7-H8	1.10	1.11	-0.01	C5-C10-H11	108.7	110.5	-1.8
C7-H9	1.10	1.11	-0.01	C5-C10-H12	112.0	111.2	0.8
C7-C14	1.51	1.50	0.01	C5-C10-H13	111.1	111.31	-0.2
C14=C15	1.40	1.40	0.00	C7-C14-C23	119.8	119.8	0.0
C15-H16	1.09	1.08	0.01	C7-C14-C15	120.7	120.6	0.1
C15-C17	1.40	1.40	0.00	C14-C15-C17	120.2	120.4	-0.2
C17-H18	1.09	1.08	0.01	C14-C15-H16	120.0	120.2	-0.2
C17=C19	1.40	1.40	0.00	C15-C17-C19	120.1	119.6	0.5
C19-H20	1.09	1.08	0.01	C15-C17-H18	119.8	120.5	-0.7
C19-C21	1.40	1.40	0.00	C17-C19-C21	120.0	120.00	0.0
C21-H22	1.09	1.08	0.01	C17-C19-H20	120.1	119.8	0.3
C21=C23	1.39	1.40	-0.01	C23-C19-C21	120.0	119.7	0.3
C23-H24	1.09	1.08	0.01	C21-C14-C23	120.2	120.3	-0.1
C23-C14	1.40	1.40	0.00	C14-C23-H24	120.2	120.4	-0.2
Dihedrals/º				Impropers/º			
H2-N1-C5-C7	62.7	59.7	3.0	H2-C5-N1-H3	119.6	119.7	-0.1
N1-C5-C7-C14	174.4	175.5	-1.1	H2-C5-N1-H4	-119.4	-119.6	0.2
N1-C5-C10-H11	-178.5	-176.7	-1.8	C10-N1-C5-H6	-118.2	-119.6	1.4
C5-C7-C14-C23	-75.8	-71.6	-4.2	C14-C5-C7-H8	-121.0	-118.3	-2.7
C7-C14-C15-C17	-178.3	-179	0.7	H11-C5-C10-H12	119.2	118.9	0.3
C14-C15-C17-C19	0.5	0.1	0.4	H11-C5-C10-H13	-118.4	-119.7	1.3
C15-C17-C19-C21	-0.7	-0.2	-0.5	C23-C7-C14-C15	177.9	178.9	-1.0
				C17-C14-C15-H16	-178.5	-179.4	0.9
				C19-C15-C17-H18	-179.9	-179.8	-0.1
				C21-C17-C19-H20	-179.1	-179.5	0.4
				C23-C19-C21-H22	-179.7	-179.6	-0.1
				C21-C14-C23-H24	178.0	178.7	-0.7
				C7-N1-C5-C10	-124.5	-119.7	-4.8
				H8-C5-C7-H9	-117.8	-120.9	3.1

b) Dopamine

The MP2 optimized geometry was used as a target to optimize the CGenFF structure. The bond length and angle differences were less than 0.04Å and 3.2°, respectively, except for the C2-C3-O4, which was 7.2°. All the dihedral differences were less than 3.5°, except for C6-C9-C12-N15 and C9-C12-N15-H16, which were 6.1 and 13.2°, respectively. The impropers were less than 4.4°, except for the C6-C8-C7-H7, which was 6.4° (Table 2.17).

Table 2.17. CGenFF equilibrium geometry of dopamine compared to MP2 level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/°			
O1-H1	0.98	0.96	0.02	H1-O1-C2	108.4	107.2	1.2
O1-C2	1.35	1.39	-0.04	O1-C2-C8	119.7	117.50	2.2
C2-C3	1.41	1.41	0	O1-C2-C3	120.8	122.3	-1.5
C3-O4	1.37	1.40	-0.03	C2-C3-C5	120.6	119.6	1.0
O4-H4	0.97	0.96	0.01	C2-C3-O4	113.9	121.0	-7.2
C3=C5	1.39	1.40	-0.01	C3-O4-H4	110.8	109.06	1.7
C5-H5	1.09	1.08	0.01	C3-C5-C6	120.1	120.2	-0.1
C5-C6	1.41	1.40	0.01	C3-C5-H5	119.5	120.6	-1.1
C6=C7	1.40	1.40	0.00	C2-C8-C7	119.9	119.9	0
C7-H7	1.09	1.08	0.01	C2-C8-H8	118.5	118.9	-0.4
C7-C8	1.40	1.40	0.00	C6-C8-C7	120.9	120.2	0.7
C8-H8	1.09	1.08	0.01	C8-C7-H7	118.8	119.7	-1.1
C2=C8	1.39	1.40	-0.01	C7-C5-C6	119.0	119.9	-0.9
C6-C9	1.51	1.51	0.00	C5-C6-C9	120.1	119.9	0.2
C9-H10	1.10	1.11	-0.01	C6-C9-C12	110.6	112.3	-1.7
C9-H11	1.10	1.11	-0.01	C6-C9-H10	110.6	109.6	1.0
C9-C12	1.53	1.54	-0.01	C6-C9-H11	111.6	108.4	3.2
C12-H13	1.09	1.10	-0.01	C9-C12-N15	108.5	110.6	-2.1
C12-H14	1.09	1.10	-0.01	C9-C12-H13	111.1	111.5	-0.4
C12-N15	1.51	1.48	0.03	C9-C12-H14	113.7	110.7	3.0
N15-H15	1.04	1.04	0.00	C12-N15-H16	113.0	111.1	1.9
N15-H16	1.03	1.04	-0.01	C12-N15-H17	111.4	109.5	1.9
N15-H17	1.03	1.04	-0.01	C12-N15-H15	107.9	108.9	-1.0
Dihedrals/°				Improper dihedrals/°			
H1-O1-C2-C8	176.0	177.1	-1.1	C8-O1-C2-C3	-177.7	-179.2	2.5
O1-C2-C3-C5	-178.5	-179.6	1.1	C5-C2-C3-O4	177.5	180.1	-2.6
C2-C3-O4-H4	171.0	167.5	3.5	C6-C3-C5-H5	184.3	181.0	3.3
C2-C3-C5-C6	-4.0	-0.68	-3.3	C7-C2-C8-H8	180.3	178.7	1.6
O1-C2-C8-C7	179.4	180.1	-0.7	C6-C8-C7-H7	172.7	179.1	-6.4
C5-C6-C9-C12	-79.6	-77.8	-1.8	C7-C5-C6-C9	173.7	176.3	-2.6
C6-C9-C12-N15	-55.2	-61.3	6.1	C12-C6-C9-H10	-122.1	-123.5	1.4
C9-C12-N15-H16	163.6	176.8	-13.2	H10-C6-C9-H11	-118.4	-115.8	-2.6
				N15-C9-C12-H13	116.3	120.5	-4.2
				H13-C9-C12-H14	124.1	119.7	4.4
				H16-C12-N15-H17	121.5	120.9	0.6
				H16-C12-N15-H15	-121.2	-120.5	-0.7

c) Serotonin

Table 2.18. CGenFF equilibrium geometry of serotonin compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/°			
N16-H17	1.03	1.04	-0.01	H1-O1-C1	109.6	107.4	2.2
N16-H18	1.03	1.04	-0.01	O1-C1-C2	116.6	117.9	-1.3
N16-H19	1.04	1.04	0.00	O1-C1-C6	122.2	120.7	1.5
N16-C13	1.51	1.50	0.01	C1-C2-C3	118.5	119.1	-0.6
C13-H14	1.09	1.10	-0.01	C1-C2-H2	118.0	118.2	-0.2
C13-H15	1.09	1.10	-0.01	C2-C3-C7	133.7	132.1	1.6
C13-C10	1.53	1.55	-0.02	C2-C3-C4	119.3	120	-0.7
C10-H11	1.10	1.11	-0.01	C1-C6-C5	121.3	119.2	2.1
C10-H12	1.10	1.11	-0.01	C1-C6-H6	119.4	120.1	-0.7
C10-C7	1.50	1.51	-0.01	C4-C6-C5	117.6	119.3	-1.7
C7-C3	1.44	1.44	0.00	C6-C5-H5	120.6	122.2	-1.6
C3=C4	1.42	1.40	0.02	C3-C7-C10	126.2	123.3	2.9
C4-N9	1.38	1.36	0.02	C3-C7-C8	106.6	107.2	-0.6
N9-H9	1.01	1.01	0.00	C7-C8-N9	109.4	107.5	1.9
N9-C8	1.37	1.38	-0.01	C7-C8-H8	129.8	128.6	1.2
C8-H8	1.08	1.08	0.00	C4-C8-N9	109.9	111.3	-1.4
C8=C7	1.38	1.37	0.01	C8-N9-H9	124.8	126.6	-1.8
C3-C2	1.41	1.39	0.02	C7-C10-C13	111.2	114.0	-2.8
C2-H2	1.09	1.08	0.01	C7-C10-H11	110.9	108.9	2
C2=C1	1.39	1.39	0.00	C7-C10-H12	111.1	108.7	2.4
C1-O1	1.37	1.40	-0.03	C10-C13-N16	109.1	111.6	-2.5
O1-H1	0.97	0.96	0.01	C10-C13-H14	110.6	110.0	-0.6
C6-C1	1.42	1.39	0.03	C10-C13-H15	113.7	111.0	2.7
C6-H6	1.09	1.08	0.01	H17-C13-N16	112.7	111.6	1.1
C5=C6	1.39	1.41	-0.02	C13-N16-H19	109.2	107.8	1.4
C5-H5	1.09	1.08	0.01	C13-N16-H18	111.5	110.9	0.6
C5-C4	1.40	1.38	0.02				
Dihedrals/°				Impropers/°			
H1-O1-C1-C2	173.6	167.9	5.7	C2-O1-C1-C6	178.8	180.5	-1.3
O1-C1-C2-C3	180.1	179.6	0.5	C3-C1-C2-H2	172.7	175.1	-2.4
C1-C2-C3-C7	178.9	182.2	3.3	C7-C2-C3-C4	180.8	180.4	0.4
O1-C1-C6-C5	178.9	178.2	0.7	C5-C1-C6-H6	182.0	181.2	10.8
C2-C3-C7-C10	5.4	7.6	-2.2	C4-C6-C5-H5	179.0	180.3	-1.3
C3-C7-C8-N9	-0.2	-1.5	1.3	C10-C3-C7-C8	175.8	174.4	1.4
C3-C7-C10-C13	83.2	77.7	5.5	N9-C7-C8-H8	178.6	179.6	-1.0
C7-C10-C13-N16	-55.8	-60.8	5	C4-C8-N9-H9	177.0	174.5	2.5
C10-C13-N16-H17	165.1	160.1	5.0	C13-C7-C10-H11	-122.8	-123.7	0.9
				H11-C7-C10-H12	-118.0	-115.7	-2.3
				N16-C10-C13-H14	116.3	119.7	-3.4
				H14-C10-C13-H15	123.9	119.4	4.5
				H17-C13-N16-H18	121.3	122.4	-1.1
				H17-C13-N16-H19	-121.7	-119.8	-1.9

The differences in bond lengths and angles between the MP2 and CGenFF were less than 0.03Å and 2.9°, respectively; whereas the differences of dihedrals and impropers were

less than 5.7º and 10.8º, respectively (Table 2.18). This data shows that CGenFF structure was optimized and matches well with the QM optimized structure.

d) Methylphenidate

In the case of methylphenidate, the CGenFF geometry fits well with the QM geometry. The bond length and bond angle deviations were below 0.03Å and 2.82º, respectively. The dihedral differences were below 5.1º, except for C5-C7-C8-O9 and C7-C8-O9-C10, which were 10.1º and 7.26º, respectively. Improper deviations were below 4.49º, except for the O9-C7-C8-O8, which was 10.8º (Table 2.19).

Table 2.19. CGenFF equilibrium geometry of methylphenidate compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/°		
C1-H1	1.09	1.08	0.01	C2-C6-C1	120.5	120.0	0.46
C1-C2	1.40	1.40	0.00	C6-C1-H1	119.5	120.8	-1.3
C2-H2	1.09	1.08	0.01	C1-C6-C5	119.9	120.5	-0.62
C2=C3	1.40	1.40	0.00	C1-C6-H6	119.4	118.7	0.65
C3-H3	1.09	1.08	0.01	C6-C5-C7	122.4	122.7	-0.26
C3-C4	1.40	1.40	0.00	C6-C5-C4	119.4	118.8	0.51
C4-H4	1.09	1.08	0.01	C5-C4-C3	120.5	120.6	-0.12
C4=C5	1.40	1.41	-0.01	C5-C4-H4	120.0	120.3	-0.32
C5-C6	1.40	1.41	-0.01	C2-C4-C3	119.9	129.0	-0.09
C6-H6	1.09	1.08	-0.01	C4-C3-H3	119.8	120.2	-0.40
C6=C1	1.40	1.40	0.00	C3-C1-C2	119.8	119.9	-0.14
C5-C7	1.51	1.52	-0.01	C1-C2-H2	120.1	120.0	0.1
C7-H7	1.10	1.11	-0.01	C5-C7-C11	109.5	107.5	2.01
C7-C8	1.53	1.53	0.00	C5-C7-C8	113.9	113.3	0.59
C8=O8	1.21	1.21	0.00	C5-C7-H7	107.4	108.7	-1.34
C8-O9	1.38	1.35	0.03	C7-C8-O9	110.1	110.0	0.11
O9-C10	1.46	1.45	0.01	C7-C8-O8	127.0	126.4	0.62
C10-H101	1.09	1.11	-0.02	C8-O9-C10	113.9	111.2	2.74
C10-H102	1.09	1.11	-0.02	O9-C10-H101	105.0	107.7	-2.68
C10-H103	1.09	1.11	-0.02	O9-C10-H102	109.4	108.5	0.92
C7-C11	1.54	1.51	0.03	O9-C10-H103	109.8	108.6	1.25
C11-H11	1.10	1.12	-0.02	C7-C11-N12	110.1	109.3	0.76
C11-N12	1.52	1.51	0.01	C7-C11-C24	115.1	114.5	0.61
N12-H13	1.03	1.01	0.02	C7-C11-H11	107.6	110.4	-2.82
N12-H14	1.04	1.02	0.02	C11-N12-C15	114.2	116.2	-2.04
N12-C15	1.51	1.51	0.00	C11-N12-H13	109.0	108.8	0.24
C15-H16	1.09	1.10	-0.02	C11-N12-H14	106.1	103.5	2.63
C15-H17	1.09	1.10	-0.02	N12-C15-C18	109.7	110.8	-1.16
C15-C18	1.52	1.53	-0.01	N12-C15-H16	106.4	105.1	1.24
C18-H19	1.10	1.11	-0.01	N12-C15-H17	107.1	105.3	1.75
C18-H20	1.09	1.11	-0.02	C15-C18-C21	110.8	109.7	1.07
C18-C21	1.53	1.54	-0.01	C15-C18-H20	108.2	109.8	-1.62
C21-H22	1.09	1.11	-0.02	C15-C18-H19	109.6	110.2	-0.65
C21-H23	1.10	1.11	-0.01	C24-C18-C21	110.7	110.9	-0.25
C21-C24	1.53	1.54	-0.01	C18-C21-H22	110.0	109.7	0.31
C24-H25	1.09	1.11	-0.02	C18-C21-H23	109.8	109.8	0.0
C24-H26	1.10	1.11	-0.01	C21-C11-C24	111.9	112.9	-0.96
C24-C11	1.52	1.55	-0.03	C11-C24-H26	109.3	109.9	-0.63
	Dihedrals/°				Impropers/°		
H1-C1-C6-C5	-179.9	-179.0	-0.9	C2-C6-C1-H1	-179.9	-179.4	-0.5
C1-C6-C5-C7	-176.6	-177.2	0.6	C5-C1-C6-H6	176.8	178.8	-2.05
C6-C5-C4-C3	0.3	0.3	0	C7-C6-C5-C4	176.1	176.8	-0.67
C6-C5-C7-C11	83.9	88.6	-4.67	C3-C5-C4-H4	178.5	179.1	-0.6
C5-C7-C8-O9	174.7	164.6	10.1	C2-C4-C3-H3	179.6	179.5	0.1
C7-C8-O9-C10	176.4	169.1	7.26	C3-C1-C2-H2	-178.9	-179.3	0.4
C5-C7-C11-N12	161.4	162.7	1.31	C11-C5-C7-C8	-127.3	-122.8	-4.49
C7-C11-N12-C15	182.7	177.6	5.1	C8-C5-C7-H7	-116.0	-116.9	0.9
C11-N12-C15-C18	-57.0	-54.5	-2.5	O9-C7-C8-O8	177.7	166.8	10.8
N12-C15-C18-C21	55.7	55.9	-0.2	N12-C7-C11-C24	123.0	124.0	-1.06
C8-O9-C10-H101	181.2	179.5	1.73	H101-O9-C10-H102	119.3	119.4	-0.18
	Impropers/°			C24-C7-C11-H11	122.9	123.5	-0.64

Coordinate				Coordinate			
C21-C15-C18-H20	122.2	120.7	1.5	C15-C11-N12-H13	123.0	126.6	-3.61
C21-C15-C18-H19	-122.0	-120.8	-1.2	C18-N12-C15-H16	120.8	121.9	-1.1
C24-C18-C21-H22	-121.4	-121.1	-0.3	H16-N12-C15-H17	116.3	115.8	0.5
C15-C11-N12-H14	-123.8	-121.8	-2.0	H101-O9-C10-H103	-119.8	-119.7	-0.06
C21-C11-C24-H26	121.8	121.6	0.15	H22-C18-C21-H23	-117.2	-117.6	0.37
				H26-C11-C24-H25	116.1	117.4	-1.33

e) Bupropion

Table 2.20. CGenFF equilibrium geometry of bupropion compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/º			
C1-H1	1.09	1.08	0.01	C2-C1-C6	119.6	120.1	-0.5
C1=C2	1.40	1.40	0.00	C6-C1-H1	120.7	120.1	0.6
C2-CL2	1.73	1.74	-0.01	C1-C2-C3	120.8	120.0	0.8
C2-C3	1.39	1.40	-0.01	C1-C2-CL2	119.7	120.1	-0.4
C3-H3	1.09	1.08	0.01	C2-C3-C4	119.1	120.6	-1.5
C3=C4	1.41	1.40	0.01	C2-C3-H3	121.0	118.5	2.5
C4-C5	1.40	1.41	-0.01	C1-C6-C5	120.6	120.1	0.5
C5-H5	1.09	1.08	0.01	C1-C6-H6	119.4	119.5	-0.1
C5=C6	1.39	1.40	-0.01	C4-C5-C6	119.2	120.3	1.1
C6-H6	1.09	1.08	0.01	C6-C5-H5	118.9	117.4	1.5
C6=C1	1.39	1.40	-0.01	C5-C4-C3	120.7	118.8	1.9
C4-C7	1.47	1.50	-0.03	C3-C4-C7	116.9	118.5	-1.6
C7=O7	1.24	1.22	0.02	C4-C7-C8	119.9	121.8	-1.9
C7-C8	1.54	1.52	0.02	C4-C7-O7	123.7	121.0	2.7
C8-H8	1.09	1.11	-0.02	C7-C8-N13	104.6	107.2	-2.6
C8-C9	1.52	1.54	-0.02	C7-C8-C9	110.6	111.8	-1.2
C9-H10	1.09	1.11	-0.02	C7-C8-H8	112.4	111.3	1.1
C9-H11	1.09	1.11	-0.02	C8-C9-H10	113.3	111.4	1.9
C9-H12	1.09	1.11	-0.02	C8-C9-H11	109.6	110.8	-1.2
C8-N13	1.50	1.52	-0.02	C8-C9-H12	110.6	111.7	-1.1
N13-H14	1.04	1.02	0.02	C8-N13-C16	114.5	117.3	-2.8
N13-H15	1.03	1.01	0.02	C8-N13-H14	103.9	104.6	-0.7
N13-C16	1.50	1.51	-0.01	C8-N13-H15	110.6	109.0	1.6
C16-H17	1.09	1.11	-0.02	N13-C16-H17	108.0	106.6	1.4
C16-H18	1.09	1.11	-0.02	N13-C16-H18	108.8	106.8	2.0
C16-H19	1.09	1.11	-0.02	N13-C16-H19	108.6	106.6	2.0
Dihedrals/º				Impropers/º			
C6-C1-C2-C3	-0.1	0.06	-0.16	C2-C6-C1-H1	180.0	180	0.0
C1-C2-C3-C4	0.4	0.1	0.3	C3-C1-C2-CL2	-179.9	-179.9	0.0
C2-C1-C6-C5	-0.1	-0.04	-0.06	C4-C2-C3-H3	179.7	179.8	-0.1
C3-C4-C7-C8	-168.6	-174.5	5.9	C5-C1-C6-H6	179.7	179.9	-0.2
C4-C7-C8-N13	-164.3	-166.9	2.6	C4-C6-C5-H5	179.4	179.9	-0.5
C7-C8-C9-H10	177.8	178.8	-1.0	C5-C3-C4-C7	-179.2	-179.2	0.0
C7-C8-N13-C16	85.0	88.0	-3.0	C8-C4-C7-O7	176.2	178.9	-2.7
C8-N13-C16-H17	-57.2	-57.94	0.74	N13-C7-C8-C9	-118.6	-119.3	0.7
Impropers/º				C9-C7-C8-H8	-125.5	-124.6	-0.9
H17-N13-C16-H19	-119.3	-119.6	0.4	H10-C8-C9-H11	119.4	119.1	0.3
H17-N13-C16-H18	120.5	120.3	0.2	H10-C8-C9-H12	-121.8	-121.4	-0.4
H14-C8-N13-H15	-116.9	-112.9	-4	C16-C8-N13-H14	-117.6	-119.8	2.2

The optimized CGenFF geometry was in good correlation with the QM geometry. The bond length and angle deviations were below 0.03Å and 2.8º, respectively. All the dihedral and improper differences did not exceed 5.9º and 4º, respectively (Table 2.20).

f) Benztropine

The CGenFF geometry fits reasonably well with the QM geometry. Bond length and bond angle deviations were below 0.02Å and 2.6º, respectively; whereas, the dihedral differences were below 5.1º, except for C16-C19-O24-C25, O24-C25-C27-C28, and O24-C25-C38-C39, which were 9.8º, 9.1º, and 8.8º, respectively. All the improper deviations did not exceed 5.8º (Table 2.21).

Table 2.21. CGenFF equilibrium geometry of benztropine compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/º		
C1-H2	1.09	1.11	-0.02	H2-N5-C1	108.6	106.5	2.1
C1-H3	1.09	1.11	-0.02	N5-C1-H3	109.2	108.1	1.1
C1-H4	1.09	1.11	-0.02	N5-C1-H4	108.7	106.6	2.1
C1-N5	1.49	1.50	-0.01	C1-N5-C14	114.0	114.5	-0.5
N5-H5	1.03	1.04	-0.01	C1-N5-C6	114.1	114.0	0.1
N5-C6	1.52	1.50	0.02	C1-N5-H5	108.1	108.5	-0.4
C6-H7	1.09	1.11	-0.02	N5-C6-C21	107.2	109.4	-2.2
C6-C8	1.53	1.54	-0.01	N5-C6-C8	102.0	102	0
C8-H9	1.09	1.10	-0.01	N5-C6-H7	108.1	107.6	0.5
C8-H10	1.09	1.10	-0.01	C6-C8-C11	105.3	104.8	0.5
C8-C11	1.55	1.56	-0.01	C6-C8-H10	111.2	111.5	-0.3
C11-H12	1.09	1.10	-0.01	C11-C6-C8	105.4	104.8	0.6
C11-H13	1.09	1.10	-0.01	C6-C8-H9	109.5	111.8	-2.3
C11-C14	1.54	1.54	0.00	C14-C8-C11	105.4	104.8	0.6
C14-H15	1.10	1.11	-0.01	C8-C11-H13	111.2	110.5	0.7
C14-N5	1.52	1.50	0.02	C8-C11-H12	112.6	112.4	0.2
C14-C16	1.53	1.54	-0.01	C11-N5-C14	101.9	101.8	0.1
C16-H17	1.09	1.11	-0.02	N5-C14-C16	107.0	109.3	-2.3
C16-H18	1.10	1.11	-0.01	N5-C14-H15	108.2	107.8	0.4
C16-C19	1.53	1.54	-0.01	C14-C16-C19	111.6	112.9	-1.3
C19-H20	1.10	1.11	-0.01	C14-C16-H18	110.6	109.4	1.2
C19-C21	1.53	1.54	-0.01	C14-C16-H17	109.4	110.0	-0.6
C21-H22	1.09	1.11	-0.02	C21-C16-C19	110.9	111.6	-0.7
C21-H23	1.10	1.11	-0.01	C16-C19-O24	106.8	105.3	1.5
C21-C6	1.52	1.53	-0.01	C16-C19-H20	109.2	109.9	-0.7
C6-H7	1.09	1.10	-0.02	C19-C6-C21	111.6	112.9	-1.3
C19-O24	1.42	1.42	0	C6-C21-H22	109.0	109.2	-0.2
C25-O24	1.45	1.43	0.02	C6-C21-H23	110.7	109.7	1.0
C25-H26	1.10	1.11	-0.01	C19-O24-C25	114.0	114.0	0.0
C25-C38	1.51	1.53	-0.02	O24-C25-C27	107.7	110.3	-2.6
C38-C39	1.4	1.41	0.01	O24-C25-C38	110.5	109.3	1.2
C39-H40	1.09	1.08	0.01	O24-C25-H26	107.9	108.1	-0.2
C39-C41	1.39	1.40	-0.01	C25-C27-C28	118.3	118.9	-0.6
C41-H42	1.09	1.08	0.01	C25-C27-C36	122.0	122.6	-0.6
C41=C43	1.4	1.40	0	C27-C28-C30	120.3	120.8	-0.5
C43-H44	1.09	1.08	0.01	C27-C28-H29	119.8	119.8	0
C43-C45	1.40	1.40	0	C28-C30-C32	120.0	120.0	0
C45-H46	1.09	1.08	0.01	C28-C30-H31	119.8	120.0	-0.2
C45=C47	1.4	1.40	0	C30-C32-C34	119.8	120.0	-0.2
C47-H48	1.09	1.08	0.01	C30-C32-H33	120.1	120	0.1
C47-C38	1.4	1.41	-0.01	C36-C32-C34	120.4	120.0	0.4
C25-C27	1.51	1.53	-0.02	C32-C34-H35	120.0	119.8	0.2
C27=C28	1.4	1.41	-0.01	C34-C27-C36	119.9	120.7	-0.8
C28-H29	1.09	1.08	0.01	C27-C36-H37	119.7	120.0	-0.3
C28-C30	1.39	1.40	-0.01	C25-C38-C39	120.0	121.0	-1
C30-H31	1.09	1.08	0.01	C38-C39-C41	120.2	120.6	-0.4
C30-C32	1.4	1.40	0	C38-C39-H40	119.3	119.5	-0.2
C32-H33	1.09	1.08	0.01	C39-C41-C43	120.2	120	0.2
C32-C34	1.4	1.40	0	C39-C41-H42	119.8	120.0	-0.2
C34-H35	1.09	1.08	0.01	C41-C43-C45	120.0	120	0
C34-C36	1.4	1.40	0	C41-C43-H44	120.0	119.9	0.1

C36-H37	1.09	1.08	0.01	C47-C43-C45 119.9 120.0 0.1	
C36-C27	1.4	1.41	-0.01	C43-C45-H46 120.1 119.8 0.3	
				C45-C38-C47 120.4 120.5 -0.1	
				C38-C47-H48 119.7 120.1 -0.4	

Dihedrals/°				Impropers/°		
H2-C1-N5-C14	-178.6	-179.1	0.5	H2-N5-C1-H3	120.0	120.2	-0.2
C1-N5-C6-C21	163.7	164.8	-1.1	H2-N5-C1-H4	-119.8	-119.5	-0.3
N5-C6-C8-C11	-27.9	-27.8	-0.1	C14-C1-N5-C6	116.6	118.4	-1.8
C19-O24-C25-C27	-179.0	-175.5	-3.5	C14-C1-N5-H5	-121.7	-121.3	-0.4
C16-C19-O24-C25	138.4	128.6	9.8	C21-N5-C6-C8	119.1	117.1	2.0
O24-C25-C27-C28	163.2	172.3	-9.1	C8-N5-C6-H7	120.3	120.8	-0.5
C25-C27-C28-C30	177.7	180.7	-3.0	C11-C6-C8-H10	120.5	119.6	0.9
C27-C28-C30-C32	-0.4	0.2	-0.6	C11-C6-C8-H9	-121.5	-122.2	0.7
O24-C25-C38-C39	71.6	62.8	8.8	C14-C8-C11-H13	120.4	120.2	0.2
C25-C38-C39-C41	-173.6	-178.7	5.1	C14-C8-C11-H12	-119.9	-121.8	1.9
C38-C39-C41-C43	0.1	-0.06	0.16	C11-N5-C14-C16	119.4	117.2	2.2
C39-C41-C43-C45	-1.1	-0.2	-0.9	C11-N5-C14-H15	-120.2	-1207	0.5

Impropers/°						
C30-C27-C28-H29	179.8	180.8	-1	N5-C14-C16-C19	-62.2	-58.1	-4.1
C28-C30-C32-C34	0.3	-0.06	0.36	C19-C14-C16-H18	121.2	121.6	-0.4
C34-C30-C32-H33	179.5	180.1	-0.6	C19-C14-C16-H17	-122.0	-121.5	-0.5
C36-C32-C34-H35	-179.8	-179.9	0.1	C21-C16-C19-O24	120.8	115.0	5.8
C34-C27-C36-H37	181.5	179.7	1.8	C21-C16-C19-H20	-120.5	-123.5	3.0
C39-C25-C38-C47	-174.2	-179.1	4.8	C19-C6-C21-H22	122.2	121.3	0.9
C41-C38-C39-H40	-179.0	-179.3	0.3	H22-C6-C21-H23	117.0	116.7	0.3
C43-C39-C41-H42	-179.4	-179.1	-0.3	C27-O24-C25-C38	-123.4	-121.3	-2.1
C47-C43-C45-H46	-179.3	-179.1	-0.2	C38-O24-C25-H26	-118.8	-121.2	2.4
				C28-C25-C27-C36	177.0	181.1	-4.1
				C45-C41-C43-H44	-178.2	-178.9	0.7

g) Cocaine

The CGenFF geometry of cocaine was optimized by fitting it against the QM optimized geometry. The differences in bond lengths and bond angles were all below 0.03Å and 3º, respectively. The dihedral deviations did not exceed 4.5º, except for C10-C15-C19-O21, which was 18º. The improper deviations were all below 6.2º (Table 2.22), and these deviations indicate the quality of CGenFF structure.

Table 2.22. CGenFF equilibrium geometry of cocaine compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/°		
C7-H7	1.09	1.08	0.01	C9-C4-C2	122.1	119.7	2.4
C7=C6	1.40	1.40	0.00	C2-C4-C5	117.3	119.8	-2.5
C6-H6	1.09	1.08	0.01	C4-C5-C6	119.6	120.1	-0.4
C6-C5	1.39	1.40	-0.01	C4-C5-H5	119.2	120.6	-1.4
C5-H5	1.09	1.08	0.01	C5-C6-C7	120.1	120.1	0.0
C5=C4	1.40	1.41	-0.01	C5-C6-H6	119.8	120.1	-0.3
C4-C9	1.40	1.41	-0.01	C4-C9-C8	119.3	120.1	-0.8
C9-H9	1.09	1.08	0.01	C4-C9-H9	120.3	120.4	-0.1
C9=C8	1.39	1.40	-0.01	C8-C7-C6	120.1	120.1	0.0
C8-H8	1.09	1.08	0.01	C6-C7-H7	120.0	119.9	0.0
C8-C7	1.40	1.40	0.00	C7-C9-C8	120.3	120.1	0.2
C4-C2	1.48	1.49	-0.01	C9-C8-H8	119.7	120.0	-0.3
C2=O3	1.22	1.22	0.00	C4-C2-O1	112.3	110.4	2.3
C2-O1	1.38	1.36	0.02	C4-C2-O3	126.2	124.2	2.2
O1-C10	1.43	1.42	0.01	C2-O1-C10	113.7	113.2	0.5
C10-H10	1.09	1.12	-0.03	O1-C10-C15	105.1	106.6	-1.5
C10-C11	1.53	1.52	0.01	O1-C10-C11	110.9	109.4	1.5
C11-H111	1.09	1.12	-0.03	O1-C10-H10	108.4	111.2	-2.8
C11-H112	1.09	1.11	-0.02	C10-C15-C14	110.7	113.2	-2.5
C11-C12	1.52	1.52	0.00	C10-C15-C19	109.9	107.0	2.9
C12-H12	1.09	1.11	-0.01	C10-C15-H15	110.1	109.4	0.7
C12-N13	1.53	1.51	0.02	C15-C19-O21	110.4	107.5	2.9
N13-H13	1.03	1.01	0.02	C15-C19-O20	123.9	123.3	0.6
N13-C18	1.49	1.51	-0.02	C19-O21-C22	114.4	113.8	0.6
C18-H181	1.09	1.11	-0.02	O21-C22-H221	104.8	106.9	-2.1
C18-H182	1.09	1.11	-0.02	O21-C22-H222	109.9	108.0	1.9
C18-H183	1.09	1.11	-0.02	O21-C22-H223	109.6	107.8	1.8
N13-C14	1.52	1.52	0.00	O1-C10-C11	110.9	109.4	1.5
C14-H14	1.09	1.11	-0.02	C10-C11-H112	109.7	109.3	0.4
C14-C15	1.53	1.54	-0.01	C10-C11-H111	108.9	108.6	0.3
C15-H15	1.09	1.11	-0.02	C10-C11-C12	111.1	114.0	-2.9
C15-C10	1.52	1.51	0.01	C11-C12-N13	109.0	112.0	-3.0
C14-C16	1.53	1.53	0.00	C11-C12-C17	113.2	111.0	2.1
C16-H161	1.09	1.10	-0.01	C11-C12-H12	111.7	113.7	2.6
C16-H162	1.09	1.10	-0.01	C14-C12-N13	102.0	103.6	-1.6
C16-C17	1.56	1.59	-0.03	C12-N13-C18	117.6	115.7	1.8
C17-H171	1.09	1.10	-0.01	C12-N13-H13	106.2	106.5	-0.3
C17-H172	1.09	1.10	0.01	N13-C15-C14	108.7	111.5	-2.9
C17-C12	1.54	1.53	0.01	C15-C14-C16	113.4	112.9	0.5
C15-C19	1.52	1.51	0.01	C15-C14-H14	110.9	113.2	0.3
C19=O20	1.22	1.22	0.00	C17-C14-C16	105.3	105.1	0.2
C19-O21	1.34	1.34	0.00	C14-C16-H161	110.4	112.4	-2.0
O21-C22	1.46	1.49	-0.03	C14-C16-H162	110.2	112.0	-1.8
C22-H221	1.09	1.11	-0.02	C16-C12-C17	105.5	105.0	0.5
C22-H222	1.09	1.11	-0.02	C12-C17-H171	110.6	111.3	-0.7
C22-H223	1.09	1.11	-0.02	C12-C17-H172	110.4	111.9	-1.5
				N13-C18-H181	108.3	106.6	1.7
				N13-C18-H183	110.2	107.8	2.4
				N13-C18-H182	108.2	106.5	1.7
	Dihedrals/°				Impropers/°		
C2-C4-C5-C6	180.2	180.1	-0.1	C9-C2-C4-C5	179.7	179.8	0.1

C4-C5-C6-C7	-0.1	0.04	0.14	C6-C4-C5-H5	179.7	179.8	0.1
C2-C4-C9-C8	179.8	179.9	0.1	C7-C5-C6-H6	180.0	179.8	-0.2
C5-C4-C2-O1	-177.2	-177.9	-0.7	C8-C4-C9-H9	180.2	-179.9	-0.1
C4-C2-O1-C10	-175.9	-177.4	-1.5	C8-C6-C7-H7	179.9	179.8	-0.1
C2-O1-C10-C15	157.7	161.0	3.3	C7-C9-C8-H8	-179.9	-179.9	0
O1-C10-C15-C14	168.3	167.3	-1	O1-C4-C2-O3	179.6	179.4	-0.2
C10-C15-C19-O21	147.9	165.9	18	C15-O1-C10-C11	121.7	119.3	-2.4
C15-C19-O21-C22	179.2	174.7	-4.5	C11-O1-C10-H10	121.2	121.1	-0.1
C19-O21-C22-H221	179.4	178.8	-0.6	C14-C10-C15-C19	127.1	124.5	-2.6
O1-C10-C11-C12	-164.7	-162.6	2.1	C14-C10-C15-H15	-125.8	-122.4	3.4
C10-C11-C12-N13	60.4	57.5	-2.9	O21-C15-C19-O20	178.0	171.8	-6.2
Impropers/⁰				H221-O21-C22-H222	119.7	119.5	-0.2
H161-C14-C16-H162	-118.4	-118.5	-0.1	H221-O21-C22-H223	-119.5	-119.2	0.3
C16-C12-C17-H171	120.9	119.4	-1.5	C12-C10-C11-H112	124.2	123.8	-0.4
C16-C12-C17-H172	-121.8	-121.8	0	H112-C10-C11-H111	116.4	115.1	-1.3
C12-N13-C18-H181	179.4	180.5	1.1	N13-C11-C12-C17	-111.0	-109.1	1.9
H181-N13-C18-H183	120.4	120.0	-0.4	N13-C11-C12-H12	119.1	121.2	2.1
H181-N13-C18-H182	-119.4	-119.6	-0.2	C14-C12-N13-C18	130.4	128.3	-2.1
N13-C15-C14-H14	-118.9	-119.4	-0.5	C14-C12-N13-H13	-110.6	-111.5	-0.9
C17-C14-C16-H161	-119.6	-119.1	0.5	N13-C15-C14-C16	111.3	110.4	-0.9

h) Clomipramine

The geometry optimization of tricyclic antidepressants is complicated because of the three rings in the structure. Despite multiple iterations, only desipramine fits fairly well with the QM structure. In the case of clomipramine and imipramine, a few dihedrals/torsions in the tricyclic region still need some optimization.

In the case of clomipramine, all the bond length and bond angle deviations were less than 0.03Å and 4.7⁰, respectively. All the impropers fit fairly well with the QM data (difference <6.1⁰), except for the CD2-NZ-CG3-CD3, whose difference was 17.4⁰ (Table 2.23). The dihedrals in the tail region match well with the QM angles, but several dihedrals in the tricyclic rings deviated significantly.

Table 2.23. CGenFF equilibrium geometry of clomipramine compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/º			
CD1=CG2	1.41	1.41	0	CA-NB-CG	111.4	109.3	2.0
CG2-CD4	1.40	1.40	0	CG-NB-CB	112.6	111.4	1.2
CD4-HD4	1.09	1.08	0.01	CG-NB-HB	106.5	107.0	-0.6
CD4=CE6	1.39	1.40	-0.01	NB-CA-HA11	109.0	107.2	2.2
CE6-CL	1.74	1.74	0.00	NB-CA-HA12	108.8	107.2	1.6
CE6-CZ2	1.39	1.40	-0.01	NB-CA-HA13	108.1	106.9	1.1
CZ2-HZ2	1.09	1.08	0.01	NB-CB-HB21	108.6	107.0	1.6
CZ2=CE5	1.40	1.40	0	NB-CB-HB22	108.0	107.3	0.9
CE5-HE5	1.09	1.08	0.01	NB-CB-HB23	109.3	107.0	2.2
CE5-CD1	1.40	1.40	0	NB-CG-CD	112.5	112.3	0.2
CD1=CE1	1.50	1.51	-0.01	NB-CG-HG1	105.8	106.0	0.2
CE1-HE11	1.10	1.11	-0.01	NB-CG-HG2	106.2	105.5	0.6
CE1-HE12	1.10	1.11	-0.01	CG-CD-CE	110.9	110.3	0.6
CE1-CE2	1.53	1.55	-0.02	CG-CD-HD1	110.1	110.7	-0.6
CE2=HE21	1.10	1.11	-0.01	CG-CD-HD2	111.8	111.5	0.2
CE2-HE22	1.10	1.11	-0.01	CD-CE-NZ	110.2	112.5	-2.3
CE2-CD2	1.52	1.52	0	CD-CE-HE1	110.0	108.5	1.4
CD2=CG3	1.41	1.42	-0.01	CD-CE-HE2	109.1	108.7	0.4
CG3-NZ	1.44	1.43	0.01	CE-NZ-CG2	114.0	117.3	-3.3
NZ-CG2	1.43	1.43	0.00	CE-NZ-CG3	116.1	114.2	2.2
CD2-CE3	1.41	1.41	0.00	NZ-CG2-CD1	121.7	117.0	4.7
CE3-HE3	1.09	1.08	0.01	NZ-CG2-CD4	117.5	118.9	-1.4
CE3=CZ1	1.39	1.40	-0.01	CG2-CD1-CE1	119.5	119.4	0.1
CZ1-HZ1	1.09	1.08	0.01	CG2-CD1-CE5	118.3	120.1	-1.8
CZ1-CE4	1.40	1.40	0.00	CD1-CE1-CE2	110.6	111.9	-1.3
CE4-HE4	1.09	1.08	0.01	CD1-CE1-HE12	111.8	108.2	3.6
CE4=CD3	1.39	1.40	-0.01	CD1-CE1-HE11	109.1	109.3	-0.2
CD3-HD3	1.09	1.08	0.01	CE1-CE2-CD2	116.6	115.3	1.3
CD3-CG3	1.41	1.41	0.01	CE1-CE2-HE22	108.1	107.1	0.9
NZ-CE	1.46	1.49	-0.03	CE1-CE2-HE21	108.8	110.7	-1.9
CE-HE1	1.09	1.12	-0.03	CG2-CD4-CE6	119.6	120.9	-1.4
CE-HE2	1.10	1.11	-0.01	CG2-CD4-HD4	119.4	119.2	0.2
CE-CD	1.53	1.54	-0.01	CD4-CE6-CL	119.3	119.4	-0.1
CD-HD1	1.10	1.11	-0.01	CD4-CE6-CZ2	120.7	119.7	0.6
CD-HD2	1.09	1.11	-0.02	CE5-CE6-CZ2	119.0	119.9	-0.9
CD-CG	1.52	1.53	-0.01	CE6-CZ2-HZ2	120.1	120.2	-0.1
CG-HG1	1.09	1.10	-0.02	CZ2-CD1-CE5	121.6	120.1	1.5
CG-HG2	1.09	1.10	-0.01	CD1-CE5-HE5	119.3	119.9	-0.6
CG-NB	1.51	1.51	0	CD2-NZ-CG3	125.5	122.3	3.2
NB-HB	1.03	1.05	-0.02	NZ-CG3-CD3	114.6	117.1	-2.5
NB-CA	1.50	1.51	-0.01	CG3-CD3-CE4	121.6	121.3	0.3
CA-HA11	1.09	1.11	-0.02	CG3-CD3-HD3	117.9	119.5	-1.6
CA-HA12	1.09	1.11	-0.02	CZ1-CD3-CE4	119.1	119.7	-0.6
CA-HA13	1.09	1.11	-0.02	CD3-CE4-HE4	120.2	120.4	-0.2
NB-CB	1.50	1.51	-0.01	CE4-CZ1-CE3	119.4	120.0	-0.6
CB-HB21	1.09	1.11	-0.02	CE4-CZ1-HZ1	120.5	120.2	0.3
CB-HB22	1.09	1.11	-0.02	CD2-CZ1-CE3	122.6	120.8	0.6
CB-HB23	1.09	1.11	-0.02	CZ1-CE3-HE3	119.1	119.4	-0.3
Dihedrals/º				Impropers/º			
CG-NB-CA-HA11	-176.7	-178.5	-1.8	CA-CG-NB-CB	125.5	121.4	-4.1
CG-NB-CB-HB21	-178	-60.3	117.7	CA-CG-NB-HB	-117.2	-118.8	-1.6

HB-NB-CG-CD	54.8	46.6	-8.2	HA11-NB-CA-HA12	120.3	120.0	-0.3
NB-CG-CD-CE	-167.1	-168.3	-1.2	HA11-NB-CA-HA13	-119.8	-120.3	-0.5
CG-CD-CE-NZ	-182	-62.3	119.7	HB21-NB-CB-HB22	119.6	120.1	0.5
CD-CE-NZ-CG2	-71.3	109.1	180.4	HB21-NB-CB-HB23	-120.1	-120.0	0.1
CG2-CE-NZ-CG3	215.2	144.2	-71	CD-NB-CG-HG1	-121.2	-121.3	-0.1
CE-NZ-CG2-CD1	-77.9	59.7	137.6	HG1-NB-CG-HG2	-115.1	-115.9	-0.8
NZ-CG2-CD1-CE1	-9.3	33.4	42.7	CE-CG-CD-HD1	119.7	118.4	-1.3
CG2-CD1-CE1-CE2	-64.6	44.7	109.3	HD1-CG-CD-HD2	119.8	120.9	1.1
CD1-CE1-CE2-CD2	71.3	-28.2	-99.5	NZ-CD-CE-HE1	118.4	123.0	4.6
NZ-CG2-CD4-CE6	183.3	144.9	-38.4	HE1-CD-CE-HE2	-117.6	-114.1	3.5
CG2-CD4-CE6-CL	179.7	190.5	10.8	CD1-NZ-CG2-CD4	176.3	206.3	30
NZ-CG3-CD3-CE4	178.7	198.2	19.5	CE1-CG2-CD1-CE5	-173.6	-179.7	-6.1
CG3-CD3-CE4-CZ1	0.9	-1.2	-2.1	CE2-CD1-CE1-HE12	122.5	118.9	-3.6
CD3-CE4-CZ1-CE3	-176.7	-178.5	-1.8	CE2-CD1-CE1-HE11	-119.9	-123.7	-3.8
Impropers/°				CD2-CE1-CE2-HE22	-123.4	-120.4	-3.0
CE4-CG3-CD3-HD3	177.5	179.8	2.3	CD2-CE1-CE2-HE21	121.8	122.9	1.1
CZ1-CD3-CE4-HE4	178.8	180.6	1.8	CE6-CG2-CD4-HD4	183.4	180.0	-3.4
CE3-CE4-CZ1-HZ1	179.3	181.5	2.2	CL-CD4-CE6-CZ2	178.9	173.3	-5.6
CD2-CZ1-CE3-HE3	-178.2	-178.9	-0.7	CE5-CE6-CZ2-HZ2	182.1	176.7	-5.4
CD2-NZ-CG3-CD3	181.3	163.9	-17.4	CZ2-CD1-CE5-HE5	-179.0	-179.5	-0.5

i) Imipramine

The CGenFF geometry fits fairly well with the QM structure, but some angles and dihedrals in the rings still deviated significantly. All the bond length deviations were less than 0.03Å, whereas most of the angle deviations were less than 4°, except five angles in the rings (NZ-CG2-CD4, CG2-CD1-CE1, CD1-CE1-CE2, CG3-CE2-CD2, and CE2-CD2-CE3), which varied between 4-9° (Table 2.24). All the dihedrals and impropers match well with the QM torsions, except a few dihedrals (CE-NZ-CG2-CD1, NZ-CG2-CD1-CE1, CG2-CD1-CE1-CE2, and NZ-CG2-CD4-CE6) in the rings.

