Identification of Potential Inhibitors for Important Protein Targets in Alzheimer's Disease Using Pharmacophore Modeling and Molecular Dynamics Simulations

THESIS

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

By

RAVICHAND PALAKURTI

ID. No: 2010PHXF015H

Under the supervision of

Prof. Ramakrishna Vadrevu

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

HYDERABAD CAMPUS

2015

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE,

PILANI - HYDERABAD CAMPUS

CERTIFICATE

This is to certify that the thesis entitled ***"Identification of Potential Inhibitors for Important Protein Targets in Alzheimer's Disease Using Pharmacophore Modeling and Molecular Dynamics Simulations"*** was submitted by **Ravichand Palakurti, ID.No. 2010PHXF015H** for the award of Ph. D. degree of the Institute embodies the original work done by him under my supervision.

Signature in full of the supervisor: _____

Name in capital block letters: **RAMAKRISHNA VADREVU**

Designation: **Associate Professor,**

Dept. of Biological Sciences

Date: _____

Acknowledgement

The work was carried out during 2010-2015 at the BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE-Pilani, Hyderabad Campus.

Firstly, I would like to express my gratitude to my advisor ***Prof. Ramakrishna Vadrevu****, for his valuable guidance and supervision throughout my research. He have been a tremendous mentor for me. I would like to thank him for encouraging me in my research and for allowing me to mature as a researcher.*

My deep sense of gratitude to ***Prof. B. N. Jain****, Ex-Vice Chancellor (BITS) and* ***Prof. V.S. Rao****, Acting Vice Chancellor & Director (BITS-Pilani, Hyderabad Campus), for permitting me to carry out my research work in the campus.*

I would also like to thank ***Dr. Debashree Bandyopadhyay****,* ***Prof. S Swaminathan****,* ***Prof. D. Sriram*** *and* ***Dr. Savitha Govardhan*** *for serving as my committee members.*

I also acknowledge ***Dr. Sajeli Begum*** *and* ***Dr. K. Naga Mohan*** *for their valuable time and resources that helped me in progressing in my research.*

*I am grateful to all the Faculty and HOD****,*** *Biological Sciences Department,* ***Prof. Suman Kapur****, BITS, Pilani, Hyderabad campus for their valuable support. I would also like to thank* ***Prof. Vidhya Rajesh, Dr. Jayati Ray Datta (ARD)*** *for their support.*

I would like to express my gratitude for ***Prof. P. Yogeeswari*** *for being my Co-supervisor and mentor and helped me throughout my time in BITS Pilani Hyderabad Campus. Thank you*

I would like to acknowledge all my friends who have taken part in this journey and in some way helped me along. Especially, ***Raja Shekar Varma****,* ***Imran Khan****,* ***Nandini Chowdary****,* ***Komal Sharma, Priyanka Sharma, Reshma Chowdary*** *and* ***C.N. Rahul****. I would also like to acknowledge all the lab technicians for their backing.*

I acknowledge BITS, Pilani, Hyderabad campus for funding my fellowship.

At the end, a special thanks to my family. Words cannot express how grateful I am to my ***mother (Vijaya)*** *and* ***father (Rajendra Prasad)*** *for all of the sacrifices that you have made on my behalf.*

Date: *Ravichand P*

Abstract

Efforts to combat Alzheimer's disease (AD) are focused predominantly on inhibiting the activity of enzyme(s) that have been identified to be responsible for the production of the main causative Aβ amyloid forming peptide. Drug candidates targeting different aspects of the disease failed due to various reasons including toxicity, inability to permeate to the blood brain barrier, lack of efficacy in humans. Research efforts by many groups to decipher the reasons for the failure of different lead compounds, and successfully deliver a viable therapy are being pursued around the world. Added to this, is the inherent complexity associated with the network of pathways leading to the progress of the disease may involve additional targets for designing effective therapies.

Keeping this in view, we have attempted to identify lead compounds for inhibiting two important protein targets that play a very important role in AD. A combination of modeling and experimental studies have been utilized in our efforts to identify lead molecules and their interactions/stability with the targets in order to gain insights into the interactions that can be modulated for the desired result. A recent approach to develop multiple pharmacophore models combining energy based assessment of interactions has been extensively used in the present work. Rigorous analysis of the performance of the different pharmacophore models provided detailed insights into its efficacy to identify diverse set of potential leads. Further, in addition to targeting the proteins involved in AD, the amyloids that cause the disease were targeted for disruption as a complementary approach.

Table of contents

Abbreviations

A	Acceptor
Aβ	Amyloid beta
AD	Alzheimer's disease
ADME	Absorption, distribution, metabolism and excretion
APP	Amyloid precursor protein
BACE1	β-site Amyloid Precursor Protein Cleaving Enzyme 1
BBB	Blood-brain barrier
c-Abl	Abelson's Tyrosine Kinase
CADD	Computer-Assisted Drug Design
CML	Chronic Myeloid Leukemia
CNS	Central nervous system
D	Donor
DMSO	Dimethyl sulfoxide
fAb	Fibrillar Amyloid-b
FRET	Fluorescence Resonance Energy Transfer
H	Hydrophobic
H-bond	Hydrogen-bonding
HTVS	High-throughput virtual screening
IC_{50}	Mean inhibitory concentration
LBDD	Ligand-Based Drug Design
MD	Molecular Dynamics
μM	Micro molar