Table 2.24. CGenFF equilibrium geometry of imipramine compared to MP2 level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/°		
CD1=CG2	1.41	1.44	-0.03	CA-CG-NB	111.4	110.97	0.4
CG2-CD4	1.40	1.41	-0.01	CG-NB-CB	112.5	111.96	0.5
CD4-HD4	1.09	1.08	0.01	CG-NB-HB1	106.4	106.84	-0.4
CD4=CE6	1.40	1.40	0.00	NB-CA-HA11	109.0	106.85	3.8
CE6-HE6	1.09	1.08	0.01	NB-CA-HA12	108.8	106.88	1.9
CE6-CZ2	1.40	1.39	0.01	NB-CA-HA13	108.1	106.94	1.2
CZ2-HZ2	1.09	1.08	0.01	NB-CB-HB21	108.6	106.79	1.8
CZ2=CE5	1.40	1.39	0.01	NB-CB-HB22	108	107.04	0.9
CE5-HE5	1.09	1.08	0.01	NB-CB-HB23	109.3	107.26	2.0
CE5-CD1	1.40	1.40	0.00	NB-CG-CD	112.5	111.09	1.4
CD1-CE1	1.50	1.53	-0.03	NB-CG-HG1	105.8	105.77	0.1
CE1-HE11	1.10	1.11	-0.01	NB-CG-HG2	106.1	106.18	-0.1
CE1-HE12	1.10	1.11	-0.01	CG-CD-CE	110.6	110.75	-0.1
CE1-CE2	1.53	1.55	-0.02	CG-CD-HD1	111.9	111.27	0.6
CE2-HE21	1.10	1.11	-0.01	CG-CD-HD2	110.3	110.4	-0.1
CE2-HE22	1.10	1.11	-0.01	CD-CE-NZ	110.1	108.33	1.8
CE2-CD2	1.52	1.51	0.01	CD-CE-HE1	109.0	106.65	2.3
CD2=CG3	1.41	1.42	-0.01	CD-CE-HE2	109.9	112.57	-2.7
CG3-NZ	1.44	1.43	0.01	CE-NZ-CG2	114.1	116.28	-2.2
NZ-CG2	1.43	1.44	-0.01	CE-NZ-CG3	115.8	113.65	2.1
CD2-CE3	1.41	1.40	0.01	NZ-CG2-CD1	121.4	118.75	2.6
CE3-HE3	1.09	1.08	0.01	NZ-CG2-CD4	118.0	122.57	-4.6
CE3=CZ1	1.39	1.40	-0.01	CG2-CD1-CE1	119.5	123.74	-4.2
CZ1-HZ1	1.09	1.08	0.01	CG2-CD1-CE5	118.5	120.23	-1.7
CZ1-CE4	1.40	1.40	0.00	CD1-CE1-CE2	110.4	119.20	-8.8
CE4-HE4	1.09	1.08	0.01	CD1-CE1-HE12	112.1	108.46	3.6
CE4=CD3	1.39	1.40	-0.01	CD1-CE1-HE11	111.5	109.89	1.6
CD3-HD3	1.09	1.08	0.01	CE1-CE2-CD2	116.7	113.04	3.7
CD3-CG3	1.41	1.41	0.00	CE1-CE2-HE22	108.8	108.91	-0.1
NZ-CE	1.46	1.43	0.03	CE1-CE2-HE21	108.0	107.41	0.6
CE-HE1	1.10	1.11	-0.01	CG2-CD4-CE6	120.3	121.61	-1.3
CE-HE2	1.10	1.12	-0.02	CG2-CD4-HD4	118.4	120.78	-2.4
CE-CD	1.53	1.55	-0.02	CG3-CE2-CD2	126.1	117.27	8.8
CD-HD1	1.09	1.11	-0.02	CE2-CD2-CE3	116.5	121.76	-5.3
CD-HD2	1.10	1.11	-0.01	CD2-CE3-HE3	118.2	120.05	-1.8
CD-CG	1.52	1.53	-0.01	CE3-CZ1-HZ1	120	119.85	0.1
CG-HG1	1.09	1.10	-0.01	CZ1-CE4-HE4	120.6	119.85	0.7
CG-HG2	1.09	1.10	-0.01	CD4-CE6-HE6	120.1	120.32	-0.2
NB-CA	1.50	1.51	-0.01	CD4-CE6-CZ2	119.7	119.83	-0.1
CA-HA11	1.09	1.11	-0.02	CE5-CE6-CZ2	120	119.82	0.2
CA-HA12	1.09	1.11	-0.02	CE6-CZ2-HZ2	120.2	120.23	-0.1
CA-HA13	1.09	1.11	-0.02	CZ2-CD1-CE5	121.1	120.81	0.3
NB-CB	1.50	1.51	-0.01	CD1-CE5-HE5	119.2	119.55	-0.3
CB-HB21	1.09	1.11	-0.02	CG3-CD3-CE4	121.6	120.58	1.0
CB-HB22	1.09	1.11	-0.02	CZ1-CD3-CE4	119.2	119.92	-0.7
CB-HB23	1.09	1.11	-0.02	CE4-CZ1-CE3	119.4	120.07	-0.7
CG-NB	1.51	1.51	0.00	CD2-CZ1-CE3	122.6	119.83	2.8
NB-HB1	1.03	1.04	-0.01				
	Dihedrals/°				Impropers/°		
CG-NB-CA-HA11	-176.9	-178.0	1.1	CA-CG-NB-CB	125.6	123.9	1.7
CG-NB-CB-HB21	180.9	178.6	2.3	CA-CG-NB-HB1	-117.2	-118.1	0.9

HB1-NB-CG-CD	54.4	41.9	12.5	HA11-NB-CA-HA12	120.3	119.5	0.8
NB-CG-CD-CE	-167.3	-178.6	11.3	HA11-NB-CA-HA13	-119.9	-120.3	0.4
CG-CD-CE-NZ	177.6	190.4	12.8	HB21-NB-CB-HB22	119.6	120.2	-0.6
CD-CE-NZ-CG2	-70.7	-72.7	2.0	HB21-NB-CB-HB23	-120.2	-119.4	-0.8
CE-NZ-CG2-CD1	283.5	149.7	133.8	CD-NB-CG-HG1	-121.2	-120.2	-1
NZ-CG2-CD1-CE1	-9.7	27.5	-37.2	HG1-NB-CG-HG2	-115.1	-116.2	1.1
CG2-CD1-CE1-CE2	-64.5	-22.2	-42.3	CE-CG-CD-HD1	-120.4	-119.1	-1.3
CD1-CE1-CE2-CD2	71.3	62.6	8.7	HD1-CG-CD-HD2	-120.1	-119.6	-0.5
NZ-CG2-CD4-CE6	182.7	161.3	21.4	NZ-CD-CE-HE1	-124.1	-126.0	1.9
CG2-CD4-CE6-CZ2	1.2	2.1	-0.9	HE1-CD-CE-HE2	-117.5	-116.7	-0.8
CE3-CZ1-CE4-CD3	-0.2	-0.6	0.4	CD1-NZ-CG2-CD4	176.8	193.2	16.4
CE2-CD2-CE3-CZ1	183.6	187.4	-3.8	CE1-CG2-CD1-CE5	187.0	170.6	9.4
CD2-CE3-CZ1-CE4	-1.3	0.7	-2	CE2-CD1-CE1-HE12	122.5	122.4	0.1
Impropers/°				CE2-CD1-CE1-HE11	-119.7	-122.0	2.3
CG3-CE4-CD3-HD3	182.7	179.1	3.6	CD2-CE1-CE2-HE22	121.8	121.8	0
CD3-CZ1-CE4-HE4	180.1	179.4	0.7	CD2-CE1-CE2-HE21	236.6	236.6	0
CE4-CE3-CZ1-HZ1	180.6	179.0	1.6	CE6-CG2-CD4-HD4	183.4	178.5	4.9
CG2-CE-NZ-CG3	-144.1	-133.3	-10.8	CZ2-CD1-CE5-HE5	180.4	182.5	-2.1
CG3-CE2-CD2-CE3	181.7	176.8	4.9	CE5-CE6-CZ2-HZ2	182.2	177.0	5.2
CZ1-CD2-CE3-HE3	178.3	179.1	-0.8	CZ2-CD4-CE6-HE6	180.0	178.3	1.7

j) **Desipramine**

Desipramine CGenFF geometry fits better with the QM geometry compared to clomipramine and imipramine. All the bond length and angle deviations were less than 0.03Å and 3.6º, respectively. Most of the dihedral and improper deviations were less than 9.4º, except for the CE-NZ-CG3-CD2 and the NZ-CG2-CD4-CE6 in the tricyclic region, which were 19.1º and 24.1º (Table 2.25), respectively.

Table 2.25. CGenFF equilibrium geometry of desipramine compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/º		
CD1=CG2	1.41	1.41	0	CA-NB-CG	115.1	116.15	1.05
CG2-CD4	1.4	1.4	0	CG-NB-HB1	109.0	108.50	-0.5
CD4-HD4	1.09	1.08	-0.01	CG-NB-HB2	108.4	107.54	-0.86
CD4=CE6	1.39	1.4	0.01	NB-CA-HA11	108.6	106.71	-1.89
CE6-HE6	1.09	1.08	-0.01	NB-CA-HA12	108.9	106.87	-2.03
CE6-CZ2	1.4	1.39	-0.01	NB-CA-HA13	108.1	106.73	-1.37
CZ2-HZ2	1.09	1.08	-0.01	NB-CG-CD	111.5	112.12	0.62
CZ2=CE5	1.39	1.4	0.01	NB-CG-HG1	106.2	105.19	-1.01
CE5-HE5	1.09	1.08	-0.01	NB-CG-HG2	106.0	105.78	-0.22
CE5-CD1	1.41	1.4	-0.01	CG-CD-CE	110.7	109.15	-1.55
CD1-CE1	1.51	1.52	0.01	CG-CD-HD1	111.0	110.74	-0.26
CE1-HE11	1.1	1.11	0.01	CG-CD-HD2	110.7	110.54	-0.16
CE1-HE12	1.1	1.11	0.01	CD-CE-NZ	109.9	106.7	-3.2
CE1-CE2	1.53	1.55	0.02	CD-CE-HE1	110.0	110.1	0.1
CE2-HE21	1.1	1.11	0.01	CD-CE-HE2	109.1	108.79	-0.31
CE2-HE22	1.1	1.11	0.01	CE-NZ-CG2	115.7	115.9	0.2
CE2-CD2	1.5	1.51	0.01	CE-NZ-CG3	114.2	115.7	1.5
CD2=CG3	1.41	1.4	-0.01	NZ-CG3-CD2	121.5	120.6	-0.9
CG3-NZ	1.44	1.44	0	NZ-CG2-CD1	125.5	121.9	-3.6
NZ-CG2	1.43	1.43	0	CG3-CD2-CE2	119.5	121.9	2.4
CD2-CE3	1.4	1.4	0	CG3-CD2-CE3	118.4	119.3	0.9
CE3-HE3	1.09	1.08	-0.01	CD2-CE2-CE1	110.4	109.33	-1.07
CE3=CZ1	1.4	1.4	0	CE2-CE1-HE12	108.8	106.67	-2.13
CZ1-HZ1	1.09	1.08	-0.01	CE2-CE1-HE11	108.0	109.83	1.83
CZ1-CE4	1.4	1.4	0	CD2-CE2-HE22	112.1	111.91	-0.19
CE4-HE4	1.09	1.08	-0.01	CD2-CE2-HE21	109.2	110.34	1.14
CE4=CD3	1.4	1.4	0	CE2-CE1-CD1	116.8	114.27	-2.53
CD3-HD3	1.09	1.08	-0.01	CD1-NZ-CG2	122.1	121.35	-0.75
CD3-CG3	1.4	1.39	-0.01	NZ-CG2-CD4	114.6	115.9	1.3
NZ-CE	1.46	1.46	0.00	CG2-CD4-CE6	121.6	121.4	-0.2
CE-HE1	1.09	1.12	0.03	CG2-CD4-HD4	117.8	119.7	1.9
CE-HE2	1.1	1.12	0.02	CD4-CE6-CZ2	119.1	119.54	0.44
CE-CD	1.53	1.54	0.01	CD4-CE6-HE6	120.2	120.45	0.25
CD-HD1	1.09	1.11	0.02	CE6-CZ2-CE5	119.4	119.75	0.35
CD-HD2	1.1	1.11	0.01	CE6-CZ2-HZ2	120.6	120.26	-0.34
CD-CG	1.52	1.53	0.01	CD1-CZ2-CE5	122.6	120.95	-1.65
CG-HG1	1.09	1.1	0.01	CZ2-CE5-HE5	119.1	119.27	0.17
CG-HG2	1.09	1.1	0.01	CG3-CD3-HD3	118.4	118.98	0.58
NB-CA	1.5	1.51	0.01	CD3-CE4-CZ1	119.6	119.67	0.07
CA-HA11	1.09	1.11	0.02	CG3-CD3-CE4	120.3	120.5	0.2
CA-HA12	1.09	1.11	0.02	CD3-CE4-HE4	120.1	120.32	0.22
CA-HA13	1.09	1.11	0.02	CE3-CE4-CZ1	120.0	119.9	-0.1
CG-NB	1.52	1.51	-0.01	CE4-CZ1-HZ1	120.2	120.15	-0.05
NB-HB1	1.03	1.01	-0.02	CZ1-CD2-CE3	121.2	120.40	-0.8
NB-HB2	1.03	1.01	-0.02	CD2-CE3-HE3	119.2	119.95	0.75
	Dihedrals/º				Impropers/º		
CG-NB-CA-HA11	176.1	176	-0.1	CA-CG-NB-HB1	122.6	122.9	0.3
HB1-NB-CG-CD	189.1	188.9	-0.2	HB1-CG-NB-HB2	114.8	113.6	-1.2
NB-CG-CD-CE	177.3	177	-0.3	HA11-NB-CA-HA12	120.3	120.1	-0.2
CG-CD-CE-NZ	174.6	174	-0.6	HA11-NB-CA-HA13	-119.8	-119.9	-0.1
CD-CE-NZ-CG3	-70.7	-82.4	-11.7	CD-NB-CG-HG1	237.3	238.6	1.3

CE-NZ-CG3-CD2	-76	-56.9	19.1	HG1-NB-CG-HG2	-115.5	-115.4	0.1
NZ-CG3-CD2-CE2	-9.8	-19.2	-9.4	CE-CG-CD-HD1	119.7	120.1	0.4
CG3-CD2-CE2-CE1	-64.3	-55.8	8.5	HD1-CG-CD-HD2	120	119.9	-0.1
CD2-CE2-CE1-CD1	71.2	74	2.8	NZ-CD-CE-HE1	118.4	122.3	3.9
NZ-CG2-CD4-CE6	178.7	154.6	-24.1	HE1-CD-CE-HE2	117.6	116.2	-1.4
CG2-CD4-CE6-CZ2	0.7	2	1.3	CD2-NZ-CG3-CD3	176.9	165.8	-11.1
NZ-CG3-CD3-CE4	-177.3	-164.3	13	CE2-CG3-CD2-CE3	187	182.8	-4.2
CG3-CD3-CE4-CZ1	1.1	-0.3	-1.4	HE11-CE2-CE1-HE12	-114.7	-114.6	0.1
CD-CG-NB-CA	66	66	0	CZ1-CD3-CE4-HE4	180	180.6	0.6
CD4-CE6-CZ2-CE5	-0.1	1.8	1.9	CE4-CG3-CD3-HD3	183.4	180.3	3.1
CG2-NZ-CE-CD	145	145	0	CD1-CE2-CE1-HE11	-123.4	-123.4	0
Impropers/º				CE1-CD2-CE2-HE21	-119.7	-118.1	-1.6
CZ1-CD2-CE3-HE3	180.2	178.6	1.6	CD1-NZ-CG2-CD4	-178.5	-159.5	19
CD1-CZ2-CE5-HE5	181.7	177.0	4.7	CE6-CG2-CD4-HD4	177.4	181.1	3.7
CE3-CE4-CZ1-HZ1	182.1	181.2	0.9	CE5-CE6-CZ2-HZ2	179.3	176.7	-2.6
CG2-CE-NZ-CG3	143.7	132.6	8.0	CZ2-CD4-CE6-HE6	178.9	178.5	-0.4

k) Fluoxetine (R)

The optimized CGenFF geometry was in excellent correlation with the QM optimized geometry. The distance deviations were all less than 0.04Å, and all the angles deviated less than 3.1º; whereas the dihedral and improper deviations were all less than 7.2º and 5.1º, respectively (Table 2.26).

Table 2.26. CGenFF equilibrium geometry of fluoxetine (R) compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/º		
C1-F12	1.35	1.34	-0.01	C5-C1-F12	111.6	111.4	-0.2
C1-F13	1.35	1.34	-0.01	F12-C1-F13	107.4	106.8	-0.6
C1-F14	1.35	1.34	-0.01	F12-C1-F14	107.3	106.9	-0.4
C1-C5	1.50	1.54	0.04	C1-C5-C6	120.0	120.7	0.7
C5=C6	1.39	1.41	0.02	C1-C5-C13	119.7	120.8	1.1
C6-H7	1.09	1.08	-0.01	C5-C6-C8	120.5	120.8	0.3
C6-C8	1.40	1.40	0	C5-C6-H7	120.0	120.7	0.7
C8-H9	1.08	1.08	0	C6-C8-H9	119.6	117.6	-2
C8=C10	1.40	1.41	0.01	C8-C10-O15	125.6	124.1	-1.5
C10-C11	1.40	1.40	0	C8-C10-C11	120.4	118.6	-1.8
C11-H12	1.09	1.08	-0.01	C13-C11-C10	120.1	121.2	1.1
C11-C13	1.39	1.40	0.01	C10-C11-H12	118.7	118.9	0.2
C13-H14	1.09	1.08	-0.01	C11-C13-C5	119.6	120.5	0.9
C13-C5	1.40	1.41	0.01	C5-C13-H14	120.3	120.8	0.5
C10-O15	1.39	1.42	0.03	C10-O15-C16	119.3	120.1	0.8
O15-C16	1.44	1.43	-0.01	O15-C16-C18	113.0	115.7	2.7
C16-H17	1.10	1.12	0.02	O15-C16-C29	109.7	109.4	-0.3
C16-C18	1.51	1.52	0.01	O15-C16-H17	101.6	104.7	3.1
C18-C19	1.40	1.41	0.01	C16-C29-C32	109.1	108.1	-1
C19-H20	1.09	1.08	-0.01	C16-C29-H30	110.4	109.3	-1.1
C19-C21	1.40	1.40	0	C16-C29-H31	108.7	108.7	0
C21-H22	1.09	1.08	-0.01	C29-C32-N35	109.8	109.0	-0.8
C21=C23	1.40	1.40	0	C29-C32-H33	112.1	112.3	0.2
C23-H24	1.09	1.08	-0.01	C29-C32-H34	112.6	111.2	-1.4
C23-C25	1.40	1.40	0	C32-N35-H36	111.9	110.1	-1.8
C25-H26	1.09	1.08	-0.01	C32-N35-H37	111.6	109.1	-2.5
C25-C27	1.39	1.40	0.01	C32-N35-H38	110.7	108.0	-2.7
C27-H28	1.09	1.08	-0.01	C16-C18-C27	118.0	118.0	0
C27-C18	1.40	1.41	0.01	C16-C18-C19	122.4	123.3	0.9
C16-C29	1.54	1.55	0.01	C18-C19-C21	120.2	120.4	0.2
C29-H30	1.09	1.11	0.02	C18-C19-H20	121.3	121.5	0.2
C29-H31	1.10	1.11	0.01	C19-C21-C23	120.2	120.1	-0.1
C29-C32	1.52	1.53	0.01	C19-C21-H22	119.7	120.2	0.5
C32-H33	1.09	1.10	0.01	C21-C23-C25	119.7	119.9	0.2
C32-H34	1.09	1.10	0.01	C21-C23-H24	120.1	120.1	0
C32-N35	1.52	1.51	-0.01	C27-C23-C25	120.3	120.0	-0.3
N35-H36	1.03	1.04	0.01	C23-C25-H26	120.0	120.1	0.1
N35-H37	1.03	1.04	0.01	C25-C18-C27	120.1	120.8	0.7
N35-H38	1.03	1.04	0.01	C18-C27-H28	119.6	119.4	-0.2
	Dihedrals/º				Impropers/º		
F12-C1-C5-C6	90.7	86.9	-3.8	C5-F12-C1-F13	122.4	123.1	0.7
C1-C5-C6-C8	-177.5	-176.9	0.6	F13-F12-C1-F14	115.4	113.6	-1.8
C5-C6-C8-C10	0.6	0.5	-0.1	C6-C1-C5-C13	-176.1	-175.2	0.9
C6-C8-C10-O15	-179.5	-183.7	-4.2	C8-C5-C6-H7	-179.7	-181.3	-1.6
C8-C10-O15-C16	21.4	17.9	-3.5	C10-C6-C8-H9	176.5	177.5	1
C10-O15-C16-C18	-70.6	-68.2	2.4	O15-C8-C10-C11	179.3	184.4	5.1
O15-C16-C29-C32	42.5	45.5	3	C13-C10-C11-H12	177.8	177.8	0
C16-C29-C32-N35	177.5	181.7	4.2	C11-C5-C13-H14	178.8	181.3	2.5
C29-C32-N35-H36	179.4	183.1	3.7	C18-O15-C16-C29	130	132.3	2.3
O15-C16-C18-C27	-34.5	-41.7	-7.2	C29-O15-C16-H17	113.7	113.6	-0.1
C16-C18-C19-C21	177.9	178.9	1	C32-C16-C29-H30	-121	-120.8	0.2

94

C18-C19-C21-C23	-0.6	-0.3	0.3	H30-C16-C29-H31	-118.4	-118.6	-0.2
C19-C21-C23-C25	0.05	-0.11	-0.16	N35-C29-C32-H33	117.8	118.9	1.1
Impropers/º				H33-C29-C32-H34	123.6	121.1	2.5
C25-C18-C27-H28	-179	-178.9	0.1	H36-C32-N35-H37	120.7	121.2	0.5
C23-C19-C21-H22	-179.6	-178.8	0.8	H36-C32-N35-H38	-119.8	-119.8	0
C25-C21-C23-H24	179.9	180.1	0.2	C27-C16-C18-C19	-176.8	-178.2	-1.4
C27-C23-C25-H26	180.3	180	-0.3	C21-C18-C19-H20	179.7	179.9	0.2

l) Fluoxetine (S)

Similar to the R-fluoxetine, the CGenFF geometry correlates well with the QM geometry.
All the bond length and bond angle deviations were below 0.02Å and 2.7º, respectively.
Except for the dihedrals C16-C29-C32-N35 and O15-C16-C18-C27 (8.5º and 13.2º,
respectively), the remaining dihedral deviations were all less than 2.8º. The deviations of
all the impropers were below 3.8º (Table 2.27).

Table 2.27. CGenFF equilibrium geometry of fluoxetine (S) compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/°			
C1-F12	1.36	1.34	0.02	C5-C1-F12	110.8	110.8	0
C1-F13	1.35	1.34	0.01	F12-C1-F13	107.2	106.7	-0.5
C1-F14	1.35	1.34	0.01	F12-C1-F14	107.2	106.7	-0.5
C1-C5	1.50	1.52	-0.02	C1-C5-C6	119.9	120.8	0.9
C5=C6	1.39	1.41	-0.02	C1-C5-C13	119.6	120.8	1.2
C6-H7	1.09	1.08	0.01	C5-C6-C8	120.2	120.8	0.6
C6-C8	1.40	1.40	0.00	C5-C6-H7	120.0	120.6	0.6
C8-H9	1.09	1.08	0.01	C6-C8-H9	119.3	118.3	-1
C8=C10	1.40	1.40	0.0	C8-C10-O15	124.3	122.8	-1.5
C10-C11	1.40	1.40	0	C8-C10-C11	120.6	119.0	-1.6
C11-H12	1.09	1.08	0.01	C13-C11-C10	119.9	120.9	1
C11-C13	1.39	1.40	-0.01	C10-C11-H12	118.9	119.2	0.3
C13-H14	1.09	1.08	0.01	C11-C13-C5	119.7	120.8	1.1
C13-C5	1.40	1.41	-0.01	C5-C13-H14	120.2	120.8	0.67
C10-O15	1.38	1.40	-0.01	C10-O15-C16	117.9	118.8	0.9
O15-C16	1.43	1.43	0.00	O15-C16-C18	107.8	109.0	1.2
C16-H17	1.10	1.12	-0.02	O15-C16-C29	109.1	108.2	-0.9
C16-C18	1.51	1.52	-0.01	O15-C16-H17	110.9	111.2	0.3
C18-C19	1.40	1.41	-0.01	C16-C29-C32	108.9	108.3	-0.6
C19-H20	1.09	1.08	0.01	C16-C29-H30	108.3	110.1	1.8
C19-C21	1.39	1.40	-0.01	C16-C29-H31	110.1	108.3	-1.8
C21-H22	1.09	1.08	0.01	C29-C32-N35	110.1	109.2	-0.92
C21=C23	1.40	1.40	0	C29-C32-H33	111.8	111.7	-0.1
C23-H24	1.09	1.08	0.01	C29-C32-H34	111.9	112.1	0.2
C23-C25	1.40	1.40	0	C32-N35-H36	112.0	109.9	-2.05
C25-H26	1.09	1.08	0.01	C32-N35-H37	111.1	108.7	-2.4
C25-C27	1.40	1.40	0.00	C32-N35-H38	111.0	108.3	-2.7
C27-H28	1.09	1.08	0.01	C16-C18-C27	121.9	123	1.1
C27-C18	1.40	1.41	-0.01	C16-C18-C19	118.4	118.5	0.15
C16-C29	1.54	1.54	0.00	C18-C19-C21	120.3	120.7	0.44
C29-H30	1.10	1.11	-0.01	C18-C19-H20	120.2	120.0	-0.2
C29-H31	1.10	1.11	-0.01	C19-C21-C23	120.0	120.0	0
C29-C32	1.52	1.53	-0.01	C19-C21-H22	119.9	120.1	0.24
C32-H33	1.09	1.10	-0.01	C21-C23-C25	119.8	119.9	0.16
C32-H34	1.09	1.10	-0.01	C21-C23-H24	120.0	120.0	0.05
C32-N35	1.52	1.51	0.01	C27-C23-C25	120.4	120.0	-0.39
N35-H36	1.03	1.04	-0.01	C23-C25-H26	120.0	119.9	-0.05
N35-H37	1.03	1.04	-0.01	C25-C18-C27	119.8	120.7	0.89
N35-H38	1.03	1.04	-0.01	C18-C27-H28	119.9	120.0	0.1
Dihedrals/°				Impropers/°			
F12-C1-C5-C6	-90.4	-88.5	1.9	C5-F12-C1-F13	122	123.2	1.2
C1-C5-C6-C8	176.1	178	1.9	F13-F12-C1-F14	115.8	113.5	-2.3
C5-C6-C8-C10	-0.6	-1.44	-0.84	C6-C1-C5-C13	176.6	176.4	-0.2
C6-C8-C10-O15	179.3	176.5	-2.8	C8-C5-C6-H7	178.7	178	-0.7
C8-C10-O15-C16	29.6	28.8	-0.8	C10-C6-C8-H9	177.9	178.3	0.4
C10-O15-C16-C18	182.9	184	1.1	O15-C8-C10-C11	181.7	184.8	3.1
O15-C16-C29-C32	50.6	51.7	1.1	C13-C10-C11-H12	177.6	178.4	0.8
C16-C29-C32-N35	173	181.5	8.5	C11-C5-C13-H14	180.4	182	1.6
C29-C32-N35-H36	179.1	179.7	0.6	C18-O15-C16-C29	-120.7	-116.7	4
O15-C16-C18-C27	-16.6	-3.4	13.2	C29-O15-C16-H17	240	236.2	-3.8
C16-C18-C19-C21	177.8	178.8	1	C32-C16-C29-H30	120.9	122.7	1.8

C18-C19-C21-C23	-0.6	0.43	1.03	H30-C16-C29-H31	118.3	117.4	-0.9
C19-C21-C23-C25	1	0.5	-0.5	N35-C29-C32-H33	118.1	119.3	1.2
Impropers/º				H33-C29-C32-H34	123.2	121.3	1.9
C25-C18-C27-H28	-176.8	-179.0	-2.2	H36-C32-N35-H37	120.3	120.6	0.3
C23-C19-C21-H22	179.6	178.9	-0.7	H36-C32-N35-H38	-120.3	-120.2	0.1
C25-C21-C23-H24	178.5	178.5	0.0	C27-C16-C18-C19	-177.3	-180.2	-2.9
C27-C23-C25-H26	179.8	178.7	-1.1	C21-C18-C19-H20	177.4	178.5	1.1

m) Citalopram (R)

The CGenFF optimized structure matches reasonably well with the QM optimized structure. The deviations of all the bond lengths were less than 0.03Å. The bond angle deviations were less than 2.7º, except for the angles C6-C9-C24 and C6-C9-C12, which were 4.8º and 4.7º, respectively. The dihedrals and impropers fit reasonably well with the QM, except one dihedral angle C6-C9-C24-C29 (deviation was 33.7º)(Table 2.28).

Table 2.28. CGenFF equilibrium geometry of citalopram (R) compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/°		
N1Ξ C2	1.18	1.18	0.00	N1-C2-C3	178.9	178.8	-0.1
C2-C3	1.43	1.44	-0.01	C2-C3-C4	119.4	120.2	0.8
C3=C4	1.41	1.40	0.01	C2-C3-C8	119.4	120.0	0.6
C4-H4	1.09	1.08	0.01	C3-C4-C5	120.1	120.4	0.3
C4-C5	1.40	1.40	0.00	C3-C4-H4	119.3	120.2	0.9
C5-H5	1.09	1.08	0.01	C4-C5-C6	118.7	119.9	1.2
C5=C6	1.40	1.39	0.01	C4-C5-H5	119.7	119.3	-0.4
C6-C7	1.40	1.41	-0.01	C5-C6-C9	131.2	132.6	1.4
C7=C8	1.39	1.39	0.00	C5-C6-C7	120.8	119.2	-1.6
C8-H8	1.09	1.08	0.01	C7-C8-C3	117.8	119.5	1.7
C8-C3	1.41	1.40	0.01	C6-C9-C24	111.0	106.2	-4.8
C6-C9	1.51	1.52	-0.01	C6-C9-C12	115.7	111.0	-4.7
C9-O11	1.45	1.43	0.02	C6-C9-O11	102.0	103.3	1.3
O11-C10	1.44	1.43	0.01	C9-O11-C10	106.7	109.3	2.6
C10-H30	1.09	1.1	-0.01	C7-C10-O11	102.9	103.6	0.7
C10-H31	1.1	1.1	0.00	O11-C10-H31	110.5	108.9	-1.6
C10-C7	1.5	1.52	-0.02	O11-C10-H30	107.8	107.6	-0.2
C9-C24	1.52	1.54	-0.02	C9-C12-C13	115.1	116.9	1.8
C24-C25	1.4	1.41	-0.01	C9-C12-H33	108.2	108.9	0.7
C25-H25	1.09	1.08	0.01	C9-C12-H32	108.4	110.2	1.8
C25=C26	1.39	1.40	-0.01	C12-C13-C14	113.2	111.7	-1.5
C26-H26	109	1.08	0.01	C12-C13-H34	107.7	108.4	0.7
C26-C27	1.39	1.39	0.00	C12-C13-H35	107.6	107.2	-0.4
C27-F27	1.35	1.36	-0.01	C12-C13-C14	113.2	111.7	-1.5
C27=C28	1.39	1.39	0.00	C13-C14-N15	112.2	111.6	-0.6
C28-H28	1.09	1.08	0.01	C13-C14-H36	113.1	113.4	0.3
C28-C29	1.40	1.40	0.00	C13-C14-H37	110.4	110.7	0.3
C29-H29	1.09	1.08	0.01	C14-N15-C16	110.7	108.9	-1.8
C29-C24	1.40	1.42	-0.02	C14-N15-C20	113.3	110.6	-2.7
C9-C12	1.53	1.53	0.00	C14-N15-H15	106.3	107.1	0.8
C12-H32	1.10	1.11	-0.01	N15-C16-H17	109.1	108.4	-0.7
C12-C13	1.54	1.55	-0.02	N15-C16-H18	108.7	107.7	-1
C13-H34	1.1	1.11	-0.01	N15-C16-H19	108.2	107.9	-0.3
C13-H35	1.09	1.11	-0.02	C14-N15-C20	113.3	110.6	-2.7
C13-C14	1.52	1.53	-0.01	N15-C20-H21	108.7	108.4	-0.3
C14-H36	1.10	1.1	0.00	N15-C20-H22	108.0	108.1	0.1
C14-H37	1.09	1.1	-0.01	N15-C20-H23	109.4	107.8	-1.6
C14-N15	1.52	1.50	0.02	C9-C24-C29	121.9	121.4	-0.5
N15-H15	1.03	1.06	-0.03	C9-C24-C25	118.6	121.3	2.7
N15-C20	1.50	1.50	0.00	C24-C25-C26	120.6	121.4	0.8
C20-H21	1.09	1.11	-0.02	C24-C25-H25	119.5	120.9	1.4
C20-H22	1.09	1.11	-0.02	C25-C26-C27	118.6	119.1	0.5
C20-H23	1.09	1.11	-0.02	C25-C26-H26	121.7	120.7	-1
N15-C16	1.50	1.50	0.00	C24-C29-C28	120.6	121.3	0.7
C16-H17	1.09	1.11	-0.02	C24-C29-H29	120.7	121.4	0.7
C16-H18	1.09	1.11	-0.02	C28-C27-C26	122.1	121.7	-0.4
C16-H19	1.09	1.11	-0.02	C26-C27-F27	118.9	119.4	0.5
C12-H33	1.09	1.11	-0.01	C27-C28-C29	118.7	119.1	0.4
				C29-C28-H28	121.7	120.4	-1.3
	Dihedrals/°				Impropers/°		
N1-C2-C3-C4	99.1	101.4	2.3				

98

C2-C3-C4-C5	181.6	181.4	-0.2	C4-C2-C3-C8	181.6	182.1	0.5
C3-C4-C5-C6	0.8	-1.5	-2.3	C5-C3-C4-H4	-179.6	-179.7	-0.1
C4-C5-C6-C9	179.3	176.3	-3	C6-C4-C5-H5	-180.7	-178.5	2.2
C5-C6-C9-C24	83.6	74.52	-9.0	C9-C5-C6-C7	-180.4	-172.3	8.1
C6-C9-O11-C10	325.2	334.5	9.3	C7-C3-C8-H8	180	179.8	-0.2
C6-C9-C12-C13	284.4	271.1	-13.3	C24-C6-C9-C12	-126.9	-116.6	10.3
C9-C12-C13-C14	56.7	67.1	10.4	C12-C6-C9-O11	-115.8	-120.8	-5
C12-C13-C14-N15	-173	-178.7	-5.7	C7-O11-C10-H31	121.8	123.7	1.9
C13-C14-N15-C16	169.2	162.4	-6.8	C7-O11-C10-H30	-119.2	-120.7	-1.5
C14-N15-C16-H17	-176.1	-179.7	-3.6	C13-C9-C12-H33	-120.1	-119.5	0.6
C14-N15-C20-H21	-178.7	-180.3	-1.6	H33-C9-C12-H32	-114.8	-116.3	-1.5
C6-C9-C24-C29	320.8	287.1	-33.7	C14-C12-C13-H34	-120.7	-122.1	-1.4
C9-C24-C25-C26	177.1	179.3	2.2	H34-C12-C13-H35	-116.4	-116.1	0.3
C24-C25-C26-C27	-1	-0.06	0.9	N15-C13-C14-H36	-120.2	-120	0.2
C9-C24-C29-C28	-176.6	-179.6	-3	C16-C14-N15-H15	-116.5	-120.9	-4.4
Impropers/º				H17-N15-C16-H18	120.4	119.7	0.17
H21-N15-C20-H23	-120.2	-119.7	0.5	H17-N15-C16-H19	-119.9	-121	-1.1
C29-C9-C24-C25	-176	-178.5	-2.5	H21-N15-C20-H22	119.7	121	1.3
C26-C24-C25-H25	-181.3	-177.9	3.4	C28-C26-C27-F27	179.5	180.4	0.9
C27-C25-C26-H26	-179.7	-179.7	0	C28-C24-C29-H29	181.6	181.6	0

n) Citalopram (S)

The CGenFF optimized structure is in reasonable agreement with the QM structure. All the bond length and angle deviations were below 0.03Å and 2.94º, respectively. All the dihedral deviations were less than 3.1º, except for the N1-C2-C3-C4, the C6-C9-C12-C13 and the C13-C14-N15-C16, which were 15º, 7.5º, and 7.9º, respectively. Impropers were below 4.2º, except for the C24-C6-C9-C12 and the C12-C6-C9-O11, which were 11.6º and 7.7º, respectively (Table 2.29).

Table 2.29. CGenFF equilibrium geometry of citalopram (S) compared to MP2 Level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
	Bond lengths/Å				Angles/°		
N1≡C2	1.18	1.18	0.00	N1-C2-C3	179.8	179.78	0.02
C2-C3	1.44	1.44	0.00	C2-C3-C4	119.3	120.22	-0.92
C3=C4	1.41	1.40	0.01	C2-C3-C8	119.5	120.11	-0.61
C4-H4	1.09	1.08	0.01	C3-C4-C5	120.1	120.46	-0.36
C4-C5	1.40	1.40	0.00	C3-C4-H4	119.4	120.20	-0.80
C5-H5	1.09	1.08	0.01	C4-C5-C6	118.8	119.57	-0.77
C5=C6	1.39	1.39	0.00	C4-C5-H5	119.3	119.46	-0.16
C6-C7	1.40	1.41	-0.01	C5-C6-C9	130.9	133.20	-2.3
C7=C8	1.39	1.39	0.00	C5-C6-C7	120.9	119.83	1.07
C8-H8	1.09	1.08	0.01	C7-C8-C3	117.8	119.54	-1.74
C8-C3	1.41	1.40	0.01	C6-C9-C24	111.8	108.86	2.94
C6-C9	1.51	1.54	-0.03	C6-C9-C12	113.2	112.19	1.01
C9-O11	1.44	1.43	0.01	C6-C9-O11	103.3	104.65	-1.35
O11-C10	1.44	1.43	0.01	C9-O11-C10	108.5	107.91	0.59
C10-H30	1.09	1.1	-0.01	C7-C10-O11	103.7	103.76	-0.06
C10-H31	1.1	1.1	0.00	O11-C10-H31	110.3	108.28	1.95
C10-C7	1.5	1.52	-0.02	O11-C10-H30	107.5	110.34	-2.84
C9-C24	1.52	1.54	-0.02	C9-C12-C13	111.0	112.79	-1.79
C24-C25	1.4	1.42	-0.02	C9-C12-H33	107.6	106.33	1.27
C25-H25	1.09	1.08	0.01	C9-C12-H32	109.0	109.87	-0.87
C25=C26	1.39	1.4	-0.01	C12-C13-C14	110.3	110.11	0.19
C26-H26	109	1.08	0.01	C12-C13-H34	109.0	109.64	-0.64
C26-C27	1.39	1.39	0.00	C12-C13-H35	107.7	107.98	-0.28
C27-F27	1.35	1.36	-0.01	C12-C13-C14	110.3	110.11	0.19
C27=C28	1.39	1.39	0.00	C13-C14-N15	112.6	111.54	1.06
C28-H28	1.09	1.08	0.01	C13-C14-H36	112.4	113.38	-0.98
C28-C29	1.40	1.40	0.00	C13-C14-H37	111.1	110.52	0.48
C29-H29	1.09	1.08	0.01	C14-N15-C16	111.4	109.11	2.29
C29-C24	1.40	1.42	-0.02	C14-N15-C20	112.6	110.50	2.10
C9-C12	1.53	1.56	-0.03	C14-N15-H15	106.5	107.17	-0.67
C12-H32	1.1	1.11	0.00	N15-C16-H17	109.0	108.39	0.61
C12-C13	1.53	1.55	-0.02	N15-C16-H18	108.9	107.75	1.15
C13-H34	1.1	1.11	-0.01	N15-C16-H19	108.1	107.96	0.14
C13-H35	1.09	1.11	-0.02	C14-N15-C20	112.6	110.49	2.11
C13-C14	1.52	1.53	-0.01	N15-C20-H21	108.6	108.35	0.25
C14-H36	1.09	1.1	-0.01	N15-C20-H22	107.9	108.11	-0.21
C14-H37	1.09	1.1	-0.01	N15-C20-H23	109.4	107.80	1.6
C14-N15	1.52	1.50	0.02	C9-C24-C29	120.3	119.83	0.47
N15-H15	1.03	1.06	-0.03	C9-C24-C25	120.6	121.82	-1.22
N15-C20	1.50	1.50	0.00	C24-C25-C26	120.8	120.81	-0.01
C20-H21	1.09	1.11	-0.02	C24-C25-H25	120.4	120.98	-0.58
C20-H22	1.09	1.11	-0.02	C25-C26-C27	118.6	119.12	-0.52
C20-H23	1.09	1.11	-0.02	C25-C26-H26	121.7	120.73	0.97
N15-C16	1.50	1.50	0.00	C24-C29-C28	120.7	120.98	-0.28
C16-H17	1.09	1.11	-0.02	C24-C29-H29	119.3	120.49	-1.19
C16-H18	1.09	1.11	-0.02	C28-C27-C26	122.0	121.80	0.2
C16-H19	1.09	1.11	-0.02	C26-C27-F27	118.9	119.33	-0.43
C12-H33	1.10	1.11	-0.01	C27-C28-C29	118.7	119.03	-0.33
				C29-C28-H28	121.7	120.41	1.29
	Dihedrals/°				Impropers/°		
N1-C2-C3-C4	-135.2	-120.2	-15				

100

C2-C3-C4-C5	179.7	179.8	-0.1	C4-C2-C3-C8	179.3	179.2	0.1
C3-C4-C5-C6	0.1	1.1	-1	C5-C3-C4-H4	180.5	180.1	0.4
C4-C5-C6-C9	179.7	182.1	-2.4	C6-C4-C5-H5	180.8	179.9	0.9
C5-C6-C9-C24	-79.3	-74.3	-5	C9-C5-C6-C7	179.0	174.8	4.2
C6-C9-O11-C10	27.6	27.5	0.1	C7-C3-C8-H8	179.0	179.5	0.5
C6-C9-C12-C13	171.7	179.2	-7.5	C24-C6-C9-C12	127.0	115.4	11.6
C9-C12-C13-C14	-175.9	-175.6	-0.3	C12-C6-C9-O11	114.3	122.0	-7.7
C12-C13-C14-N15	-173.0	-176.1	3.1	C7-O11-C10-H31	120.8	120.3	0.5
C13-C14-N15-C16	170.7	162.8	7.9	C7-O11-C10-H30	-121.3	-123.5	2.2
C14-N15-C16-H17	-177.2	-179.7	2.5	C13-C9-C12-H33	-121.0	-119.1	-1.9
C14-N15-C20-H21	180.9	179.5	1.4	H33-C9-C12-H32	-117.0	-116.8	-0.2
C6-C9-C24-C29	-126.6	-124.2	-2.4	C14-C12-C13-H34	-121.2	-120.5	-0.7
C9-C24-C25-C26	178.1	176.5	1.6	H34-C12-C13-H35	-117.0	-117.6	0.6
C24-C25-C26-C27	0.1	0.0	0.1	N15-C13-C14-H36	-119.7	-120.0	-1.0
C9-C24-C29-C28	-178.4	-176.5	-1.9	C16-C14-N15-H15	-117.1	-120.9	3.8
Impropers/°				H17-N15-C16-H18	120.4	119.7	0.7
H21-N15-C20-H23	-120.3	-119.7	-0.6	H17-N15-C16-H19	-119.9	-121.0	1.1
C29-C9-C24-C25	-178.0	-176.7	-1.3	H21-N15-C20-H22	119.6	121.0	-1.4
C26-C24-C25-H25	178.3	180.2	-0.5	C28-C26-C27-F27	179.9	180.6	-0.7
C27-C25-C26-H26	179.1	179.8	-0.7	C28-C24-C29-H29	180.9	179.1	1.8

o) **Sertraline**

The optimized CGenFF geometry of sertraline fits well with the QM structure. All the bond length and bond angle differences were below 0.03Å and 3.1°, respectively. All the dihedral and improper deviations were below 6.3° and 4.6°, respectively (Table 2.30).

Table 2.30. CGenFF equilibrium geometry of sertraline compared to MP2 level.

Coordinate	MP2	CGenFF	Diff	Coordinate	MP2	CGenFF	Diff
Bond lengths/Å				Angles/°			
C1-H2	1.09	1.11	-0.02	H2-C1-N5	108.7	106.8	-1.9
C1-H3	1.09	1.11	-0.02	N5-C1-H3	108.5	106.6	-1.9
C1-H4	1.09	1.11	-0.02	N5-C1-H4	108.8	107.1	-1.7
C1-N5	1.50	1.49	0.01	C1-N5-C8	116.4	119.5	3.1
N5-H6	1.03	1.01	0.02	C1-N5-H6	109.5	111.4	1.9
N5-H7	1.04	1.01	0.03	C1-N5-H7	110.4	109.9	-0.5
N5-C8	1.53	1.51	0.02	N5-C8-C19	108.6	105.9	-2.7
C8-H9	1.10	1.12	0.02	N5-C8-C10	111.5	109.1	-2.4
C8-C10	1.52	1.54	0.02	N5-C8-H9	104.8	105.8	1
C10-H11	1.10	1.11	-0.01	C8-C10-H11	112.2	112.0	-0.2
C10-H12	1.09	1.12	-0.03	C8-C10-H12	104.8	107.8	3
C10-C13	1.54	1.53	0.01	C10-C13-C16	114.7	115.7	1
C13-H14	1.10	1.11	-0.01	C10-C13-H14	109.6	108.2	-1.4
C13-H15	1.09	1.11	-0.02	C10-C13-H15	108.1	107.7	-0.4
C13-C16	1.56	1.56	0.00	C13-C16-C28	110.8	110.0	-0.8
O15-C16	1.43	1.43	0.00	C13-C16-C18	113.0	112.3	-0.7
C16-H17	1.10	1.12	-0.02	C13-C16-H17	107.8	105.8	-2
C16-C18	1.51	1.53	-0.02	C18-C19-C8	117.8	118.7	0.9
C16-C28	1.53	1.53	0.00	C8-C19-C20	121.2	120.0	-1.2
C18=C19	1.41	1.42	-0.01	C19-C20-C22	119.9	120.0	0.1
C19-C8	1.50	1.49	0.01	C19-C20-H21	120.1	121.1	1
C19-C20	1.40	1.41	-0.01	C20-C22-C24	119.7	120.1	-1.9
C20-H21	1.09	1.08	0.01	C20-C22-H23	120.0	120.2	-1.9
C20=C22	1.40	1.40	0.00	C22-C24-C26	120.2	120.0	-1.7
C22-H23	1.09	1.08	0.01	C22-C24-H25	120.0	120.0	3.1
C22-C24	1.40	1.40	0.00	C18-C26-C24	120.8	120.6	1.9
C24-H25	1.09	1.08	0.01	C24-C26-H27	119.8	119.5	-0.5
C24=C26	1.40	1.40	0.00	C16-C28-C29	119.4	117.6	-2.7
C26-H27	1.09	1.08	0.01	C16-C28-C37	122.5	124.3	-2.4
C26-C18	1.40	1.41	-0.01	C28-C29-C31	120.9	121.0	1
C28=C29	1.41	1.41	0.00	C28-C29-H30	120.2	119.89	-0.2
C29-H30	1.09	1.08	0.01	C29-C31-C33	120.7	120.3	3
C29-C31	1.39	1.40	-0.01	C29-C31-H32	120.5	120.0	1
C31-H32	1.09	1.08	0.01	C31-C33-C35	119.0	119.5	-1.4
C31=C33	1.40	1.40	0.00	C31-C33-CL34	119.3	118.5	-0.4
C33-CL34	1.72	1.74	-0.02	C37-C35-C33	120.0	119.7	-0.8
C33-C35	1.40	1.41	-0.01	C33-C35-CL36	121.6	122.1	-0.7
C35-CL36	1.73	1.74	-0.01	C28-C37-H38	120.7	120.6	-2
C35=C37	1.40	1.40	0.00				
C37-H38	1.09	1.08	0.01				
C37-C28	1.40	1.41	-0.01				
Dihedrals/°				Impropers/°			
H2-C1-N5-C8	52.8	59.0	6.2	H2-N5-C1-H3	120.0	119.9	-0.1
C1-N5-C8-C19	178.7	181.6	2.9	H2-N5-C1-H4	-120.3	-120.4	-0.1
N5-C8-C10-C13	59.9	53.6	-6.3	C8-C1-N5-H6	-121.3	-125.9	-4.6
C8-C10-C13-C16	32.5	36.2	3.7	H6-C1-N5-H7	-114.6	-115.3	-0.7
C10-C13-C16-C28	-111.4	-119.6	-8.2	C19-N5-C8-C10	-120.4	-116.1	4.3
C20-C22-C24-C26	0.8	-0.2	-1	C10-N5-C8-H9	-120.6	-123.1	-2.5
C13-C16-C28-C29	-55.6	-57.2	-1.6	C13-C8-C10-H11	-125.5	-124.9	0.6
C16-C28-C29-C31	180.5	179.4	-1.1	H11-C8-C10-H12	-114.7	-116.4	-1.7
C28-C29-C31-C33	-0.9	-0.01	0.89	C16-C10-C13-H14	-122.8	-125.6	-2.8

C29-C31-C33-C35	1.5	1.5	0	H14-C10-C13-H15	-114.7	-113.5	1.2
Impropers/°				C28-C13-C16-C18	125.5	126.9	-0.5
C35-C31-C33-CL34	179.5	179.8	0.3	C18-C13-C16-H17	118.6	117.2	-1.4
C37-C33-C35-CL36	180.0	180.7	0.7	C18-C8-C19-C20	184.5	-176.9	-1.4
C35-C28-C37-H38	183.2	181.5	-1.7	C8-C19-C20-C22	176.5	178.6	2.1
C18-C24-C26-H27	180.2	180.2	0	C22-C19-C20-H21	-178.9	-179.8	-0.9
C29-C16-C28-C37	181.1	181.9	0.8	C19-C20-C22-C24	-1.2	-0.8	0.4
C31-C28-C29-H30	178.0	179.0	1	C24-C20-C22-H23	-178.6	-179.2	-0.6
C33-C29-C31-H32	180.0	179.8	-0.2	C26-C22-C24-H25	180.0	179.3	-0.7

2.3.3. Dihedral energy scans

Since the CGenFF geometries of all the compounds fit reasonably well with the QM optimized structures, dihedral/torsion angles were not fitted against the dihedral energy scans.

2.3.4. Lennard-Jones parameter optimization

According to the CGenFF parameterization philosophy, LJ parameters are to be optimized only when the new atom types are created. Since the new atom types were not created in our study, this step was not performed.

2.4. Conclusions

Good force field parameters for the ligands are important for the accurate estimation of protein-ligand binding affinities. MATs are key therapeutic drug targets in humans. The unavailability of the crystal structures of MATs until recently necessitated the use of computational techniques for drug discovery as well as for structural and functional studies. More importantly, MD simulations and free energy calculations stand at the forefront among other computational methods, and utilization of good force field parameters ensures the reliability of the results. For that reason, parameters for fifteen MAT ligands using the CGenFF philosophy were developed. The CGenFF parameters

correlated well with the QM data, which was used as a target for parameter optimization. However, in this study dihedral fitting using the potential energy scans was not pursued since the CGenFF charges and equilibrium geometry matched well with the QM data. Also, the LJ parameters were not optimized, since no new LJ parameters were generated in this study. The CGenFF parameters for the ligands in conjugation with the CHARMM force field for the proteins improve the reliability of computational results and lessen the dependence on experiments; therefore, research efforts can be more economical and less time-consuming.

2.5. References

1. Immadisetty, K., L.M. Geffert, C.K. Surratt, and J.D. Madura, *New design strategies for antidepressant drugs.* Expert Opin Drug Discov, 2013. **8**(11): p. 1399-414.

2. Amara, S.G. and M.S. Sonders, *Neurotransmitter transporters as molecular targets for addictive drugs.* Drug and Alcohol Dependence, 1998. **51**(1,Äì2): p. 87-96.

3. Andersen, J., A.S. Kristensen, B. Bang-Andersen, and K. Stromgaard, *Recent advances in the understanding of the interaction of antidepressant drugs with serotonin and norepinephrine transporters.* Chemical Communications, 2009(25): p. 3677-3692.

4. Gether, U., P.H. Andersen, O.M. Larsson, and A. Schousboe, *Neurotransmitter transporters: molecular function of important drug targets.* Trends in Pharmacological Sciences, 2006. **27**(7): p. 375-383.

5. Iversen, L., *Neurotransmitter transporters and their impact on the development of psychopharmacology.* British Journal of Pharmacology, 2006. **147**(S1): p. S82-S88.

6. Kristensen, A.S., J. Andersen, T.N. Jorgensen, L. Sorensen, J. Eriksen, C.J. Loland, K. Stromgaard, and U. Gether, *SLC6 neurotransmitter transporters: structure, function, and regulation.* Pharmacol Rev, 2011. **63**(3): p. 585-640.

7.   Masson, J., C. Sagne, M. Hamon, and S.E. Mestikawy, *Neurotransmitter transporters in the central nervous system.* Pharmacological Reviews, 1999. **51**(3): p. 439-464.

8.   Immadisetty, K. and J.D. Madura, *A review of monoamine transporter-ligand interactions.* Curr Comput Aided Drug Des, 2013.

9.   Manepalli, S., C.K. Surratt, J.D. Madura, and T.L. Nolan, *Monoamine transporter structure, function, dynamics, and drug discovery: a computational perspective.* AAPS J, 2012. **14**(4): p. 820-31.

10.   Indarte, M., J.D. Madura, and C.K. Surratt, *Dopamine transporter comparative molecular modeling and binding site prediction using the LeuTAa leucine transporter as a template.* Proteins: Structure, Function, and Bioinformatics, 2008. **70**(3): p. 1033-1046.

11.   Manepalli, S., L.M. Geffert, C.K. Surratt, and J.D. Madura, *Discovery of novel selective serotonin reuptake inhibitors through development of a protein-based pharmacophore.* J Chem Inf Model, 2011. **51**(9): p. 2417-26.

12.   Nolan, T.L., D.J. Lapinsky, J.N. Talbot, M.n. Indarte, Y. Liu, S. Manepalli, L.M. Geffert, M.E. Amos, P.N. Taylor, J.D. Madura, and C.K. Surratt, *Identification of a novel selective serotonin reuptake inhibitor by coupling monoamine transporter-based virtual screening and rational molecular hybridization.* ACS Chemical Neuroscience, 2011. **2**(9): p. 544-552.

13.   Nolan, T.L., L.M. Geffert, B.J. Kolber, J.D. Madura, and C.K. Surratt, *Discovery of novel-scaffold monoamine transporter ligands via in silico screening with the S1 pocket of the serotonin transporter.* ACS Chem Neurosci, 2014. **5**(9): p. 784-92.

14.   Indarte, M.n., Y. Liu, J.D. Madura, and C.K. Surratt, *Receptor-based discovery of a plasmalemmal monoamine transporter inhibitor via high-throughput docking and pharmacophore modeling.* ACS Chemical Neuroscience, 2010. **1**(3): p. 223-233.

15.   Mackerell, A.D., Jr., *Empirical force fields for biological macromolecules: overview and issues.* J Comput Chem, 2004. **25**(13): p. 1584-604.

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

17. Wang, J., R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case, *Development and testing of a general amber force field.* J Comput Chem, 2004. **25**(9): p. 1157-74.

18. MacKerell, A.D., D. Bashford, Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* The Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

19. Foloppe, N. and J.A.D. MacKerell, *All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data.* Journal of Computational Chemistry, 2000. **21**(2): p. 86-104.

20. Hegazy, L. and N.G. Richards, *Optimized CGenFF force-field parameters for acylphosphate and N-phosphonosulfonimidoyl functional groups.* J Mol Model, 2013. **19**(11): p. 5075-87.

21. Yu, W., X. He, K. Vanommeslaeghe, and A.D. MacKerell, Jr., *Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations.* J Comput Chem, 2012. **33**(31): p. 2451-68.

22. Vanommeslaeghe, K. and A.D. MacKerell, Jr., *Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing.* J Chem Inf Model, 2012. **52**(12): p. 3144-54.

23. Vanommeslaeghe, K., E.P. Raman, and A.D. MacKerell, Jr., *Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges.* J Chem Inf Model, 2012. **52**(12): p. 3155-68.

24. Francl, M.M., W.J. Pietro, W.J. Hehre, J.S. Binkley, M.S. Gordon, D.J. DeFrees, and J.A. Pople, *Self‐consistent molecular orbital methods. XXIII. A polarization‐type basis set for second‐row elements.* The Journal of Chemical Physics, 1982. **77**(7): p. 3654-3665.

25. Head-Gordon, M., J.A. Pople, and M.J. Frisch, *MP2 energy evaluation by direct methods.* Chemical Physics Letters, 1988. **153**(6): p. 503-506.

26. Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, *Comparison of simple potential functions for simulating liquid water.* The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.

27.   Frisch, M.J., J.A. Pople, and J.S. Binkley, *Self-consistent molecular orbital methods 25. Supplementary functions for gaussian basis sets.* The Journal of Chemical Physics, 1984. **80**(7): p. 3265-3269.

28.   Ditchfield, R., W.J. Hehre, and J.A. Pople, *Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules.* The Journal of Chemical Physics, 1971. **54**(2): p. 724-728.

29.   Frisch, M.J., G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M.J. Bearpark, J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, and D.J. Fox, *Gaussian 09*, 2009, Gaussian, Inc.: Wallingford, CT, USA.

30.   Brooks, B.R., C.L. Brooks, 3rd, A.D. Mackerell, Jr., L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus, *CHARMM: the biomolecular simulation program.* J Comput Chem, 2009. **30**(10): p. 1545-614.

3. CHAPTER 3

CALCULATION OF RELATIVE BINDING FREE ENERGIES OF LIGANDS TO LEUCINE TRANSPORTER

3.1. Introduction
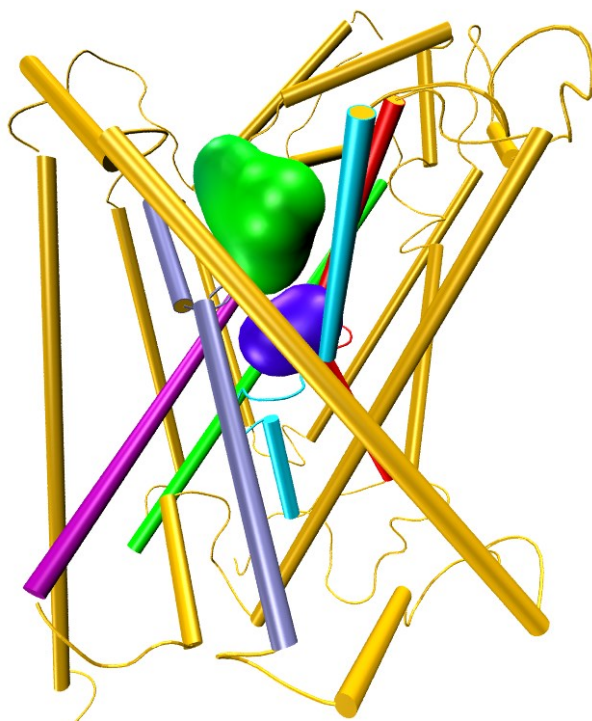
Figure 3.1. 3D-model of leucine transporter with S1 (blue) and S2 (green) pockets highlighted. TM helices 1, 3, 6, 8, and 10 are colored red, purple, cyan, green, and blue, respectively; rest of the helices are colored yellow.

Leucine transporter (LeuT) is a membrane protein of the NSS family. LeuT was the first member of the NSS family that was crystallized [1]. NSS family also consists of other

therapeutically important membrane proteins such as the human serotonin transporter (hSERT), the human dopamine transporter (hDAT) and the human norepinephrine transporter (hNET). All three are famously addressed as monoamine transporters (MATs) because they transport monoamines serotonin, dopamine and norepinephrine, respectively, across the presynaptic cell membrane. Until the crystal structure of dDAT recently became available, LeuT was the only member of NSS family that was crystallized, and to date 38 crystal structures of LeuT are available with several amino acids and drugs [1-5]. LeuT shares approximately 20-25% sequence identity with MATs [6]. LeuT also shares other similarities with MATs such as: 12 transmembrane alpha-helices (Fig. 3.1), Na^+/Cl^- gradient to transport the neurotransmitters, and a hypothesized similar transport mechanism (alternating access) [1, 7]. LeuT has two unique binding sites, an S1 site that accommodates substrate leucine and other amino acids [1, 2, 8] and an S2 site which harbors the TCAs [3, 4] and selective serotonin reuptake inhibitors (SSRIs) [5] (Fig. 3.1). However, according to another hypothesis substrate leucine binds in both S1 and S2 sites [9]; this is still debatable. LeuT was extensively used to study the structure, function and dynamics aspects of the MATs through various computational and experimental approaches [6, 10-14]. Computer models of MATs generated using LeuT were successfully used for drug discovery purposes as well, especially for addiction and depression-related research [6, 10]. Here, we chose LeuT as an ideal system to test the accuracy of FEP method in estimating the RBEs of amino acids and TCAs.

Binding free energy is a quantitative measure of how well a ligand binds with the protein/receptor. Since the determination of binding free energies of ligands for a target through experimental binding assays is an expensive process, free energy calculations are

gaining prominence.
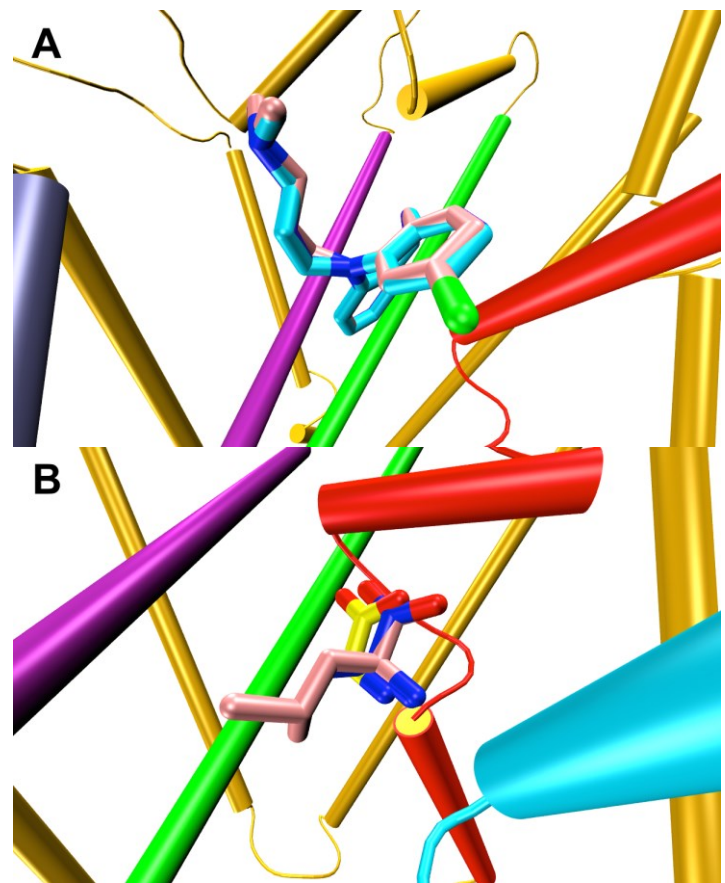
Figure 3.2. TCAs and amino acids binding in LeuT.
(A) Clomipramine (pink), imipramine (cyan) and
desipramine (purple) in the S2 site; (B) leucine (pink),
alanine (blue) and glycine (yellow) in the S1 site. All
ligands are represented as sticks.

Computing the RBEs of ligands with a protein is a topic of interest, especially in drug

discovery. RBEs can be calculated using ABFE and RBFE methods; however, because of

the methodological challenges associated with ABFE methods, RBFE methods have been

extensively used for this purpose. RBFE methods are also used to calculate the impact of

protein mutation on the binding of a ligand and also for the lead optimization [15] in

rational drug design. We used FEP method to calculate the RBEs [16]. Alchemical methods have been used since the 1980s to calculate RBEs [17-20].

FEP is an MD-based method and explicitly treats protein, ligand and solvent using a force field. The binding affinities estimated through FEP method are quantitatively comparable to the experimental binding affinities. Most often, the calculated RBEs are within 1 kcal/mol of the experimental values. These methods offer detailed molecular insights by sampling in all degrees of freedom [21, 22]. Since the role of water in ligand binding was illustrated through several studies, inclusion of water explicitly makes the calculations more realistic [5]. The major advantage of RBFE calculations is the cancelation of errors, given that both the ligands have a similar binding mode and protein conformational changes have a similar impact on their binding. Limitations include (a) a need for prior knowledge of at least one complex and (b) can handle only small differences between ligands. Other factors affecting the accuracies of explicit free energy calculations are poor convergence due to insufficient sampling and lack of good force field parameters for drug-like molecules. Availability of better sampling techniques and high-performance supercomputers solved the sampling issues to some extent. Also, General force fields such as amber (GAFF) [23] and CHARMM (CGenFF) [24] can be extended to organic molecules to get reasonable accuracy in free energy calculations involving drug-like molecules. We used CGenFF parameters for calculations involving TCAs.

FEP method was initially validated through estimation of RBEs of leucine-to-alanine and alanine-to-glycine perturbations in LeuT. We chose these two perturbations as a test case because crystal structures of all three amino acid ligands leucine, alanine and glycine

with LeuT [2] were available, and they binded in the same pocket and shared a similar binding mode (Fig. 3.2B). Also, the experimental inhibition constants (K_i) of all three amino acid ligands with LeuT were available [2], which enabled us to compare the experimental and calculated RBEs and also their trends, apart from the availability of well characterized force field parameters for the ligands. The calculated RBEs in both these cases were in excellent agreement with the experimental values. The calculated and experimental $\Delta\Delta$Gs for the alanine-to-glycine perturbation were 3.26 ± 0.05 and 2.85 ± 0.22 kcal/mol, respectively; whereas in the case of leucine-to-alanine they were 3.34 ± 0.07 and 3.17 ± 0.30 kcal/mol, respectively (Table 3.1). The calculated results were not only in quantitative agreement with the experimental values. They also reproduced the experimental binding trend: leucine binded stronger, followed by alanine and glycine with LeuT [2]. These results show the reliability of FEP method in calculating the RBEs, given that good force field parameters for the ligands and ample computational resources for sufficient sampling are available. These results also demonstrate the applicability of explicit free energy methods to membrane proteins such as LeuT.

With the confidence gained through amino acid perturbations, RBEs of TCAs (clomipramine, imipramine and desipramine) with LeuT were calculated. All three TCAs were co-crystallized with LeuT, and they bind in the S2 site (Fig. 3.1) in a similar orientation (Fig. 3.2A). The structural differences between the three TCAs are minimal (Fig. 3.7) and inhibit the protein in a similar manner [3, 4]. However, there is an ambiguity with respect to the binding affinities and binding trends of the three TCAs with LeuT. TCAs are weak binders and either experimental K_i or K_d values with LeuT are not available. Half maximal inhibitory concentration (IC_{50}) values are available, but not for

all three of them. Singh et al. reported only IC_{50} values of just clomipramine and imipramine with LeuT (IC_{50} of clomipramine is 8 fold lower than imipramine), but not desipramine [3]. However, Noskov et al. sought personal communication with Singh et al. and reported that imipramine inhibits stronger than desipramine [25]. Therefore, the experimental binding trend of TCAs with LeuT according to Singh et al. is that clomipramine binds stronger followed by imipramine and desipramine. On the contrary, Zhou et al. reported the IC_{50} values of imipramine and desipramine but not clomipramine. According to Zhou et al., desipramine is a potent inhibitor of leucine transport compared to imipramine (three-fold) [4]. Zhao et al. through ABFE and RBFE calculations reported a similar trend in binding affinities as Singh et al. However, they did not mention about the binding trend of Zhou et al., which was opposite of Singh et al. [25]. In this paper, they calculated the $\Delta\Delta$Gs of clomipramine-to-imipramine and imipramine-to-desipramine perturbations using the ABFE and RBFE methods; the calculated $\Delta\Delta$Gs resulting from both the methods were in reasonable agreement with each other. ABEs and RBEs were estimated using the FEP methodology; however, in the ABFE calculations reduced GSBP systems were used, and in the RBFE calculations the entire system was treated explicitly. For the clomipramine-to-imipramine perturbation, the $\Delta\Delta$Gs estimated through the ABFE and RBFE methods were 0.2 and 0.4 kcal/mol, respectively; whereas the experimental $\Delta\Delta$G was 1.26 kcal/mol. The calculated $\Delta\Delta$Gs for the perturbation of imipramine-to-desipramine using the ABFE and RBFE methods were 1.3 and 1.0 kcal/mol, respectively, and the experimental $\Delta\Delta$G in this case was > 0 kcal/mol (The IC_{50} of desipramine was not reported; it was only reported that imipramine was more potent than desipramine). Since the clomipramine-to-desipramine transformation was not considered for RBFE

calculations, the ΔΔGs resulting from ABFE and RBFE calculations cannot be compared. The ΔΔG of clomipramine-to-desipramine perturbation would have facilitated the estimation of reliability of RBFE calculations. The sum of ΔΔGs for the clomipramine-to-imipramine, imipramine-to-desipramine and clomipramine-to-desipramine should add to zero; because they did not do clomipramine-to-desipramine calculation we were not able to test the validity of their results. Also, error bars were not reported for the RBFE calculations. For all these reasons, we repeated the RBFE calculations of TCAs using FEP method. Apart from clomipramine-to-imipramine and imipramine-to-desipramine perturbations, we also performed clomipramine-to-desipramine perturbation. Theoretically, the sum of all three calculated RBFEs should add up to zero, and this acts as an extra check on the calculations. Our results do support Singh et al. and Zhao et al. with regards to the binding trend of TCAs at LeuT and not Zhou et.al. The calculated RBEs were also in excellent agreement with the experimental numbers.

3.2. Methods

3.2.1. Relative binding energy calculations

We used FEP [26] method to calculate the RBEs of ligands with the protein. RBFE calculations are based on the thermodynamic cycle shown in Figure 3.3. Relative binding (ΔΔG) of two different ligands A and B to a protein P is estimated using equation 1.

$$\Delta\Delta G = \Delta G_4 - \Delta G_3 = \Delta G_2 - \Delta G_1 \qquad \text{- Equation 1}$$

ΔG_3 and ΔG_4 are the free energy changes of perturbing ligand A to B in the solvent and the solvated protein P, respectively. ΔG_1 and ΔG_2 are the ABEs of ligands A and B,

respectively, with a protein P. ΔG_3 and ΔG_4 are estimated using FEP method. In FEP method, free energy change (ΔG) between states A and B is calculated using equation 2.

$$\Delta G(A \rightarrow B) = G_B - G_A = -K_B T \ln \left\langle \exp\left[(U_B - U_A)/K_B T \right] \right\rangle_A \qquad \text{- Equation 2}$$

G_A and G_B are the free energies of states A and B, respectively; U_B and U_A are the potentials of states B and A, respectively; K_B is Boltzmann constant; T is the temperature.
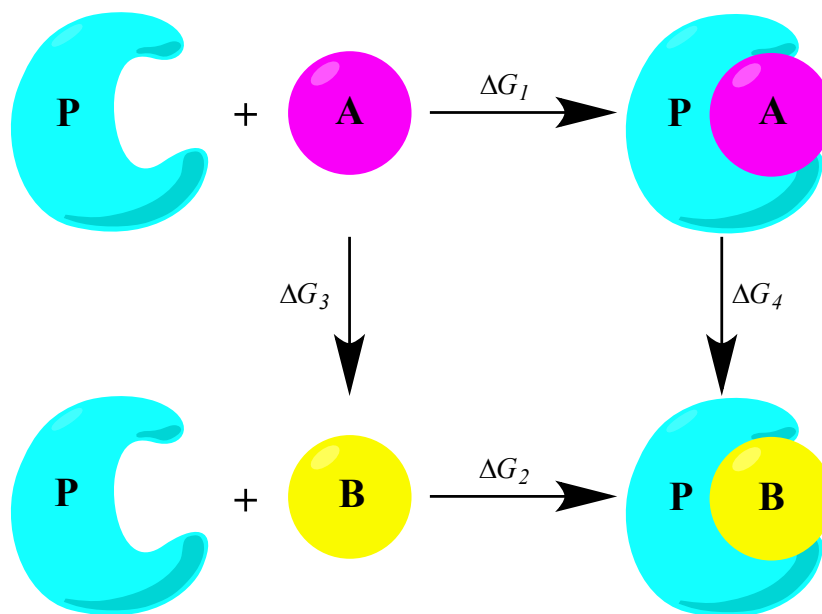
Figure 3.3. Thermodynamic cycle for computing relative binding energies. A and B are two different ligands and P is a protein.

Experimental RBEs ($\Delta\Delta G_{Expt}$) were estimated using either experimental K_i (in the case of amino acids) or IC_{50} values (in the case of TCAs) as shown in equations 3 and 4.

$$\Delta\Delta G_{Expt}(A \rightarrow B) = RT \ln\left(\frac{K_i^B}{K_i^A} \right) \qquad \text{- Equation 3}$$

$$\Delta\Delta G_{Expt}(A \rightarrow B) = RT \ln\left(\frac{IC_{50}^B}{IC_{50}^A} \right) \qquad \text{- Equation 4}$$

K_i^B and K_i^A are the inhibition constants of ligands B and A, respectively; IC_{50}^B and IC_{50}^A are the half maximal inhibitory concentrations of ligands B and A, respectively; R is the gas constant; T is the temperature.

3.2.2. System preparation and simulation details

All six crystal structures of LeuT with clomipramine (PDB ID: 2Q6H), imipramine (PDB ID: 2Q72), desipramine (PDB ID: 2QB4), leucine (PDB ID: 3F4J), alanine (PDB ID: 3F3E) and glycine (PDB ID: 3F48) were downloaded from Protein Data Bank (www.pdb.org) [27]. Crystal complexes were reprocessed using MOE. Appropriate protonation states were assigned for the residues at pH 7, and hydrogens were added. Further, they were solvated using VMD [28]. Final size of the simulation box was approximately 89.0Å X 103.0Å X 87.0Å in all six cases. The complexes were then minimized using NAMD 2.9 [29] for 50,000s steps to remove steric clashes. Systems were treated under periodic boundary condition. The minimized complexes were used for calculating ΔG_4. NAMD 2.9 was used to perform free energy calculations. For the calculation of ΔG_3, crystal conformations of the ligands were placed in the water boxes and minimized for 50,000s steps prior to calculation. Dual topology approach was followed for perturbing the ligands. Protein, water, (TIP3P model [30]) and ions [31] were treated using CHARMM force field. For TCA perturbations, CHARMM general force field (CGenFF) was used, [24] and in the case of amino acid mutations CHARMM force field parameters were used. Both ΔG_3 and ΔG_4 simulations were performed in forward and backward directions. This helps to assess the hysteresis and thereby improves the statistical precision in free energy calculation. Reliability and efficiency of

free energy calculations are enhanced using bidirectional simulations. Perturbations were carried out by just turning off the intermolecular interactions of the outgoing atoms with the environment and simultaneously turning on the intermolecular interactions of the incoming atoms. Bonded interactions and masses of the perturbed atoms were not altered in these perturbations. The transformation between initial and final states was divided into multiple windows to get sufficient overlap between configurational ensembles and to get reasonable accuracy [32]. Bennett acceptance ratio (BAR) method [33] was used to calculate the free energy differences by combining the forward and backward simulations. Analysis of the free energy simulations, binding energy and error bars was carried using ParseFEP [34] plugin in VMD. The error bars are statistical errors in the binding energy estimation. Probability density plots produced by ParseFEP plugin helped assess the convergence between subsequent windows. Windows with poor convergence were further split to get good convergence and to reduce the error in the calculation. The transformation between initial and target states is a function of a coupling parameter (λ). λ varies from 0 (initial state) to 1 (target state). End-point catastrophes were dealt with by introducing soft-core potentials [35]. Electrostatics and van der Waals interactions were turned off at λ=0.5 and λ=1.0, respectively. Simulations were carried out under NPT ensemble (temperature was 310K and pressure was 1 bar). A 2fs time step was used. Particle mesh Ewald (PME) method [36] was used to treat long-range electrostatics, whereas van der Waals and short-range electrostatics were cut off at 12Å.

3.2.2.1. Amino acid perturbations

For estimating ΔG_3 and ΔG_4, bidirectional simulations were carried, and the binding energy was estimated using BAR method. In both cases, a 23ns simulation was performed in each direction, and the simulation in each direction was split into 46 windows. In each window, the system was equilibrated for 0.1ns and data was collected for 0.4ns.

3.2.2.2. TCA perturbations

Similar to the amino acid perturbations, the TCA perturbations were carried bi-directionally, and a BAR estimator was used to calculate the binding energy.

3.2.2.2.1. Clomipramine-to-imipramine

For calculating ΔG_4, a 1.5ns simulation in each direction was carried, and the entire simulation was bifurcated into 293 windows. In each window, system was equilibrated for 0.001ns and data was collected for 0.004ns. For ΔG_3 calculation, in each direction a 1.3ns simulation was performed and 263 windows were used. In each window, 0.001ns equilibration and 0.004ns data were collected.

3.2.2.2.2. Imipramine-to-desipramine

In both ΔG_3 and ΔG_4 calculations, in each direction a 1.3ns simulation was performed and 263 windows were used. In each window, 0.001ns equilibration and 0.004ns data collection were carried.

3.2.2.2.3. Clomipramine-to-desipramine

In this perturbation, a 5.2ns simulation was carried for estimating ΔG_3 and ΔG_4 in either direction, and the simulation in each direction was split into 263 windows. In each window, 0.004ns equilibration and 0.016ns data collection were carried.

3.3. Results and Discussion

3.3.1. Relative binding of amino acid ligands leucine, alanine, and glycine with LeuT

As stated earlier, we initially validated the FEP method through alanine-to-glycine and leucine-to-alanine perturbations. Chemical structures of leucine, alanine and glycine are shown in Figure 3.4. In the alanine-to-glycine perturbation, the methyl side chain (highlighted in blue in Fig. 3.4B) attached to the alpha carbon was converted to a hydrogen atom (highlighted in pink in Fig. 3.4C). In the leucine-to-alanine perturbation, the isobutyl side chain (highlighted in red in Fig. 3.4A) was converted to a methyl group (highlighted in blue in Fig. 3.4B).
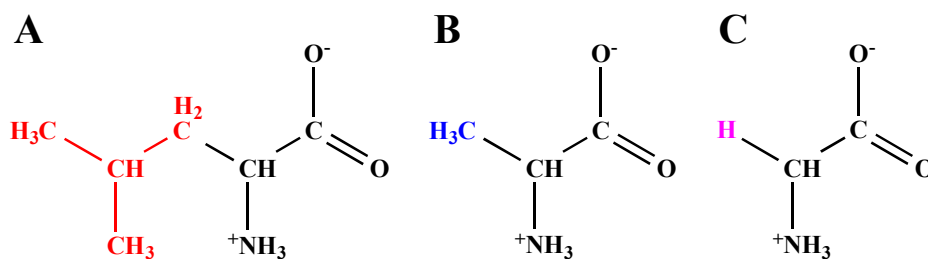
Figure 3.4. Chemical structures of leucine (A), alanine (B) and glycine (C). Structural differences are highlighted in red, blue, and pink.

ΔG_3 and ΔG_4 for the alanine-to-glycine perturbation were -5.00 ± 0.02 and -1.74 ± 0.03 kcal/mol, respectively. The calculated $\Delta\Delta G$ was 3.26 ± 0.05 kcal/mol, whereas the experimental $\Delta\Delta G$ was 2.85 ± 0.22 kcal/mol (Table 3.1). Simulations converged quickly and perfectly in the water box (Table 3.1 and Fig. 3.5A) unlike in the protein (Table 3.1 and Fig. 3.5B). The hysteresis between the forward and backward simulations in ΔG_3 calculation was minimal (0.05 kcal/mol, Fig. 3.5A); whereas in the case of ΔG_4 it was 0.84 kcal/mol (Fig. 3.5B). Hydrogen is more soluble compared to a methyl group in water, and our result (-5.00 kcal/mol) in mutating the alanine-to-glycine in water supports this fact. The impact of mutating the alanine-to-glycine in water (-5.00 ± 0.02 kcal/mol) was more compared to the protein (-1.74 ± 0.03 kcal/mol), and this shows that the methyl group prefers to stay in the protein compared to the solvent. This might be because pocket environment is hydrophobic compared to bulk solvent. The calculated $\Delta\Delta G$ of alanine-to-glycine was positive (3.26 ± 0.05 kcal/mol) and unfavorable (Fig. 3.5C and Table 3.1), which indicates that LeuT prefers to bind with alanine compared to glycine. This is due to the more hydrophobic nature of alanine compared to glycine. Since the pocket is predominantly hydrophobic, it favors alanine than glycine. Also, because of its size alanine fits better in the pocket than glycine and indulges in favorable van der Waals interactions with the pocket residues. The other important factor might be the desolvation penalty for the ligand to enter the hydrophobic pocket from the bulk solvent. Since glycine is less hydrophobic than alanine, it is more soluble in water compared to alanine and has to pay more penalty to enter the hydrophobic pocket of LeuT compared to alanine. The calculated $\Delta\Delta G$ was slightly overestimated compared to the experimental

ΔΔG; however, the calculated ΔΔG was within 0.41 kcal/mol of the experimental ΔΔG (Table 3.1).

ΔG_3 and ΔG_4 for the leucine-to-alanine transformation were -2.63 ± 0.03 kcal/mol and 0.71 ± 0.04 kcal/mol, respectively; whereas the calculated and experimental RBEs were 3.34 ± 0.07 and 3.17 ± 0.30 kcal/mol, respectively. Interestingly, simulations converged better in the case of ΔG_4 (Fig. 3.5E and Table 3.1) rather than ΔG_3 (Fig. 3.5D and Table 3.1). This might be because isobutyl to methyl perturbation is huge and this perturbation samples huge volume in the water box compared to the protein pocket as it is packed and needs more simulation time in the water box than in the protein to get reasonable convergence. The hysteresis between the forward and backward simulations in ΔG_3 calculation was 0.64 kcal/mol, whereas in the case of ΔG_4 it was 0.52 kcal/mol. Compared to an isobutyl group, a methyl group is less hydrophobic and expected to solubilize more in water, and our result (-2.63 ± 0.03 kcal/mol) supports this fact. The calculated ΔΔG was positive indicating that leucine binds stronger with LeuT compared to alanine. This is because leucine is a natural substrate of LeuT, and it is tightly packed compared to alanine in the protein pocket. Also, leucine is more hydrophobic compared to alanine, and LeuT pocket is predominantly hydrophobic; therefore, leucine can form more favorable van der Waals interactions compared to alanine with LeuT. Another factor might be the lower desolvation penalty of leucine compared to alanine to bind in the hydrophobic pocket. Since leucine is more hydrophobic than alanine, the desolvation penalty for leucine to enter the hydrophobic pocket should be lower compared to alanine. Similar to the alanine-to-glycine mutation, the calculated ΔΔG was slightly overestimated compared to the experimental ΔΔG, but it falls within the experimental error. This might

121

be because we used the crystal structure of LeuT:leucine for this calculation; when the crystal structures of alanine:LeuT and leucine:LeuT were compared, some of the pocket residues (F259 and I359) in the alanine:LeuT crystal structure moved into the pocket compared to the leucine:LeuT crystal structure to compensate for the smaller size of alanine[2]. Since we used the crystal structure of LeuT:Leucine for this calculation, the sampling time might not be sufficient for the leucine pocket in LeuT to adjust or collapse into alanine pocket. Therefore, alanine might not interact with the pocket as intended, making the perturbation more unfavorable than expected.

RBEs of leucine-to-alanine and alanine-to-glycine mutations were estimated earlier by Sergei Noskov as the difference of ABEs [37]. FEP method was used to calculate the ABEs in this paper. For the sake of efficiency, only atoms within 20Å of the ligand were treated explicitly in these calculations, and the rest of the system was treated using reduced generalized solvent boundary potential (GSBP). In this study only the crystal structure of LeuT:Leucine complex was used since LeuT:Alanine and LeuT:Glycine crystal structures were not available at the time of this publication. Therefore, they modeled LeuT:Alanine and LeuT:Glycine complexes using LeuT:Leucine crystal structure for the free energy calculations. Also, error bars for ABEs and RBEs (calculated and experimental) were not provided in Noskov's paper. In our RBFE calculations, we considered the crystal structures of all three amino acids with LeuT.

Table 3.1. Relative binding energies of amino acids leucine, alanine, and glycine with LeuT.

	ΔG_3			ΔG_4			$\Delta\Delta G_{Calc}$	$\Delta\Delta G_{Expt}$
	For	Back	BAR	For	Back	BAR		
Ala → Gly	-5.03	4.98	-5.00 ± 0.02	-1.29	2.13	-1.74 ± 0.03	3.26 ± 0.05	2.85 ± 0.22[a]
Leu → Ala	-3.01	2.37	-2.63 ± 0.03	0.92	-0.40	0.71 ± 0.04	3.34 ± 0.07	3.17 ± 0.30[b]

All energies are in kcal/mol. $\Delta\Delta G_{Calc}$ and $\Delta\Delta G_{Expt}$ are the calculated and experimental RBEs. Ala, Leu, and Gly refer to leucine, alanine, and glycine, respectively. For and Back refer to forward and backward simulations. BAR is free energy estimated through Bennett acceptance ratio method. All energies are in kcal/mol. [a,b]The experimental RBEs were calculated as the difference of experimental K_i values of the amino acids with LeuT [2].
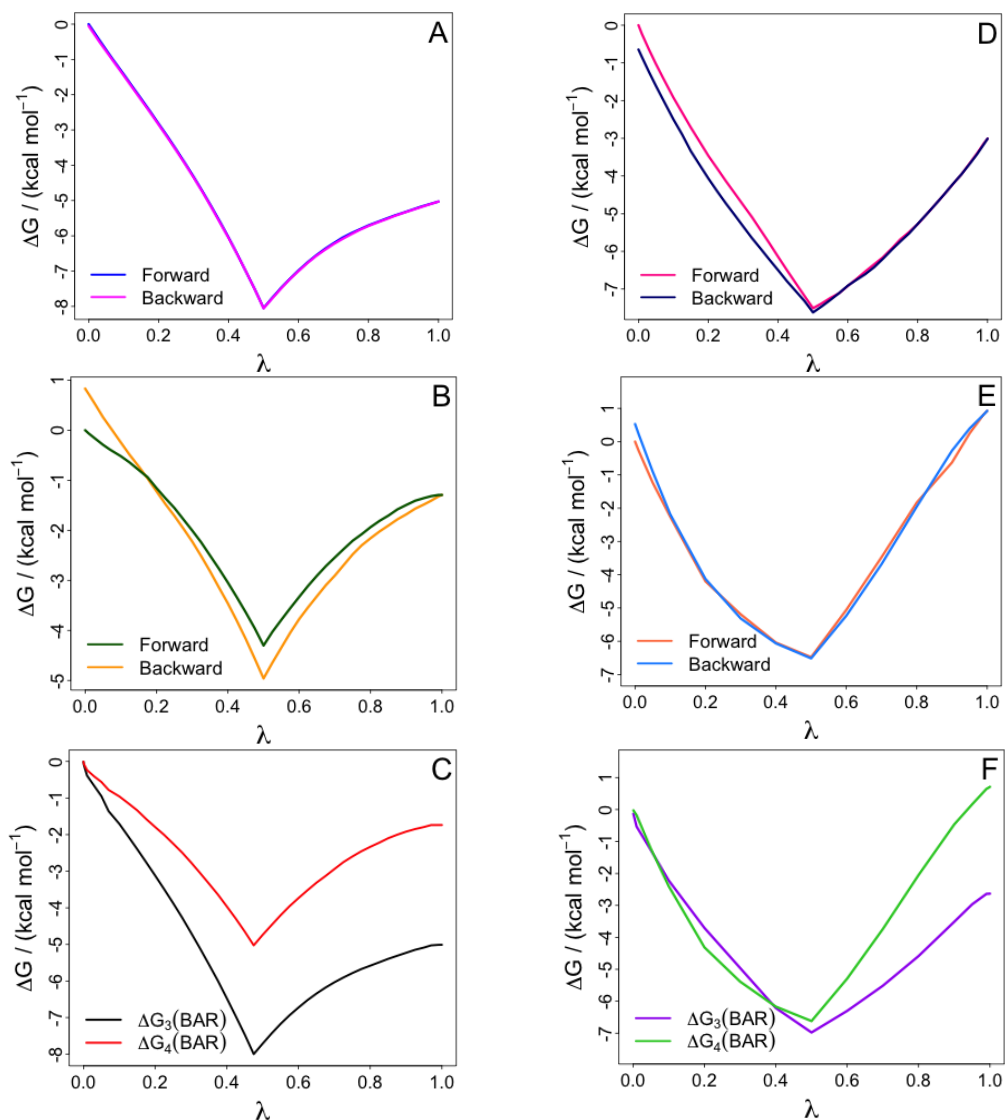
Figure 3.5. Convergence of the simulations in the amino acid ligand perturbations. Left and right columns represent the alanine-to-glycine and leucine-to-alanine transformations, respectively. A, B, and C are free energy changes vs. λ□□□□e□□□□□□□□□ ΔG_3, ΔG_4, and BAR values in the alanine-to-glycine transformation; D, E, and F are corresponding values for the leucine-to-alanine transformation.

The earlier reported RBFE for this transformation (estimated as the difference of ABEs) was 2.3 kcal/mol, whereas our calculated $\Delta\Delta G$ was 3.34 ± 0.07 kcal/mol, and the experimental $\Delta\Delta G$ was 3.17 ± 0.30 kcal/mol. The earlier reported $\Delta\Delta G$ for the alanine-to-

glycine mutation (3.2 kcal/mol) was in good agreement with our calculated $\Delta\Delta G$ (3.26 ± 0.05 kcal/mol) and the experimental $\Delta\Delta G$ (2.85 ± 0.22 kcal/mol). The experimental $\Delta\Delta G$s reported in Noskov's paper were slightly different from ours because they used experimental K_d values, and we considered experimental K_i values for calculating the experimental $\Delta\Delta G$s. K_d values for all three amino acid ligands were not available, but K_i values were.

Our results were not only in quantitative agreement with the experimental values, but they also reproduced the experimental binding trend, that is leucine binds stronger followed by alanine and glycine with LeuT. These calculations not only validated the FEP method, but also acted as a test for the CHARMM force field parameters of the amino acids. We used the ParseFEP [34] plugin in VMD [28] to analyze the free energy simulations and to calculate the BAR values. Convergence of the simulations was carefully monitored using the probability density plots generated by ParseFEP plugin.

3.3.2. Relative binding of clomipramine, imipramine and desipramine with LeuT

Next, we calculated the RBEs of TCAs with LeuT using the FEP method and compared RBEs to the experimental numbers. Since experimental K_d or K_i values were not available, we used IC_{50} values to calculate the experimental $\Delta\Delta G$s. Dual topology approach was adopted for TCA calculations as well. The three TCAs are structurally very similar, and the differences are highlighted in Figure 3.6. We performed three perturbations calculations; clomipramine-to-imipramine, imipramine-to-desipramine, and clomipramine-to-desipramine.
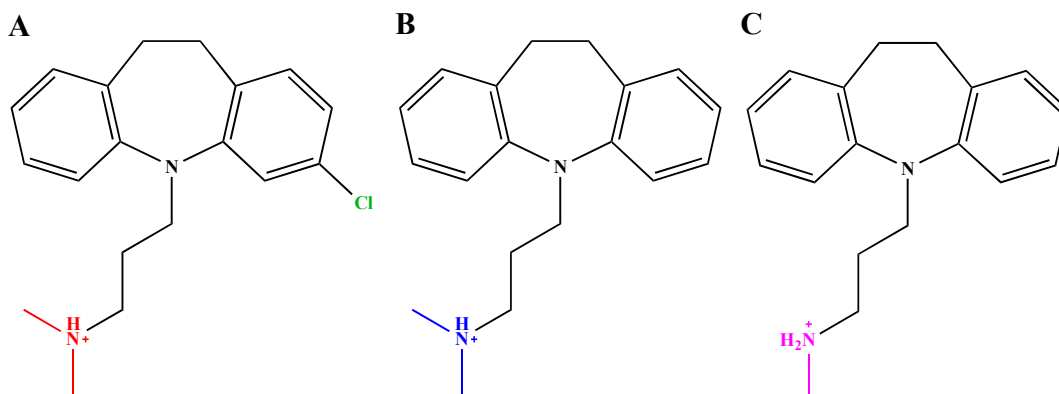
Figure 3.6. Structures of (A) clomipramine, (B) imipramine, and (C) desipramine. Differences in structures are highlighted in green, red, blue, and pink.

In the clomipramine-to-imipramine perturbation, the chlorine atom of clomipramine (colored green in Fig. 3.6A) was mutated to a hydrogen atom; in the imipramine-to-desipramine perturbation, the methyl group attached to the protonated nitrogen (colored blue in the Fig. 3.6B) was mutated to a hydrogen atom; in the clomipramine-to-desipramine mutation, the chlorine atom attached to the tricyclic rings (colored green in Fig. 3.6A) and the methyl group attached to the protonated nitrogen (colored red in the Fig. 3.6A) were mutated to hydrogens at respective positions. ΔG_3 and ΔG_4 for the clomipramine-to-imipramine perturbation were 0.40 ± 0.11 and 1.42 ± 0.07, respectively. Simulations reasonably converged in both ΔG_3 (Fig. 3.7A) and ΔG_4 (Fig. 3.7B) calculations. The hystereses in ΔG_3 and ΔG_4 calculations were minimal (0.08 kcal/mol in either case) (Table 3.2).

Table 3.2. Relative binding energies of the three TCAs with LeuT estimated using FEP method.

	ΔG_3			ΔG_4			$\Delta\Delta G_{Calc}$	$\Delta\Delta G_{Expt}$
	For	Back	BAR	For	Back	BAR		
Clo → Imi	0.37	-0.29	0.40 ± 0.11	1.26	-1.18	1.42 ± 0.07	1.02 ± 0.18	1.26 ± 0.17[a]
Imi → Des	-2.78	3.48	-3.47 ± 0.12	-1.44	2.67	-2.20 ± 0.15	1.27 ± 0.27	>0.0[b]
Clo → Des	-2.29	2.42	-2.77 ± 0.10	0.07	1.22	-0.56 ± 0.12	2.21 ± 0.22	>1.26[c]

All energies are in kcal/mol. $\Delta\Delta G_{Calc}$ and $\Delta\Delta G_{Expt}$ are the calculated and experimental RBEs. Clo, Imi, and Des refer to clomipramine, imipramine, and desipramine, respectively. [a]The experimental RBFE of clomipramine-to-imipramine mutation [3]; [b,c]The experimental RBEs of imipramine-to-desipramine and clomipramine-to-desipramine mutations [3, 25, 38]. The error bars associated with the calculated RBEs are the statistical errors in binding energy estimation.

The calculated ΔΔG was 1.02 ± 0.18 kcal/mol (Fig. 3.7C and Table 3.2), and the experimental ΔΔG was 1.26 ± 0.17 kcal/mol. The positive value of ΔΔG indicates that the clomipramine-to-imipramine perturbation was unfavorable, which means clomipramine binds stronger with LeuT compared to the imipramine. The chlorine of clomipramine indulges in a polar interaction with the side chain amide of glu34. Imipramine lacking this chlorine cannot form this interaction and that partially explains the stronger potency of clomipramine compared to imipramine with LeuT [3]. Zhao et al. estimated ΔΔG for this perturbation through the ABFE and RBFE calculations, which were 0.2 kcal/mol and 0.4 kcal/mol, respectively [25].

In the case of the imipramine-to-desipramine mutation, ΔG_3 and ΔG_4 were estimated as -3.47 ± 0.12 and -2.20 ± 0.15 kcal/mol, respectively. As shown in Figures 3.7D and 3.7E, simulations were well converged in both ΔG_3 and ΔG_4 calculations. Hysteresis between the forward and backward simulations in ΔG_3 calculation was smaller (0.7 kcal/mol) compared to the ΔG_4 calculation (1.23 kcal/mol) (Table 3.2).

Our result (ΔG_3 = -3.47 ± 0.12 kcal/mol) supports the fact that imipramine with an extra methyl moiety is less soluble compared to desipramine in water. The calculated ΔΔG was 1.27 ± 0.27 kcal/mol, which indicates that the perturbation is unfavorable and also that imipramine is a strong binder with LeuT compared to desipramine. We cannot quantitatively compare our calculated ΔΔG to the experimental value since the IC_{50} of desipramine with LeuT was not reported by Singh et al. However, it was reported that desipramine was a less potent inhibitor of LeuT compared to imipramine, and our result supports it[25, 38]. The calculated ΔΔGs of Zhao et al. for this mutation were 1.3 kcal/mol (through ABFE calculations) and 1.0 kcal/mol (through RBFE calculations)

[25], which were in good agreement with our calculated $\Delta\Delta G$ (1.27 ± 0.27 kcal/mol). Imipramine because of the extra methyl group is more hydrophobic compared to desipramine and can form more favorable van der Waals interactions with the pocket residues (F320 and A319) than desipramine. This partially explains the increased potency of imipramine compared to desipramine. Also, the desolvation penalty required to enter the pocket from the bulk should be less in the case of imipramine compared to desipramine because imipramine is more hydrophobic than desipramine.

Finally, we estimated the $\Delta\Delta G$ of clomipramine-to-desipramine perturbation. ΔG_3 and ΔG_4 for this calculation were -2.77 ± 0.10 and -0.56 ± 0.12 kcal/mol, respectively. The hysteresis between the forward and backward calculation for estimating ΔG_3 and ΔG_4 were 0.13 and 1.29 kcal/mol, respectively (Table 3.2). Simulations were reasonably converged in both ΔG_3 and ΔG_4 calculations (Fig. 3.7G,H). This calculation allows us to check the reliability and consistency of our RBFE calculations. Theoretically, the calculated $\Delta\Delta G$ of clomipramine-to-desipramine perturbation should be equal to the sum of $\Delta\Delta G$s of clomipramine-to-imipramine and imipramine-to-desipramine perturbations. The calculated $\Delta\Delta G$ (2.21 ± 0.22 kcal/mol) was in good agreement with the sum of $\Delta\Delta G$s of clomipramine-to-imipramine and imipramine-to-desipramine perturbations (2.29 ± 0.45 kcal/mol), and this shows the consistency of RBFE methodology. Also, the sum of ΔG_3 (-3.07 kcal/mol) and ΔG_4 (-0.78 kcal/mol) values for the clomipramine-to-imipramine and imipramine-to-desipramine perturbations were in good agreement with the direct clomipramine-to-desipramine perturbation (-2.77 and -0.56 kcal/mol). The positive $\Delta\Delta G$ indicates that the transformation was unfavorable, and as expected,

clomipramine was a potent inhibitor compared to desipramine. Since the IC_{50} of desipramine was not available, we were not able to compare quantitatively.

However, since the experimental $\Delta\Delta G$ of clomipramine-to-desipramine perturbation is equal to the sum of experimental $\Delta\Delta G$s of clomipramine-to-imipramine and imipramine-to-desipramine perturbations, $\Delta\Delta G_{Expt}$ for clomipramine-to-desipramine should be greater than 1.26 kcal/mol. $\Delta\Delta G_{Expt}$ was in good agreement with our $\Delta\Delta G_{Calc}$ (2.21 ± 0.22 kcal/mol). Since Zhao et al. did not perform this calculation, we cannot compare the calculated $\Delta\Delta G\square$ in this case. Unlike desipramine, clomipramine has an extra chlorine atom on the aromatic ring and can form a polar interaction with the side chain amide of glu34. Clomipramine also has an extra methyl group attached to the tail and can form favorable van der Waals interactions with the pocket residues. These reasons might explain the stronger potency of clomipramine compared to the desipramine.

Theoretically, the sum of $\Delta\Delta G$s of clomipramine-to-imipramine and imipramine-to-desipramine perturbations is quantitatively equal and opposite in sign to the $\Delta\Delta G$ of clomipramine-to-desipramine perturbation; therefore, the sum of all three $\Delta\Delta G$s should add up to zero as shown below. The sum of all three calculated $\Delta\Delta G$s was 0.08 kcal/mol and proves the reliability of our calculations.

$$\Delta\Delta G^{Clo \text{ -> } Imi} + \Delta\Delta G^{Imi \text{ -> } Desi} + (-\Delta\Delta G^{Clo \text{ -> } Desi}) = 0$$

$$1.02 + 1.27 - 2.21 = 0.08$$
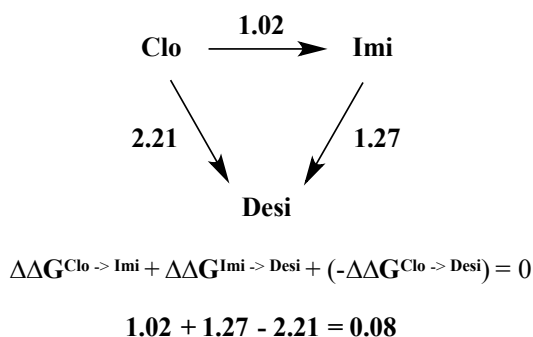
Overall, our results indicate that the order of binding of TCAs with LeuT is clomipramine followed by imipramine and desipramine.

3.4. Conclusions

We estimated the relative binding affinities of TCAs clomipramine, imipramine and desipramine with LeuT through relative binding energy calculations. We used the FEP method to calculate the relative binding affinities. Our results indicate that clomipramine binds stronger than imipramine, followed by desipramine. This binding trend was in agreement with the experimental binding trend of Singh et al. Zhao et al. also supported this binding trend through binding energy calculations. Also, the calculated relative binding energies were in excellent agreement with the experimental binding energies. We used the ParseFEP tool in VMD to assist in accurately estimating the binding energies. ParseFEP facilitated the analysis of FEP calculations and application of good practices in free energy calculations. We initially tested the ability of relative binding energy method to reproduce the experimental binding energies through leucine-to-alanine and alanine-to-glycine transformations with LeuT. The calculated relative binding energies in both the cases were within 0.4 kcal/mol of the experimental values. Our results demonstrate that, provided good force field parameters and adequate sampling, reasonable accuracy in estimation of binding free energies is achievable.
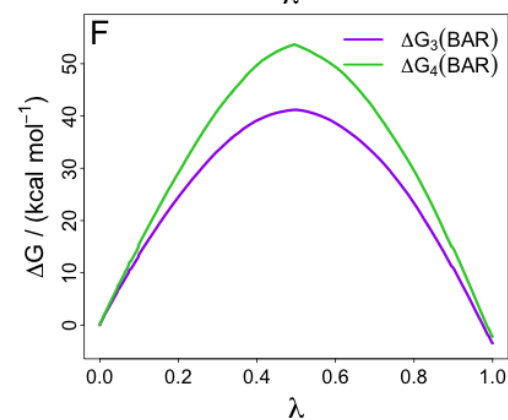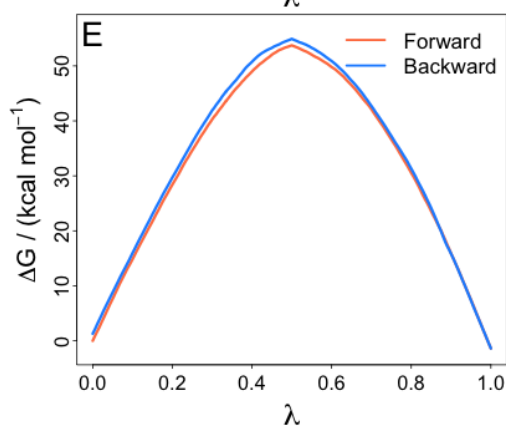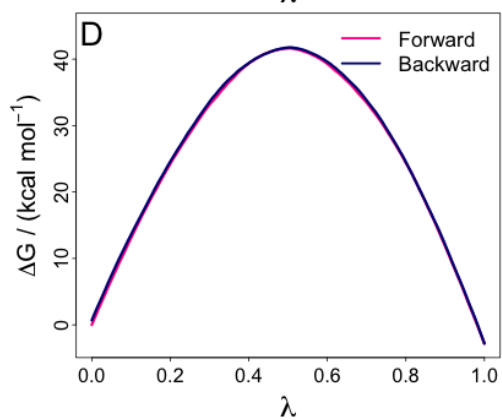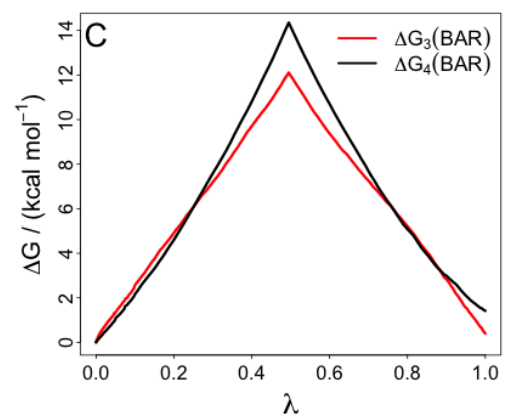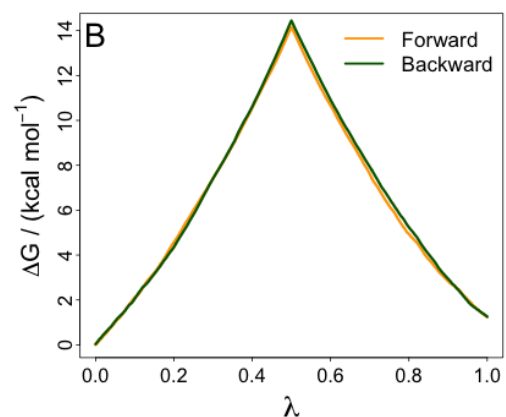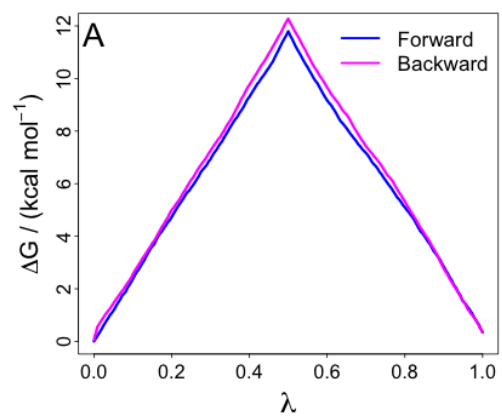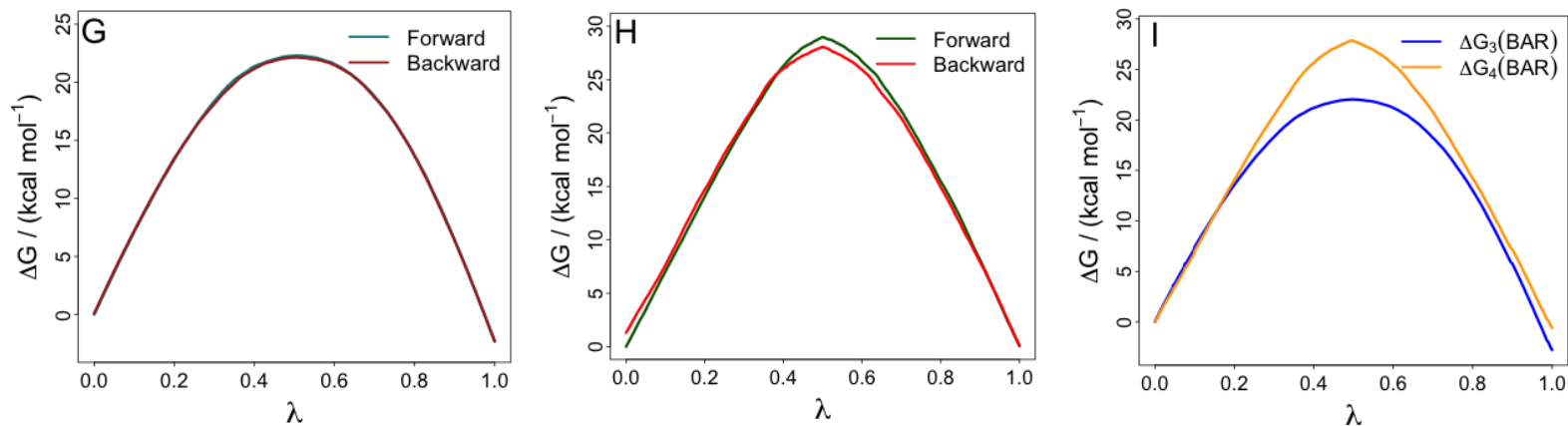
Figure 3.7. Convergence of free energy simulations in the clo-to-imi, imi-to-desi and clo-to-desi transformations. Clo, imi, and desi refer to clomipramine, imipramine, and desipramine, respectively. A, B, and C are simulations of ΔG_3, ΔG_4, and BAR values in clo-to-imi estimation; D, E, and F are simulations of ΔG_3, ΔG_4, and BAR values in imi-to-desi perturbation; G, H, and I are simulations of ΔG_3, ΔG_4, and BAR values in clo-to-desi calculations.

3.5. References

1.   Yamashita, A., S.K. Singh, T. Kawate, Y. Jin, and E. Gouaux, *Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters.* Nature, 2005. **437**(7056): p. 215-223.

2.   Singh, S.K., C.L. Piscitelli, A. Yamashita, and E. Gouaux, *A competitive inhibitor traps LeuT in an open-to-out conformation.* Science, 2008. **322**(5908): p. 1655-1661.

3.   Singh, S.K., A. Yamashita, and E. Gouaux, *Antidepressant binding site in a bacterial homologue of neurotransmitter transporters.* Nature, 2007. **448**(7156): p. 952-956.

4.   Zhou, Z., J. Zhen, N.K. Karpowich, R.M. Goetz, C.J. Law, M.E.A. Reith, and D.-N. Wang, *LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake.* Science, 2007. **317**(5843): p. 1390-1393.

5.   Zhou, Z., J. Zhen, N.K. Karpowich, C.J. Law, M.E.A. Reith, and D.-N. Wang, *Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures.* Nat Struct Mol Biol, 2009. **16**(6): p. 652-657.

6.   Immadisetty, K., L.M. Geffert, C.K. Surratt, and J.D. Madura, *New design strategies for antidepressant drugs.* Expert Opin Drug Discov, 2013. **8**(11): p. 1399-414.

7.   Gouaux, E., *Review. The molecular logic of sodium-coupled neurotransmitter transporters.* Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1514): p. 149-54.

8.   Piscitelli, C.L., H. Krishnamurthy, and E. Gouaux, *Neurotransmitter/sodium symporter orthologue LeuT has a single high-affinity substrate site.* Nature, 2010. **468**(7327): p. 1129-32.

9.   Shi, L., M. Quick, Y. Zhao, H. Weinstein, and J.A. Javitch, *The mechanism of a neurotransmitter:sodium symporter—inward release of Na+ and substrate is triggered by substrate in a second binding site.* Molecular Cell, 2008. **30**(6): p. 667-677.

10.   Manepalli, S., C.K. Surratt, J.D. Madura, and T.L. Nolan, *Monoamine transporter structure, function, dynamics, and drug discovery: a computational perspective.* AAPS J, 2012. **14**(4): p. 820-31.

11.   Koldsø, H., K. Severinsen, T.T. Tran, L. Celik, H.H. Jensen, O. Wiborg, B. Schiøtt, and S. Sinning, *The two enantiomers of citalopram bind to the human*

serotonin transporter in reversed orientations. Journal of the American Chemical Society, 2010. **132**(4): p. 1311-1322.

12.   Kristensen, A.S., J. Andersen, T.N. Jorgensen, L. Sorensen, J. Eriksen, C.J. Loland, K. Stromgaard, and U. Gether, *SLC6 neurotransmitter transporters: structure, function, and regulation.* Pharmacol Rev, 2011. **63**(3): p. 585-640.

13.   Wang, H., A. Goehring, K.H. Wang, A. Penmatsa, R. Ressler, and E. Gouaux, *Structural basis for action by diverse antidepressants on biogenic amine transporters.* Nature, 2013. **503**(7474): p. 141-5.

14.   Andersen, J., N. Stuhr-Hansen, L.G. Zachariassen, H. Koldso, B. Schiott, K. Stromgaard, and A.S. Kristensen, *Molecular basis for selective serotonin reuptake inhibition by the antidepressant agent fluoxetine (Prozac).* Mol Pharmacol, 2014. **85**(5): p. 703-14.

15.   Jorgensen, W.L., *Efficient drug lead discovery and optimization.* Acc Chem Res, 2009. **42**(6): p. 724-33.

16.   Simonson, T., *Free energy calculations. Computational Biochemistry and Biophysics*, ed. M.A. Becker OM, Roux B, Watanabe M 2001: Marcel Dekker, Inc.

17.   Tembre, B.L. and J.A. Mc Cammon, *Ligand-receptor interactions.* Computers & Chemistry, 1984. **8**(4): p. 281-283.

18.   Warshel, A., F. Sussman, and G. King, *Free energy of charges in solvated proteins: microscopic calculations using a reversible charging process.* Biochemistry, 1986. **25**(26): p. 8368-8372.

19.   Shirts, M.R., D.L. Mobley, and J.D. Chodera, *Chapter 4 Alchemical free energy calculations: ready for prime time?*, in *Annual Reports in Computational Chemistry*, D.C. Spellmeyer and R. Wheeler, Editors. 2007, Elsevier. p. 41-59.

20.   Bash, P., U. Singh, F. Brown, R. Langridge, and P. Kollman, *Calculation of the relative change in binding free energy of a protein-inhibitor complex.* Science, 1987. **235**(4788): p. 574-576.

21.   Mackerell, A.D., *Empirical force fields for biological macromolecules: overview and issues.* Journal of Computational Chemistry, 2004. **25**(13): p. 1584-1604.

22.   Guvench, O. and A.D. MacKerell, *Comparison of protein force fields for molecular dynamics simulations molecular modeling of proteins*, A. Kukol, Editor. 2008, Humana Press. p. 63-88.

23.   Wang, J., R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case, *Development and testing of a general amber force field.* J Comput Chem, 2004. **25**(9): p. 1157-74.

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

25.   Zhao, C., D.A. Caplan, and S.Y. Noskov, *Evaluations of the absolute and relative free energies for antidepressant binding to the amino acid membrane transporter LeuT with free energy simulations.* Journal of Chemical Theory and Computation, 2010. **6**(6): p. 1900-1914.

26.   Zwanzig, R.W., *High‐ temperature equation of state by a perturbation method. I. Nonpolar gases.* The Journal of Chemical Physics, 1954. **22**(8): p. 1420-1426.

27.   Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *The protein data bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

28.   Humphrey, W., A. Dalke, and K. Schulten, *Visual molecular dynamics.* Journal of Molecular Graphics, 1996. **14**: p. 33-38.

29.   Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten, *Scalable molecular dynamics with NAMD.* Journal of Computational Chemistry, 2005. **26**(16): p. 1781-1802.

30.   Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, *Comparison of simple potential functions for simulating liquid water.* The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.

31.   MacKerell, A.D., D. Bashford, Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* The Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

32.   Lu, N., D.A. Kofke, and T.B. Woolf, *Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods.* Journal of Computational Chemistry, 2004. **25**(1): p. 28-40.

33.   Bennett, C.H., *Efficient estimation of free energy differences from monte carlo data.* Journal of Computational Physics, 1976. **22**(2): p. 245-268.

34.   Liu, P., F. Dehez, W. Cai, and C. Chipot, *A toolkit for the analysis of free-energy perturbation calculations.* Journal of Chemical Theory and Computation, 2012. **8**(8): p. 2606-2616.

35.   Beutler, T.C., A.E. Mark, R.C. van Schaik, P.R. Gerber, and W.F. van Gunsteren, *Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations.* Chemical Physics Letters, 1994. **222**(6): p. 529-539.

36.   Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An N⊕log(N) method for Ewald sums in large systems.* The Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.

37.   Noskov, S.Y., *Molecular mechanism of substrate specificity in the bacterial neutral amino acid transporter LeuT.* Proteins, 2008. **73**(4): p. 851-63.

38.   Zdravkovic, I., C. Zhao, B. Lev, J.E. Cuervo, and S.Y. Noskov, *Atomistic models of ion and solute transport by the sodium-dependent secondary active transporters.* Biochim Biophys Acta, 2012. **1818**(2): p. 337-47.

4. CHAPTER 4

MODELING THE BINDING OF INHIBITORS DJLDU-3-79 AND SSA-426 IN THE HUMAN SEROTONIN TRANSPORTER

4.1. Introduction

Monoamine transporters (MATs) such as the human serotonin transporter (hSERT), the human norepinephrine transporter (hNET), and the human dopamine transporter (hDAT) are membrane proteins whose function is to reuptake serotonin, norepinephrine and dopamine from the synapse into the presynaptic neuron, thereby maintaining homeostasis in the body. MATs play a key role in modulating sleeps, appetite, reward, fear, sexual drive, and motivation [1]. MATs are implicated in several CNS disorders such as depression, substance abuse and addiction, attention deficit hyperactivity disorder (ADHD), orthostatic hypertension, obsessive-compulsive disorder, Parkinson's disease, and schizophrenia [2-7]. MATs belongs to the NSS family [8] and use sodium gradient to drive the neurotransmitters across the membrane [9]. hSERT requires symport of one Na^+, one Cl^- and antiport of one K^+ ion to transport one substrate molecule across the membrane [10, 11].

Depression is a major psychological disorder affecting 121 million people around the world, and in the USA alone approximately 30 million people are suffering every year [12]. About one million people commit suicide yearly because of depression [13]. Already depression is the second-highest cause of DALYs (Disability Adjusted Life Years) in the age group of 15-44 today [14]. According to the world health organization,

depression is projected as the second largest health disorder by 2020 [13]. Health care costs per person diagnosed with depression in the USA was projected as \$23,000 per year in 2012 [12]. Despite the seriousness of the condition, there is no sophisticated treatment for depression. Although several drugs are available for treating depression, they have several limitations such as adverse effects, a latency phase of 3-4 weeks and a high percentage of non-respondents (approximately 30%) [15]. This necessitates the development of novel medication to treat depression.

hSERT is the primary target for several antidepressants and has been the target of antidepressant drug discovery for over five decades. Several of the currently prescribed antidepressants primarily act at hSERT, including tricyclic antidepressants (TCAs), selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs). Since the first class of antidepressants (e.g., TCAs) act on adrenergic, muscarinic acetylcholine and histamine receptors along with hSERT and hNET, thereby causing several adverse effects [16, 17], more selective antidepressants popularly known as SSRIs were discovered [18]. Although SSRIs are better than TCAs, they are not completely free of adverse effects. For example, SSRIs are known for causing insomnia or hypersomnia, somnolence, weight gain, gastrointestinal disturbances, cardiovascular problems and sexual dysfunction. These adverse effects are the result of secondary interactions of these drugs with several serotonin receptors. SSRIs are the most effective and highly prescribed antidepressants to date in the market and act by selectively inhibiting the hSERT. However, the molecular basis for their selectivity and efficacy, binding site location and drug binding mode is unknown and is highly debated [19]. Although several critical residues for antidepressant selectivity and affinity were

identified through mutagenesis experiments [20-24], non-availability of the crystal structure of hSERT hampers the development of better therapeutics for treating depression.

Knowledge of the binding site and interaction of the drugs with hSERT prove critical for the potential development of future therapeutic agents. However, there is no clarity with regards to the binding site and the binding mode of antidepressants in hSERT. Although hSERT has been extensively studied with several experimental and computational methods, several predictions remain controversial because of the non-availability of crystal structure of hSERT. For example, several studies proposed that antidepressants bind with high-affinity to the hSERT S1 pocket (Fig. 4.1) and the method of inhibition was competitive [20, 22-36]. This was further supported by studies showing that antidepressants depend on ions to bind with hSERT [37, 38]. On the contrary, there is a second hypothesis that has evidence to show that antidepressants bind in the S2 pocket (Fig. 4.1) of hSERT and inhibit the protein in a noncompetitive manner [39-41]. The third hypothesis was that antidepressants bind to both S1 and S2 pockets in hSERT and inhibit the transporter by allosterism [42-46]. Lately, studies focused on the inhibitor ibogaine predict that it binds at a site other than S1 and S2 in hSERT [47].

To better understand the structural and functional aspects of hSERT and to gain insight into the molecular basis of drug affinity, selectivity and the inhibition process, computer models of hSERT were developed based on leucine transporter (LeuT) as a template [20, 23, 34, 35, 48-50]. LeuT is a bacterial membrane protein from *Aquifex aeolicus,* belongs to the NSS family and shares 20-25% sequence identity with hSERT [51].
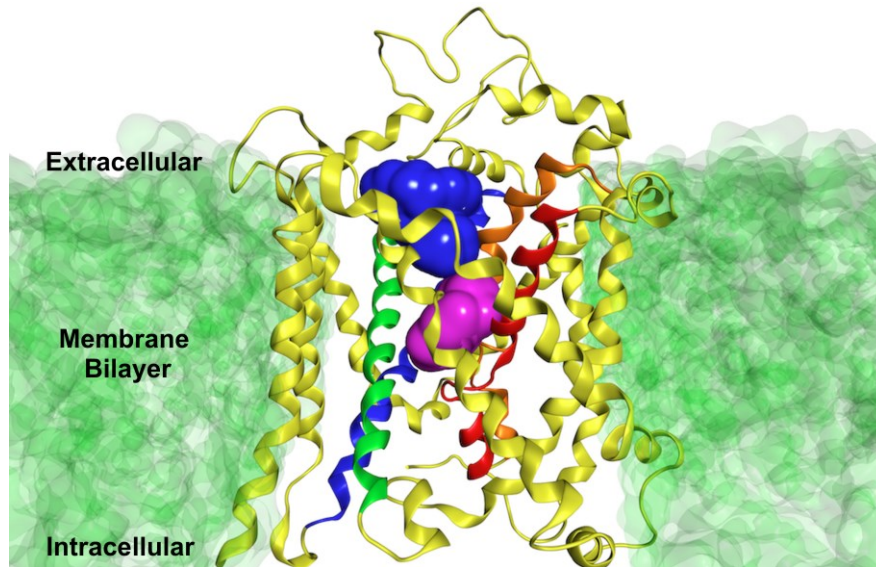
Figure 4.1. Serotonin transporter embedded in a membrane bilayer. The S1 or primary binding site (colored pink) is located halfway through the protein, and the S2 or secondary binding site (colored blue) is positioned towards the extracellular side. Transmembrane (TM) helices 1, 3, 6, and 8 are colored orange, green, red, and blue, respectively; remaining TM helices are colored yellow. The membrane bilayer is colored green.

LeuT was the first member of the NSS family that has been crystallized, and, to date, over 38 crystal structures of LeuT are available with various amino acids [52, 53], TCAs [41, 54] and SSRIs [39]. Several aspects of LeuT include the following: binding pockets of various substrates and antidepressants, binding modes, mechanism of binding, transport and inhibition, different conformation states of protein, key structural determinants for ligand binding, and transport and inhibition. These aspects were extensively studied using various pharmacological, biophysical and computational experiments. With regards to the binding sites, similar to hSERT, LeuT has two distinguished binding pockets, S1 and S2. Substrate leucine [53] and other amino acids like alanine, glycine and tryptophan [52] bind in the S1 site, whereas antidepressants

TCAs [41, 54] and SSRIs [39] bind in the S2 site of LeuT. However, there is controversy with regards to the substrate-binding site in LeuT. Shi et al. proposed that there are two high-affinity substrate binding sites in LeuT (i.e., S1 and S2), and leucine occupies both the S1 and S2 sites simultaneously and substrate in the S2 site allosterically triggers the release of substrate from S1 site into the intracellular side [55]. However, this was later contradicted by Piscitelli et al. through various studies, which concluded that LeuT has only one high-affinity substrate-binding site (i.e., S1) [56]. Computer models of hSERT developed using LeuT were useful in guiding site-directed mutagenesis experiments [23], structure-based drug design to identify novel lead compounds [49, 57], exploration of binding sites [58] and for understanding the substrate translocation mechanism [59].

In view of this ambiguity and significance of hSERT in depression treatment, we explored the binding sites and binding modes of two antidepressant lead compounds SSA-426 and DJLDU-3-79 [40] in hSERT using a computational protocol involving active site detection, docking, scoring, MD simulations and ABFE calculations. Active site detection is a process of identifying the binding sites in a protein. Binding pockets are the regions where drugs bind and interact with the protein. Therefore, identifying binding pockets on proteins is key for the rational design of novel therapeutics in the structure-based drug design. Docking is a process of generating all possible conformations of the ligand in the binding site of the protein. Scoring functions play a key role in ranking compounds in the computational drug screening approaches and also to rank various poses of a compound in docking simulations. Today in drug discovery, docking and scoring protocols are playing a significant role in lead identification and optimization. We utilized a unique protocol called AADS, which stands for automated active site

detection, docking and scoring [60], to identify the active site and bioactive conformation of the ligand. Further, we employed MD simulations and ABFE calculations to confirm or validate AADS results, thereby identifying the most probable bioactive conformation of the ligand and its binding site in the protein. We used the FEP technique to calculate ABEs. Identifying the binding mode allows optimizing the inhibitor for better affinity and improving selectivity.

We initially validated the AADS protocol on five different LeuT crystal structures, where it successfully identified the active sites and reproduced the bioactive conformation of the ligands with reasonable accuracy. In all five cases, the active site was among the top 10 cavities identified, and the bioactive conformation of the ligand was within the top five AADS poses, except in the case of 2A65, where it was ranked 8th. Identification of a ligand pose is nearly superimposable on the bioactive conformation within the top 10 docking poses, which improved our confidence in the scoring function. These results prove that it is applicable to huge membrane proteins such as LeuT and hSERT. We also validated the ABFE method with LeuBAT:Clomipramine crystal complex [36] (PDB ID: 4MMA). The calculated binding affinity (-7.69 ± 0.34 kcal/mol) correlated well with the experimental binding value (-8.25 ± 0.05 kcal/mol).

Finally, we applied this protocol to the hSERT homology model to explore the binding sites and the binding modes of antidepressant lead compounds SSA-426 and DJLDU-3-79. We used the homology model of hSERT developed by Manepalli et al. for this purpose [49]. SSA-426 is a dual hSERT and 5-HT$_{1A}$-receptor antagonist and possesses a higher affinity for hSERT (K$_i$ = 2.34 ± 0.59nM). DJLDU-3-79 is our in-house designed antidepressant lead compound and a molecular hybrid of SSA-426 and MI-17,

which is a virtual screening hit compound. The binding affinity (K_i) of DJLDU-3-79 with hSERT is $37 \pm 4nM$ [48]. Since DJDLU-3-79 and SSA-426 are structurally similar, Nolan et al. proposed that both these inhibitors bind in the same pocket (i.e., S2) and in a similar orientation in hSERT [48]. However, our results do not correlate with the earlier predictions. Instead, we observed that these two inhibitors bind in the S1 site of hSERT, but in a similar orientation as proposed by Nolan et.al. We also identified a pose of DJLDU-3-79 binding in the S2 site, which is stable in the MD simulations and energetically less favorable than the pose binding in the S1 site. We propose this as the metastable pose and the S2 site as the low energy binding metastable site of DJLDU-3-79 in hSERT. Our results also indicate that it is not always wise to use one static structure of the protein for docking purposes, as it restricts the sampling of ligand conformations in the protein.

4.2. Materials and Methods

4.2.1. Computational protocol

The computational protocol we followed for identifying the binding site and binding mode of the inhibitors is shown in Figure 4.2. The first step of our protocol was to develop the hSERT homology model since the crystal structure is not yet available. However, we used the homology model of hSERT developed by Manepalli et al. [48] for this study. In the next step, binding sites in the protein were identified to dock the ligands. Then we docked the ligands and ranked the poses via scoring functions. For active site identification, docking and scoring purposes we used a web server, known as AADS [60]. We used AADS and molecular operating environment (MOE) software to

144

obtain docked poses [61]. The poses of interest were selected and MD simulations were performed to test the stability and also to understand the various interactions playing a key role in the binding of a ligand in the pocket. The last step is computing binding energies of various poses using the ABFE method. The ABFE method depends on the FEP technique; although it is computationally demanding, it is rigorous and accurate compared to scoring functions. Other advantages of the FEP method over scoring functions are the inclusion of protein flexibility and explicit water in the simulations. The pose whose calculated binding affinity matches with the experimental binding energy was considered the most probable experimental binding pose, and the site in which it is binding is considered the possible binding pocket of the ligand in hSERT.

4.2.1.1.   Automated active site identification and scoring protocol (AADS)

AADS [60] is a robust automated protocol for active site detection, docking and scoring in proteins. It is three individual computational tools tied together to make the process easier. AADS methodology can be divided into three steps. In step 1, when a nascent protein is given to the AADS server, it identifies all the cavities in the protein and ranks them based on physiochemical properties of the functional groups circling the cavities. In step 2, the ligand of interest is docked into the top 10 ranked cavities and generates ≈1000 conformations and finally saves eight low energy docked poses at each site. In step 3, all the 80 low energy docked poses corresponding to the 10 sites are energy minimized and ranked using a free energy scoring function and outputs ten best-ranked

complexes. Expectations of AADS are that it should predict the active site within the top 10 cavities and bioactive conformation of the ligand within the top 5 poses.
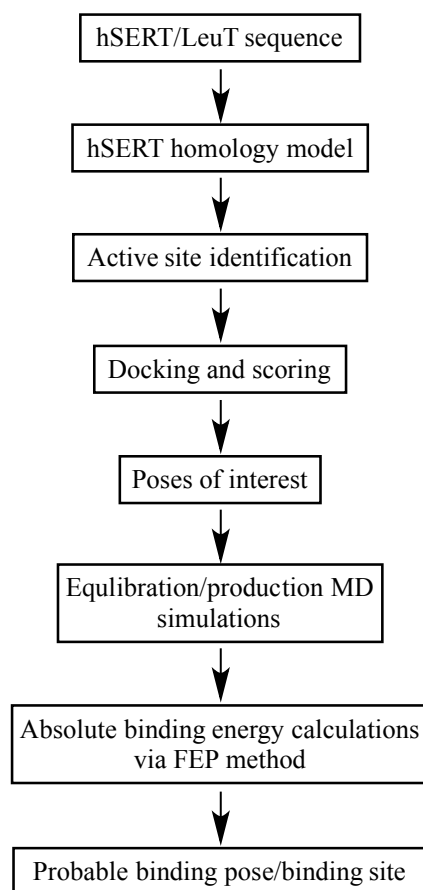
```
┌─────────────────────────┐
│   hSERT/LeuT sequence   │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│  hSERT homology model   │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Active site identification │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Docking and scoring   │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    Poses of interest    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Equlibration/production MD │
│       simulations       │
└─────────────────────────┘
            │
            ▼
┌──────────────────────────────┐
│ Absolute binding energy calculations │
│        via FEP method        │
└──────────────────────────────┘
            │
            ▼
┌──────────────────────────────┐
│ Probable binding pose/binding site │
└──────────────────────────────┘
```

Figure 4.2. Computational protocol for elucidating the binding site and binding mode of inhibitors.

4.2.1.1.1.   Active site finder

Active site finder [60] is a geometric-based algorithm and typically works in 3 steps. In the first-step, average coordinates of the grid points in the 4Å cluster are determined. A 4Å cluster is a collection of grid points within a 4Å radius and encircled by protein atoms from two sides. In step 2, the average points surrounded by greater than 150 atoms of

protein are identified and clustered into 10Å radii, and a reference point is generated to represent the 10Å cavity position in the protein. Finally, all the cavities are ranked using equation 1. In scoring the cavities, physicochemical properties of the residues lining the cavities such as the number of hydrogen bond donors, acceptors, aromatic rings, and hydrophobic atoms were considered along with the volume of the cavity.

$$Score_C = \frac{\left(\sum_{P=1..N} \left(\frac{X_{PC}}{X_P^{\max}} \right) \right)}{N}$$
- Equation 1

$Score_C$ is the score of cavity C; N is the total number of properties considered in calculating the score and in this case it is 5; P refers to properties such as the hydrogen bond donor, acceptor, aromatic rings, hydrophobic atoms and volume of the cavity; X_{PC} is the value of the property P in the cavity C; $X_P^{\max}$ is the maximum value of that property in the protein. It is believed that the higher the score of a cavity, the higher the chances of that cavity being the active site in the protein.

4.2.1.1.2.   Docking

Docking is carried through a program called Pardock (Parallel Dock). Docking of the candidate molecule at the top 10 sites and generation of conformations occurs in 5 stages [62]. In stage 1, the protein and the candidate molecule are prepared for docking calculations. Input conformation of the ligand is considered for docking purposes. Hydrogens are added to the ligand using AMBER [48] and an appropriate ionization state is assigned. Then the ligand is geometrically optimized through the AM1 procedure, and partial charges are derived through the AM1-BCC procedure [63]. Atom types, bond angle, dihedral and van der Waals parameters for the ligand are assigned using the GAFF

force field [64]. In stage 2, ligand is translated to the top 10 cavities identified by the active site finder. In stage 3, a cubic grid of 1Å resolution is generated around the protein, and the translational points with less number of clashes are identified for docking. In stage 4, at each of the ten binding sites around 1000 conformations are generated using a six-dimensional rigid body Monte Carlo methodology. All the conformations are ranked through a scoring function, which calculates the interaction energy of ligand conformations with a protein based on equation 2.

$$E = \sum \left(E_{ele} + E_{vdw} + E_{hpb} \right)$$

- Equation 2

E is the interaction energy between the ligand atoms and proteins atoms, and Σ is the sum over all ligand atoms; E_{ele}, E_{vdw} and E_{hpb} are the electrostatic, van der Waals and hydrophobic components.

In stage 5, the eight lowest energy conformations at each binding site are collected for the next stage of the calculation.

4.2.1.1.3. Scoring function

The eighty lowest energy conformations collected in stage 5 of the docking are energy minimized through AMBER [48] and ranked based on an empirical scoring function called BAPPL [65]. Change in binding energy (ΔG) is calculated using equation 3.

$$\Delta G = \alpha \left(E_{ele} \right) + \beta \left(E_{vdw} \right) + \left(\sum_{A=1}^{22} \sigma_A \Delta A_{LSA} \right) + \lambda \left(\Delta S_{CR} \right) + \delta$$

- Equation 3

E_{ele} and E_{vdw} are the electrostatic and van der Waals components; ΔA_{LSA} is the loss in surface area; σ_A is the atomic desolvation parameter; ΔS_{CR} is the loss in conformational

entropy. α, β, and λ are the empirical coefficients for electrostatics, van der Waals and conformational entropy, respectively; δ is a constant.

4.2.1.2. AADS validation on LeuT crystal structures

We tested the AADS capability to explore binding sites and to retrieve the bioactive conformation of the ligand on five different LeuT crystal structures. All the five crystal structures (2A65, 2Q72, 3F3A, 3GWU, 3GWV) of LeuT were downloaded from the Protein Data Bank (www.pdb.org). Water, ions, and ligands were removed from the crystal structures using MOE software. The protonation states were assigned using the protonate 3D capability in MOE and hydrogens were added. All the crystal structures were then saved in standard pdb format and given as input to the AADS. Initially, AADS identified all possible cavities in the protein and ranked them using the scoring function as described earlier. AADS output consisted of all the cavities identified in the order of their ranks along with the residues forming the cavities, coordinates of the atoms surrounding the cavities, cavity points and volume of the cavities. Then the ligand of interest was provided to the AADS in standard pdb format and the total charge of the ligand was specified. Ligands were initially built in MOE using the molecule builder utility, protonation states were assigned using protonate 3D, hydrogens were added, energy minimized, and finally saved as pdb. The top 10 poses of the ligand in complex with the protein along with their binding energies (in kcal/mol) were output by AADS.

The AADS results were summarized in Table 4.1. Cavities of LeuT identified by AADS were validated by comparing them against the pockets in the crystal structures. To identify a ligand pose closest to the bioactive conformation, all the ten AADS ligand

149

poses were superposed onto the crystal conformation of the ligand and calculated the RMSD.

4.2.1.3.   Application of AADS to hSERT

AADS was applied to the hSERT homology model to explore the binding sites and the binding modes of SSA-426 and DJLDU-3-79. The development of hSERT homology model was previously described [66]. The rest of the procedure is the same as explained in section 1.2.1.2.

4.2.1.4.   hSERT docking using MOE docking protocol

Induced-fit docking protocol in MOE was used to dock the ligands in the hSERT. Initially, the protein-ligand complex for docking was prepared in MOE. Ligand coordinates were generated using MOE. Two Na^+ ions and one Cl^- ion were placed at appropriate positions in the protein. Appropriate protonation states for the protein residues were assigned using the protonate 3D capability in MOE. The CHARMM force field was used to assign hydrogens and to add partial charges to the protein atoms, ligand atoms, and ions. The alpha site-finder in MOE was used to define the docking site in the protein. Wall constraints (radius 8Å) were placed around the docking site to restrict the docking poses. The proxy triangle placement method in combination with the London dG scoring function was employed to dock the ligands. Further, all the poses were refined via the force field refinement scheme and ranked through the affinity dG function.

4.2.1.5. System preparation and protocol for MD simulations

Protein:Ligand complexes resulting from the AADS server were reprocessed using MOE. Two Na^+ ions and one Cl^- ion were placed at appropriate places in the protein, and the titratable residues were protonated appropriately at pH 7. The system prepared in MOE was hydrated using VMD [67]. The size of the simulation box was 90Å X 100Å X 96Å. The final system contains one hSERT protein, one ligand, 24,428 waters, two sodium ions and one chloride ion, to create a total of 81,764 atoms. Molecular dynamics simulations were performed using NAMD [68] in periodic boundary conditions. The system was initially energy minimized for 50,000 steps to remove steric clashes. Following minimization, systems were equilibrated in NPT ensemble (temperature = 310K and pressure = 1 bar) for 2ns. A 2fs time step was used; van der Waals and short-range electrostatics were cut off at 12Å. Long-range electrostatics were treated using the particle mesh Ewald (PME) method [69]. Following equilibration, each system was simulated for 50ns, including 2ns equilibration. Trajectories were saved every 0.5ps. The CHARMM force field [70] was used for protein, water (TIP3P model [71]) and ions. Inhibitors were treated with the CHARMM general force field (CGenFF) [72]. Poses 4, 5, and M of DJLDU-3-79, and pose-1 of SSA-426 were simulated using this procedure. The RMSDs of the protein backbone and the ligand non-hydrogen atoms were calculated using an RMSD trajectory tool in VMD.

4.2.1.6. Absolute binding energy calculations

We used the double annihilation approach to compute ABFE. The double annihilation approach depends on the thermodynamic cycle shown in Figure 4.3, and free energy

change (ΔG) was computed using the FEP [73] principle. In the FEP method, ΔG between the initial and the final states (0 and 1) is calculated using equation 4.

$$\Delta G = G_1 - G_0 = -K_B T \left\langle \exp\left[(U_1 - U_0)/K_B T\right]\right\rangle_0 \qquad \text{- Equation 4}$$

ΔG is the change in Gibbs free energy; G_1 and G_0 are the free energies of initial and final states 0 and 1, respectively; K_B is the Boltzmann constant; T is the temperature; U_1 and U_0 are the potentials of states 1 and 0, respectively.

In the double annihilation approach, ABFE (ΔG_{Abs}) of a ligand (L) with a protein (P) is computed as

$$\Delta G_{Abs} = \Delta G_2 - \Delta G_1 \qquad \text{- Equation 5}$$

ΔG_2 and ΔG_1 are the free energy changes involved in annihilating the ligand L from the solvent and the solvated protein P.

To prevent the ligand from escaping the binding site while calculating ΔG_1, positional restraints were imposed on the ligand and the bias introduced by the restraints was estimated using equation 6 [74, 75].

$$\Delta G^{Restraint} = -RT \ln\left(\frac{Vol^{eff}}{Vol^0}\right) \qquad \text{- Equation 6}$$

$\Delta G^{Restraint}$ is the restraint free energy; R is the gas constant; T is the temperature. Vol^{eff} is the effective volume, which is the volume accessible by the ligand in the pocket in the presence of the restraints; Vol^0 is the volume accessible by the ligand at standard concentration of 1M, which is equal to 1660 Å^3. Vol^{eff} in turn was calculated using equation 7.

$$Vol^{eff} = \left(\frac{2\Pi RT}{K}\right)^{3/2} \qquad \text{- Equation 7}$$

K is the force constant, which was used to restrain a ligand in the binding site of a protein. $\Delta G^{Restraint}$ estimated through equation 6 was added to ΔG_{Abs} calculated through equation 5 to compensate for the restraints, and the final expression for calculating ABFE is shown in equation 8.

$$\Delta G_{Abs} = \Delta G_2 - \Delta G_1 + \Delta G^{Restraint}$$
- Equation 8
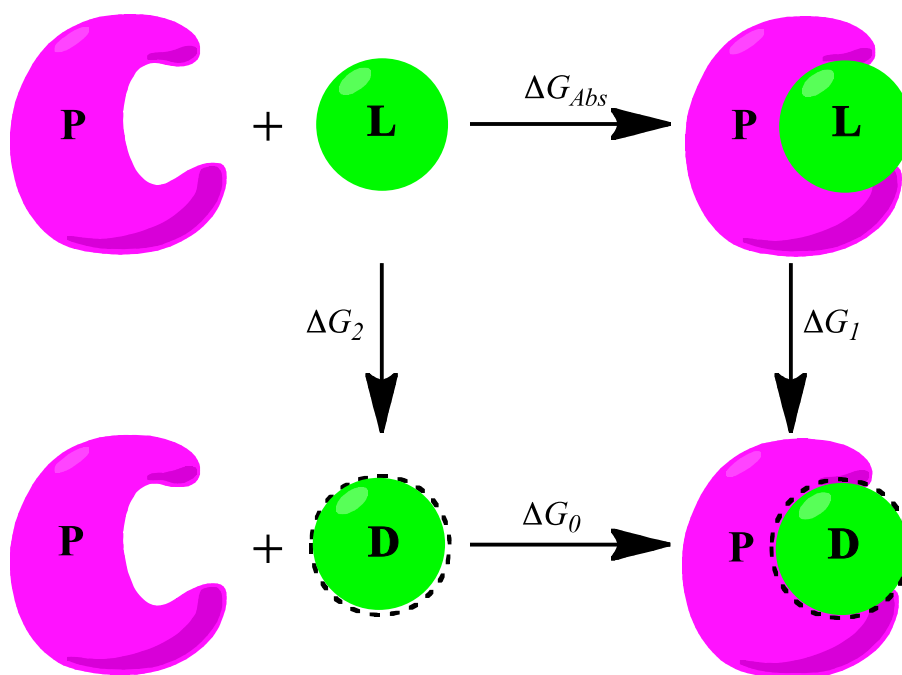
Figure 4.3. Double annihilation approach for estimating absolute binding energy. P, L, and D refer to protein, ligand, and dummy, respectively. ΔG_{Abs} is the absolute binding energy of the L-P complex. ΔG_2 and ΔG_1 are the free energy changes of perturbing L to D in the solvent and in the P, respectively. ΔG_0 is the absolute binding energy of the D-P complex, which is equal to zero.

The experimental binding energy (ΔG_{Exp}) was estimated from the experimental inhibition constant (K_i) using equation 9.

$$\Delta G_{Exp} = RT \ln K_i$$
- Equation 9

R is the gas constant; T is the temperature; K_i is the inhibition constant.

4.2.1.6.1.   System preparation and simulation details

In the case of hSERT, final structures of the protein-ligand complexes from MD simulations were used for estimating ΔG_1. For the estimation of ΔG_2, ligand conformation from the MD simulation was placed in the water box and energy minimized for 50,000 steps prior to decoupling with the solvent. The LeuBAT:Clomipramine complex for ΔG_1 estimation, as well as clomipramine in the water box for ΔG_2 calculation, were prepared following the similar procedure as in the case of hSERT. The LeuBAT:Clomipramine complex and the clomipramine in the water box were energy minimized for 50,000 steps prior to decoupling calculations. Free energy calculations were performed using NAMD. A coupling parameter (λ) was introduced, and the transformation between the initial (0) and final states (1) is a function of λ. All simulations were carried in both forward and backward directions to monitor hysteresis and to improve the statistical precision. Free energy differences were estimated using the Bennett acceptance ratio (BAR) method [76]. Bidirectional simulations enhance the reliability and efficiency of the free energy estimation. The ParseFEP [77] plugin in VMD was used to analyze the free energy simulations and to compute the binding energy along with error bars. The error bars represent the statistical error in the binding energy calculation. Convergence of the binding energies due to the insufficient sampling is a bigger problem in the FEP method. Therefore, the entire transformation was divided into several windows, to ensure sufficient overlap between configurational ensembles representing the reference and target states, and to get reasonable accuracy [78]. Probability density plots generated using the ParseFEP plugin were used to assess the

convergence between subsequent windows, and to optimize them to reduce the statistical as well as systematic errors in the calculation. A soft-core potential [79] was introduced to handle the end-point catastrophes. Electrostatics were turned off from $\lambda=0$ to $\lambda=0.5$ and van der Waals were turned off from $\lambda=0$ to $\lambda=1.0$. A 2fs time step was used in all the calculations. The CHARMM force field for the protein, water (TIP3P model) and ions [70], and CGenFF [72] parameters for the inhibitors (obtained from the ParamChem website) were used. No restraints were placed on the ligands in the calculation of ΔG_2. All simulations were run under NPT ensemble (temperature = 310K and pressure = 1 bar).

LeuBAT:Clomipramine

ΔG_2 calculation: In each direction, a 10ns simulation was carried, and the total transformation was divided into 263 windows. A 0.008ns equilibration and a 0.03ns data collection were carried in each window.

ΔG_1 calculation: A 1.0ns forward and a 1.0ns backward simulation were performed. The total calculation was split into 150 windows. A force constant of 0.1 kcal/mol.Å^2 was used to confine clomipramine to the binding site of LeuBAT. In each window, the system was equilibrated for 0.0022ns and data was collected for 0.0044ns.

hSERT:SSA-426 (Pose-1)

ΔG_2 calculation: A 4.5ns simulation in each direction was performed, and we used 150 windows in either direction. In each window, a 0.005ns equilibration followed by a 0.025ns data collection was performed.

ΔG_1 calculation: The simulation time and number of windows were similar to the ΔG_2 calculation. The ligand was restricted in the binding site using a force constant of 0.5 kcal/mol.Å^2. The system was equilibrated for 0.01ns in each window followed by a 0.02ns data collection.

hSERT:DJLDU-3-79 (Poses M and 5)

ΔG_2 calculation: A 15.6ns simulation was performed in each direction, and the total simulation was split into 130 windows in either direction. In each window, the system was equilibrated for 0.024ns and data was collected for 0.096ns.

ΔG_1 calculation: In each direction a 4.5ns simulation was performed and 150 windows were used. The ligand was restricted in the binding site with a force constant of 0.5 kcal/mol.Å^2. In every window, a 0.01ns equilibration and a 0.02ns data collection was carried out.

4.3. Results and Discussion

4.3.1. Validation of AADS protocol with LeuT crystal structures

After careful inspection, we selected five high-resolution LeuT crystal structures to test the AADS method. Details of all the five crystal structures were provided in Table 4.1. We included all relevant variables into the data set. Since the goal was to test the ability of AADS in identifying the active site and reproducing the bioactive conformation, we included two well-established binding sites, S1 and S2, into the dataset. Similar to hSERT, the S1 pocket of LeuT is located halfway through the protein and harbors substrate leucine and inhibitors such as tryptophan; the S2 pocket is located 11Å above

the S1 on the extracellular side of the protein and shelters antidepressants such as imipramine, R-fluoxetine and sertraline. To see the impact of protein conformation in identifying the binding site, the two identified conformational states, open outside (C_o) and occluded (O_{cc}), were included in the dataset. In C_o, the transporter is open toward the extracellular side, whereas in the case of O_{cc}, the transporter is obscured both from the extracellular and intracellular side. Since our goal is to see where and how potential antidepressants bind in hSERT, we included two different classes of antidepressants, such as TCAs (e.g., Imipramine) and SSRIs (e.g., Sertraline and R-fluoxetine) into the dataset, along with inhibitor tryptophan and substrate leucine. The AADS performed reasonably well in identifying the binding pockets and bioactive conformations of ligands in all the crystal structures.

4.3.1.1. Binding site identification

AADS successfully identified the active sites in all five LeuT structures, although active site rank varies. In the case of 2Q72, 3GWU, and 3GWV, the active site rank was one; whereas it was nine in 2A65 and 3F3A. Interestingly, in all the three crystal structures (2Q72, 3GWU and 3GWV) ligand binds in the S2 pocket, and was ranked one by AADS in all three cases. In the case of 2A65 and 3F3A, ligand binds in the S1 pocket, and it was ranked nine by AADS in both cases. This might be because the S2 pocket is huge compared to the S1 pocket; moreover, it is located on the extracellular side, making it readily accessible to the ligands. Whereas the S1 pocket is small compared to the S2 site, and it is located deep in the protein making it inaccessible to the ligands. Therefore, it is no surprise to see that irrespective of the conformation of the protein, S2 is the highest

ranked site in all five of the crystal structures. The total number of pockets identified in each structure by AADS was shown in Table 4.1. The difference in the total pockets among five structures was due to the repeated identification of the same pockets in some structures. The conservation of S1 and S2 pockets in all five of the crystal structures shows the functional importance of those sites in LeuT. The results demonstrated that AADS was successful in identifying the active sites irrespective of the conformation of the protein and location of the binding sites. It is worth mentioning that all the active sites were within the top 10 sites identified by AADS. Since AADS docks the ligand only in the top 10 binding sites to identify the bioactive conformation, it is crucial that the active site is within the top 10 binding sites.

4.3.1.1. Binding site validation

After successful identification of the active sites in LeuT by AADS, we validated the AADS pockets by comparing them to the pockets in the crystal structures. Of all the AADS identified binding pockets in LeuT, only two pockets were well established, S1 and S2. These two pockets were found in all five of the crystal structures, thereby attributing druggability and functional relevance.

The S1 pocket is the harbor for the substrate leucine and the inhibitor tryptophan. We validated the S1 pocket identified by AADS by comparing it to the S1 pocket of 2A65. 2A65 is a crystal structure of LeuT with substrate leucine bound in the S1 pocket. LeuT attains occluded conformation by accommodating leucine. In Figure 4.4A, we superposed cavity-9 identified by AADS onto 2A65 with leucine bound in the S1 pocket. Cavity-9 of AADS correlates well with the S1 pocket of 2A65 (Fig. 4.4). This validates

the S1 pocket of LeuT identified by AADS. We validated the S2 pocket by comparing it

against the S2 pocket of 2Q72. 2Q72 is a crystal structure of LeuT with the

noncompetitive inhibitor imipramine bound to the S2 pocket. LeuT was stabilized in an

occluded conformation by the imipramine. In Figure 4.4B, we superposed cavity-1 of the

AADS onto 2Q72 with the imipramine bound to the S2 pocket. Cavity-1 of AADS

matches well with the S2 pocket of 2Q72. The correlation between AADS cavities and

crystal pockets demonstrate the ability of AADS to detect the active sites. Further, we

tested the AADS capability to retrieve the bioactive conformation of a ligand.
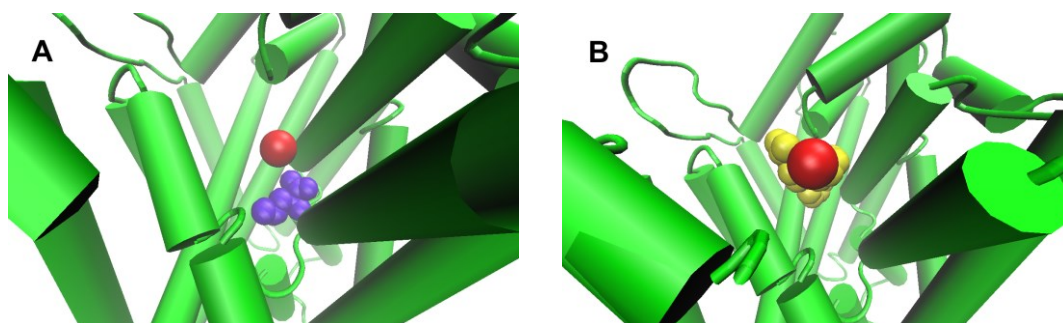
Figure 4.4. AADS cavity points vs. S1 and S2 sites in the LeuT crystal structures.
A is the S1 site of 2A65 with leucine (colored blue) bound, and B is the S2 site of
2Q72 with imipramine (colored yellow) bound. Cavity points are represented as red
spheres in both A and B. Cavity point nine matches with the S1 site of 2A65 and
cavity point one matches with the S2 site of 2Q72.

Table 4.1. Validation of AADS protocol on different LeuT crystal structures.

Pdb code	Ligand	Resolution	Conformation	Binding site	Site rank	T.N.P	RMSD	Pose rank
2A65	Leucine	1.65	O_{cc}	S1	9	30	3.35	8
2Q72	Imipramine	1.70	O_{cc}	S2	1	31	2.47	2
3F3A	Tryptophan	2.00	C_o	S1	9	28	2.75	1
3GWU	Sertraline	2.14	O_{cc}	S2	1	29	3.75	4
3GWV	R-fluoxetine	2.35	O_{cc}	S2	1	29	1.93	3

O_{cc} and C_o refer to occluded and outward facing conformations of LeuT; T.N.P refers to total number of pockets identified by AADS; site rank is the rank of the active site identified by AADS; pose rank is rank of the AADS pose that is closest to the bioactive conformation of the ligand.

4.3.1.2. Pose validation

Identification of the bioactive conformation of ligands is a crucial aspect in drug discovery. Knowledge of the bioactive conformation of ligands helps in the rational design of new drugs. Several efficient docking tools are available to reproduce the bioactive conformation of ligands, but they lack an efficient scoring function. We tested the AADS docking and scoring protocol for its robustness and accuracy in reproducing the bioactive conformation.

We self docked ligands leucine, imipramine, tryptophan, sertraline and R-fluoxetine in the top 10 binding sites identified by AADS in the respective crystal structures. The top 10 docking poses of each ligand were analyzed for the closest match to the bioactive conformation. In all 5 cases, AADS was able to retrieve the bioactive conformation of the ligands. Except in the case of 2A65, the bioactive conformation was among the top 5 poses. As shown in Table 4.1, pose ranks were 8, 2, 1, 4 and 3 in 2A65, 2Q72, 3F3A, 3GWU, and 3GWV. From the ranking of poses, it was clear that AADS performed equally well in retrieving the bioactive conformations in both the S1 and S2 pockets of LeuT, except in the case of 2A65. Although ligand binds in the S1 pocket in both 2A65 and 3F3A, the difference in pose ranking might be because 2A65 was in occluded conformation making the pocket more compact and inaccessible unlike 3F3A, where the protein was in open to out conformation. Still, AADS performed reasonably well in reproducing the bioactive conformation of the ligand in 2A65. The RMSD between the closest AADS pose and the pose in the crystal structure was considerably

low (<3.8Å) as shown in Table 4.1. Comparison of the closest AADS poses with the

poses from the crystal structures are shown in Figure 4.5.
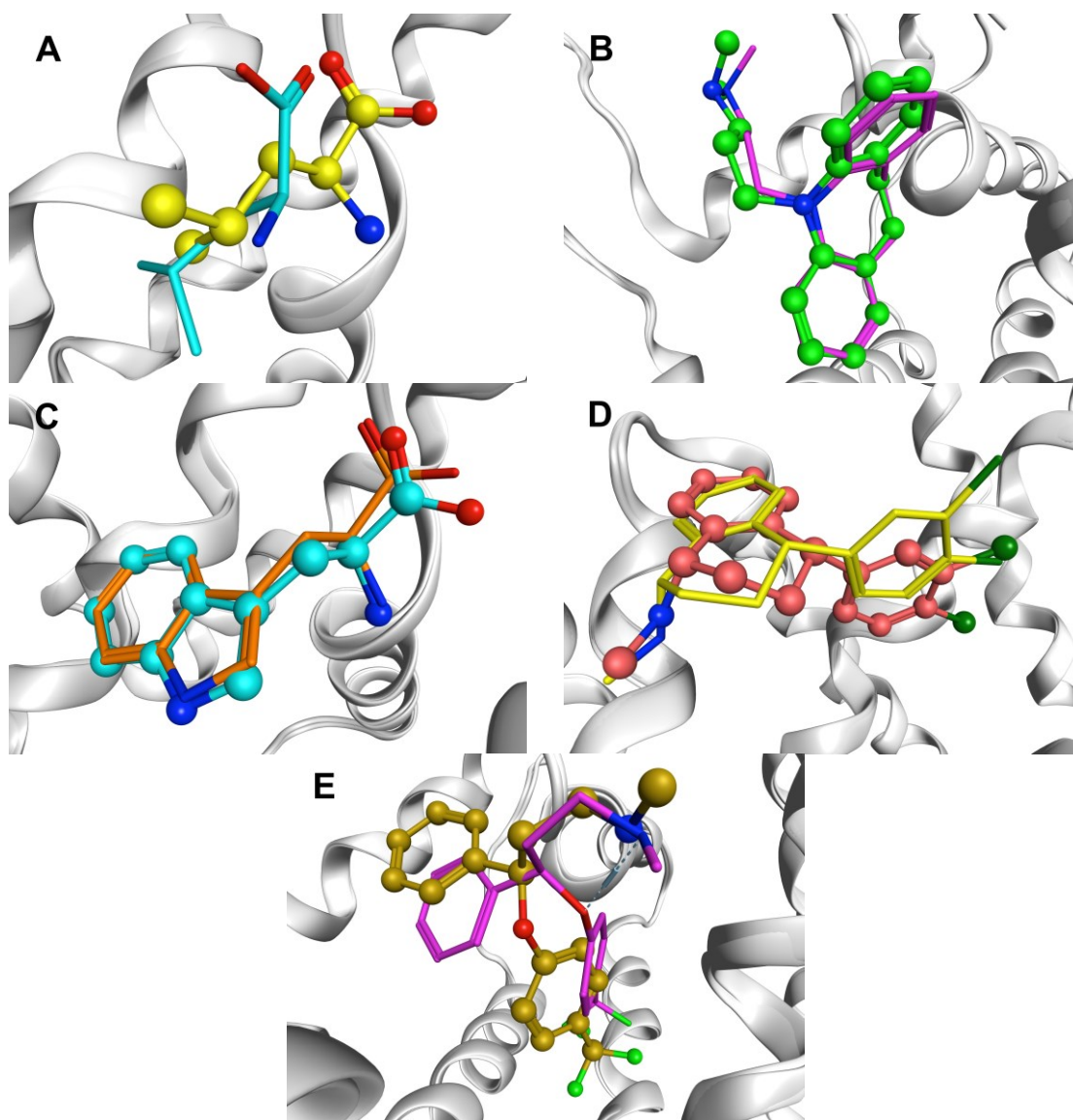
Figure 4.5. Closest AADS pose vs. conformation of ligands in LeuT crystal structures. (A) Leucine (2A65), (B) Imipramine (2Q72), (C) Tryptophan (3F3A), (D) Sertraline (3GWU), and (E) R-fluoxetine (3GWV). The crystal poses are depicted as balls and sticks and the AADS poses are depicted as sticks.

It is evident from the above results that AADS performed reasonably well in the active site identification and in retrieving the native conformations of the ligands. AADS identified the active sites in 2A65, 2Q72, 3F3A, 3GWU, and 3GWV, and also reproduced the crystal conformation of the ligands with reasonable accuracy. AADS performed better in identifying the bioactive conformation of the ligands than identifying the active sites. Retrieval of the native conformation of the ligands within the top 10 poses was outstanding. After successful validation of AADS with LeuT crystal structures, we used it to explore the binding sites of SSA-426 and DJLDU-3-79 in hSERT.

4.3.2. DJLDU-3-79 and SSA-426 binding in hSERT

4.3.2.1. Application of AADS protocol to hSERT homology model
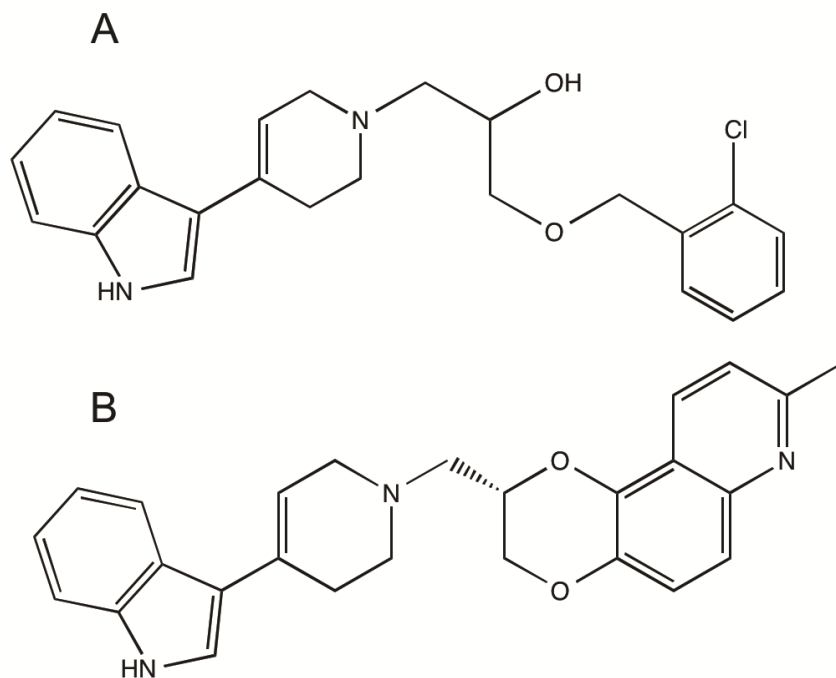
Figure 4.6. Chemical structures of DJLDU-3-79 (A) and SSA-426 (B).

After successful validation of AADS with LeuT, we investigated the binding site and binding mode of DJLDU-3-79 and SSA-426 in hSERT. AADS identified 34 binding pockets in hSERT; some of them were repeated twice. The S1 and S2 pockets, similar to LeuT, were conserved in the hSERT and were among the top 10 binding sites. Next, inhibitors were docked into the top 10 pockets of hSERT using the AADS docking and scoring protocol.

The top 10 docking poses of both the inhibitors resulting from AADS were analyzed. In the case of SSA-426, only two poses (ranked 1 and 3) ended up in the sites of interest (S1 and S2), and rest of the eight poses were binding elsewhere in the protein. Poses 1 and 3 were binding between the S2 and S1 pockets in a similar fashion (Fig. 4.7). These two poses were accommodated between TMs 1, 3, 6, 8 and 10; the indole group was binding in the S1 pocket, the 2-methyl quinolone group was binding in the S2 pocket facing the extracellular side of the protein, and the tetrahydropyridine ring was caught between the S1 and S2 pockets. Both poses were passing through the extracellular hydrophobic gate formed by residues Tyr-176 and Phe-335. The indole moieties of both the poses were surrounded by the S1 (Asp-98, Ser-438, Tyr-95, Ile-172, and Ala-173) and S2 pocket residues (Tyr-175, Tyr-176, Phe-335, Thr-497, and Gly-442). Most importantly, the indole nitrogen was interacting with Asp-98, which was demonstrated to be key for the binding of various antidepressants with hSERT [20, 21, 23, 31-33, 36]. The tetrahydropyridine ring was caught between the extracellular gate (formed by Arg-104 and Glu-493) and the extracellular hydrophobic gate (formed by Tyr-176 and Phe-335), and also surrounded by Gly-100, Leu-99, Phe-335, and Tyr-175. Whereas the 2,3-dioxypropyl 2-methyl quinolone was packed with residues Arg-104, Glu-493, Trp-103,

Ile-179, Val-489, Lys-490, Tyr-107, Trp-182, Gly-402, Ala-401, Pro-403, Tyr-232, Tyr-487, and Gln-238 of the S2 site. It is worth mentioning here that the S1 pocket residues Tyr-95, Asp 98, Ile-172, Asn-177, Phe-341, and Ser-438 [23], and the S2 pocket residues Val-489, Lys-490, Glu-493, Ala-401, Ile-179, Tyr-175, and Gly-100 [39] were established as important for the affinity of various antidepressants with hSERT through several studies. Orientation of poses 1 and 3 was similar to what was proposed earlier by Nolan et al.; however, the AADS poses are binding between the S1 and S2 sites, whereas Nolan et al. proposed that they bind completely in the S2 site [48].
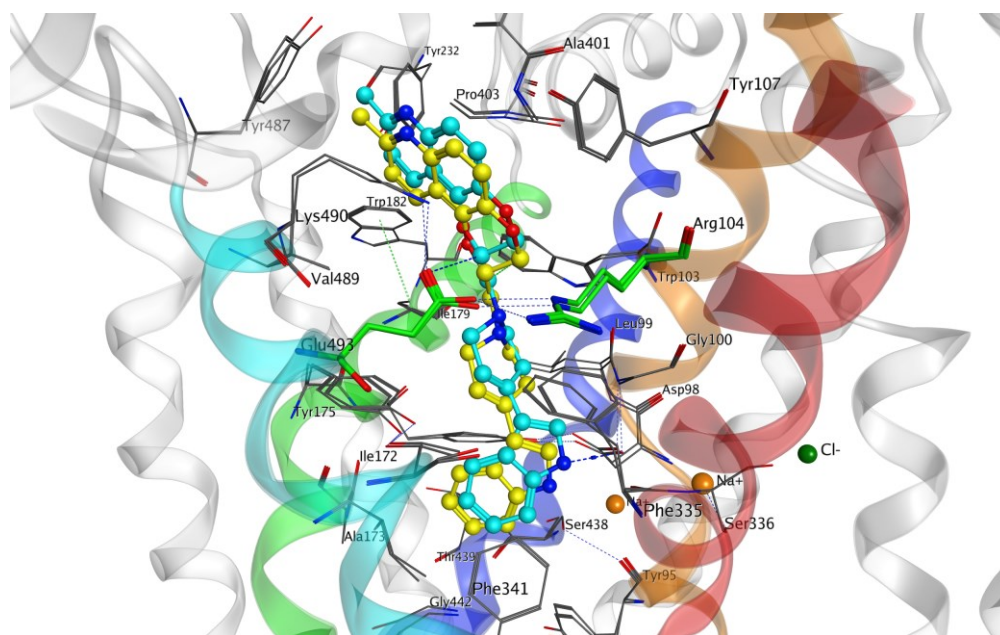
Figure 4.7. Poses 1 and 3 of SSA-426 binding between the S1 and S2 sites of hSERT. Poses 1 and 3 (depicted as balls and sticks in yellow and cyan) bind in a similar fashion between the S1 and S2 sites. The two extracellular gated residues are depicted as sticks (green) and rest of the residues around the ligand are represented as lines (grey). The sodium and chloride ions are represented as orange and green spheres, respectively. TM helices 1, 3, 6, 8, and 10 are colored orange, green, red, blue and cyan, respectively.

With regards to DJLDU-3-79, only two poses (4 and 5) ended up in the functionally relevant regions (i.e., the S1 and S2 sites). Pose-4 was completely binding in the S1 site of hSERT (Fig. 4.8A). The indole group was surrounded by S1 pocket residues Glu-493, Arg-104, Ile-179, Trp-103, Tyr-175, and Leu-99; the tetrahydropyridine ring was packed with Tyr-176, Phe-335, Thr-497, Asp-98, and Ile-172; the 2-hydroxypropyl 2-chlorobenzyl ether moiety was accommodated by residues Gly-442, Tyr-95, Ala-96, Leu-337, Asn-368, Ser-336, Asn-101, and Ser-438. The indole nitrogen was interacting with the gated residue Glu-493 through a hydrogen bond, and benzene was interacting with Tyr-175 through hydrogen-pi interactions. Pose-5 was binding between the S2 and S1 sites similar to poses 1 and 3 of SSA-426, but in a reverse orientation (Fig. 4.8B).

The indole moiety and the tetrahydropyridine ring were binding in the S2 site, whereas the 2-hydroxypropyl-2-chlorobenzyl ether was binding in the S1 site. The indole group was packed with residues Tyr-487, Tyr-232, Lys-490, Trp-182, Phe-407, Pro-403, Gly-402, and Leu-406, and the tetrahydropyridine ring was surrounded by Ile-179, Trp-103, Glu-493, Trp-182, Phe-407, and Arg-104. The 2-hydroxypropyl-2-chlorobenzyl ether moiety was accommodated by residues Arg-104, Glu-493, Tyr-175, Tyr-176, Gly-100, Leu-99, Phe-335, Asp-98, Ala-173, Ile-172, Ser-336, and Ser-438. These results were not in agreement with the earlier predictions (1) that the inhibitor binds completely in the S2 site and (2) that it was oriented with the indole group bound deep into the S1 site while the chlorobenzene group was oriented towards the extracellular side [48]. None of these poses were binding completely in the S2 site and their orientation was not similar to the earlier prediction. For further studies, we generated a pose of DJLDU-3-79 similar

to poses 1 and 3 of SSA-426 using MOE induced fit docking protocol (since AADS failed to predict this orientation).
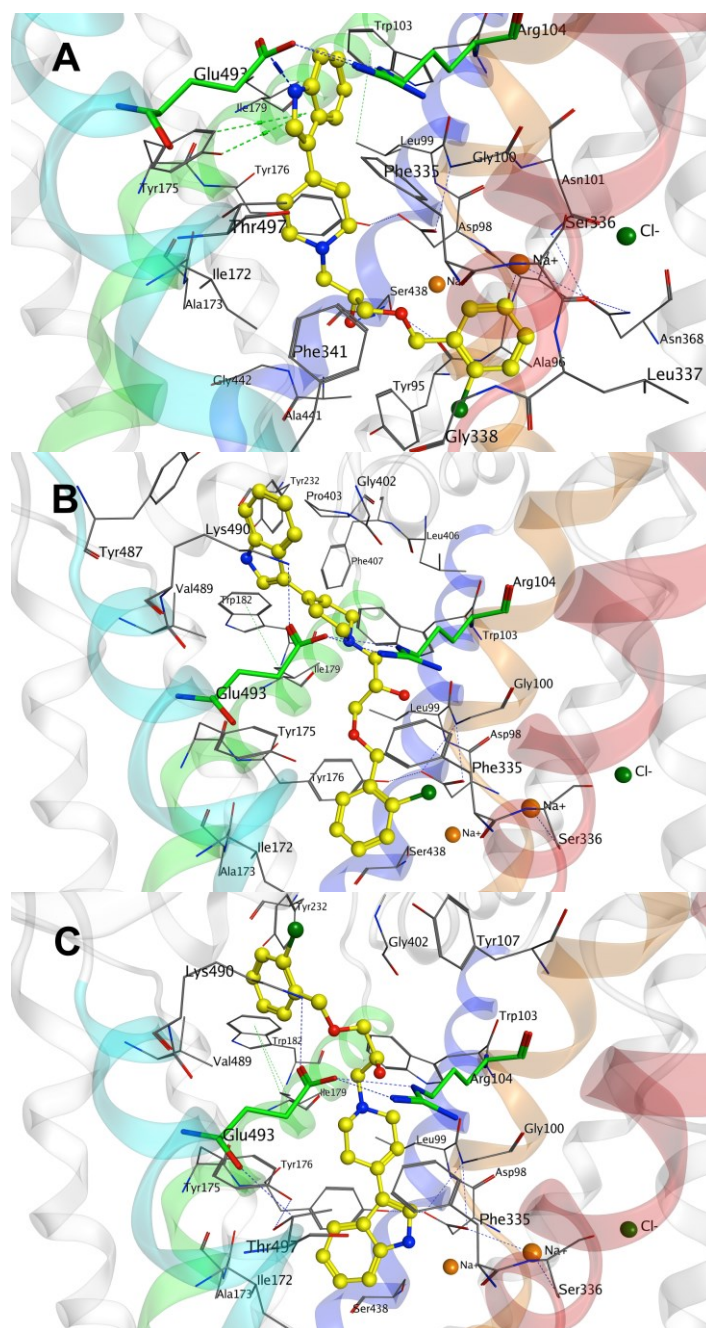
Figure 4.8. Docking poses of DJLDU-3-76 considered for this study. A, B, and C are poses 4, 5, and M of DJLDU-3-76, respectively. All the poses are highlighted as yellow balls and sticks. The extracellular gate (represented as green sticks) acts as a barrier between the S1 and S2 sites.

4.3.2.2.   Molecular dynamics simulations of poses of DJLDU-3-79 and

SSA-426

We performed MD simulations of pose-1 of SSA-426, poses 4, 5 and M of DJLDU-3-79 with hSERT to test their stability in the binding pocket and to investigate the residues playing a key role in the inhibitor binding in hSERT. We simulated each complex for 50ns and collected data for analysis. During the course of MD simulation, pose-1 of SSA-426, which was binding between the S1 and S2 sites, drifted after the first few ns into the S1 site for the most part. The ligand RMSD stabilized after 3ns and remained in the same conformation for another 24ns. After 27ns, RMSD of the ligand significantly changed (from 3 to 4.5Å) and remained stable for the remaining part of the simulation (Fig. 4.9A). The average structure of SSA-426 from the MD simulation was shown in Figure 4.10. The protein backbone RMSD fluctuated significantly for the first 25ns and stabilized around 6Å for the rest of the simulation (Fig. 4.9B). Pose-1 was passing through hydrophobic gate (Tyr-176 and Phe-335) into the S1 site. The 2-methyl quinolone was partially binding in the S2 site, while the rest of the molecule was binding in the S1 site. It was binding between the TM helices 1, 3, 6, 8, and 10 in the S1 site. The tetrahydropyridine was stuck between Tyr-176 and Phe-335, and its protonated nitrogen was interacting with the gated residues Glu-493 and Tyr-175 through a water molecule. There were studies indicating that Ser-438 plays a key role in the binding of antidepressants S-citalopram and TCAs [20]. Asp-98, although binding close to the protonated nitrogen, did not interact directly. The tetrahydropyridine was also interacting

168

with Tyr-176 through a hydrogen-pi interaction. The indole moiety was stacked between Ser-438 and Phe-341, and interacted with them through a hydrogen-pi interaction.
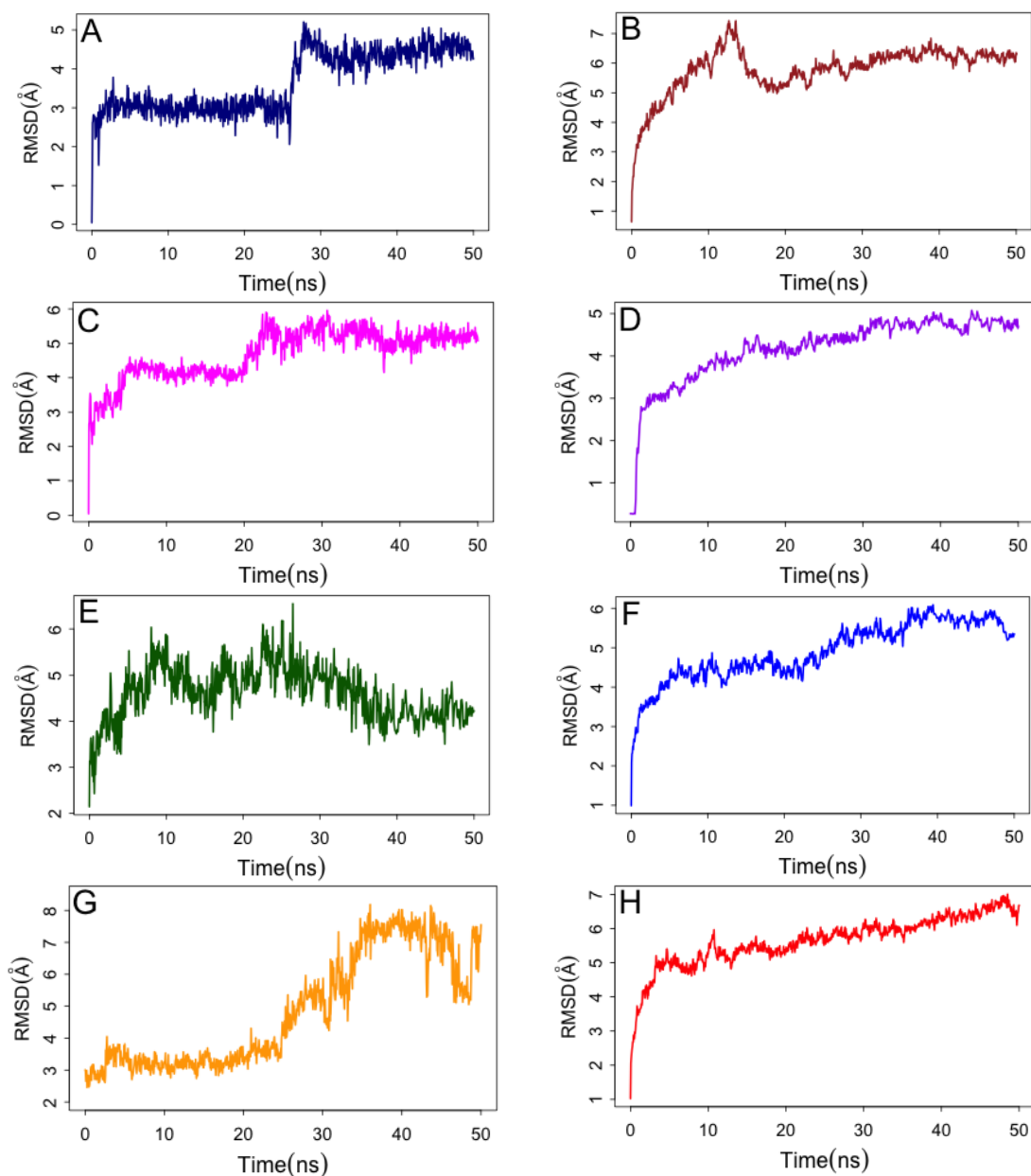
Figure 4.9. Ligand and protein RMSD changes vs. time in MD simulations.
The left and right columns represent the ligand and protein RMSD changes, respectively. The first row (A and B) is pose-1 of SSA-426, the second row (C and D) is pose-M of DJLDU-3-79, the third row (E and F) is pose-5 of DJLDU-3-79, and the fourth row (G and H) is pose-4 of DJLDU-3-79. All non-hydrogen atoms of the ligand, and backbone atoms of the protein, were considered for RMSD calculation. The docking pose was used as a reference for calculating the protein and ligand RMSDs.
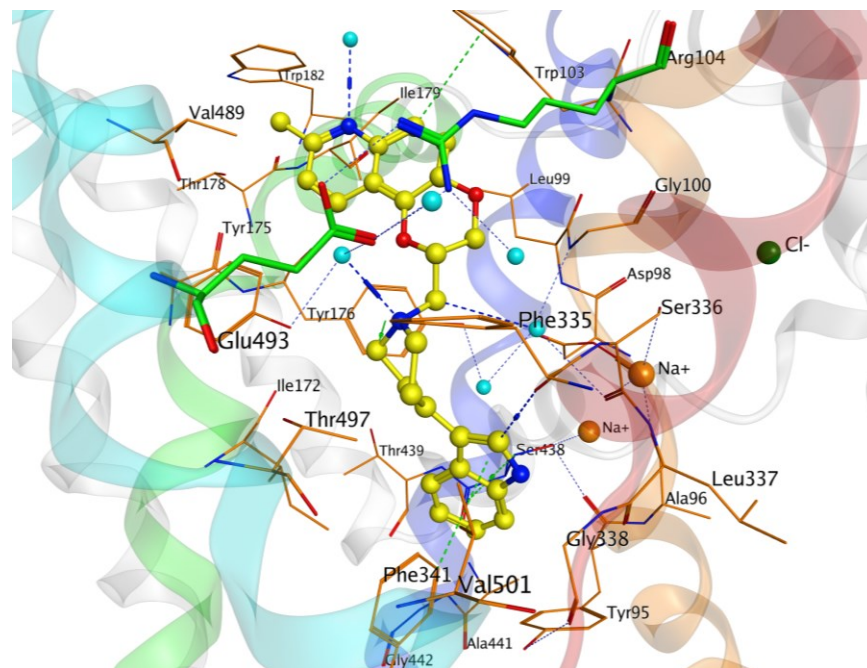
Figure 4.10. Snapshot of pose-1 of SSA-426 from MD simulations. The ligand is depicted as ball and stick (yellow) and the pocket residues are represented as lines (orange). Oxygens of the water molecules are displayed as spheres (cyan). Blue and green dotted lines represent hydrogen bonds and hydrogen-pi interactions, respectively.

The pyrrole ring of the indole was also interacting with Phe-335 through a hydrogen bond. In another study, Tyr-95 and Ile-175 were proved to be critical for the binding of antidepressants [22]. MD simulation results strongly support S1 as the most probable binding site of SSA-426 in hSERT, and not S2. Similar to pose-1 of SSA-426, pose-M of DJLDU-3-79 also migrated into the S1 site after the first few ns of the simulation. The ligand was stabilized at an RMSD of 4.5Å between 5-20ns. Around 20ns, the RMSD of the ligand jumped to 5.5Å and remained stable from 22-50ns (Fig. 4.9C). The protein backbone RMSD increased until 30ns and stabilized at 4.8Å for the rest of the simulation (Fig. 4.9D).

The representative snapshot of the ligand from the MD simulations was shown in Figure 4.11. The salt bridge (Glu-493 and Arg-104) separating the S1 and S2 sites was moved apart and was not interacting directly, but rather through a water molecule. This was in contrast to SSA-426 pose-1, where the salt bridge was still intact and did not fall apart during the course of the simulation. The chlorobenzene moiety slightly protruded into the S2 site and bound in the vicinity of the halogen-binding region, which was comprised of residues Leu-99, Gly-100, Trp-103, Arg-104, Tyr-176, Ile-179 and Phe-335. The halogen-binding region was identified as key for the binding and the selectivity of SSRIs in hSERT [39]. The hydroxyl group of pose-M was interacting with the gated charged residue Glu-493 through a hydrogen bond. The indole group on both sides was interacting with the Ser-438 and Ile-172 through a hydrogen-pi interaction. Although Phe-341 was in the vicinity, unlike in the case of SSA-426, indole was interacting directly with Ile-172 rather than Phe-341. Similar to pose-1 of SSA-426, the protonated nitrogen of tetrahydropyridine was binding near the Asp-98, but not interacting with it directly. The conformations of pose-1 of SSA-426 and pose-M of DJLDU-3-79 that resulted from MD simulations were compared in Figure 4.12. They oriented similarly in the S1 pocket and also overlapped to the maximum extent. In particular, the orientation of indole groups of both the inhibitors in the pocket was exact. Also, similar residues of the TM helices 1, 3, 6, 8 and 10 were interacting with both the poses.

Pose-5 of DJLDU-3-79, which was binding in the opposite orientation of pose-M and binding between the S2 and S1 sites, moved completely into the S2 site during the first 10ns of the simulation. This was in contrast to pose-M of DJLDU-3-79 and pose-1 of SSA-426, where they migrated to the S1 site. The ligand RMSD attained a maximum of

5.9Å at 10ns, fluctuated between 4 and 5.9Å for the next 30ns, and stabilized at 4.3Å in the last 10ns simulation (Fig. 4.9E). The protein backbone stabilized at an RMSD of 4.5Å between 7-22ns, increased from 22-37ns, and stabilized at 6Å for the rest of the 13ns simulation.
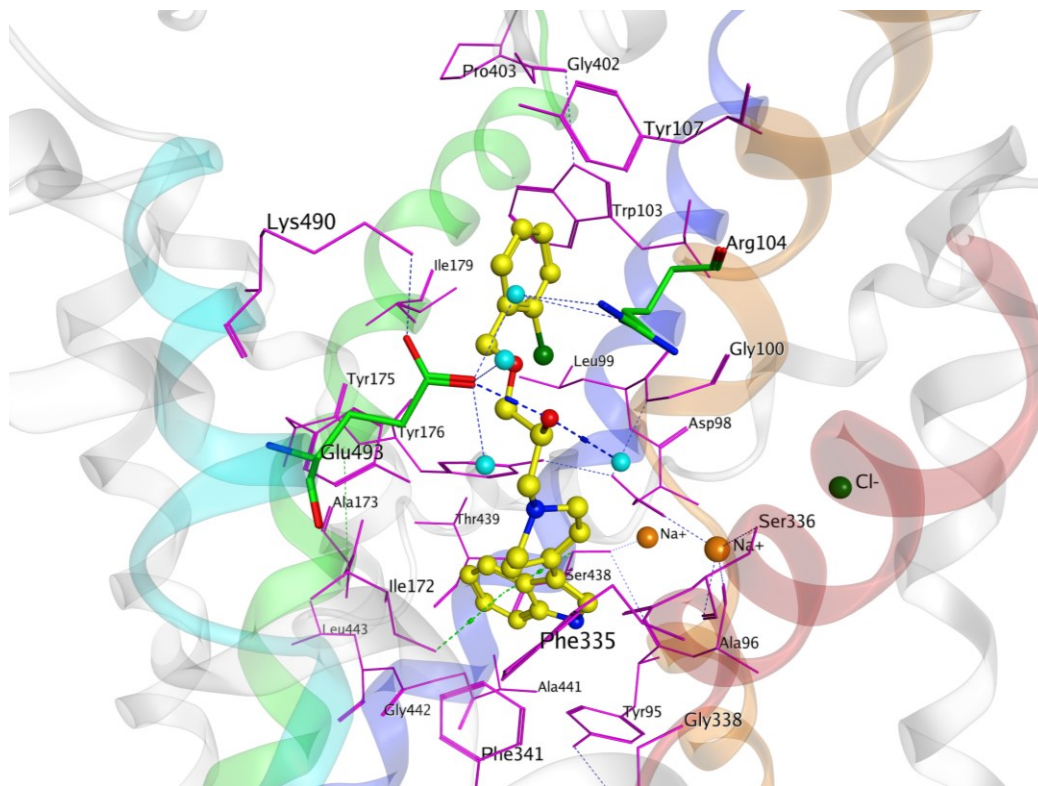

Figure 4.11. Average structure of pose-M of DJLDU-3-79 from MD simulations. The ligand is colored yellow and depicted as ball and stick, and the pocket residues are depicted as pink colored lines.

The ligand snapshot from the last 10ns of the simulation was shown in Figure 4.13. During the entire simulation, the protonated nitrogen of the tetrahydropyridine was interacting with the gated residue Glu-493 through a salt bridge (Fig. 4.14). Also, the chlorobenzene moiety that was binding earlier in the S1 site moved completely into the S2 site and accommodated in the halogen-binding site (Fig. 4.14). Earlier it was proved

with X-ray crystallographic studies that the SSRIs R-fluoxetine, S-fluoxetine and sertraline bind in the S2 site of LeuT and their halogens bind in the halogen-binding site. Along the same lines, it was also speculated through comparative modeling and mutagenesis experiments that halogens of SSRIs bind in the halogen-binding site, which was key for their affinity and selectivity for hSERT [39]. The indole group was pointing towards the extracellular side and exposed to the solvent.
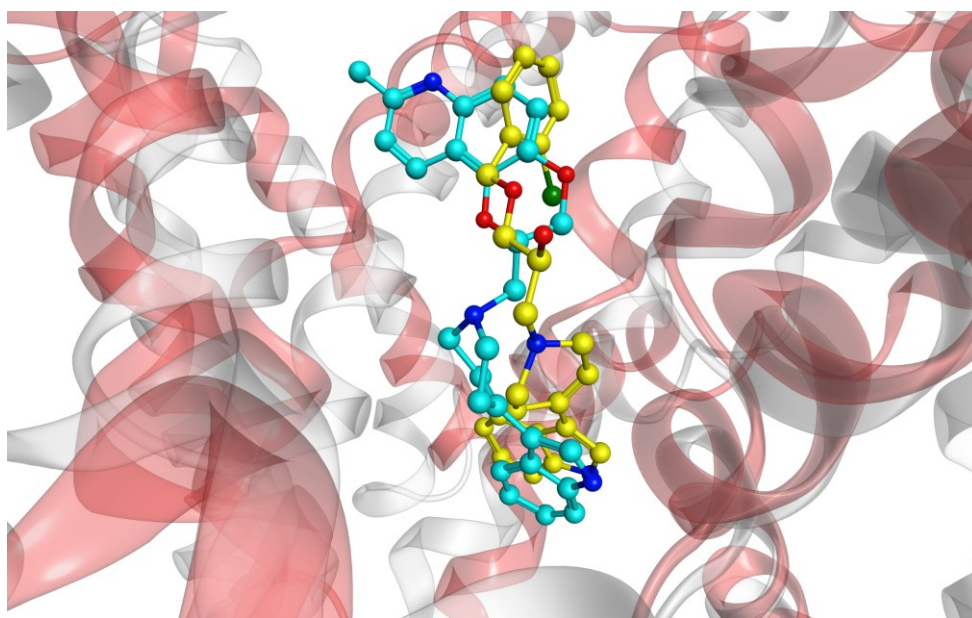
Figure 4.12. MD snapshots of pose-M of DJLDU-3-79 and pose-1of SSA-426. Pose-M (yellow) and pose-1 (blue) are depicted as balls and sticks and their corresponding protein TM helices are represented as white and peach colored ribbons, respectively.

Now we were left with two poses of DJLDU-3-79 (poses 5 and M) that are stable in the MD simulations. These two poses are binding in two different sites (pose-M in the S1 site and pose-5 in the S2 site) in opposite orientations. Since our ultimate goal was to identify the binding site and binding mode of these inhibitors in hSERT, we resolved the ambiguity between these two poses (M and 5) through ABFE calculations.
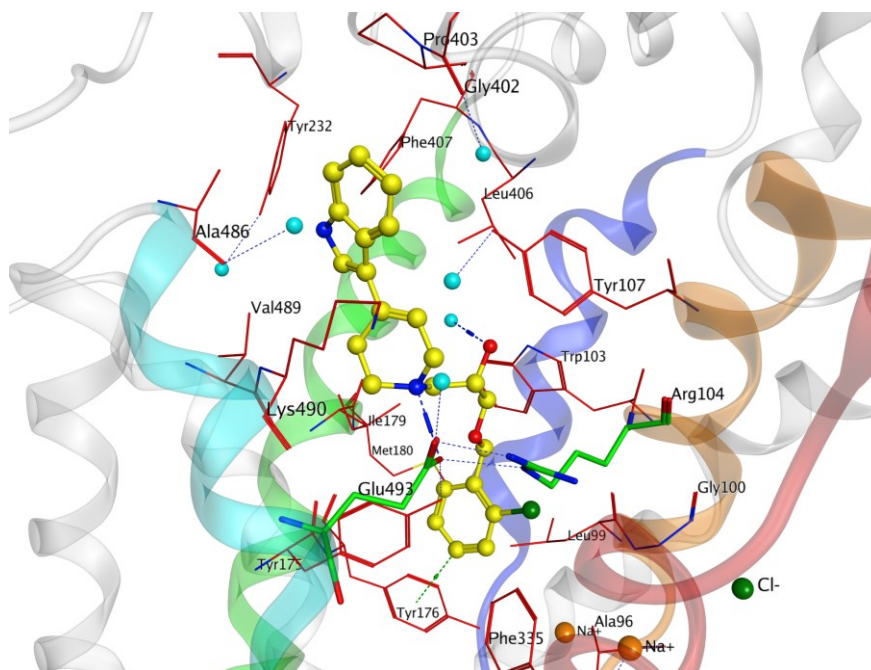
Figure 4.13. Snapshot of pose-5 of DJLDU-3-79 from MD simulations. The ligand is depicted as ball and stick (yellow), and pocket residues are depicted as pink lines.
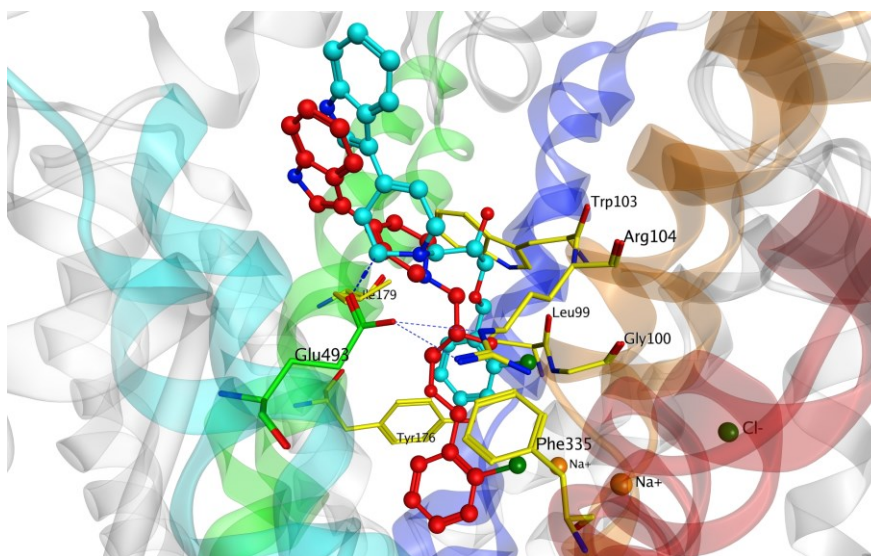
Figure 4.14. Docking vs. MD snapshot of pose-5 of DJLDU-3-79. The residues of halogen-binding site are depicted as sticks (yellow). The docking (red) and MD snapshots (cyan) of the ligand are depicted as balls and sticks.
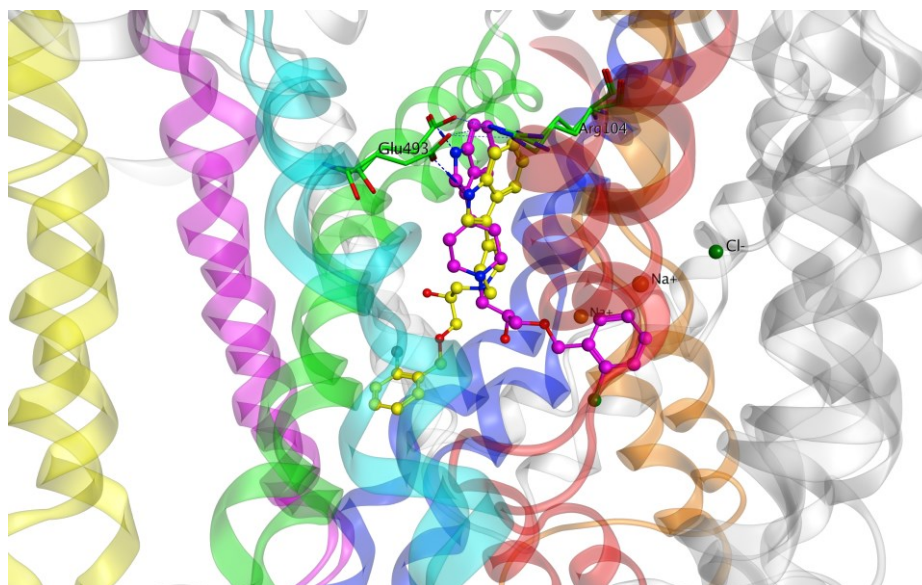
174

Figure 4.15. Docking pose and MD snapshot of pose-4 of DJLDU-3-79. The TM helices 1, 3, 6, 8, 9, 10, and 12 are colored orange, green, red, blue, pink, cyan, and yellow, respectively. Ligands colored yellow and pink represent dynamics and docking poses, respectively.

Pose-4 of DJLDU-3-79 binding originally in the S1 site was not stable (Fig. 4.9G) in the entire 50ns simulation. It tried to escape from the pocket through the TMs 3, 9, and 12 (Fig. 4.15), and in that process it tried to destabilize the protein as well. Although the ligand RMSD was stabilized at 3.5Å between 3-20ns, it drastically increased to 7.5Å between 20-35ns. For a brief period (35-42ns) it was stable at 7.5Å and dropped to 5.5Å in the last 8ns simulation. In Figure 4.15, we compared the initial and the final snapshots of the ligand from the MD simulation. The protein backbone RMSD did not reach a plateau in the entire 50ns simulation. The protein and ligand RMSDs indicate that pose-4 was not the right conformation of DJLDU-3-79 in the S1 site of hSERT. Through MD simulations, we ruled out pose-4 as the bioactive conformation of DJLDU-3-79 in hSERT.

4.3.2.3.   Role of electrostatics in drifting of ligands in MD simulations

Electrostatics played a major role in the migration of ligands either completely into the S1 or S2 sites in MD simulations. Pose-5 of DJLDU-3-79, initially binding between the S1 and S2 sites, shifted into the S2 site in MD simulations. The pocket environment was electrostatically unfavorable for the docking pose. As shown in Figure 4.16C, the chlorine atom was binding in the strong electronegative region. This repulsion might be a major reason for the ligand moving completely into the S2 site. Also, the protonated nitrogen of tetrahydropyridine was binding in the partially electropositive/electronegative region of the protein, and the region around the indole ring was not totally favorable either. In the MD snapshot (Fig. 4.16B) the chlorine atom was binding in the halogen-binding site that was partially electropositive and the protonated nitrogen of tetrahydropyridine was binding in the strong electronegative region of the protein. The indole nitrogen was exposed to the electronegative region and the hydroxyl oxygen was exposed to the strong electropositive region of the pocket. These observations indicate that the S2 site was electrostatically more favorable for pose-5 of DJLDU-3-79. The other possible reason for shifting might be the interaction of indole nitrogen with water on the extracellular side.

Pose-M of DJLDU-3-79 binding initially between the S1 and S2 sites migrated into the S1 site in the MD simulation. In the docking pose (Fig. 4.16E), the protonated nitrogen of tetrahydropyridine was binding in the partially electronegative/positive region. Although the indole nitrogen was near the strong electronegative region, it was not perfectly accommodated. Similarly, the hydroxyl and ester oxygens were also not

favorably accommodated. In the dynamics snapshot (Fig. 4.16F), both the nitrogens of indole and tetrahydropyridine groups, and the oxygens of hydroxyl and ester groups were binding in the strong electronegative and electropositive regions of the protein. The chlorine atom was binding close to the halogen-binding region. These factors demonstrate that S1 site is the right place for the inhibitor to bind in hSERT (Fig. 4.16F).

Similar to pose-M of DJLDU-3-79, pose-1 of SSA-426, which was initially binding between the S1 and S2 sites, migrated into the S1 site. The nitrogen atoms were binding in the electronegative regions in the docking pose (Fig. 4.16A). However, the indole nitrogen was not adequately accommodated near the electronegative region, and the protonated nitrogen of tetrahydropyridine was binding in the partially electronegative region. The rest of the ligand was binding in the S2 site, and this region was more electropositive. In the dynamics snapshot, both the nitrogens were binding in the strong electronegative region of the S1 site, and the other part of the ligand was binding in the partial electronegative/positive region of the S1 site (Fig. 4.16B). Also, the aromatic quinolone might not prefer to stay in the S2 site because of the repulsion for water. Therefore, it is the electrostatics and repulsion of the aromatic region of the ligand for water that guided the drifting of the ligand into the S1 site in the MD simulations.
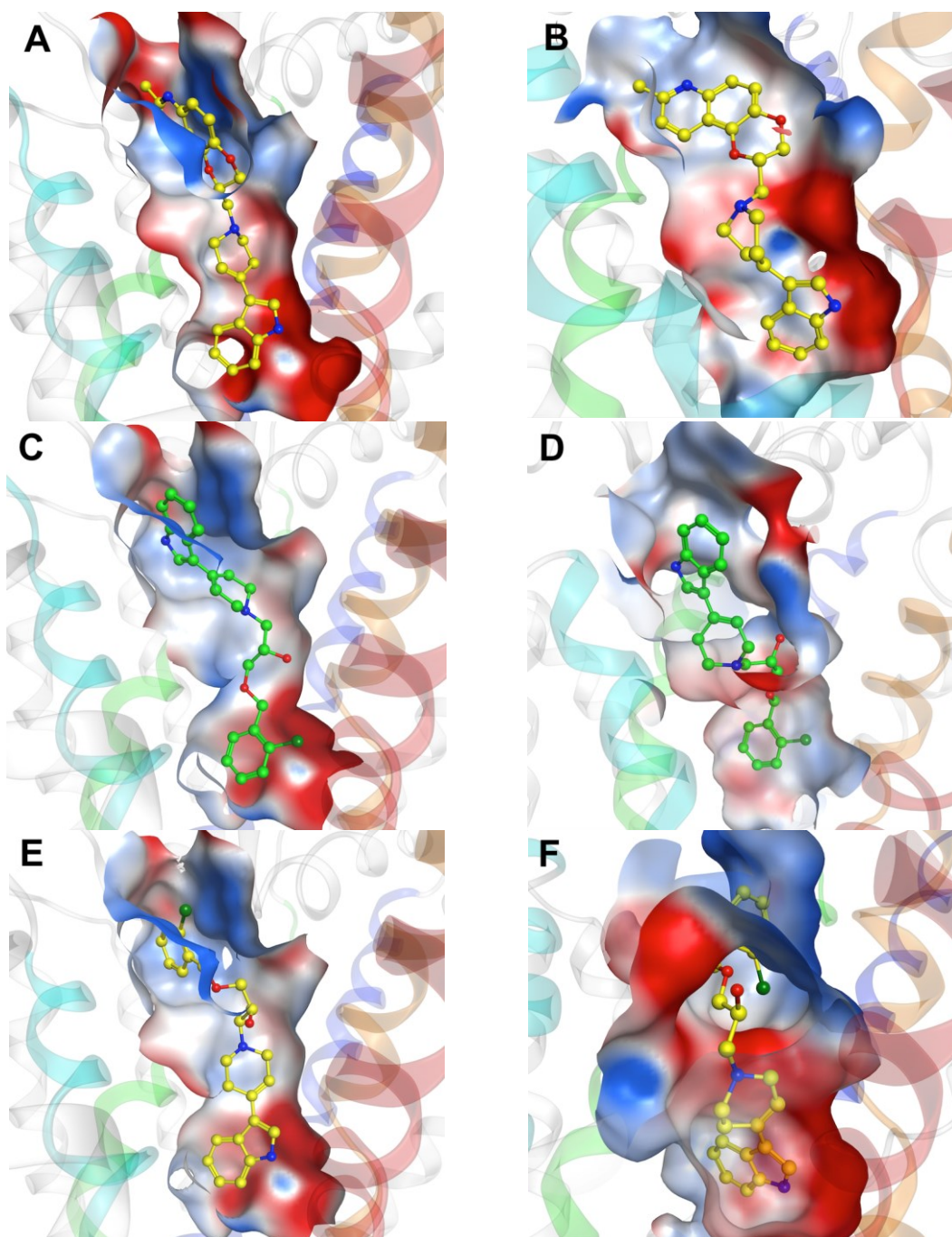
Figure 4.16. Electrostatic potential surfaces of binding sites of docking poses (left column) and MD (right column) snapshots of SSA-426 and DJLDU-3-79 in hSERT. A, C, and E are the electrostatic potential surfaces of docking structures of pose-1 of SSA-426, poses 5 and M of DJLDU-3-79; B, D, and F are the electrostatic potential surfaces of their respective structures resulting from the MD simulations. The red and blue surfaces represent the electronegative and electropositive regions of the protein, respectively. Ligands are depicted in a ball and stick model. The electrostatic potential surfaces were generated using MOE.

4.3.2.4.   Estimation of absolute binding energies through FEP method

We calculated the ABEs of poses M and 5 of DJLDU-3-79 and pose-1 of SSA-426 in hSERT using the FEP method. We did not consider pose-4 of DJLDU-3-79 for FEP calculations since it was not stable in the MD simulations. Before applying to hSERT, we validated the FEP method on LeuBAT:Clomipramine crystal complex (PDBID: 4MMA). Annihilation of the ligand in the solvent and in the protein was carried bidirectionally to monitor hysteresis in the free energy estimation. Data from the forward and backward simulations were combined to estimate the free energy change using the BAR method. In estimating ΔG_1, restraints were imposed on the ligand in the pocket to prevent it from wandering while turning off the non-bonded interactions. The restraints were also used to speed up the convergence of the free energy calculation. As shown in Table 4.2, the calculated binding energy of clomipramine with LeuBAT (-7.69 ± 0.34 kcal/mol) was within 1 kcal/mol of the experimental value (-8.25 ± 0.05 kcal/mol), and this validates the FEP method. However, there was a huge hysteresis between the forward and backward simulations in estimating ΔG_1 (Table 4.2 and Fig. 4.17B), which might be due to inadequate sampling of the configurational space. However, the forward and backward simulations in estimating ΔG_2 were perfectly overlapped and the hysteresis was minimal (Table 4.2 and Fig. 4.17A), which indicates adequate sampling.

Further, we calculated the ABEs of poses M and 5 of DJLDU-3-79 and pose-1 of SSA-426 at hSERT. For this purpose, final snapshots of the protein-ligand complexes from the MD simulations (Fig. 4.12, 4.13 and 4.15) were considered. The computed binding energy of pose-1 of SSA-426 in the S1 site (-12.4 ± 0.47 kcal/mol) was within

179

the error margin of the experimental binding energy (-12.39 ± 0.17 kcal/mol) (Table 4.3). Surprisingly, the hysteresis between the forward and backward simulations was minimal both in the solvent (Fig. 4.18A) and in the protein (Fig. 4.18B). FEP results support the MD simulation data and confirm (1) S1 as the site of SSA-426 and (2) pose-1 as the right orientation of SSA-426 in hSERT.

The calculated binding energies of poses M and 5 of DJLDU-3-79 were -9.93 ± 0.39 kcal/mol and -4.9 ± 0.4 kcal/mol, respectively (Table 4.4). The experimental binding energy of DJLDU-3-79 with hSERT was -10.70 ± 0.07 kcal/mol. Since the calculated binding energy of pose-M correlates well with the experimental value, we conclude pose-M as the probable bioactive conformation of DJLDU-3-79 in hSERT, rather than pose-5. Binding energy calculations clear the ambiguity in MD simulation data. Rather interestingly, the hysteresis was huge between the forward and backward simulations in the bulk solvent (Fig. 19A) unlike in the protein (Fig. 19B,C) even after 15.6ns simulation, indicating inadequate sampling. ABFE calculations clearly indicate that SSA-426 and DJLDU-3-79 bind in the S1 site of hSERT in a similar orientation.

These results demonstrate the usefulness of explicit binding free energy calculations in identifying a near-native conformation of a ligand in a protein.

Table 4.2. Validation of absolute FEP method with the LeuBAT:Clomipramine crystal complex.

ΔG_2			ΔG_1			ΔG^{Rest}	ΔG_{Abs}	ΔG_{Expt}
For	Back	BAR	For	Back	BAR			
33.41	-33.46	32.97 ± 0.11	41.4	-44.6	41.85 ± 0.23	1.19	-7.69 ± 0.34	-8.25 ± 0.05[a]

All energies are in kcal/mol. ΔG^{Rest}, ΔG_{Abs} and ΔG_{Expt} are the restraint free energy, the calculated ABFE, and the experimental binding energy, respectively. [a]The experimental binding affinity of clomipramine with LeuBAT [80]. The error bars associated with the calculated binding energies represent the statistical error, which was estimated through the BAR estimator.

Table 4.3. Calculated binding affinity of pose-1 of SSA-426 with hSERT.

ΔG_2			ΔG_1			ΔG^{Rest}	ΔG_{Abs}	ΔG_{Expt}
For	Back	BAR	For	Back	BAR			
42.05	-41.72	40.12 ± 0.28	55.42	-56.83	55.19 ± 0.19	2.67	-12.4 ± 0.47	-12.39 ± 0.17[a]

All energies are in kcal/mol. [a]The experimental binding affinity of SSA-426 [65].
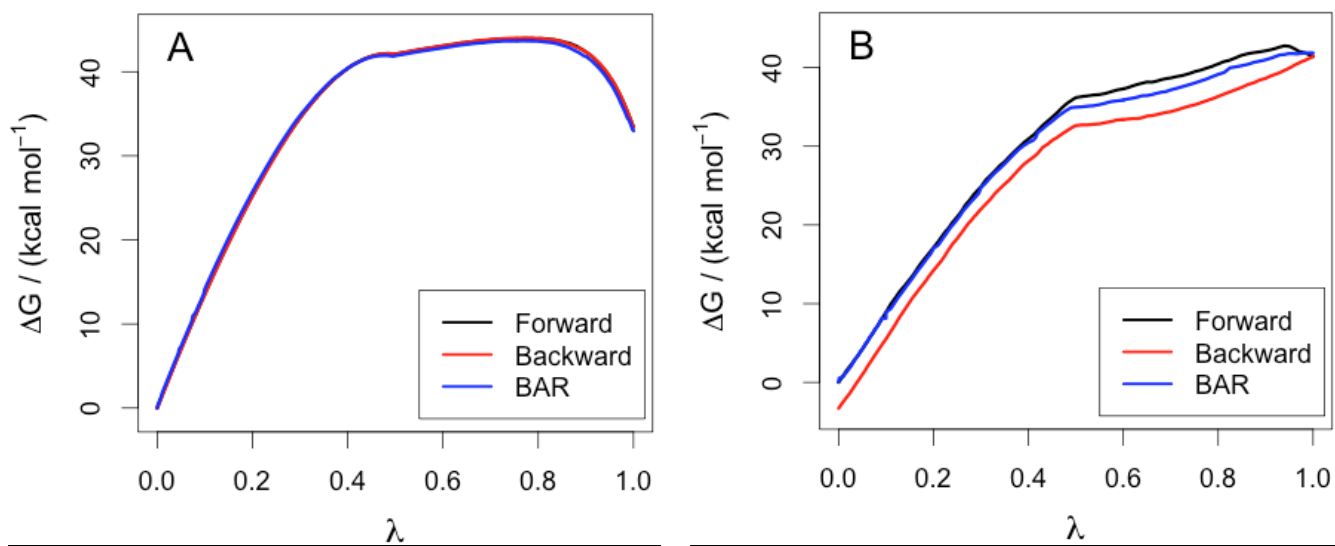
Figure 4.17. Validation of free energy perturbation method with the LeuBAT:Clomipramine complex. Free energy changes as a function of λ for perturbing clomipramine to dummy in the solvent (ΔG_2) and in the protein (ΔG_1) are shown in A and B, respectively. Forward, backward, and BAR estimations are shown in black, red, and blue, respectively.
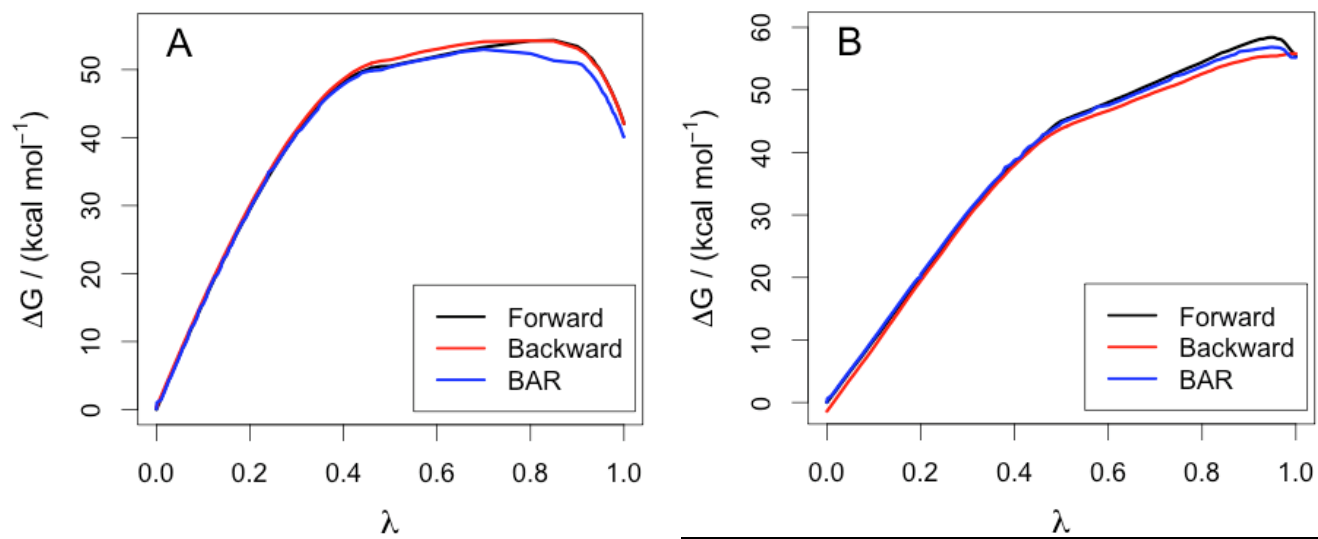
Figure 4.18. Estimation of absolute binding energy of pose-1 of SSA-426 with hSERT using FEP method. Free energy change versus λ for annihilation of SSA-426 in (A) water (ΔG_2) and in (B) hSERT (ΔG_1).
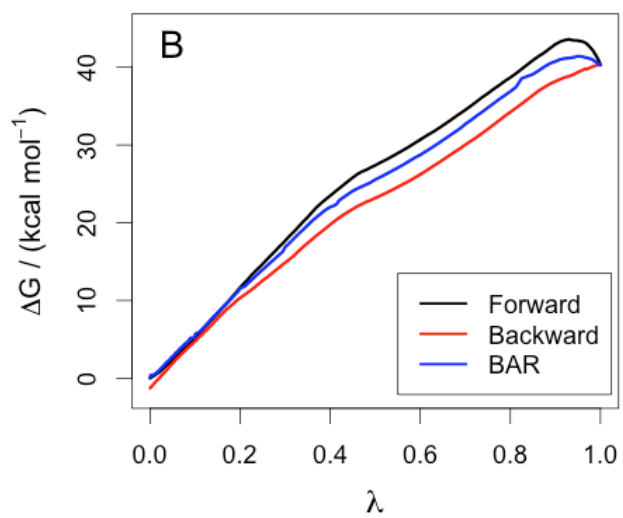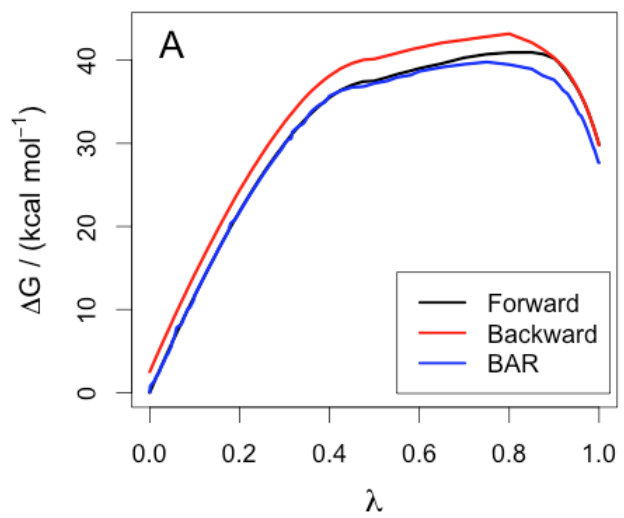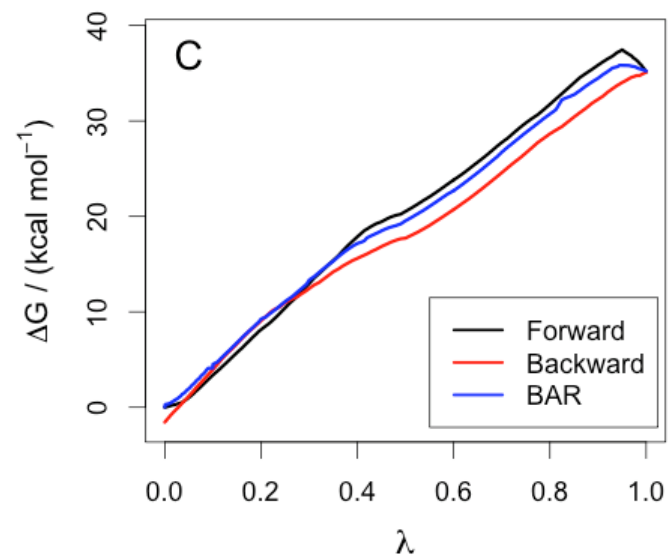
Figure 4.19. Calculating absolute binding energies of poses M and 5 of DJLDU-3-79 with hSERT. A is free energy change vs. λ in perturbing DJLDU-3-79 in water (ΔG_2); B and C are changes in free energies vs. λ for perturbing poses M and 5 of DJLDU-3-79 in hSERT (ΔG_1).

Table 4.4. Absolute binding energies of poses 5 and M of DJLDU-3-79 with hSERT.

	ΔG_2			ΔG_1			ΔG^{Rest}	ΔG_{Abs}	ΔG_{Expt}
	For	Back	BAR	For	Back	BAR			
Pose-5				35.12	36.68	35.24 ± 0.1	2.67	-4.9 ± 0.4	
	29.8	-27.3	27.67 ± 0.3						-10.70 ± 0.07[a]
Pose-M				40.36	-41.65	40.27 ± 0.09	2.67	-9.93 ± 0.39	

Pose-M was generated using MOE. Pose-5 was from the AADS webserver. All energies are in kcal/mol. [a]The experimental binding affinity of DJLDU-3-79 with hSERT[48].

4.4. Discussion and Conclusions

The present study was aimed at determining the binding modes and the binding sites of a dual SERT/5HT-1A antagonist SSA-426 and a selective hSERT blocker DJLDU-3-79 in hSERT. For this purpose, we adopted a computational protocol that incorporates active site detection, docking, scoring, MD simulations, and ABFE calculations. For binding site identification, docking and scoring purposes, we used an automated web server called AADS. For computing ABEs, we used FEP method. To our knowledge, these are the first explicit free energy calculations ever carried on hSERT. Before applying to hSERT, we tested AADS on five different LeuT crystal structures for its active site identification and ligand's near-native conformation retrieval capabilities. AADS performed extremely well in identifying the active sites; in all five cases active site was among the top ten cavities identified, irrespective of the conformation of the protein and location of the binding site. However, AADS performed better in identifying the S2 site rather than S1 because the S1 site is located deep in the protein and occluded conformation of the protein makes AADS difficult to access from both the ends. In contrast, S2 site is located on the extracellular side, and because of its size it is readily accessible to the ligands, making it the highest ranked site in all the crystal structures. With regards to the retrieval of the bioactive conformation of the ligand, AADS performed reasonably well. At least one among the top ten AADS poses was close to the bioactive conformation of the ligand in all five cases. The RMSD of the closest AADS pose to the bioactive conformation was less than 3.75Å in all 5 cases. As in the case of active site detection, AADS was more accurate in identifying the bioactive conformation in the S2 pocket than in the S1 pocket for similar

reasons as mentioned before. Overall, the docking and scoring protocol in AADS performed better than the active site finder. We also validated the FEP method on LeuBAT:Clomipramine crystal complex. The calculated binding affinity (-7.69 ± 0.34 kcal/mol) was in reasonable agreement with the experimental value (-8.25 ± 0.05 kcal/mol).

Finally, we explored the binding sites and binding modes of SSA-426 and DJLDU-3-79 in hSERT. The hSERT homology model developed using the crystal structure of LeuT was used for this purpose. Nolan et al. proposed that both these inhibitors would (1) bind in the S2 site and (2) orient similarly in the binding site. According to Nolan et al. the indole groups of both the ligands bind deep into the S2 site and remaining parts of the ligands face the extracellular side in a similar fashion. Initially we used AADS to generate the poses of DJLDU-3-79 and SSA-426 in hSERT. Either in the case of SSA-426 or DJLDU-3-79, none of the AADS poses bound exclusively in the S2 site with similar orientation as proposed by Nolan et al. In the case of SSA-426, orientation of poses 1 and 3 was similar to what was proposed by Nolan et al.; however, they were binding between the S1 and S2 sites. In the case of DJLDU, pose-5 was binding between the S1 and S2 sites similar to poses 1 and 3 of SSA-426, but in reverse orientation. Pose-4 was binding completely in the S1 site. We generated a pose of DJLDU-3-79 (Pose-M) with a similar orientation as poses 1 and 3 of SSA-426 using MOE for further studies; its orientation was exactly opposite of pose-5 of DJLDU-3-79.

We performed MD simulations to test the stability of the poses in the protein and to understand the factors impacting the binding of these ligands. Pose-1 of SSA-426 drifted into the S1 site for the most part, and this partially ruled out the possibility of S2

being the binding site of SSA-426. Pose-M of DJLDU-3-79 migrated into the S1 site similar to pose-1 of SSA-426. Also, they oriented similarly with indole bound deep in the S1 site and remaining parts of the ligands extended towards the extracellular gate separating the S1 and S2 sites. This is in agreement with the earlier hypothesis by Nolan et al. with regards to the orientation, but not with the binding site. Pose-5 of DJLDU-3-79 moved completely into the S2 site. The halogen atom was accommodated in the halogen-binding site and the protonated nitrogen of tetrahydropyridine was interacting with Glu-493. It is worth mentioning here that these two features are characteristic of SSRIs binding in LeuT. However, TM helices on the extracellular side have to come closer to accommodate the ligand in the S2 site. Pose-4 was not stable in the entire 50ns simulation. Therefore, it might not be the bioactive conformation of the ligand in hSERT. So, we were left with two different poses of DJLDU-3-79 binding in two different pockets (Pose-M in the S1 site and Pose-5 in the S2 site) in two different orientations.

To clear this ambiguity, we performed FEP calculations to compute the ABFE and compared against the experimental binding affinity. Free energy simulations favor the S1 site rather than S2. Pose-M of DJDLU-3-79 in the S1 site binds stronger (-9.93 ± 0.39 kcal/mol) and correlates well with the experimental binding affinity (-10.70 ± 0.07 kcal/mol) than pose-5 (-4.9 ± 0.4 kcal/mol) in the S2 site. We believe that pose-5 might be a less energetically favorable metastable state for DJLDU-3-79 in the S2 site. We also calculated the binding affinity of pose-1 of SSA-426 in the S1 site, which was in excellent agreement with the experimental value (-12.4 ± 0.47 VS -12.39 ± 0.17 kcal/mol). MD simulations and ABFE calculations clearly favor the S1 site in the case of SSA-426.

Our results demonstrate that we need a full computational protocol (site finder,

docking, MD simulations and free energy calculations) to clearly elucidate the binding site and orientation of ligands in the protein. Although AADS placed the poses between the S1 and S2 sites, MD simulations pushed them either into the S1 or S2 sites completely. We think this was partially because AADS used a rigid docking protocol, and one needs multiple protein conformations (especially with the homology models) to accurately predict the binding site and orientation of the ligand. For example, in the case of LeuT, AADS was highly successful either in the binding site or pose prediction. This was because we used the right conformations of the protein, i.e., conformations of the protein that were determined in the presence of the same ligands that were docked in the protein (self-docking). This was not surprising and most docking programs performed reasonably well in self-docking, but they failed miserably when they were cross-docked into alternate conformations of the proteins [81, 82]. Also, AADS does not handle ions. However, several studies established the importance of ions (sodium and chlorine) in the binding of inhibitors/antidepressants in hSERT. In the MD simulations, we incorporated flexibility to the protein and also added sodium and chloride ions at appropriate places in the protein; these factors might contribute to the drifting of ligands into S1 or S2 sites completely during MD simulations. Already we have established the role of electrostatics in ligand migration in MD simulations. Another important factor was water; multiple studies suggested the role of water in the binding of ligands to protein pockets, and we incorporated water explicitly both in the MD simulations and the ABFE calculations. Our results show that one cannot completely depend on docking and scoring algorithms in identifying a native pose of ligands in proteins. Even though computationally expensive, MD simulations and ABFE calculations are more reliable in discriminating binders from

non-binders.

In summary, we have proposed that S1 is the binding site for SSA-426 and DJLDU-3-79, and not S2 as predicted earlier by Nolan et al. With regards to the orientation, both the inhibitors overlap and share a similar orientation, and this is in agreement with the Nolan et al. hypothesis. Our calculated free energy results are within 0.8kcal/mol of the experimental values.

4.5. References

1.   Gether, U., P.H. Andersen, O.M. Larsson, and A. Schousboe, *Neurotransmitter transporters: molecular function of important drug targets.* Trends in Pharmacological Sciences, 2006. **27**(7): p. 375-383.

2.   Ozaki, N., D. Goldman, W.H. Kaye, K. Plotnicov, B.D. Greenberg, J. Lappalainen, G. Rudnick, and D.L. Murphy, *Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype.* Mol Psychiatry, 0000. **8**(11): p. 933-936.

3.   Hahn, M.K., D. Robertson, and R.D. Blakely, *A mutation in the human norepinephrine transporter gene (SLC6A2) associated with orthostatic intolerance disrupts surface expression of mutant and wild-type transporters.* The Journal of Neuroscience, 2003. **23**(11): p. 4470-4478.

4.   Hahn, M.K. and R.D. Blakely, *Monoamine transporter gene structure and polymorphisms in relation to psychiatric and other complex disorders.* Pharmacogenomics J, 2002. **2**(4): p. 217-235.

5.   Ramamoorthy, S., A.L. Bauman, K.R. Moore, H. Han, T. Yang-Feng, A.S. Chang, V. Ganapathy, and R.D. Blakely, *Antidepressant- and cocaine-sensitive human serotonin transporter: molecular cloning, expression, and chromosomal localization.* Proceedings of the National Academy of Sciences, 1993. **90**(6): p. 2542-2546.

6.   Amara, S.G. and M.S. Sonders, *Neurotransmitter transporters as molecular targets for addictive drugs.* Drug and Alcohol Dependence, 1998. **51**(1,Äì2): p. 87-96.

7.   Bowton, E., C. Saunders, K. Erreger, D. Sakrikar, H. Matthies, N. Sen, T. Jessen, R. Colbran, M. Caron, J. Javitch, R. Blakely, and A. Galli, *Dysregulation of*

dopamine transporters via dopamine D2 autoreceptors triggers anomalous dopamine efflux associated with attention-deficit hyperactivity disorder. The Journal of Neuroscience, 2010. **30**(17): p. 6048-57.

8.   Nelson, N., *The family of Na+/Cl− neurotransmitter transporters.* Journal of Neurochemistry, 1998. **71**(5): p. 1785-1803.

9.   Masson, J., C. Sagne, M. Hamon, and S.E. Mestikawy, *Neurotransmitter transporters in the central nervous system.* Pharmacological Reviews, 1999. **51**(3): p. 439-464.

10.   Gu, H., S.C. Wall, and G. Rudnick, *Stable expression of biogenic amine transporters reveals differences in inhibitor sensitivity, kinetics, and ion dependence.* Journal of Biological Chemistry, 1994. **269**(10): p. 7124-7130.

11.   McElvain, J.S. and J.O. Schenk, *A multisubstrate mechanism of striatal dopamine uptake and its inhibition by cocaine.* Biochemical Pharmacology, 1992. **43**(10): p. 2189-2199.

12.   *Unhappiness by the numbers: 2012 depression statistics*. Available from: http://www.healthline.com/health/depression/statistics-infographic.

13.   *Depression: a global crisis; world mental health day.* 2012; Available from: http://www.who.int/mental_health/ management/depression/en/.

14.   *Depression*. Available from: http://www.emro.who.int/health-topics/depression/index.html.

15.   Stahl, S.M., *The prescriber's guide, antipsychotics and mood stabilizers*. Third ed. 2009: Cambridge University Press, New York, NY, USA.

16.   Green, J.P. and S. Maayani, *Tricyclic antidepressant drugs block histamine H2 receptor in brain.* Nature, 1977. **269**(5624): p. 163-165.

17.   Cusack, B., A. Nelson, and E. Richelson, *Binding of antidepressants to human brain receptors: focus on newer generation compounds.* Psychopharmacology, 1994. **114**(4): p. 559-565.

18.   Frazer, A., *Pharmacology of antidepressants.* Journal of Clinical Psychopharmacology, 1997. **17**(2): p. 2S-18S.

19.   Rudnick, G., *What is an antidepressant binding site doing in a bacterial transporter?* ACS Chemical Biology, 2007. **2**(9): p. 606-609.

20. Andersen, J., O. Taboureau, K.B. Hansen, L. Olsen, J. Egebjerg, K. Stromgaard, and A.S. Kristensen, *Location of the antidepressant binding site in the serotonin transporter.* Journal of Biological Chemistry, 2009. **284**(15): p. 10276-10284.

21. Barker, E.L., K.R. Moore, F. Rakhshan, and R.D. Blakely, *Transmembrane domain I contributes to the permeation pathway for serotonin and ions in the serotonin transporter.* The Journal of Neuroscience, 1999. **19**(12): p. 4705-4717.

22. Henry, L.K., J.R. Field, E.M. Adkins, M.L. Parnas, R.A. Vaughan, M.-F. Zou, A.H. Newman, and R.D. Blakely, *Tyr-95 and Ile-172 in transmembrane segments 1 and 3 of human serotonin transporters interact to establish high affinity recognition of antidepressants.* Journal of Biological Chemistry, 2006. **281**(4): p. 2012-2023.

23. Andersen, J., L. Olsen, K.B. Hansen, O. Taboureau, F.S. Jorgensen, A.M. Jorgensen, B. Bang-Andersen, J. Egebjerg, K. Stromgaard, and A.S. Kristensen, *Mutational mapping and modeling of the binding site for (S)-citalopram in the human serotonin transporter.* Journal of Biological Chemistry, 2010. **285**(3): p. 2051-2063.

24. Andersen, J., N. Stuhr-Hansen, L. Zachariassen, S. Toubro, S.M.R. Hansen, J.N.N. Eildal, A.D. Bond, K.P. Bogeso, B. Bang-Andersen, A.S. Kristensen, and K. Stromgaard, *Molecular determinants for selective recognition of antidepressants in the human serotonin and norepinephrine transporters.* Proceedings of the National Academy of Sciences, 2011.

25. Hyttel, J., *Neurochemical characterization of a new potent and selective serotonin uptake inhibitor: Lu 10-171.* Psychopharmacology (Berl), 1977. **51**(403537): p. 225-233.

26. Talvenheimo, J., P.J. Nelson, and G. Rudnick, *Mechanism of imipramine inhibition of platelet 5-hydroxytryptamine transport.* Journal of Biological Chemistry, 1979. **254**(11): p. 4631-5.

27. Thomas, D.R., D.R. Nelson, and A.M. Johnson, *Biochemical effects of the antidepressant paroxetine, a specific 5-hydroxytryptamine uptake inhibitor.* Psychopharmacology, 1987. **93**(2): p. 193-200.

28. Apparsundaram, S., D.J. Stockdale, R.A. Henningsen, M.E. Milla, and R.S. Martin, *Antidepressants targeting the serotonin reuptake transporter act via a competitive mechanism.* Journal of Pharmacology and Experimental Therapeutics, 2008. **327**(3): p. 982-990.

29. Humphreys, C.J., J. Levin, and G. Rudnick, *Antidepressant binding to the porcine and human platelet serotonin transporters.* Molecular Pharmacology, 1988. **33**(6): p. 657-663.

30. Koe, B.K., L.A. Lebel, and W.M. Welch, *[3H] sertraline binding to rat brain membranes.* Psychopharmacology (Berl), 1990. **100**(4): p. 470-6.

31. Barker, E.L., M.A. Perlman, E.M. Adkins, W.J. Houlihan, Z.B. Pristupa, H.B. Niznik, and R.D. Blakely, *High affinity recognition of serotonin transporter antagonists defined by species-scanning mutagenesis.* Journal of Biological Chemistry, 1998. **273**(31): p. 19459-19468.

32. Sinning, S., M. Musgaard, M. Jensen, K. Severinsen, L. Celik, H. Koldso, T. Meyer, M. Bols, H.H. Jensen, B. Schiott, and O. Wiborg, *Binding and orientation of tricyclic antidepressants within the central substrate site of the human serotonin transporter.* Journal of Biological Chemistry, 2010. **285**(11): p. 8363-8374.

33. Tavoulari, S., L.R. Forrest, and G. Rudnick, *Fluoxetine (Prozac) binding to serotonin transporter is modulated by chloride and conformational changes.* The Journal of Neuroscience, 2009. **29**(30): p. 9635-9643.

34. Koldsø, H., K. Severinsen, T.T. Tran, L. Celik, H.H. Jensen, O. Wiborg, B. Schiøtt, and S. Sinning, *The two enantiomers of citalopram bind to the human serotonin transporter in reversed orientations.* Journal of the American Chemical Society, 2010. **132**(4): p. 1311-1322.

35. Jørgensen, A.M., L. Tagmose, A.M.M. Jørgensen, S. Topiol, M. Sabio, K. Gundertofte, K.P. Bøgesø, and G.H. Peters, *Homology modeling of the serotonin transporter: insights into the primary escitalopram-binding site.* ChemMedChem, 2007. **2**(6): p. 815-826.

36. Wang, H., A. Goehring, K.H. Wang, A. Penmatsa, R. Ressler, and E. Gouaux, *Structural basis for action by diverse antidepressants on biogenic amine transporters.* Nature, 2013. **503**(7474): p. 141-5.

37. Humphreys, C.J., S.C. Wall, and G. Rudnick, *Ligand binding to the serotonin transporter: equilibria, kinetics, and ion dependence.* Biochemistry, 1994. **33**(31): p. 9118-25.

38. Mann, C.D. and P.D. Hrdina, *Sodium dependence of [3H]]paroxetine binding and 5- [3H]]hydroxytryptamine uptake in rat diencephalon.* Journal of Neurochemistry, 1992. **59**(5): p. 1856-1861.

39. Zhou, Z., J. Zhen, N.K. Karpowich, C.J. Law, M.E.A. Reith, and D.-N. Wang, *Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures.* Nat Struct Mol Biol, 2009. **16**(6): p. 652-657.

40.   Sarker, S., R. Weissensteiner, I. Steiner, H.H. Sitte, G.F. Ecker, M. Freissmuth, and S. Sucic, *The high-affinity binding site for tricyclic antidepressants resides in the outer vestibule of the serotonin transporter.* Molecular Pharmacology, 2010. **78**(6): p. 1026-1035.

41.   Zhou, Z., J. Zhen, N.K. Karpowich, R.M. Goetz, C.J. Law, M.E.A. Reith, and D.-N. Wang, *LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake.* Science, 2007. **317**(5843): p. 1390-1393.

42.   Plenge, P. and O. Wiborg, *High- and low-affinity binding of S-citalopram to the human serotonin transporter mutated at 20 putatively important amino acid positions.* Neuroscience Letters, 2005. **383**(3): p. 203-208.

43.   Plenge, P., U. Gether, and S.G. Rasmussen, *Allosteric effects of R- and S-citalopram on the human 5-HT transporter: evidence for distinct high- and low-affinity binding sites.* Eur J Pharmacol, 2007. **567**(1-2): p. 1-9.

44.   Chen, F., M.B. Larsen, H.A. Neubauer, C. Sánchez, P. Plenge, and O. Wiborg, *Characterization of an allosteric citalopram-binding site at the serotonin transporter.* Journal of Neurochemistry, 2005. **92**(1): p. 21-28.

45.   Chen, F., M.B. Larsen, C. Sánchez, and O. Wiborg, *The S-enantiomer of R,S-citalopram, increases inhibitor binding to the human serotonin transporter by an allosteric mechanism. Comparison with other serotonin transporter inhibitors.* European neuropsychopharmacology, 2005. **15**(2): p. 193-198.

46.   Plenge, P., L. Shi, T. Beuming, J. Te, A.H. Newman, H. Weinstein, U. Gether, and C.J. Loland, *Steric hindrance mutagenesis in the conserved extracellular vestibule impedes allosteric binding of antidepressants to the serotonin transporter.* Journal of Biological Chemistry, 2012. **287**(47): p. 39316-39326.

47.   Bulling, S., K. Schicker, Y.W. Zhang, T. Steinkellner, T. Stockner, C.W. Gruber, S. Boehm, M. Freissmuth, G. Rudnick, H.H. Sitte, and W. Sandtner, *The mechanistic basis for noncompetitive ibogaine inhibition of serotonin and dopamine transporters.* J Biol Chem, 2012. **287**(22): p. 18524-34.

48.   Nolan, T.L., D.J. Lapinsky, J.N. Talbot, M.n. Indarte, Y. Liu, S. Manepalli, L.M. Geffert, M.E. Amos, P.N. Taylor, J.D. Madura, and C.K. Surratt, *Identification of a novel selective serotonin reuptake inhibitor by coupling monoamine transporter-based virtual screening and rational molecular hybridization.* ACS Chemical Neuroscience, 2011. **2**(9): p. 544-552.

49.   Manepalli, S., L.M. Geffert, C.K. Surratt, and J.D. Madura, *Discovery of novel selective serotonin reuptake inhibitors through development of a protein-based pharmacophore.* J Chem Inf Model, 2011. **51**(9): p. 2417-26.

50.   Celik, L., S. Sinning, K. Severinsen, C.G. Hansen, M.S. Moller, M. Bols, O. Wiborg, and B. Schiott, *Binding of serotonin to the human serotonin transporter. molecular modeling and experimental validation.* Journal of the American Chemical Society, 2008. **130**(12): p. 3853-3865.

51.   Manepalli, S., C.K. Surratt, J.D. Madura, and T.L. Nolan, *Monoamine transporter structure, function, dynamics, and drug discovery: a computational perspective.* AAPS J, 2012. **14**(4): p. 820-31.

52.   Singh, S.K., C.L. Piscitelli, A. Yamashita, and E. Gouaux, *A competitive inhibitor traps LeuT in an open-to-out conformation.* Science, 2008. **322**(5908): p. 1655-1661.

53.   Yamashita, A., S.K. Singh, T. Kawate, Y. Jin, and E. Gouaux, *Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters.* Nature, 2005. **437**(7056): p. 215-223.

54.   Singh, S.K., A. Yamashita, and E. Gouaux, *Antidepressant binding site in a bacterial homologue of neurotransmitter transporters.* Nature, 2007. **448**(7156): p. 952-956.

55.   Shi, L., M. Quick, Y. Zhao, H. Weinstein, and J.A. Javitch, *The mechanism of a neurotransmitter:sodium symporter—inward release of Na+ and substrate is triggered by substrate in a second binding site.* Molecular Cell, 2008. **30**(6): p. 667-677.

56.   Piscitelli, C.L., H. Krishnamurthy, and E. Gouaux, *Neurotransmitter/sodium symporter orthologue LeuT has a single high-affinity substrate site.* Nature, 2010. **468**(7327): p. 1129-32.

57.   Nolan, T.L., D.J. Lapinsky, J.N. Talbot, M. Indarte, Y. Liu, S. Manepalli, L.M. Geffert, M.E. Amos, P.N. Taylor, J.D. Madura, and C.K. Surratt, *Identification of a novel selective serotonin reuptake inhibitor by coupling monoamine transporter-based virtual screening and rational molecular hybridization.* ACS Chemical Neuroscience, 2011. **2**(9): p. 544-552.

58.   Ravna, A.W., M. Jaronczyk, and I. Sylte, *A homology model of SERT based on the LeuT(Aa) template.* Bioorg Med Chem Lett, 2006. **16**(21): p. 5594-7.

59.   Forrest, L.R., Y.-W. Zhang, M.T. Jacobs, J. Gesmonde, L. Xie, B.H. Honig, and G. Rudnick, *Mechanism for alternating access in neurotransmitter transporters.* Proceedings of the National Academy of Sciences, 2008.

60.   Singh, T., D. Biswas, and B. Jayaram, *AADS - An automated active site identification, docking, and scoring protocol for protein targets based on*

physicochemical descriptors. Journal of Chemical Information and Modeling, 2011. **51**(10): p. 2515-2527.

61. Chemical Computing Group, I. *Molecular operating environment (MOE), 2013.08*. 2013.

62. Sundriyal, S., S. Khanna, R. Saha, and P.V. Bharatam, *Metformin and glitazones: does similarity in biomolecular mechanism originate from tautomerism in these drugs?* Journal of Physical Organic Chemistry, 2008. **21**(1): p. 30-33.

63. Zhang, P., G. Cyriac, T. Kopajtic, Y. Zhao, J.A. Javitch, J.L. Katz, and A.H. Newman, *Structure−activity relationships for a novel series of citalopram (1-(3-(dimethylamino)propyl)-1-(4-fluorophenyl)-1,3-dihydroisobenzofuran-5-carbonitrile) analogues at monoamine transporters.* Journal of Medicinal Chemistry, 2010. **53**(16): p. 6112-6121.

64. Owens, M.J., D.L. Knight, and C.B. Nemeroff, *Second-generation SSRIs: human monoamine transporter binding profile of escitalopram and R-fluoxetine.* Biological Psychiatry, 2001. **50**(5): p. 345-350.

65. Zhou, D., G.P. Stack, J. Lo, A.A. Failli, D.A. Evrard, B.L. Harrison, N.T. Hatzenbuhler, M. Tran, S. Croce, S. Yi, J. Golembieski, G.A. Hornby, M. Lai, Q. Lin, L.E. Schechter, D.L. Smith, A.D. Shilling, C. Huselton, P. Mitchell, C.E. Beyer, and T.H. Andree, *Synthesis, potency, and in vivo evaluation of 2-piperazin-1-ylquinoline analogues as dual serotonin reuptake inhibitors and serotonin 5-HT1A receptor antagonists.* Journal of Medicinal Chemistry, 2009. **52**(15): p. 4955-4959.

66. Manepalli, S., L.M. Geffert, C.K. Surratt, and J.D. Madura, *Discovery of novel selective serotonin reuptake inhibitors through development of a protein-based pharmacophore.* Journal of Chemical Information and Modeling, 2011. **51**(9): p. 2417-2426.

67. Humphrey, W., A. Dalke, and K. Schulten, *Visual molecular dynamics.* Journal of Molecular Graphics, 1996. **14**: p. 33-38.

68. Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten, *Scalable molecular dynamics with NAMD.* Journal of Computational Chemistry, 2005. **26**(16): p. 1781-1802.

69. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: an N⊕log(N) method for Ewald sums in large systems.* The Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.

70. MacKerell, A.D., D. Bashford, Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera,

F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* The Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

71.   Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, *Comparison of simple potential functions for simulating liquid water.* The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.

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

73.   Zwanzig, R.W., *High-temperature equation of state by a perturbation method. I. Nonpolar gases.* The Journal of Chemical Physics, 1954. **22**(8): p. 1420-1426.

74.   Gilson, M.K., J.A. Given, B.L. Bush, and J.A. McCammon, *The statistical-thermodynamic basis for computation of binding affinities: a critical review.* Biophysical journal, 1997. **72**(3): p. 1047-1069.

75.   Hermans, J. and L. Wang, *Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme.* Journal of the American Chemical Society, 1997. **119**(11): p. 2707-2714.

76.   Bennett, C.H., *Efficient estimation of free energy differences from monte carlo data.* Journal of Computational Physics, 1976. **22**(2): p. 245-268.

77.   Liu, P., F. Dehez, W. Cai, and C. Chipot, *A toolkit for the analysis of free-energy perturbation calculations.* Journal of Chemical Theory and Computation, 2012. **8**(8): p. 2606-2616.

78.   Lu, N., D.A. Kofke, and T.B. Woolf, *Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods.* Journal of Computational Chemistry, 2004. **25**(1): p. 28-40.

79.   Beutler, T.C., A.E. Mark, R.C. van Schaik, P.R. Gerber, and W.F. van Gunsteren, *Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations.* Chemical Physics Letters, 1994. **222**(6): p. 529-539.

80.   Wang, H., A. Goehring, K.H. Wang, A. Penmatsa, R. Ressler, and E. Gouaux, *Structural basis for action by diverse antidepressants on biogenic amine transporters.* Nature, 2013. **503**(7474): p. 141-5.

81.   Cross, J.B., D.C. Thompson, B.K. Rai, J.C. Baber, K.Y. Fan, Y. Hu, and C. Humblet, *Comparison of several molecular docking programs: pose prediction and virtual screening accuracy.* J Chem Inf Model, 2009. **49**(6): p. 1455-74.

82.   Jain, A.N., *Effects of protein conformation in docking: improved pose prediction through protein pocket adaptation.* J Comput Aided Mol Des, 2009. **23**(6): p. 355-74.

5. CHAPTER 5

BINDING OF R-FLUOXETINE IN hSERT

5.1. Introduction

Fluoxetine, popularly known as Prozac, is a widely used drug to treat depression. It was first developed by Eli Lilly and approved by the FDA in 1987 to treat major depression [1]. Even today, it ranks third among the antidepressant sales in the United States. Also, it was the first selective serotonin reuptake inhibitor (SSRI) discovered [1]. Fluoxetine is a racemic mixture of R- and S- enantiomers, and both are equally active pharmacologically [2]. They are known to exert their function by blocking the hSERT. Several aspects of fluoxetine were well established through various studies; however, its mechanism of action, binding region, and binding mode in hSERT are still debated [3].

hSERT is a membrane protein that controls the levels of serotonin in the synaptic cleft, thereby maintaining homeostasis in the body. hSERT is the primary target for several antidepressants [4]. SSRIs are the most effective and highly prescribed antidepressants to date and act by selectively inhibiting hSERT [5]. However, the molecular basis for their selectivity, efficacy, location and structure of the binding regions is unknown and highly debated [3]. For example, it was proposed based on several studies that antidepressants bind in the S1 site of hSERT [6-10] and inhibit the protein in a competitive manner [11-15]. There were studies supporting a second hypothesis that antidepressants bind in the S2 site of hSERT and inhibit the protein in a noncompetitive manner [16, 17]. According to the third hypothesis, drugs bind in both

the sites and inhibit the protein through allosterism [18-21].

In this study, we explored the binding pocket and the binding mode of R-fluoxetine (Fig. 5.1) in hSERT using a novel computational protocol described in chapter four. Binding pockets are key for the rational design of novel therapeutics in structure-based drug design. Through this protocol, we demonstrated that R-fluoxetine binds in the S2 site of hSERT with a pose similar to that of R-fluoxetine in LeuT.
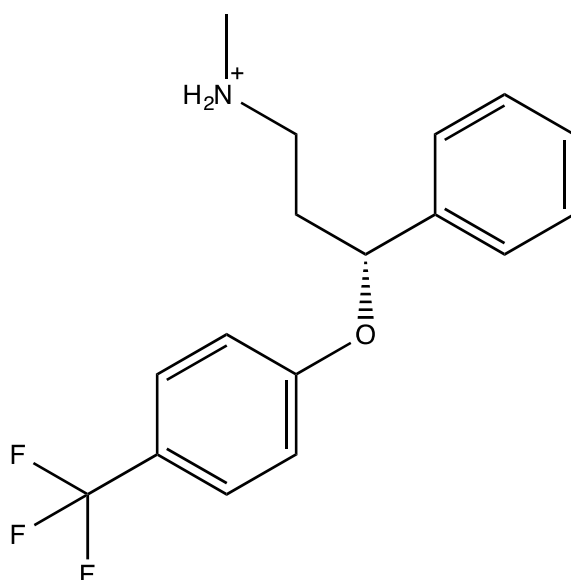
Figure 5.1. Chemical structure of R-fluoxetine.

5.2. Materials and Methods

5.2.1. Computational protocol

The computational protocol employed in chapter four was adopted for this study as well. Please refer to chapter four for complete details.

5.2.1.1.   System preparation and MD simulation details

A strategy similar to what was described in chapter four was followed. Each pose was simulated at least for 50ns in periodic boundary conditions prior to ABFE calculations. The CHARMM force field for protein, water, and ions [22, 23], and the CGenFF [24] parameters for R-fluoxetine were used.

5.2.1.2.   Absolute binding energy calculations

A method similar to what was described in chapter four was used. Please see chapter four for details. For calculating $\Delta_\square$, a 10ns simulation in each direction was performed, and the entire simulation was divided into 150 windows. A 1fs time step was employed. In the case of $\Delta_\square 1$ calculation, a 1ns simulation in each direction was carried, and the entire simulation was split into 146 windows. A 1fs time step was utilized. While calculating $\Delta_\square 1$, positional restraints (force constant = 0.5 kcal/mol.Å^2) were imposed on the ligand to prevent it from escaping from the pocket. The positional restraints were described in detail in chapter 4. The protocol was identical for all four poses of R-fluoxetine.

5.2.1.3.   Relative binding energy calculations

The RBFE change ($\Delta\Delta G$) in mutating the protein was calculated based on the thermodynamic cycle shown in Figure 5.2 and uses equation 1.

$$\Delta\Delta G = \Delta G_4 - \Delta G_3 = \Delta G_2 - \Delta G_1 \qquad \text{- Equation 1}$$

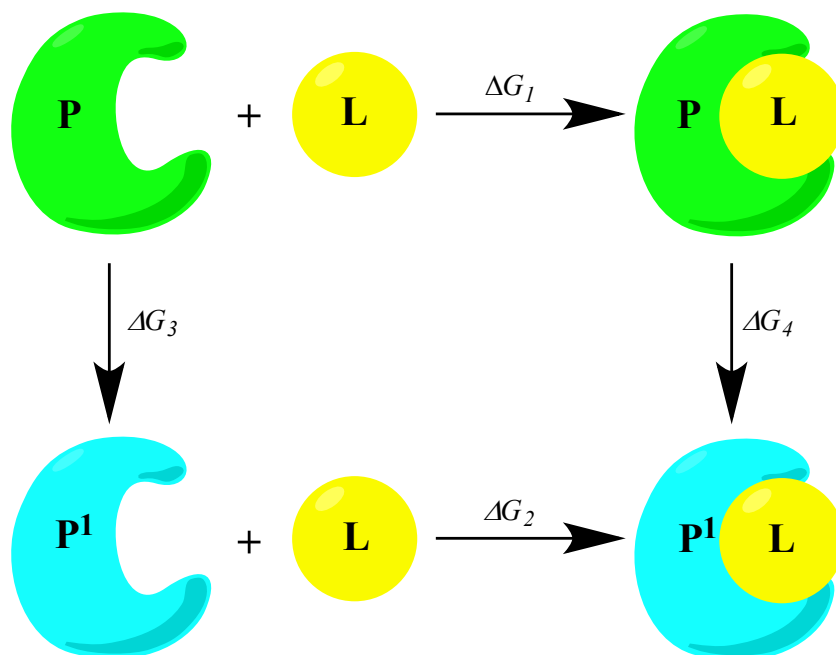All the terms in equation 1 were explained in Figure 5.2.

Figure 5.2. Thermodynamic cycle for estimating the impact of mutation in a protein. P and P^1 are the wild type and mutated proteins, and L is a ligand. ΔG_1 and ΔG_2 are the ABEs of L-P and L-P^1 complexes. ΔG_3 and ΔG_4 are the free energy changes of mutating P to P^1 in the water box and in the presence of L.

F□□ ΔG_4 calculation, in each direction a 20ns simulation was performed, and the entire simulation was split into 146 windows. For calculating ΔG_3, a 23ns simulation in either direction was performed, and 166 windows were used. A 1fs time step was used. □□□□□□□□□imental RBFE ($\Delta\Delta G_{Exp}$) was estimated using equation 2.

$$\Delta\Delta G_{Exp} = RT \ln\left(\frac{IC_{50}^{wild}}{IC_{50}^{mutated}}\right)$$ - Equation 2

IC_{50}^{wild} and $IC_{50}^{mutated}$ are the half maximal inhibitory concentrations of ligand L with the wild type and mutated proteins.

5.3. Results and Discussion

5.3.1. AADS results

After exploring the binding pockets of hSERT through AADS [25], we docked R-fluoxetine in the top 10 pockets of hSERT using AADS docking and scoring protocol. We analyzed the top 10 AADS poses of R-fluoxetine, and eight poses were binding in the regions of interest. Seven of the eight poses were binding in the S2 pocket of hSERT (Fig. 5.3), and one pose ended in the S1 site. Of the seven poses binding in the S2 site, four poses were binding deep in the S2 site and the other three poses were binding towards the extracellular side. The four poses (ranked 3, 4, 5, and 6) binding deep in the S2 pocket were binding closer to the extracellular hydrophobic gate. This is similar to R-fluoxetine binding in the LeuT crystal structure, where the drug binds deep in the S2 pocket of LeuT [16]. The trifluoromethyl groups of the four deep binding poses were binding in the HBP (circled in red in Fig. 5.3), although the orientation of the other two moieties in the poses varies. The trifluoromethyl moiety was identified as key for the specificity of R-fluoxetine at hSERT by Zhou et al. [16] The HBP is formed by residues Leu-99, Gly-100, Trp-103, Arg-104, Tyr-176, Ile-179, and Phe-335. Pose-4 was strikingly similar to R-fluoxetine orientation in LeuT (Fig. 5.4); the amine group was pointed towards the extracellular gated residues Arg-104 and Glu-493, the phenyl moiety was oriented towards the extracellular side, and the trifluoromethyl group was pointed into the HBP (Fig. 5.4). This observation was interesting, particularly when the residues in the HBP are highly conserved between LeuT and hSERT. Only one residue differs between hSERT and LeuT in the HBP; position-29 was occupied by leucine in LeuT, and the corresponding position was occupied by tryptophan (Trp-103) in hSERT. The amine

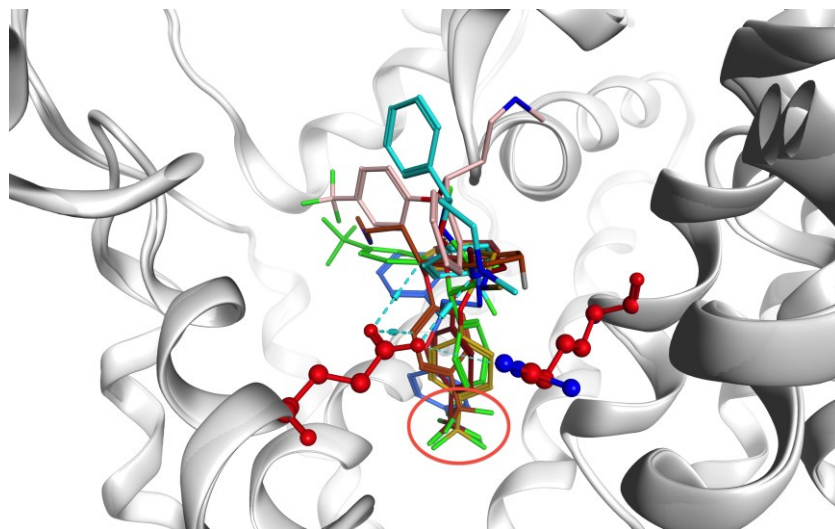nitrogen was interacting with E493, which is important for the higher affinity of R-fluoxetine with hSERT [16].

Figure 5.3. Seven AADS poses of R-fluoxetine in the S2 pocket. The trifluoromethyl moieties of four poses binding deep into the pocket are circled in red. The gated residues are represented as red balls and sticks.
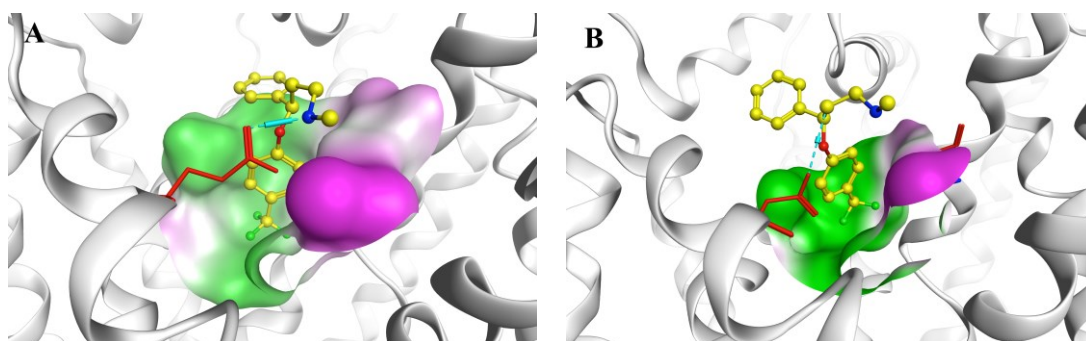

Figure 5.4. Pose-4 (A) of R-fluoxetine in hSERT vs. R-fluoxetine in LeuT (B). The R-fluoxetine is represented as a yellow ball and stick and the gated residue E493 is represented as a red stick. The HBP is represented as a multicolored surface in both hSERT and LeuT, with green and pink colors representing the hydrophobic and hydrophilic areas of the pocket, respectively. The trifluoromethyl group of R-fluoxetine is pointed into the HBP in both LeuT (B) and hSERT (A), whereas the amine nitrogen of pose-4 is interacting directly with E493 of hSERT (A).

It is worth mentioning that residues Glu-493 and Ile-179 in the S2 pocket were identified as key for the binding of R-fluoxetine in hSERT through site-directed mutagenesis and pharmacological binding assays [16]. The pose that binds closer to the S1 pocket did not interact with any residues.

To further test the validity of these results, we calculated ABEs using the FEP approach. We chose four different poses from the S2 site (poses 1, 2, 4, and 10 (Fig. 5.6)) for ABFE calculations. Prior to ABFE calculations, we performed MD simulations to test the stability of poses 1, 2, 4, and 10 in the pocket. Pose-4 in the MD simulations moved up and slightly reoriented in the S2 site with its trifluoromethyl group binding deep into the HBP, the phenyl moiety tilted 45 degrees and oriented towards the extracellular side, and the protonated amine remained interacting with the Glu-493 (Fig. 5.5).
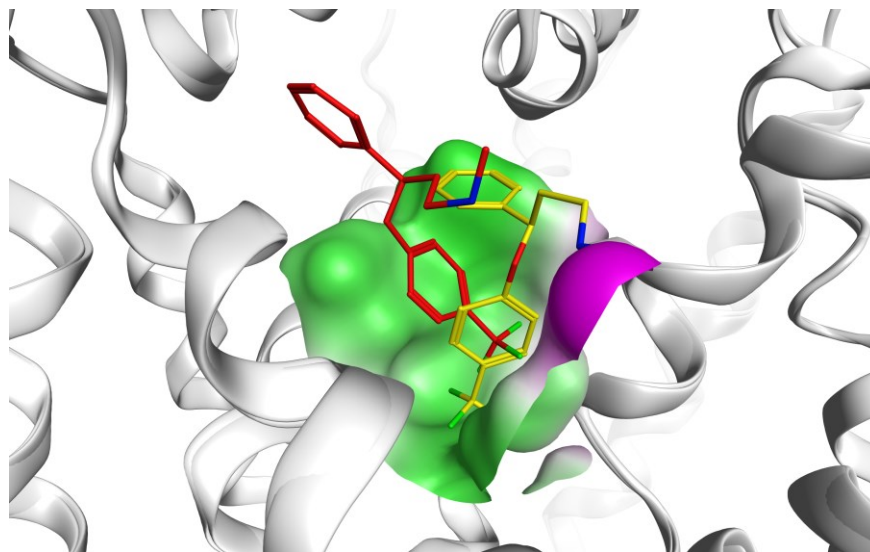
Figure 5.5. Docking pose vs. MD snapshot of Pose-4 of R-fluoxetine. The docking pose and the MD snapshot are represented as yellow and red colored sticks. The HBP is represented as a multicolored surface.
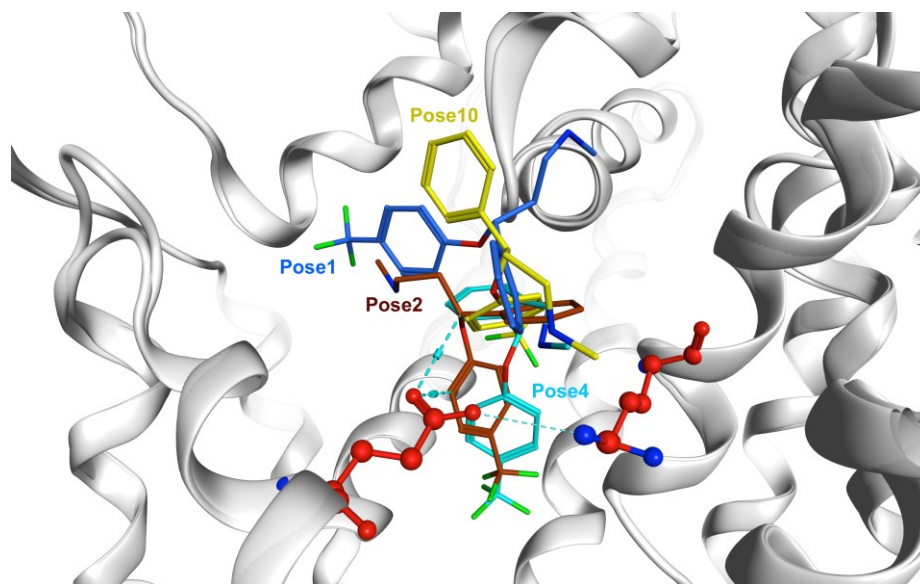
5.3.2. Absolute binding energy calculations

Figure 5.6. The four poses considered for MD simulations and FEP calculations. Poses 1, 2, 4, and 10 are represented as blue, brown, cyan, and yellow colored sticks, respectively. The gated residues are represented as red colored ball and sticks.

Next, we calculated the ABEs of poses 1, 2, 4, and 10. The calculated binding affinity of pose-4 was in good correlation with the experimental value (Table 5.1). Poses 1 and 10 were not at all favorable for binding, whereas pose-2 has minimal affinity for hSERT (Fig. 5.1).

Table 5.1. Absolute binding energies of poses 1, 2, 4, and 10 with hSERT.

Pose	Δ□₀			Δ□₁			ΔG^Rest	ΔG_Abs
	□□□	B□□□	B□□	□□□	B□□□	B□□		
1				26.95	-30.66	27.33 ± 0.20		2.04 ± 0.30
2				32.00	-34.31	32.00 ± 0.18		-2.63 ± 0.28
4	28.07	-28.87	26.70 ± 0.1	43.30	-43.78	42.53 ± 0.13	2.67	-13.16 ± 0.23
10				29.80	-27.52	27.10 ± 0.19		2.27 ± 0.29

The experimental binding energies (ΔG_{Expt}) from various sources are -12.71 ± 0.05 kcal/mol [26], -11.90 ± 0.10 kcal/mol [2], and -11.65 ± 0.08 kcal/mol [27], respectively. All energies are in kcal/mol.

Table 5.2. Mutation of E493-to-Q.

ΔG₄			ΔG₃			ΔΔG_Calc	ΔΔG_Expt
For	□□□□	□□□	For	□□□□	□□□		
89.86	-83.08	82.65 ± 0.12	82.18	-82.53	81.94 ± 0.13	0.71 ± 0.25	0.75 ± 0.29 [16]

All energies are in kcal/mol.

5.3.3. Mutation of glutamic acid 493 to glutamine

To validate the orientation of pose-4 in the S2 site, we mutated the gated residue Glu-493 to glutamine. Since Glu-493 interacts with the protonated amine through a salt bridge (Fig. 5.7), mutation of this residue to neutral glutamine should impact the binding of R-fluoxetine in the S2 site. Zhou et al. showed through experimental mutagenesis that this mutation was unfavorable for the binding of R-fluoxetine in hSERT.

We calculated the RBFE change for this mutation using the FEP method. The calculated $\Delta\Delta G$ was in good correlation with the experimental value (Table 5.2). The positive value of $\Delta\Delta G$ indicates that this mutation was not favorable for the binding of R-fluoxetine in hSERT. This shows that Glu-493 is crucial for the binding of R-fluoxetine in the S2 site and also confirmed pose-4 as the right orientation of R-fluoxetine in hSERT.
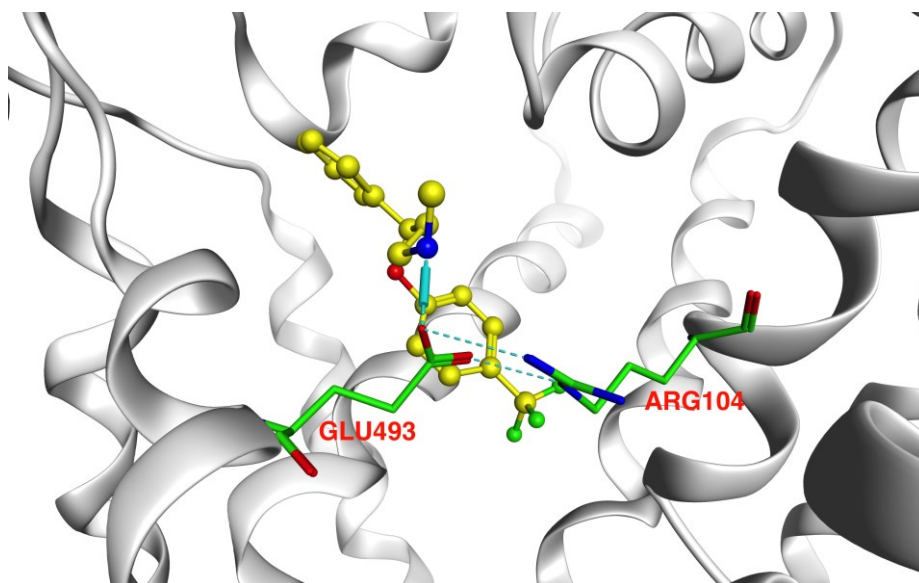

Figure 5.7. Pose-4 in the S2 site interacting with the gated residue E493. Pose-4 is represented as yellow balls and sticks and the gated residues are represented as green sticks.
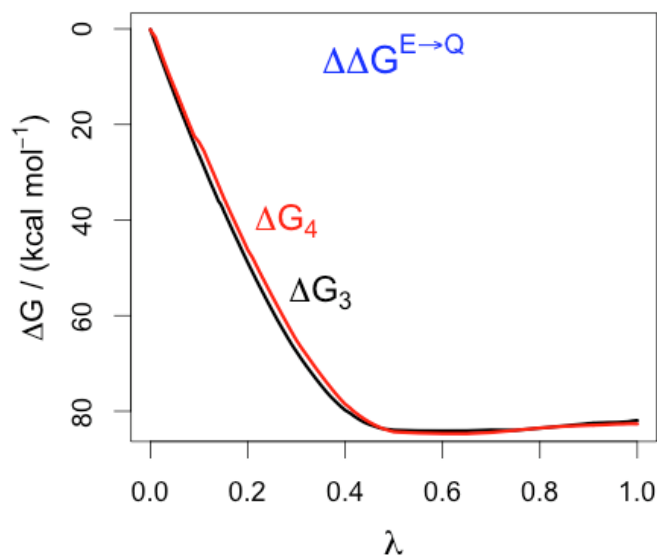
Figure 5.8. ΔG vs. λ ▯▯▯▯E493Q mutation in the presence of ▯▯▯▯-4▯
Red and black curves represent ΔG_4 and ΔG_3.

Overall, our results agree with Zhao et al. that R-fluoxetine binds in the S2 pocket of hSERT [16].

5.4. Conclusions

In this study, we explored the binding site and binding mode of R-fluoxetine using a novel computational protocol. We chose four different AADS poses of R-fluoxetine in the S2 site, performed MD simulations and finally calculated the ABEs of all four poses using FEP approach. The calculated binding energy of Pose-4 was in good agreement with the experimental value; the other three poses did not show appreciable affinity for the S2 site. In MD simulations, pose-4, slightly reoriented with its trifluoromethyl moiety, nicely accommodated into the halogen-binding pocket and the amine interacting with the charged extracellular gated residue E493. This pose resembles the R-fluoxetine

212

binding mode in LeuT crystal structure. To further validate the orientation of pose-4 in hSERT, we mutated E493-to-Q in the presence of pose-4 using FEP method. Mutation studies suggested that this change was unfavorable and confirmed pose-4 as the right orientation. Based on our findings, we conclude that R-fluoxetine binds in the S2 site of hSERT with similar orientation as in LeuT. However, recent literature reports support the alternate hypothesis that R-fluoxetine binds in the S1 site of hSERT [28-30]. Therefore, further studies are needed to clear this ambiguity regarding the binding site and orientation of R-fluoxetine in hSERT. One possible strategy is to dock the ligand in the S1 site, retrieve the poses proposed in the literature, calculate the ABEs of these poses and compare them against the experimental binding affinity. Also, S-fluoxetine can be included in this study since the experimental binding affinity is available; earlier studies suggested that it binds in the same site as R-fluoxetine.

5.5. References

1.   Immadisetty, K., L.M. Geffert, C.K. Surratt, and J.D. Madura, *New design strategies for antidepressant drugs.* Expert Opin Drug Discov, 2013. **8**(11): p. 1399-414.

2.   Hess, M., G. Höfner, and K.T. Wanner, *Cover Picture: (S)- and (R)-fluoxetine as native markers in mass spectrometry (MS) binding assays addressing the serotonin transporter.* ChemMedChem, 2011. **6**(10): p. 1753-1753.

3.   Rudnick, G., *What is an antidepressant binding site doing in a bacterial transporter?* ACS Chemical Biology, 2007. **2**(9): p. 606-609.

4.   Andersen, J., A.S. Kristensen, B. Bang-Andersen, and K. Stromgaard, *Recent advances in the understanding of the interaction of antidepressant drugs with serotonin and norepinephrine transporters.* Chemical Communications, 2009(25): p. 3677-3692.

5.   Frazer, A., *Pharmacology of antidepressants.* Journal of Clinical Psychopharmacology, 1997. **17**(2): p. 2S-18S.

6. Andersen, J., L. Olsen, K.B. Hansen, O. Taboureau, F.S. Jorgensen, A.M. Jorgensen, B. Bang-Andersen, J. Egebjerg, K. Stromgaard, and A.S. Kristensen, *Mutational mapping and modeling of the binding site for (S)-citalopram in the human serotonin transporter.* Journal of Biological Chemistry, 2010. **285**(3): p. 2051-2063.

7. Sinning, S., M. Musgaard, M. Jensen, K. Severinsen, L. Celik, H. Koldso, T. Meyer, M. Bols, H.H. Jensen, B. Schiott, and O. Wiborg, *Binding and orientation of tricyclic antidepressants within the central substrate site of the human serotonin transporter.* Journal of Biological Chemistry, 2010. **285**(11): p. 8363-8374.

8. Koldso, H., K. Severinsen, T.T. Tran, L. Celik, H.H. Jensen, O. Wiborg, B. Schiott, and S. Sinning, *The two enantiomers of citalopram bind to the human serotonin transporter in reversed orientations.* Journal of the American Chemical Society, 2010. **132**(4): p. 1311-1322.

9. Andersen, J., O. Taboureau, K.B. Hansen, L. Olsen, J. Egebjerg, K. Stromgaard, and A.S. Kristensen, *Location of the antidepressant binding site in the serotonin transporter.* Journal of Biological Chemistry, 2009. **284**(15): p. 10276-10284.

10. Andersen, J., N. Stuhr-Hansen, L. Zachariassen, S. Toubro, S.M.R. Hansen, J.N.N. Eildal, A.D. Bond, K.P. Bogeso, B. Bang-Andersen, A.S. Kristensen, and K. Stromgaard, *Molecular determinants for selective recognition of antidepressants in the human serotonin and norepinephrine transporters.* Proceedings of the National Academy of Sciences, 2011.

11. Hyttel, J., *Neurochemical characterization of a new potent and selective serotonin uptake inhibitor: Lu 10-171.* Psychopharmacology (Berl), 1977. **51**(403537): p. 225-233.

12. Talvenheimo, J., P.J. Nelson, and G. Rudnick, *Mechanism of imipramine inhibition of platelet 5-hydroxytryptamine transport.* Journal of Biological Chemistry, 1979. **254**(11): p. 4631-5.

13. Thomas, D.R., D.R. Nelson, and A.M. Johnson, *Biochemical effects of the antidepressant paroxetine, a specific 5-hydroxytryptamine uptake inhibitor.* Psychopharmacology, 1987. **93**(2): p. 193-200.

14. Apparsundaram, S., D.J. Stockdale, R.A. Henningsen, M.E. Milla, and R.S. Martin, *Antidepressants targeting the serotonin reuptake transporter act via a competitive mechanism.* Journal of Pharmacology and Experimental Therapeutics, 2008. **327**(3): p. 982-990.

15.   Humphreys, C.J., J. Levin, and G. Rudnick, *Antidepressant binding to the porcine and human platelet serotonin transporters*. Molecular Pharmacology, 1988. **33**(6): p. 657-663.

16.   Zhou, Z., J. Zhen, N.K. Karpowich, C.J. Law, M.E.A. Reith, and D.-N. Wang, *Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures*. Nat Struct Mol Biol, 2009. **16**(6): p. 652-657.

17.   Sarker, S., R. Weissensteiner, I. Steiner, H.H. Sitte, G.F. Ecker, M. Freissmuth, and S. Sucic, *The high-affinity binding site for tricyclic antidepressants resides in the outer vestibule of the serotonin transporter*. Molecular Pharmacology, 2010. **78**(6): p. 1026-1035.

18.   Chen, F., M.B. Larsen, C. Sanchez, and O. Wiborg, *The S-enantiomer of R,S-citalopram, increases inhibitor binding to the human serotonin transporter by an allosteric mechanism. Comparison with other serotonin transporter inhibitors.* European neuropsychopharmacology, 2005. **15**(2): p. 193-198.

19.   Chen, F., M.B. Larsen, H.A. Neubauer, C. Sánchez, P. Plenge, and O. Wiborg, *Characterization of an allosteric citalopram-binding site at the serotonin transporter.* Journal of Neurochemistry, 2005. **92**(1): p. 21-28.

20.   Plenge, P., U. Gether, and S.r.G. Rasmussen, *Allosteric effects of R- and S-citalopram on the human 5-HT transporter: Evidence for distinct high- and low-affinity binding sites.* European Journal of Pharmacology, 2007. **567**(1‚Äì2): p. 1-9.

21.   Plenge, P. and O. Wiborg, *High- and low-affinity binding of S-citalopram to the human serotonin transporter mutated at 20 putatively important amino acid positions.* Neuroscience Letters, 2005. **383**(3): p. 203-208.

22.   Mackerell, A.D., *Empirical force fields for biological macromolecules: overview and issues.* Journal of Computational Chemistry, 2004. **25**(13): p. 1584-1604.

23.   MacKerell, A.D., D. Bashford, Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* The Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

24.   Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A.D. Mackerell, *CHARMM general force field: a force field for drug-like molecules compatible with the*

CHARMM all-atom additive biological force fields. Journal of Computational Chemistry, 2010. **31**(4): p. 671-690.

25.   Singh, T., D. Biswas, and B. Jayaram, *AADS - An automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors.* Journal of Chemical Information and Modeling, 2011. **51**(10): p. 2515-2527.

26.   Owens, M.J., D.L. Knight, and C.B. Nemeroff, *Second-generation SSRIs: human monoamine transporter binding profile of escitalopram and R-fluoxetine.* Biological Psychiatry, 2001. **50**(5): p. 345-350.

27.   Koch, S., K.W. Perry, D.L. Nelson, R.G. Conway, P.G. Threlkeld, and F.P. Bymaster, *R-fluoxetine increases extracellular DA, NE, as well as 5-HT in rat prefrontal cortex and hypothalamus: an in vivo microdialysis and receptor binding study.* Neuropsychopharmacology, 2002. **27**(6): p. 949-959.

28.   Tavoulari, S., L.R. Forrest, and G. Rudnick, *Fluoxetine (Prozac) binding to serotonin transporter is modulated by chloride and conformational changes.* The Journal of Neuroscience, 2009. **29**(30): p. 9635-9643.

29.   Wang, H., A. Goehring, K.H. Wang, A. Penmatsa, R. Ressler, and E. Gouaux, *Structural basis for action by diverse antidepressants on biogenic amine transporters.* Nature, 2013. **503**(7474): p. 141-5.

30.   Andersen, J., N. Stuhr-Hansen, L.G. Zachariassen, H. Koldso, B. Schiott, K. Stromgaard, and A.S. Kristensen, *Molecular basis for selective serotonin reuptake inhibition by the antidepressant agent fluoxetine (Prozac).* Mol Pharmacol, 2014. **85**(5): p. 703-14.

FUTURE WORK

Future work includes

A. Validating the CGenFF parameters of the drugs through calculation of solvation energies or pKa's and comparing them against the experimental numbers.

B. Competitive binding assays are to be performed to determine the mode of inhibition of DJLDU-3-79 and SSA-426. Also, the proposed binding modes of these two compounds need to be validated by site directed mutagenesis experiments.

C. Since the binding site of R-fluoxetine in hSERT is still debated, its binding in the S1 site can be clarified by generating various poses (in the S1 site) proposed in the literature though docking, performing MD simulations to test their stability in the site, and finally calculating the ABEs and comparing them to the experimental values. Also, the binding site of S-fluoxetine in hSERT can be elucidated using a similar approach.

6. A REVIEW OF NEW DESIGN STRATEGIES FOR ANTIDEPRESSANT DRUGS

6.1. Introduction

Depression is a chronic illness affecting millions of individuals and imposing staggering economic costs [1]. Only in the last half-century has a serious effort been made to develop pharmaceuticals specifically for this disease state; for thousands of years, natural product extracts containing plant alkaloids have been used to treat depression. Such extracts contain hundreds of chemical compounds, however, with some likely to cause unintended effects. As an example, the antidepressant properties of St. John's wort are generally ascribed to inhibition of the monoamine transporter (MAT) proteins, but additional compounds in the extract are responsible for adverse effects associated with altering metabolism (*e.g.*, decreasing contraceptive efficiency) [2]. Isolating the active compound(s) was the logical solution in improving on the natural product.

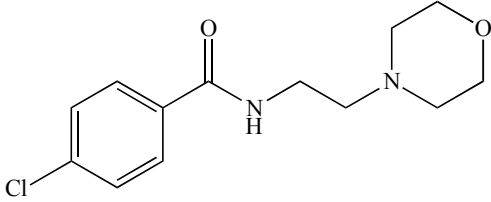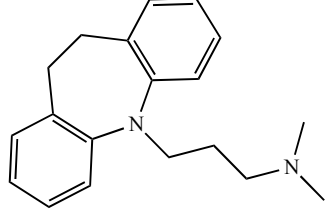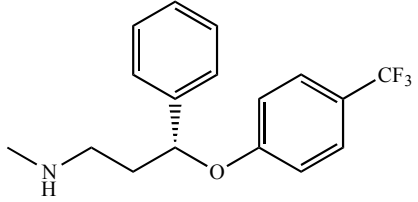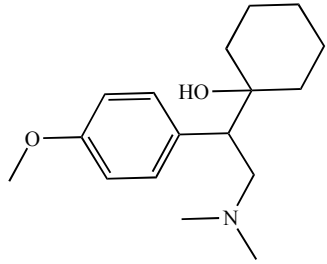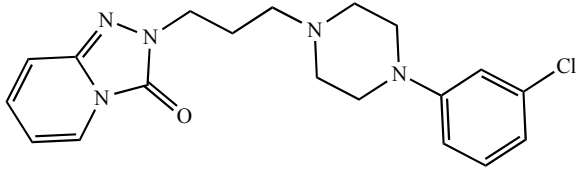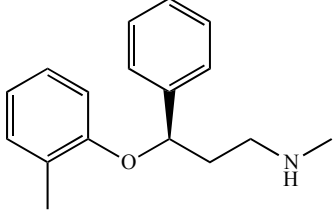### 6.1.1. First-generation antidepressant drugs

The era of antidepressant medications development began in the 1950s with iproniazid. A monoamine oxidase inhibitor (MAOI) initially developed to treat tuberculosis, iproniazid was serendipitously found to have mood enhancing properties

[3]. Later in the decade, the antidepressant actions of imipramine, the first of the tricyclic antidepressant (TCA) drug class, was also discovered through clinical practice (Fig. 6.3) [3, 4]. MAOIs extend the lifespan of the monoamine neurotransmitters serotonin (5-HT), norepinephrine (NE) and dopamine (DA) by interfering with their metabolism. TCAs extend the extraneuronal lifespan of 5-HT and NE by inhibiting their cognate neurotransmitter uptake transporter proteins. Unfortunately, these two drug classes are associated with often-intolerable adverse effects due to additional actions at cholinergic, histaminergic and adrenergic receptors. To retain the antidepressant benefits of increased monoamine levels, a new generation of antidepressant drugs was designed to inhibit either or both of the MAT proteins responsible for serotonin or norepinephrine uptake [5].

6.1.2. Second-generation antidepressant drugs

Which antidepressant drugs belong in the first-, second- or even a possible third-generation category is somewhat dependent on the literature source. The second generation is typically viewed as beginning with drugs that selectively inhibit one or more MATs (with minimal or no binding at monoaminergic receptors). Structurally diverse drugs and medications followed over the next 50 years that can be classified as selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs) or norepinephrine reuptake inhibitors (NRIs) (Table 6.1). Although probably not included in the second generation, triple-reuptake inhibitors (TRIs) are now a focus and moving toward FDA approval, due to recognition that dopamine also possesses antidepressant properties, especially in treating states involving anhedonia [6].

Table 6.1. Common classes of known antidepressants

Class	Structure Example	Generic Name	Trade Name
MAOI		Moclobemide	Aurorix
TCA		Imipramine	Tofranil
SSRI		Fluoxetine	Prozac
SNRI		Venlafaxine	Effexor
SARI		Trazodone	Desyrel
NRI		Atomoxetine	Strattera

MAOI, monoamine oxidase inhibitor; TCA, tricyclic antidepressant; SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin-norepinephrine reuptake inhibitor; SARI, serotonin antagonist and reuptake inhibitor; NRI, norepinephrine reuptake inhibitor.

Serotonin is the brain's chief mood regulator. This, coupled with the observations that 1) victims of violent suicides registered low CNS serotonin at autopsy [7, 8], and 2) serotonin biosynthesis precursors augmented MAOI antidepressant effects [9-11], drew the antidepressant focus to modulation of this neurotransmitter. Toward creating an antidepressant drug with TCA-like efficacy but without the cardiovascular effects that accompany anticholinergic or norepinephrine transporter (NET) inhibitors, Eli Lilly Corporation developed a new drug class in the 1970s: the SSRIs [12].

6.1.3. Rational design of an antidepressant drug: fluoxetine (Prozac™)

Despite its sedating effect, the histamine receptor antagonist diphenhydramine (Benadryl™) was chosen as a starting point for antidepressant rational drug design because of its enhancement of the pressor response to norepinephrine [13] (Fig. 6.1). The drug was also known to inhibit synaptic neurotransmitter uptake by at least one of the three plasma membrane MATs: hNET, hDAT and hSERT [14, 15]. To select for these properties while excluding anticholinergics lacking MAT inhibition, a diphenhydramine structure-activity series was synthesized and screened for the ability to block apomorphine-induced hypothermia in mice [13]. Compounds that activate norepinephrine or dopamine pathways block the hypothermia response; those that stimulate serotonin activity or antagonize acetylcholine do not affect the response [16, 17]. From this diphenhydramine series, the phenoxyphenylpropylamine (PPPA) compound LY86032 was found to block hypothermia, leading to identification of the highly NET-selective inhibitor (NRI) nisoxetine. The original diphenhydramine analogs that did not reverse the hypothermia were also tested for inhibition of one or more MATs, yielding the hSERT

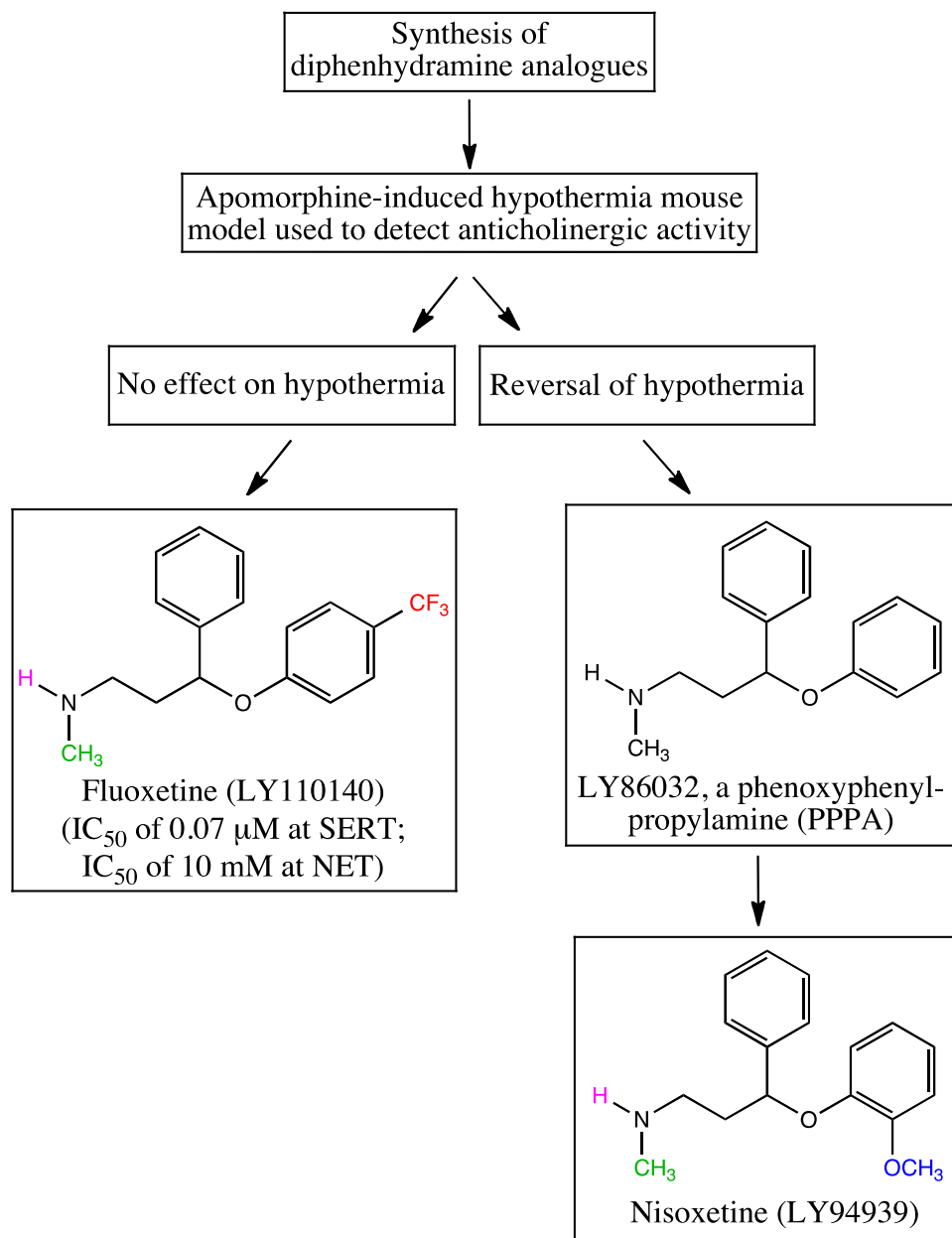inhibitor fluoxetine (Prozac™), the first SSRI FDA-approved (1987) for major depressive disorder.

Figure 6.1. Structure-activity relationship (SAR)-based drug discovery leading to fluoxetine (Prozac™), the first SSRI.

6.1.4. SSRI effects on systems served by 5-HT receptor subtypes

The advent of fluoxetine and other FDA-approved SSRIs was hailed as a profound advance in antidepressant brain receptor selectivity. Ironically, selective SERT inhibition leads to widespread side and adverse effects precisely because of nonselective serotonin actions. SSRI-induced increases in extracellular serotonin result in indiscriminant activation of all 14 discrete 5-HT receptor types; these mediate pathways in the brain, gut and other peripheral regions responsible for GI transit, sexual function, and various homeostatic functions including vascular tone, sleep/wakefulness and hunger. Thus, SSRIs are linked to cardiovascular problems, insomnia or hypersomnia, weight gain, GI distress and sexual dysfunction. The adverse effect may be severe enough to cause discontinuance of the SSRI [18]. This issue has prompted the goal of developing SSRIs or SNRIs (selective norepinephrine reuptake inhibitors) that modulate depression-relevant 5-HT receptor subtypes in a way that minimizes or eliminates the antidepressant's adverse effect. In fact, SSRI treatment-resistant outcomes may be due to excessive negative feedback on serotonin release via presynaptic $5HT_{1A}$ receptors [19, 20]. To address this, vilazodone (Viibryd^{TM}), the first FDA-approved SSRI/$5HT_{1A}$R partial agonist, was designed to boost extraneuronal serotonin levels [21]. Similarly, presynaptic $5HT_{1B}$ antagonists enhance SSRI activity [22]. While presynaptic $5HT_1$ antagonists alone have yielded antidepressant effects in animal models [23], efficacies are inferior to those seen with SSRIs [20].

Other serotonin receptor types hold promise as antidepressant drug targets. The 5-HT_2 and 5-HT_3 receptors are best characterized with respect to SSRI adverse effects. Activation of 5-HT_{2A} receptors is associated with hallucinations (LSD is an agonist at

this receptor); 5-HT$_{2C}$ mediates the vasoconstriction linked to cardiovascular and sexual dysfunction. Antagonism of these receptor subtypes would be expected to mitigate the adverse effects; the 5-HT$_{2A}$ antagonist ritanserin was observed to augment SSRI antidepressant and anxiolytic effects [24-26]. Nausea and prokinetic GI effects may accompany the 5-HT$_3$R activation via SSRIs. Ondansetron and other antagonists at this receptor have long been used as antiemetics for chemotherapy, and this drug class may have anxiolytic and antidepressant properties [20, 27, 28]. The 5-HT$_4$ receptor also mediates GI activity; however, this receptor subtype is distinct in that antidepressant effects are elicited by agonists, not antagonists [29]. The remaining serotonin receptor types are less well characterized, but 5-HT$_7$R antagonists are reported to possess antidepressant and anxiolytic activity [30, 31]. Mirtazepine (Remeron^{TM}) is touted as the only monoamine-based antidepressant on the market that does not have MAT inhibitor properties; however, many monoamine receptors are modulated, resulting in a plethora of adverse effects [32].

MAT inhibition is still desirable in an antidepressant in order to boost extracellular levels of the monoamine. Drugs lacking this capability are more likely to be used to augment a classic MAT inhibitor therapeutic [20]. The next evolution of monoamine-based antidepressant drugs will be to retain the monoamine transporter inhibition and add serotonin receptor-selective agonism or antagonism, as appropriate. Consistent with this goal, the drug vorioxetine currently in clinical trials is an SSRI and a full agonist at 5-HT$_{1A}$, partial agonist at 5-HT$_{1B}$, and antagonist at 5-HT$_{1D}$, 5-HT$_3$ and 5-HT$_7$ receptors [33]. This is not to imply that efficacies with respect to each receptor are

optimized in this drug. It is possible that the drug's intrinsic activity will need to be fine-tuned at each receptor, via structure-activity series.

6.1.5. Landmark achievements toward elucidating SERT and 5-HT receptor three-dimensional structures

Until relatively recently, a major obstacle for the design and development of new antidepressants was the limited knowledge of MAT protein structure and function. Lacking an x-ray crystal structure that could serve as a template, MAT structure had been inferential, relying on biophysical, pharmacological and molecular biological studies (*e.g.*, characterization of MAT protein site-directed mutants). Such techniques elucidated the general protein topology in the plasma membrane and indicated two or three probable substrate binding site residues, but offered little toward which transmembrane domains were juxtaposed, how substrates were translocated, or the size and shape of ligand binding pockets. In fact, the eventual availability of a credible 3D MAT computational model template revealed surprising and unforeseeable structural and functional features [34].

This SERT template was provided by the crystallization of the LeuT leucine transporter [34], a bacterial homolog of SERT. The LeuT and SERT proteins belong to the NSS family, sharing a 12 transmembrane domain structural architecture and Na^+-driven electrogenic transport of the neurotransmitter substrate. The proteins are hypothesized to have similar alternating access mechanisms [34, 35]. Similar to LeuT, hSERT is proposed to have at least two ligand binding pockets. The S1 (primary substrate) pocket is located midway through the protein and lipid bilayer. To the

extracellular side of S1 lies the S2 (secondary substrate) pocket (Fig. 6.2). Evidence supports a mechanism of the neurotransmitter substrate first binding in S2 before relocating to S1; as the S2 site is occupied by a second substrate molecule, the S1 ligand releases into the neuron [36-39]. Both pockets appear to accommodate inhibitors as well as substrates [36-40]. It should be noted, however, that the existence of the S2 site is not universally accepted [41, 42]. Several MAT protein computational models have been built based on the LeuT x-ray structure [43]. LeuT-based computational modeling is rapidly advancing the understanding of SERT ligand binding [44-52] as well as the large conformational changes that take place during substrate translocation [48, 53-57].

Our knowledge of monoamine G protein-coupled receptor (GPCR) structures has experienced a similar breakthrough. Crystal structures for the □1-adrenergic, dopamine D3, 5-HT1B and 5-HT2B receptors have been reported in the last five years [58-60]. The antagonist eticlopride was co-crystallized with the D3 receptor, providing a map for small molecule antagonist ligand binding within the seven transmembrane barrel of canonical GPCRs. The 5-HT receptor crystals included ergotamine analogs in the orthosteric agonist binding site. Thus, creation of useful monoamine receptor computational models is possible, even straightforward.
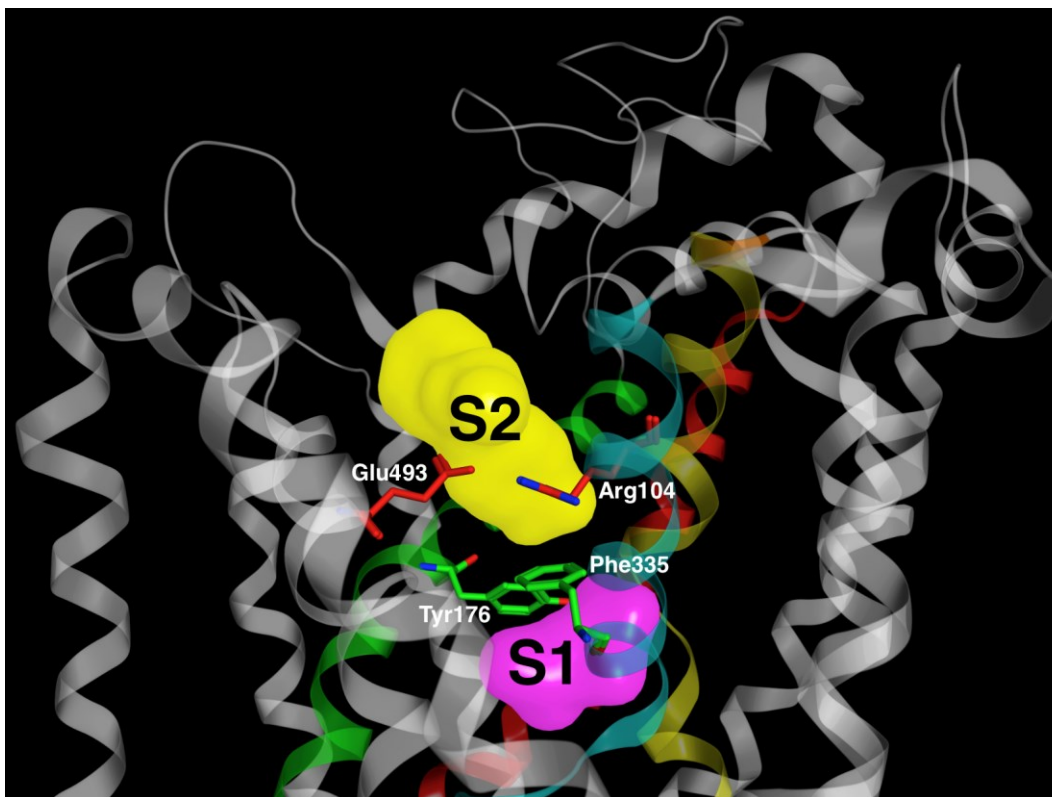
Figure 6.2. Primary (S1) and secondary (S2) substrate binding pockets of the SERT. LeuT and SERT studies support a primary substrate binding pocket (S1, pink) midway through the lipid bilayer. A secondary substrate pocket (S2, yellow) in the extracellular vestibule, to the extracellular side of S1, appears to serve as a staging area before substrate progression to S1. The pockets are separated by an extracellular gate consisting of a TM 1 (Arg104) – TM 10 (Glu493) salt bridge (red) and a TM 3 (Tyr176) – TM 8 (Phe335) aromatic (green) interaction. SERT inhibitors have been mapped to both pockets; accordingly, each SERT ligand pocket has been utilized for VS. SERT TM domains 1 (yellow), 3 (green), 6 (cyan) and 8 (red) are highlighted.

6.1.6. Antidepressant mechanisms beyond the MATs

The boost in extraneuronal monoamine neurotransmitter levels afforded by monoamine-based antidepressant drugs accounts for only a portion of their therapeutic benefit. An intriguing linkage between therapeutic antidepressant activity and hippocampal neurogenesis has been established over the last decade [61-65]. Screening for neurogenic and/or neuroprotective compounds is a new focus of the field;

interestingly, certain FDA-approved antidepressants (*e.g.*, fluoxetine) already have such properties [66-68]. The search for new antidepressant drugs extends beyond the biogenic amines. Ketamine and other NMDA glutamate receptor antagonists elicit an antidepressant response within hours, as opposed to the weeks required of monoamine-based drugs [69]. Peripheral administration of brain-derived neurotrophic factor (BDNF) displays antidepressant properties in vivo [70]. While these approaches are an important part of the future of antidepressant research, this review centers on development of therapeutics that manipulate the activity of the serotonin transporter and selected serotonin receptors.

6.2. Methods for Antidepressant Discovery and Design

In the 1970s, antidepressant discovery efforts began to include a rational (computer-aided) design strategy. Several techniques have evolved and advanced the drug discovery process, and can be broadly classified as experimental, knowledge-based, and computational methods (Fig. 6.3).

Figure 6.3. Different methods for antidepressant discovery/design. Antidepressant discovery/design methods can be classified as experimental (combinatorial chemistry and high-throughput screening), knowledge-based (SAR and molecular hybridization), or computational (structural, ligand, and hybrid (structure/ligand). There are also cases of antidepressant discovery via serendipity or clinical practices. SAR, structure activity relationship; HTS, high-throughput screening; VS, virtual screening. References: [a][2], [b,c][3], [d][71], [e][72], [f][73], [g][74], [h][75], [i][76], [j][77].

6.2.1. Experimental methods

Experimental methods include the popular techniques of combinatorial chemistry and high-throughput screening (HTS). These two methods revolutionized the lead compound discovery process in the late 1980s and early 1990s, shifting the bottleneck of the drug discovery process from lead identification to lead optimization. Combinatorial chemistry involves rapidly synthesizing a large library of structurally diverse compounds.

Advantages of combinatorial chemistry over systematic synthesis of individual compounds include rapid lead generation, structural diversity (essential for novel lead discovery), and cost effectiveness. Disadvantages of the method include low "hit" rate (sometimes precisely because of the increased diversity), complexity of structures, the presence of structurally impure compounds, occasional lack of drug-like features in a compound, and limited options for chemical modification [78-80]. The combinatorial approach has been used to identify a novel sulfonamide lead compound (designated "J20") with high binding affinity for the 5-HT$_7$ receptor (IC_{50} = 39 nM). J20 displayed significant antidepressant activity in the mouse Porsolt swimming and tail suspension tests relative to imipramine and fluoxetine [72].

HTS involves the rapid *in vitro* pharmacologic screening of a large chemical library for activity at a battery of classic drug receptor targets. HTS is the predominant process for lead compound identification. The chances of HTS yielding compounds active at one or more targets are very good, assuming that the concentration range is appropriate. The chief advantage of HTS is that large numbers of compounds can be reliably screened using an array of receptor targets. HTS limitations include the possibly prohibitive cost for an academic laboratory, the time required for large-scale screens, and that some active compounds identified may not be suitable for further drug development. Regarding antidepressant lead compounds, HTS successfully identified an entity with low nanomolar 5-HT$_{1A}$R affinity and high nanomolar SERT affinity. After SAR optimization, a derivative compound displayed increased affinity and selectivity for these targets, the desired intrinsic activity (0.2) for 5-HT$_{1A}$R agonists, and good oral bioavailability and blood-brain barrier penetration in the rat [71].

6.2.2. Knowledge-based methods

Structure-activity relationship (SAR) studies and the technique of molecular hybridization comprise the knowledge-based methods. These methods make use of information on already available chemical entities, endogenous substrates and biostructural data. SAR studies are widely used to optimize the activity of a hit or lead compound toward a specific target through generation of structural analogs, as few as 10 or as many as 1000. In this way, the combination of functional groups necessary for optimal drug activity is ascertained. As described above, fluoxetine (Prozac^{TM}) arose from diphenhydramine SAR studies [12]. Molecular hybridization involves generating a superior, chimeric bioactive compound by combining features from two or more bioactive compounds [81]. This relatively new rational drug design strategy may be used to improve receptor selectivity or create agents that can simultaneously act at multiple targets. Starting with a hDAT VS hit compound, molecular hybridization was recently used to create a high affinity SSRI lead compound [74]. The VS hit was combined with a portion of the dual SERT/5-HT$_{1A}$ inhibitor SSA-426 [82] to yield DJLDU-3-79, a compound with hSERT:hDAT and hSERT:hNET selectivity ratios of 50 and > 200, respectively [74]. Thus, molecular hybridization can be thought of as a type of directed SAR during the lead optimization stage of drug design. A common limitation of the knowledge-based methods is that, while effective in generating analogs of existing compounds, novel drug scaffolds are typically not suggested, or an outcome.

6.2.3. Computational methods

Computers are increasingly employed as an integral tool in the drug discovery process, and this trend will only accelerate as the level of drug receptor resolution increases. Because computational methods are expected to assume the forefront of drug discovery efforts, this approach is expanded upon here. High-throughput virtual screening (HTVS, or VS) is the *in silico* alternative to HTS. A computational screening technique, VS allows a database containing perhaps millions of virtual chemical structures to be scored and ranked based on predicted affinity for a specific protein or other drug target [83]. VS filters and removes obviously undesirable compounds, decreasing the burden of synthesizing trivial analogs. The use of large chemical libraries containing millions of small molecule compounds promotes structural diversity in the resultant hit compounds. The VS methods can be divided into ligand-based, structure-based, hybrid (ligand/structure) and fragment-based approaches (Fig. 6.4).

6.2.3.1. Ligand-based VS

The ligand-based VS methods are based on a similarity property principle [84]; virtual molecules are scored based on their relative similarity to a "model" ligand [85]. In the ligand-based pharmacophore approach, protein-ligand interactions are predicted based on the structures of established ligands, without knowledge of the target protein's structure. A pharmacophore is the spatial orientation of ligand features necessary for its biological activity [86, 87]. Each feature is represented as a sphere (*e.g.*, F1 - F5 in Fig. 2). The pharmacophore can be developed based on one or several ligands with the same biological activity. The ligand-based pharmacophore approach is a viable alternative for

obtaining structural information on binding sites in the absence of x-ray crystallographic structures of the drug target (*e.g.*, SERT) [88-90]. VS hit compounds that satisfy the pharmacophore filter are next confirmed as ligands via *in vitro* pharmacologic testing.

Knowledge of the bioactive conformation of at least one ligand is needed for the successful retrieval of novel scaffolds [91]. Enrichment can be improved by designing the reference ligand model based on a diverse ligand set [92]. This method has been successfully applied to discover NET [77, 93, 94], SERT [77, 93] and 5-HT$_{2C}$ [95] inhibitors. In summary, advantages of ligand-based VS are that the method is computationally less intensive compared to structure-based VS, and no knowledge of the target structure is required. Limitations include the reduced likelihood of finding a novel scaffold, performance largely depends on the initial reference ligand, and the requirement for at least one bioactive ligand conformation.

6.2.3.2. Structure-based VS

When structural information on the target protein is available, structure-based VS is the preferred computational drug discovery method. Structure-based VS relies on ligand docking to the target. Docking is a computational technique that assists in predicting the binding modes of ligands within the target protein pocket [96]. When crystal structures of the drug target are unavailable, as in the case of the MATs, homology models are substituted [97-99]. Homology modeling is used to generate high resolution 3D models of proteins with unknown structures, based primarily on alignment with one or more proteins of known 3D structure (templates) evolutionarily related to the target [100]. In the absence of an experimentally determined structure, a comparative

model can provide not only a starting point for experimentally validated research, but an evolving representation of the target [101, 102]. Generation of a reliable computational model is critical for the accuracy of the predictions to be made from the model.

Compounds from the virtual database are docked in the target protein site of interest and ranked via a scoring function; usually a force field-based measure of how well a ligand interacts with the protein. This constitutes a liability in the method, however. Currently available force fields for drug-like molecules are of limited accuracy, such that one has to rely on visual inspection for selecting the compounds for *in vitro* screening. Docking and subsequent scoring are the most important factors that determine the efficiency of any structure-based VS effort. A representative example is the discovery of lead compound PRX-93009 through structure-based VS using a 5-HT$_{1A}$ receptor 3D model [76]. This compound displayed high affinity binding at the 5-HT$_{1A}$ receptor ($K_i = 1$ nM), 65% efficacy, and a pharmacokinetic profile in rats comparable to buspirone, the only FDA approved 5-HT$_{1A}$ receptor agonist. Overall, the structure-based VS approach provides a reasonable estimate of both the ligand orientation within the receptor and the approximate binding affinity of the candidate ligand. Limitations of the method may include the absence of a target protein crystal structure, poor knowledge of the binding site, inaccuracies in the force field-based scoring functions, and variability as a function of the target protein conformation used in the docking studies [103]. Ensemble docking, which employs multiple conformations of the target protein, is suggested to improve enrichment in the structure-based VS studies [104].

6.2.3.3. Hybrid (structure/ligand) VS

In this conjugation of ligand- and structure-based VS, a pharmacophore is created within the ligand-binding pocket of the target structure. The small molecule database is screened for novel scaffolds, first using the ligand-based approach and next with the structure-based approach (Fig. 6.4). Although there are only two publications describing the application of this method to the MATs, novel hNET or hSERT inhibitors were reported [75, 105]. This method is further discussed below in the context of a specific multi-target antidepressant discovery strategy.
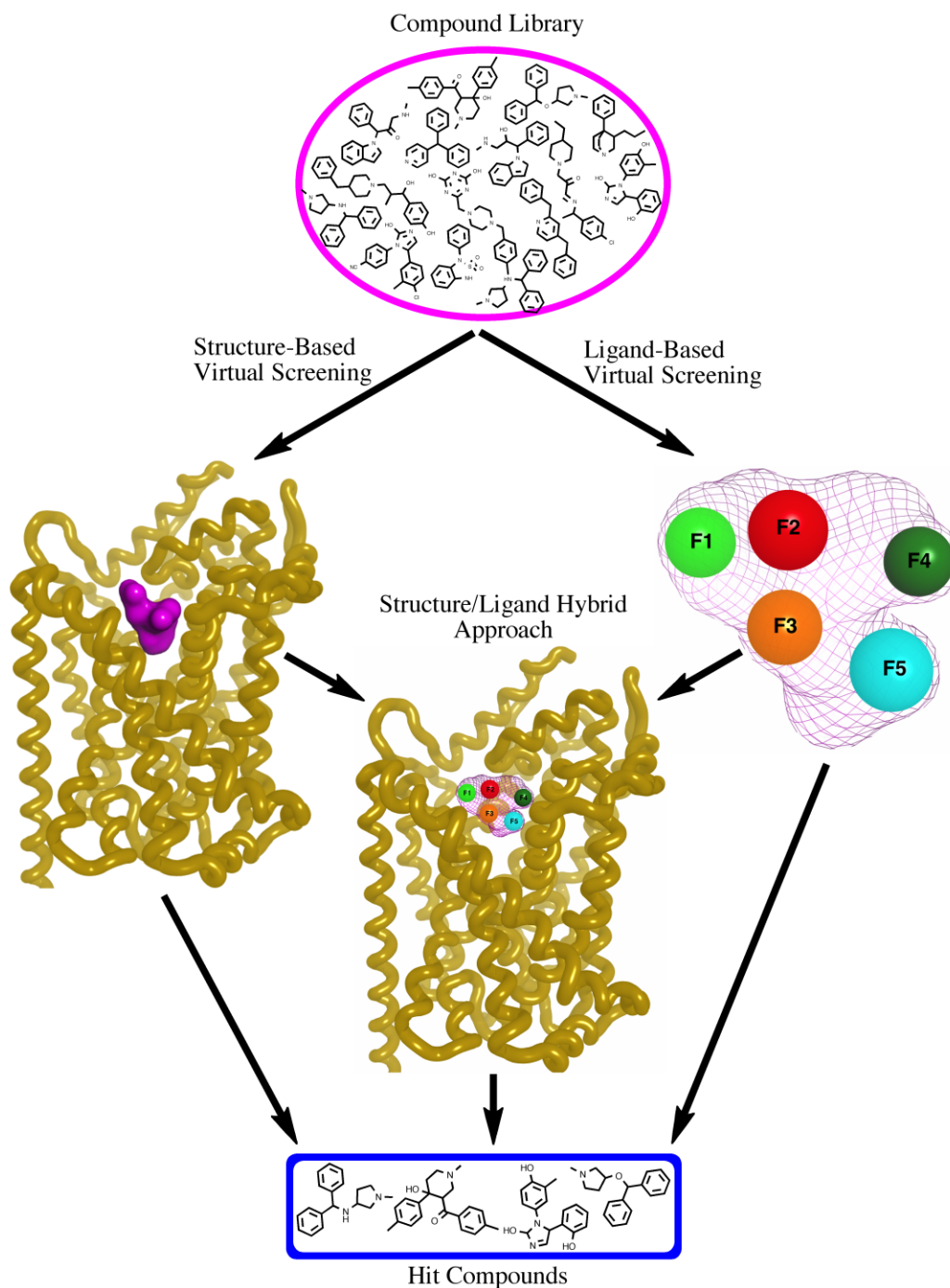
Figure 6.4. Structure- and ligand-based VS methods for discovery of novel ligands. A library of thousands to millions of compounds is mined for potential SERT ligands using as screening tools a computational SERT model (structure-based), a SERT inhibitor pharmacophore (ligand-based), or a combination of both. Features F1 – F5 of the pharmacophore indicate a requirement for ionic, hydrogen bonding, aromatic or otherwise hydrophobic interactions with SERT side chains. Compounds that align well with the pharmacophore and dock favorably within the binding pocket are next characterized with respect to in vitro SERT pharmacology.

6.2.3.4.   Fragment-based drug discovery

Perhaps the newest computational drug design strategy is fragment-based drug discovery (FBDD) [106], a powerful alternative to classic VS methods. Contrasting with the traditional approach of screening libraries of intact compounds, FBDD employs a large chemical fragment library. Fragment combinations are covalently linked within the receptor's ligand-binding pocket to create new entities, often of unique scaffold compared to the existing ligands of the target receptor [107]. Advantages of FBDD over classic VS include the ability to cover a large chemical space while offering greater chemical diversity and hit rates. FBDD also can offer smaller ligand sizes, more ligand-receptor interactions, and favorable pharmacokinetic properties [108, 109]. On the other hand, FBDD software is still in its infancy [110, 111], and FBDD hits are not always easily synthesizable [109].

6.2.4. Summary of VS comparison with HTS

To summarize the comparison of VS and HTS methods, VS is superior with respect to filtering the large chemical space, providing larger screening libraries, saving time and money, and allowing testing of theoretical compounds, perhaps containing novel scaffolds. Limitations of VS relative to HTS are overall reliability (considering the virtual nature of its elements), the considerable computational resources needed to search for 3D conformations of all molecules in the chemical library, and that compound synthesis and *in vitro* testing of VS hits are still required.

6.3. A Hybrid VS Strategy for Discovery of Novel SERT Inhibitors / 5-HT
   Receptor Modulators

The use of computational methods for antidepressant drug discovery is not only warranted but essential now that credible monoamine transporter and receptor models are finally available. Our group has undertaken a hybrid VS strategy involving the SERT and selected key 5-HTRs in the search for novel antidepressant leads (Fig. 6.5). To identify novel human SERT inhibitors, one million compounds from the ZINC structural database [112] were initially screened using the S2 pocket of our hSERT computational model (Fig. 6.2). As a first step, a structure-based pharmacophore filter developed using several established antidepressants was employed, resulting in approximately 4000 hit compounds. These hits were next allowed to dock with the hSERT model and assessed via Affinity dG scores and visual inspection. Visual inspection, a final, crucial step in the VS process, manually identifies hit compounds that display especially favorable intermolecular interactions with receptor pocket residues. Our visual inspection is based on five criteria: 1) the type and number of intermolecular interactions, 2) the binding pose, i.e. orientation and conformation, of the ligand, 3) key pocket residues are involved in the binding, 4) steric clashes, and 5) structural novelty of the VS hit. From the 68 hits obtained, the 15 most structurally diverse compounds were acquired for *in vitro* pharmacology. Two compounds, structurally distinct from current antidepressant medications, displayed low micromolar SERT binding affinities and substrate uptake inhibition potencies, and no detectable activity at the hNET or hDAT [75]. Micromolar affinities for VS hit compounds are not uncommon when homology models are employed

238

[113-119], especially given the low sequence identity (<30%) between LeuT and hSERT. Similarly, the first reported VS efforts based on a LeuT-directed hDAT homology model yielded hit compounds with low micromolar hDAT affinity [105] and slightly better (high nanomolar) SERT affinity [74]. More recently, the hybrid VS procedure described above has been conducted using the S1 SERT pocket. Several novel, *in vitro*-verified, VS hits of high nanomolar SERT affinity were obtained (unpublished data).

Compounds identified as novel SERT ligands can next be subjected to VS at 5-HT receptors relevant to the depressive state or responsible for the classic adverse effects; these include 5-HT$_{1A}$, 5HT$_{2A}$ and 5-HT$_3$ (Fig. 6.5). The resulting hit compounds would be tested for *in vitro* activity. Promising lead compounds are next optimized for target affinity and selectivity via SAR and computational FEP calculations, with the new analogs characterized *in vitro*, using animal models of depression, and hopefully one day in human clinical trials. Multi-site drugs (*e.g.*, hSERT/5-HT$_{1A}$/5-HT$_{2A}$/5-HT$_3$ ligands) developed through this approach are expected to serve as a new line of antidepressant medications carrying a more favorable side/adverse effect profile.

Figure 6.5. Multi-VS scheme toward discovery of antidepressant lead compounds. Using the structure/ligand-based hybrid approach, VS hits are pharmacologically characterized in vitro. Compounds that display specific binding to the SERT are next subjected to hybrid VS at 5-HT receptors associated with SSRI adverse effects. Resultant 5-HT receptor VS hits are tested *in vitro* for function (agonism or antagonism) as well as affinity, with promising compounds optimized via SAR and further tested in animal depression/anxiety models, and finally human clinical trials.

6.4. Conclusion

The development of antidepressants over the generations from monoamine oxidase inhibitors to selective serotonin reuptake inhibitors and beyond was discussed. Various drug design strategies were covered, with the major focus on computational VS approaches. The success of any drug discovery effort largely depends on the initial design strategy chosen. For decades, antidepressant discovery has benefited from combinatorial chemistry, HTS and SAR studies. These traditional methods must be augmented or improved upon to reduce the now-prohibitive cost of drug discovery. VS evolved as a powerful adjunct/alternative to *in vitro* HTS, and is evermore essential to the modern drug discovery project. Although there is not a single marketed antidepressant developed completely based on computer based drug design approaches, optimal use of computers, especially in lead identification, will enormously benefit drug discovery in terms of money, time, and novelty of the lead compounds. Finally, with the availability of reliable computer models of MATs and monoamine receptors, the use of computational strategies should enhance and facilitate the antidepressant drug discovery process.

6.5. Expert Opinion

Due in part to the many adverse effects associated with monoamine neurotransmitter-based antidepressant drugs, new therapeutics are in demand. In view of the astronomical cost to bring a successful drug to market, a new drug development approach is needed. Developing innovative approaches to assist in the development of novel and effective therapies to depression is a demanding research area. Searching for

therapies that will help reduce adverse effects with improved response time makes this a complex and challenging task. We are developing a novel computational and experimental approach at Duquesne University in Pittsburgh to meet the complex and challenging task. The novel approach is to derive new lead compounds by starting with VS of serotonin transporters and receptors, leading to pharmacological testing and synthesis, and ending with animal behavioral studies. This group is currently applying its workflow specifically to the serotonin transporters and receptors, with plans to extend it to other neurotransmitter transporter/receptor combinations (*e.g.,* glutamate).

VS of small molecule databases using computational drug and drug target models is a feasible, rapid and inexpensive alternative to the prevailing *in vitro* high-throughput screening approaches. Computational methodologies remain a predictive tool, but a tool that will inexorably dominate the drug discovery landscape as improvements are effected regarding specific issues such as incorporating target protein flexibility, refining water molecule placement, and improving scoring functions. The computational approach is further supplemented by pharmacological testing in order to support the computational results as well as assist in improving the computational models. Once a series of potential compounds [74, 75, 105] has been identified, two parallel approaches are undertaken. The first is the synthesis of analogs to improve affinity and selectivity [74], or to create a molecular probe by adding functional groups that can covalently attach to the transporter or receptor upon binding [120-122]. The second is to identify the specific binding site and important residues in the binding pocket using mass spectrometry methods. The final step is to test the effectiveness of the best compounds in animal models. In the authors' view, this joint computational and experimental effort in the development of future

antidepressants will hinge on close collaboration of different scientific disciplines. Finally, optimal use of the new research tools (*e.g.*, VS, FEP methods) offer greater predictive power and efficiency than the traditional reliance on serendipity and a structure-activity series. We see this team of academic scientists from diverse disciplines working on a common problem as the emerging approach for discovery of novel lead compounds. This method of drug discovery is similar to the idea behind small biotech companies that develop lead compounds and contract with large pharmaceutical companies to license their compounds.

There are several challenges that will impact the success or failure of the drug discovery process. One challenge is that a more complete and detailed understanding of the complex biochemical interactions and pathways is needed. Some of this will come as the field of systems biology develops. A second challenge is in the improvement of the computational tools. As an example, VS is heavily dependent on the filtering mechanism, comprised of a pharmacophore model and a scoring function. A ligand-based pharmacophore model depends on the alignment of substrates in three-dimensional space as well as accurate activity data. A structure-based pharmacophore model depends on the conformational state of the transporter, which at this time is not accurately known. Scoring functions are dependent on the terms used to define the scoring function as well as the data set used to train (*i.e.*, develop) the scoring function. One last challenge is mapping the transport mechanism. In this case long-time simulations with enhanced sampling methods to test the hypotheses of a one or two substrate transport mechanism. These challenges are presented because they are pursued by numerous academic [123, 124] and industrial [125-127] research groups.

Exciting developments on the drug discovery horizon include crowd computing (*e.g.*, Docking@Home [128]), cloud computing (*e.g.*, Amazon cloud [129]), and virtual collaborative organizations/research teams [130].

6.6. References

1. DEPRESSION: A Global Crisis; World Mental Health Day, 2012:[Available from: http://www.who.int/mental_health/management/depression/en/.

2. Hippius, H., *St john's wort ('hypericum perforatum') - a herbal antidepressant.* Current Medical Research and Opinion, 1998. **14**(3): p. 171-184.

3. Wong, M. and J. Licinio, *From monoamines to genomic targets: a paradigm shift for drug discovery in depression.* Nat Rev Drug Discov, 2004. **3**(2): p. 136-151.

4. Ban, T.A., *The role of serendipity in drug discovery.* Dialogues in Clinical Neuroscience, 2006. **8**(3): p. 335-344.

5. Gartlehner, G., R.A. Hansen, U. Reichenpfader, A. Kaminski, C. Kien, M. Strobelberger, M. Van Noord, P. Thieda, K. Thaler, and B. Gaynes, *Drug class review: second-generation antidepressants: final update 5 report.* 2011, Portland OR: 2011 Oregon Health & Science University.

6. Tran, P., P. Skolnick, P. Czobor, N.Y. Huang, M. Bradshaw, A. McKinney, and M. Fava, *Efficacy and tolerability of the novel triple reuptake inhibitor amitifadine in the treatment of patients with major depressive disorder: a randomized, double-blind, placebo-controlled trial.* Journal of Psychiatric Research, 2012. **46**(1): p. 64-71.

7. Shaw, D.M., F.E. Camps, and E.G. Eccleston, *5-Hydroxytryptamine in the hind-brain of depressive suicides.* Br J Psychiatry, 1967. **113**(505): p. 1407-11.

8. Bourne, H.R., W.E. Bunney, Jr., R.W. Colburn, J.M. Davis, J.N. Davis, D.M. Shaw, and A.J. Coppen, *Noradrenaline, 5-hydroxytryptamine, and 5-hydroxyindoleacetic acid in hindbrains of suicidal patients.* Lancet, 1968. **2**(7572): p. 805-8.

9. Glassman, A.H. and S.R. Platman, *Potentiation of a monoamine oxidase inhibitor by tryptophan.* Journal of Psychiatric Research, 1969. **7**(2): p. 83-88.

10. KLINE, N.S. and W. SACKS, *Relief of depression within one day using an M. A. O. inhibitor and intravenous 5-HTP.* Am. J. Psychiatry, 1963. **120**: p. 274-275.

11.   Coppen, A., D.M. Shaw, and J.P. Farrell, *Potentiation of the antidepressive effect of a monoamine-oxidase inhibitor by tryptophan.* Lancet, 1963. **1**(7272): p. 79-81.

12.   Wong, D.T., J.S. Horng, F.P. Bymaster, K.L. Hauser, and B.B. Molloy, *A selective inhibitor of serotonin uptake: lilly 110140, 3-(p-trifluoromethylphenoxy)-n-methyl-3-phenylpropylamine.* Life Sciences, 1974. **15**(3): p. 471-479.

13.   Wong, D.T., K.W. Perry, and F.P. Bymaster, *The discovery of fluoxetine hydrochloride (Prozac).* Nat Rev Drug Discov, 2005. **4**(9): p. 764-774.

14.   Molloy, B.B., D.T. Wong, and R.W. Fuller, *The discovery of fluoxetine.* Pharmaceutical News, 1994. **1**: p. 6-10.

15.   Yeh, S.Y., C. Dersch, R. Rothman, and J.L. Cadet, *Effects of antihistamines on 3, 4-methylenedioxymethamphetamine-induced depletion of serotonin in rats.* Synapse, 1999. **33**(3): p. 207-17.

16.   Schelkunov, E.L., *Efficacy of neuroleptics and antidepressants in the test of apomorphine hypothermia and some data concerning neurochemical mechanisms of the test.* Psychopharmacology, 1977. **55**(1): p. 87-95.

17.   Pawłowski, L. and H. Mazela, *Effects of antidepressant drugs, selective noradrenaline-or 5-hydroxytryptamine uptake inhibitors, on apomorphine-induced hypothermia in mice.* Psychopharmacology, 1986. **88**(2): p. 240-246.

18.   Ferguson, J.M., *SSRI antidepressant medications: adverse effects and tolerability.* Prim Care Companion J Clin Psychiatry, 2001. **3**(1): p. 22-27.

19.   Lemonde, S., G. Turecki, D. Bakish, L. Du, P.D. Hrdina, C.D. Bown, A. Sequeira, N. Kushwaha, S.J. Morris, A. Basak, X.-M. Ou, and P.R. Albert, *Impaired repression at a 5-hydroxytryptamine 1A receptor gene polymorphism associated with major depression and suicide.* The Journal of Neuroscience, 2003. **23**(25): p. 8788-8799.

20.   Artigas, F., *Serotonin receptors involved in antidepressant effects.* Pharmacology & Therapeutics, 2013. **137**(1): p. 119-131.

21.   Page, M.E., J.F. Cryan, A. Sullivan, A. Dalvi, B. Saucy, D.R. Manning, and I. Lucki, *Behavioral and neurochemical effects of 5-{4-[4-(5-cyano-3-indolyl)-butyl]-butyl]-1-piperazinyl}-benzofuran-2-carboxamide (EMD 68843): a combined selective inhibitor of serotonin reuptake and 5-hydroxytryptamine1A receptor partial agonist.* Journal of Pharmacology and Experimental Therapeutics, 2002. **302**(3): p. 1220-1227.

22. Gobert, A., A. Dekeyne, and M.J. Millan, *The ability of WAY100,635 to potentiate the neurochemical and functional actions of fluoxetine is enhanced by co-administration of SB224,289, but not BRL15572.* Neuropharmacology, 2000. **39**(9): p. 1608-1616.

23. Cryan, J.F., R.J. Valentino, and I. Lucki, *Assessing substrates underlying the behavioral effects of antidepressants using the modified rat forced swimming test.* Neuroscience & Biobehavioral Reviews, 2005. **29**(4–5): p. 547-569.

24. Bressa, G.M., S. Marini, and S. Gregori, *Serotonin S2 receptors blockage and generalized anxiety disorders. A double-blind study on ritanserin and lorazepam.* International journal of clinical pharmacology research, 1987. **7**(2): p. 111-9.

25. Shelton, R.C., G.D. Tollefson, M. Tohen, S. Stahl, K.S. Gannon, T.G. Jacobs, W.R. Buras, F.P. Bymaster, W. Zhang, K.A. Spencer, P.D. Feldman, and H.Y. Meltzer, *A novel augmentation strategy for treating resistant major depression.* American Journal of Psychiatry, 2001. **158**(1): p. 131-134.

26. Cremers, T.I.F.H., K. Rea, F.J. Bosker, H.V. Wikstrom, S. Hogg, A. Mork, and B.H.C. Westerink, *Augmentation of SSRI effects on serotonin by 5-HT2C antagonists: Mechanistic Studies.* Neuropsychopharmacology, 2007. **32**(7): p. 1550-1557.

27. Kos, T., P. Popik, M. Pietraszek, D. Schäfer, W. Danysz, O. Dravolina, E. Blokhina, T. Galankin, and A.Y. Bespalov, *Effect of 5-HT3 receptor antagonist MDL 72222 on behaviors induced by ketamine in rats and mice.* European Neuropsychopharmacology, 2006. **16**(4): p. 297-310.

28. Kelley, S.P., A.M. Bratt, and C.W. Hodge, *Targeted gene deletion of the 5-HT3A receptor subunit produces an anxiolytic phenotype in mice.* European Journal of Pharmacology, 2003. **461**(1): p. 19-25.

29. Lucas, G., V.V. Rymar, J. Du, O. Mnie-Filali, C. Bisgaard, S. Manta, L. Lambas-Senas, O. Wiborg, N. Haddjeri, G. Piñeyro, A.F. Sadikot, and G. Debonnel, *Serotonin4 (5-HT4) receptor agonists are putative antidepressants with a rapid onset of action.* Neuron, 2007. **55**(5): p. 712-725.

30. Wesołowska, A., E. Tatarczyńska, A. Nikiforuk, and E. Chojnacka-Wójcik, *Enhancement of the anti-immobility action of antidepressants by a selective 5-HT7 receptor antagonist in the forced swimming test in mice.* European Journal of Pharmacology, 2007. **555**(1): p. 43-47.

31. Bonaventure, P., L. Kelly, L. Aluisio, J. Shelton, B. Lord, R. Galici, K. Miller, J. Atack, T.W. Lovenberg, and C. Dugovic, *Selective blockade of 5-hydroxytryptamine (5-HT)7 receptors enhances 5-HT transmission, antidepressant-like behavior, and rapid eye movement sleep suppression induced*

by citalopram in rodents. Journal of Pharmacology and Experimental Therapeutics, 2007. **321**(2): p. 690-698.

32.   Anttila, S.A.K. and E.V.J. Leinonen, *A review of the pharmacological and clinical profile of mirtazapine.* CNS Drug Reviews, 2001. **7**(3): p. 249-264.

33.   Mørk, A., A. Pehrson, L.T. Brennum, S.M. Nielsen, H. Zhong, A.B. Lassen, S. Miller, L. Westrich, N.J. Boyle, C. Sánchez, C.W. Fischer, N. Liebenberg, G. Wegener, C. Bundgaard, S. Hogg, B. Bang-Andersen, and T.B. Stensbøl, *Pharmacological effects of Lu AA21004: a novel multimodal compound for the treatment of major depressive disorder.* Journal of Pharmacology and Experimental Therapeutics, 2012. **340**(3): p. 666-675.

34.   Yamashita, A., S.K. Singh, T. Kawate, Y. Jin, and E. Gouaux, *Crystal structure of a bacterial homologue of Na+/Cl−-dependent neurotransmitter transporters.* Nature, 2005. **437**(7056): p. 215-223.

35.   Gouaux, E., *The molecular logic of sodium-coupled neurotransmitter transporters.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2009. **364**(1514): p. 149-154.

36.   Shan, J., J.A. Javitch, L. Shi, and H. Weinstein, *The substrate-driven transition to an inward-facing conformation in the functional mechanism of the dopamine transporter.* PLoS ONE, 2011. **6**(1): p. e16350.

37.   Zhao, Y., D.S. Terry, L. Shi, M. Quick, H. Weinstein, S.C. Blanchard, and J.A. Javitch, *Substrate-modulated gating dynamics in a Na+-coupled neurotransmitter transporter homologue.* Nature, 2011. **474**(7349): p. 109-113.

38.   Shi, L., M. Quick, Y. Zhao, H. Weinstein, and J.A. Javitch, *The mechanism of a neurotransmitter:sodium symporter—inward release of Na+ and substrate is triggered by substrate in a second binding site.* Molecular Cell, 2008. **30**(6): p. 667-677.

39.   Quick, M., L. Shi, B. Zehnpfennig, H. Weinstein, and J.A. Javitch, *Experimental conditions can obscure the second high-affinity site in LeuT.* Nat Struct Mol Biol, 2012. **19**(2): p. 207-211.

40.   Zhou, Z., J. Zhen, N.K. Karpowich, C.J. Law, M.E.A. Reith, and D.-N. Wang, *Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures.* Nat Struct Mol Biol, 2009. **16**(2): p. 652-657.

41.   Piscitelli, C.L., H. Krishnamurthy, and E. Gouaux, *Neurotransmitter/sodium symporter orthologue LeuT has a single high-affinity substrate site.* Nature, 2010. **468**(7327): p. 1129-32.

42.   Wang, H., J. Elferich, and E. Gouaux, *Structures of LeuT in bicelles define conformation and substrate binding in a membrane-like context.* Nat Struct Mol Biol, 2012. **19**(2): p. 212-9.

43.   Manepalli, S., C. Surratt, J. Madura, and T. Nolan, *Monoamine transporter structure, function, dynamics, and drug discovery: a computational perspective.* The AAPS Journal, 2012. **14**(4): p. 820-831.

44.   Ravna, A.W., M. Jaronczyk, and I. Sylte, *A homology model of SERT based on the LeuTAa template.* Bioorganic & Medicinal Chemistry Letters, 2006. **16**(21): p. 5594-5597.

45.   Henry, L.K., L.J. DeFelice, and R.D. Blakely, *Getting the message across: a recent transporter structure shows the way.* Neuron, 2006. **49**(6): p. 791-796.

46.   Kaufmann, K.W., E.S. Dawson, L.K. Henry, J.R. Field, R.D. Blakely, and J. Meiler, *Structural determinants of species-selective substrate recognition in human and Drosophila serotonin transporters revealed through computational docking studies.* Proteins: Structure, Function, and Bioinformatics, 2009. **74**(3): p. 630-642.

47.   Celik, L., S. Sinning, K. Severinsen, C.G. Hansen, M.S. Møller, M. Bols, O. Wiborg, and B. Schiøtt, *Binding of serotonin to the human serotonin transporter. Molecular modeling and experimental validation.* Journal of the American Chemical Society, 2008. **130**(12): p. 3853-3865.

48.   Jørgensen, A.M., L. Tagmose, A.M.M. Jørgensen, S. Topiol, M. Sabio, K. Gundertofte, K.P. Bøgesø, and G.H. Peters, *Homology modeling of the serotonin transporter: insights into the primary escitalopram-binding site.* ChemMedChem, 2007. **2**(6): p. 815-826.

49.   Sarker, S., R. Weissensteiner, I. Steiner, H.H. Sitte, G.F. Ecker, M. Freissmuth, and S. Sucic, *The high-affinity binding site for tricyclic antidepressants resides in the outer vestibule of the serotonin transporter.* Molecular Pharmacology, 2010. **78**(6): p. 1026-1035.

50.   Andersen, J., L. Olsen, K.B. Hansen, O. Taboureau, F.S. Jørgensen, A.M. Jørgensen, B. Bang-Andersen, J. Egebjerg, K. Strømgaard, and A.S. Kristensen, *Mutational mapping and modeling of the binding site for (S)-citalopram in the human serotonin transporter.* Journal of Biological Chemistry, 2010. **285**(3): p. 2051-2063.

51.   Andersen, J., N. Stuhr-Hansen, L. Zachariassen, S. Toubro, S.M.R. Hansen, J.N.N. Eildal, A.D. Bond, K.P. Bøgesø, B. Bang-Andersen, A.S. Kristensen, and K. Strømgaard, *Molecular determinants for selective recognition of antidepressants in the human serotonin and norepinephrine transporters.*

Proceedings of the National Academy of Sciences, 2011. **108**(29): p. 12137-12142.

52. Andersen, J., O. Taboureau, K.B. Hansen, L. Olsen, J. Egebjerg, K. Stromgaard, and A.S. Kristensen, *Location of the antidepressant binding site in the serotonin transporter: importance of Ser-438 in recognition of citalopram and tricyclic antidepressants.* J Biol Chem, 2009. **284**(15): p. 10276-84.

53. Wenthur, C.J., G.J. Rodríguez, C.P. Kuntz, and E.L. Barker, *Conformational flexibility of transmembrane helix VII of the human serotonin transporter impacts ion dependence and transport.* Biochemical Pharmacology, 2010. **80**(9): p. 1418-1426.

54. Gabrielsen, M., A. Ravna, K. Kristiansen, and I. Sylte, *Substrate binding and translocation of the serotonin transporter studied by docking and molecular dynamics simulations.* Journal of Molecular Modeling, 2012. **18**(3): p. 1073-1085.

55. Sucic, S., S. Dallinger, B. Zdrazil, R. Weissensteiner, T.N. Jørgensen, M. Holy, O. Kudlacek, S. Seidel, J.H. Cha, U. Gether, A.H. Newman, G.F. Ecker, M. Freissmuth, and H.H. Sitte, *The N terminus of monoamine transporters is a lever required for the action of amphetamines.* Journal of Biological Chemistry, 2010. **285**(14): p. 10924-10938.

56. Koldsø, H., P. Noer, J. Grouleff, H.E. Autzen, S. Sinning, and B. Schiøtt, *Unbiased simulations reveal the inward-facing conformation of the human serotonin transporter and Na+/Cl- ion release.* PLoS Comput Biol, 2011. **7**(10): p. e1002246.

57. Jørgensen, A.M., L. Tagmose, A.M.M. Jørgensen, K.P. Bøgesø, and G.H. Peters, *Molecular dynamics simulations of Na+/Cl−-dependent neurotransmitter transporters in a membrane-aqueous system.* ChemMedChem, 2007. **2**(6): p. 827-840.

58. Warne, T., M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, and G.F. Schertler, *Structure of a beta1-adrenergic G-protein-coupled receptor.* Nature, 2008. **454**(7203): p. 486-91.

59. Chien, E.Y.T., W. Liu, Q. Zhao, V. Katritch, G. Won Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov, and R.C. Stevens, *Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist.* Science, 2010. **330**(6007): p. 1091-1095.

60. Wang, C., Y. Jiang, J. Ma, H. Wu, D. Wacker, V. Katritch, G.W. Han, W. Liu, X.-P. Huang, E. Vardy, J.D. McCorvy, X. Gao, X.E. Zhou, K. Melcher, C. Zhang, F. Bai, H. Yang, L. Yang, H. Jiang, B.L. Roth, V. Cherezov, R.C. Stevens, and

H.E. Xu, *Structural basis for molecular recognition at serotonin receptors.* Science, 2013. **340**(6132): p. 610-614.

61.   Jacobs, B.L., H. van Praag, and F.H. Gage, *Adult brain neurogenesis and psychiatry: a novel theory of depression.* Mol Psychiatry, 2000. **5**(3): p. 262-9.

62.   Malberg, J.E., A.J. Eisch, E.J. Nestler, and R.S. Duman, *Chronic antidepressant treatment increases neurogenesis in adult rat hippocampus.* The Journal of Neuroscience, 2000. **20**(24): p. 9104-9110.

63.   Santarelli, L., M. Saxe, C. Gross, A. Surget, F. Battaglia, S. Dulawa, N. Weisstaub, J. Lee, R. Duman, O. Arancio, C. Belzung, and R. Hen, *Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants.* Science, 2003. **301**(5634): p. 805-9.

64.   Malberg, J.E. and L.E. Schechter, *Increasing hippocampal neurogenesis: a novel mechanism for antidepressant drugs.* Curr Pharm Des, 2005. **11**(2): p. 145-55.

65.   Perera, T.D., J.D. Coplan, S.H. Lisanby, C.M. Lipira, M. Arif, C. Carpio, G. Spitzer, L. Santarelli, B. Scharf, R. Hen, G. Rosoklija, H.A. Sackeim, and A.J. Dwork, *Antidepressant-induced neurogenesis in the hippocampus of adult nonhuman primates.* The Journal of Neuroscience, 2007. **27**(18): p. 4894-4901.

66.   MacMillan, K.S., J. Naidoo, J. Liang, L. Melito, N.S. Williams, L. Morlock, P.J. Huntington, S.J. Estill, J. Longgood, G.L. Becker, S.L. McKnight, A.A. Pieper, J.K. De Brabander, and J.M. Ready, *Development of proneurogenic, neuroprotective small molecules.* Journal of the American Chemical Society, 2011. **133**(5): p. 1428-1437.

67.   Pieper, A.A., S. Xie, E. Capota, S.J. Estill, J. Zhong, J.M. Long, G.L. Becker, P. Huntington, S.E. Goldman, C.-H. Shen, M. Capota, J.K. Britt, T. Kotti, K. Ure, D.J. Brat, N.S. Williams, K.S. MacMillan, J. Naidoo, L. Melito, J. Hsieh, J. De Brabander, J.M. Ready, and S.L. McKnight, *Discovery of a proneurogenic, neuroprotective chemical.* Cell, 2010. **142**(1): p. 39-51.

68.   Yan, H.C., X. Cao, T.M. Gao, and X.H. Zhu, *Promoting adult hippocampal neurogenesis: a novel strategy for antidepressant drug screening.* Curr Med Chem, 2011. **18**(28): p. 4359-67.

69.   Duman, R.S. and G.K. Aghajanian, *Synaptic dysfunction in depression: potential therapeutic targets.* Science, 2012. **338**(6103): p. 68-72.

70.   Schmidt, H.D. and R.S. Duman, *Peripheral BDNF produces antidepressant-like effects in cellular and behavioral models.* Neuropsychopharmacology, 2010. **35**(12): p. 2378-2391.

71. Atkinson, P.J., S.M. Bromidge, M.S. Duxon, L.M. Gaster, M.S. Hadley, B. Hammond, C.N. Johnson, D.N. Middlemiss, S.E. North, G.W. Price, H.K. Rami, G.J. Riley, C.M. Scott, T.E. Shaw, K.R. Starr, G. Stemp, K.M. Thewlis, D.R. Thomas, M. Thompson, A.K.K. Vong, and J.M. Watson, *3,4-Dihydro-2H-benzoxazinones are 5-HT1A receptor antagonists with potent 5-HT reuptake inhibitory activity.* Bioorganic & Medicinal Chemistry Letters, 2005. **15**(3): p. 737-741.

72. Lattmann, E., I. Merino, S. Dunn, B. Parveen, P. Lattmann, D.C. Billington, Y. Bunprakob, and J. Sattayasai, *Novel 5-HT7 ligands as antidepressants: automated synthesis of nsubstituted- N-[1-methyl-3-(4-methylpiperidin-1-yl)propyl]-arylsulfonamides.* Letters in Drug Design & Discovery, 2006. **3**(1): p. 49-54.

73. Wong, D.T., F.P. Bymaster, J.S. Horng, and B.B. Molloy, *A new selective inhibitor for uptake of serotonin into synaptosomes of rat brain: 3-(p-trifluoromethylphenoxy). N-methyl-3-phenylpropylamine.* Journal of Pharmacology and Experimental Therapeutics, 1975. **193**(3): p. 804-811.

74. Nolan, T.L., D.J. Lapinsky, J.N. Talbot, M.n. Indarte, Y. Liu, S. Manepalli, L.M. Geffert, M.E. Amos, P.N. Taylor, J.D. Madura, and C.K. Surratt, *Identification of a novel selective serotonin reuptake inhibitor by coupling monoamine transporter-based virtual screening and rational molecular hybridization.* ACS Chemical Neuroscience, 2011. **2**(9): p. 544-552.

75. Manepalli, S., L.M. Geffert, C.K. Surratt, and J.D. Madura, *Discovery of novel selective serotonin reuptake inhibitors through development of a protein-based pharmacophore.* Journal of Chemical Information and Modeling, 2011. **51**(9): p. 2417-2426.

76. Becker, O.M., Y. Marantz, S. Shacham, B. Inbal, A. Heifetz, O. Kalid, S. Bar-Haim, D. Warshaviak, M. Fichman, and S. Noiman, *G protein-coupled receptors: in silico drug discovery in 3D.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(31): p. 11304-11309.

77. Kim, C.Y., P.E. Mahaney, O. McConnell, Y. Zhang, E. Manas, D.M. Ho, D.C. Deecher, and E.J. Trybulski, *Discovery of a new series of monoamine reuptake inhibitors, the 1-amino-3-(1H-indol-1-yl)-3-phenylpropan-2-ols.* Bioorganic & Medicinal Chemistry Letters, 2009. **19**(17): p. 5029-5032.

78. Hann, M.M., A.R. Leach, and G. Harper, *Molecular complexity and its impact on the probability of finding leads for drug discovery.* Journal of Chemical Information and Computer Sciences, 2001. **41**(3): p. 856-864.

79. *Handbook of combinatorial chemistry: drugs, catalysts, materials*, ed. K.C. Nicolaou, R. Hanko, and W. Hartwig. 2005: Wiley-VCH.

80.   Bleicher, K.H., H.J. Bohm, K. Muller, and A.I. Alanine, *Hit and lead generation: beyond high-throughput screening.* Nat Rev Drug Discov, 2003. **2**(5): p. 369-78.

81.   Viegas-Junior, C., A. Danuello, V.d.S. Bolzani, E.J. Barreiro, and C.A.M. Fraga, *Molecular hybridization: a useful tool in the design of new drug prototypes.* Current Medicinal Chemistry, 2007. **14**(17): p. 1829-1852.

82.   Zhou, D., G.P. Stack, J. Lo, A.A. Failli, D.A. Evrard, B.L. Harrison, N.T. Hatzenbuhler, M. Tran, S. Croce, S. Yi, J. Golembieski, G.A. Hornby, M. Lai, Q. Lin, L.E. Schechter, D.L. Smith, A.D. Shilling, C. Huselton, P. Mitchell, C.E. Beyer, and T.H. Andree, *Synthesis, potency, and in vivo evaluation of 2-piperazin-1-ylquinoline analogues as dual serotonin reuptake inhibitors and serotonin 5-HT1A receptor antagonists.* Journal of Medicinal Chemistry, 2009. **52**(15): p. 4955-4959.

83.   Bajorath, J., *Integration of virtual and high-throughput screening.* Nat Rev Drug Discov, 2002. **1**(11): p. 882-894.

84.   *Concepts and applications of molecular similarity*, ed. M.A. Johnson and G.M. Maggiora. 1990, New York: John Wiley & Sons.

85.   Willett, P., *Similarity-based approaches to virtual screening.* Biochemical Society Transactions, 2003. **31**: p. 603-606.

86.   Martin, Y.C., *3D database searching in drug design.* Journal of Medicinal Chemistry, 1992. **35**(12): p. 2145-2154.

87.   Martin, Y., M. Bures, E. Danaher, J. DeLazzer, I. Lico, and P. Pavlik, *A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists.* Journal of Computer-Aided Molecular Design, 1993. **7**(1): p. 83-102.

88.   Orús, L., S. Pérez-Silanes, A.-M. Oficialdegui, J. Martínez-Esparza, J.-C. Del Castillo, M. Mourelle, T. Langer, S. Guccione, G. Donzella, E.M. Krovat, K. Poptodorov, B. Lasheras, S. Ballaz, I. Hervías, R. Tordera, J. Del Río, and A. Monge, *Synthesis and molecular modeling of new 1-aryl-3-[4-arylpiperazin-1-yl]-1-propane derivatives with high affinity at the serotonin transporter and at 5-HT1A receptors.* Journal of Medicinal Chemistry, 2002. **45**(19): p. 4128-4139.

89.   Zhang, S., F. Fernandez, S. Hazeldine, J. Deschamps, J. Zhen, M.E.A. Reith, and A.K. Dutta, *Further structural exploration of trisubstituted asymmetric pyran derivatives (2S,4R,5R)-2-benzhydryl-5-benzylamino-tetrahydropyran-4-ol and their corresponding disubstituted (3S,6S) pyran derivatives: a proposed pharmacophore model for high-affinity interaction with the dopamine, serotonin, and norepinephrine transporters.* Journal of Medicinal Chemistry, 2006. **49**(14): p. 4239-4247.

90. Chang, C., S. Ekins, P. Bahadduri, and P.W. Swaan, *Pharmacophore-based discovery of ligands for drug transporters.* Advanced Drug Delivery Reviews, 2006. **58**(12–13): p. 1431-1450.

91. Zhang, Q. and I. Muegge, *Scaffold hopping through virtual screening using 2D and 3D similarity descriptors: ranking, voting, and consensus scoring.* Journal of Medicinal Chemistry, 2006. **49**(5): p. 1536-1548.

92. Krüger, D.M. and A. Evers, *Comparison of structure- and ligand-based virtual screening protocols considering hit list complementarity and enrichment factors.* ChemMedChem, 2010. **5**(1): p. 148-158.

93. Wang, S., S. Sakamuri, I.J. Enyedy, A.P. Kozikowski, O. Deschaux, B.C. Bandyopadhyay, S.R. Tella, W.A. Zaman, and K.M. Johnson, *Discovery of a novel dopamine transporter inhibitor, 4-hydroxy-1-methyl-4-(4-methylphenyl)-3-piperidyl 4-methylphenyl ketone, as a potential cocaine antagonist through 3d-database pharmacophore searching. Molecular modeling, structure−activity relationships, and behavioral pharmacological studies.* Journal of Medicinal Chemistry, 2000. **43**(3): p. 351-360.

94. Brust, A., E. Palant, D.E. Croker, B. Colless, R. Drinkwater, B. Patterson, C.I. Schroeder, D. Wilson, C.K. Nielsen, M.T. Smith, D. Alewood, P.F. Alewood, and R.J. Lewis, *χ-conopeptide pharmacophore development: toward a novel class of norepinephrine transporter inhibitor (Xen2174) for pain.* Journal of Medicinal Chemistry, 2009. **52**(22): p. 6991-7002.

95. Ahmed, A., H. Choo, Y.S. Cho, W.-K. Park, and A.N. Pae, *Identification of novel serotonin 2C receptor ligands by sequential virtual screening.* Bioorganic & Medicinal Chemistry, 2009. **17**(13): p. 4559-4568.

96. Moustakas, D.T., P.T. Lang, S. Pegg, E. Pettersen, I.D. Kuntz, N. Brooijmans, and R.C. Rizzo, *Development and validation of a modular, extensible docking program: DOCK 5.* J Comput Aided Mol Des, 2006. **20**(10-11): p. 601-619.

97. Indarte, M., J.D. Madura, and C.K. Surratt, *Dopamine transporter comparative molecular modeling and binding site prediction using the LeuTAa leucine transporter as a template.* Proteins: Structure, Function, and Bioinformatics, 2008. **70**(3): p. 1033-1046.

98. Oshiro, C., E.K. Bradley, J. Eksterowicz, E. Evensen, M.L. Lamb, J.K. Lanctot, S. Putta, R. Stanton, and P.D.J. Grootenhuis, *Performance of 3D-database molecular docking studies into homology models.* Journal of Medicinal Chemistry, 2004. **47**(3): p. 764-767.

99.   Kairys, V., M.X. Fernandes, and M.K. Gilson, *Screening drug-like compounds by docking to homology models: a systematic study.* Journal of Chemical Information and Modeling, 2005. **46**(1): p. 365-379.

100.   Petrey, D. and B. Honig, *Protein structure prediction: inroads to biology.* Molecular Cell, 2005. **20**(6): p. 811-819.

101.   Hillisch, A., L.F. Pineda, and R. Hilgenfeld, *Utility of homology models in the drug discovery process.* Drug Discovery Today, 2004. **9**(15): p. 659-669.

102.   Dunbrack Jr, R.L., *Sequence comparison and protein structure prediction.* Current Opinion in Structural Biology, 2006. **16**(3): p. 374-384.

103.   Sheridan, R.P., *Alternative global goodness metrics and sensitivity analysis: heuristics to check the robustness of conclusions from studies comparing virtual screening methods.* Journal of Chemical Information and Modeling, 2008. **48**(2): p. 426-433.

104.   Klebe, G., *Virtual ligand screening: strategies, perspectives and limitations.* Drug Discovery Today, 2006. **11**(13–14): p. 580-594.

105.   Indarte, M.n., Y. Liu, J.D. Madura, and C.K. Surratt, *Receptor-based discovery of a plasmalemmal monoamine transporter inhibitor via high-throughput docking and pharmacophore modeling.* ACS Chemical Neuroscience, 2010. **1**(3): p. 223-233.

106.   Schneider, G. and U. Fechner, *Computer-based de novo design of drug-like molecules.* Nat Rev Drug Discov, 2005. **4**(8): p. 649-663.

107.   Jahnke, W. and D.A. Erlanson, *Fragment-based approaches in drug discovery.* Methods and Principles in Medicinal Chemistry, ed. R. Mannhold, H. Kubinyi, and G. Folkers. 2006: Wiley-VCH.

108.   Carr, R.A.E., M. Congreve, C.W. Murray, and D.C. Rees, *Fragment-based lead discovery: leads by design.* Drug Discovery Today, 2005. **10**(14): p. 987-992.

109.   Zoete, V., A. Grosdidier, and O. Michielin, *Docking, virtual high throughput screening and in silico fragment-based drug design.* Journal of Cellular and Molecular Medicine, 2009. **13**(2): p. 238-248.

110.   Konteatis, Z.D., *In silico fragment-based drug design.* Expert Opinion on Drug Discovery, 2010. **5**(11): p. 1047-1065.

111.   Hubbard, R., L. Chen, and B. Davis, *Informatics and modeling challenges in fragment-based drug discovery.* Current Opinion in Drug Discovery and Development, 2007. **10**(3): p. 289-297.

112. Irwin, J.J. and B.K. Shoichet, *ZINC − A Free database of commercially available compounds for virtual screening.* Journal of Chemical Information and Modeling, 2004. **45**(1): p. 177-182.

113. Anand, K., J. Ziebuhr, P. Wadhwani, J.R. Mesters, and R. Hilgenfeld, *Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs.* Science, 2003. **300**(5626): p. 1763-1767.

114. Enyedy, I.J., S.-L. Lee, A.H. Kuo, R.B. Dickson, C.-Y. Lin, and S. Wang, *Structure-based approach for the discovery of bis-benzamidines as novel inhibitors of matriptase.* Journal of Medicinal Chemistry, 2001. **44**(9): p. 1349-1355.

115. Enyedy, I.J., Y. Ling, K. Nacro, Y. Tomita, X. Wu, Y. Cao, R. Guo, B. Li, X. Zhu, Y. Huang, Y.-Q. Long, P.P. Roller, D. Yang, and S. Wang, *Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening.* Journal of Medicinal Chemistry, 2001. **44**(25): p. 4313-4324.

116. Li, R., X. Chen, B. Gong, P.M. Selzer, Z. Li, E. Davidson, G. Kurzban, R.E. Miller, E.O. Nuzum, J.H. McKerrow, R.J. Fletterick, S.A. Gillmor, C.S. Craik, I.D. Kuntz, F.E. Cohen, and G.L. Kenyon, *Structure-based design of parasitic protease inhibitors.* Bioorganic & Medicinal Chemistry, 1996. **4**(9): p. 1421-1427.

117. Rajnarayanan, R.V., S. Dakshanamurthy, and N. Pattabiraman, *"Teaching old drugs to kill new bugs": structure-based discovery of anti-SARS drugs.* Biochemical and Biophysical Research Communications, 2004. **321**(2): p. 370-378.

118. Selzer, P.M., X. Chen, V.J. Chan, M. Cheng, G.L. Kenyon, I.D. Kuntz, J.A. Sakanari, F.E. Cohen, and J.H. McKerrow, *Leishmania major:molecular modeling of cysteine proteases and prediction of new nonpeptide inhibitors.* Experimental Parasitology, 1997. **87**(3): p. 212-221.

119. Zuccotto, F., M. Zvelebil, R. Brun, S.F. Chowdhury, R. Di Lucrezia, I. Leal, L. Maes, L.M. Ruiz-Perez, D. Gonzalez Pacanowska, and I.H. Gilbert, *Novel inhibitors of trypanosoma cruzi dihydrofolate reductase.* European Journal of Medicinal Chemistry, 2001. **36**(5): p. 395-405.

120. Lapinsky, D.J., S. Aggarwal, T.L. Nolan, C.K. Surratt, J.R. Lever, R. Acharya, R.A. Vaughan, A. Pandhare, and M.P. Blanton, *(±)-2-(N-tert-Butylamino)-3'-[125I]-iodo-4'-azidopropiophenone: a dopamine transporter and nicotinic acetylcholine receptor photoaffinity ligand based on bupropion (Wellbutrin, Zyban).* Bioorganic & Medicinal Chemistry Letters, 2012. **22**(1): p. 523-526.

121. Lapinsky, D.J., R. Velagaleti, N. Yarravarapu, Y. Liu, Y. Huang, C.K. Surratt, J.R. Lever, J.D. Foster, R. Acharya, R.A. Vaughan, and H.M. Deutsch, *Azido-iodo-N-benzyl derivatives of threo-methylphenidate (Ritalin, Concerta): rational design, synthesis, pharmacological evaluation, and dopamine transporter photoaffinity labeling.* Bioorganic & Medicinal Chemistry, 2011. **19**(1): p. 504-512.

122. Lapinsky, D.J., N. Yarravarapu, T.L. Nolan, C.K. Surratt, J.R. Lever, M. Tomlinson, R.A. Vaughan, and H.M. Deutsch, *Evolution of a compact photoprobe for the dopamine transporter based on (±)-threo-methylphenidate.* ACS Medicinal Chemistry Letters, 2012. **3**(5): p. 378-382.

123. Trott, O. and A.J. Olson, *AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* Journal of Computational Chemistry, 2010. **31**(2): p. 455-461.

124. *DOCK*. Available from: http://dock.compbio.ucsf.edu/.

125. Halgren, T.A., R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, and J.L. Banks, *Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1750-1759.

126. Friesner, R.A., J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, and P.S. Shenkin, *Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.

127. Jones, G., P. Willett, R.C. Glen, A.R. Leach, and R. Taylor, *Development and validation of a genetic algorithm for flexible docking.* Journal of Molecular Biology, 1997. **267**(3): p. 727-748.

128. *Docking@Home*. Available from: http://docking.cis.udel.edu/.

129. Moldover, B., A. Solidar, C. Montgomery, H. Miziorko, J. Murphy, and G.J. Wyckoff, *ChemVassa: a new method for identifying small molecule hits in drug discovery.* The Open Medicinal Chemistry Journal, 2012. **6**: p. 29-34.

130. Hardy, B. and R. Affentranger, *Collaborative virtual organisation and infrastructure for drug discovery.* Drug Discovery Today, 2013. **18**(13–14): p. 681-686.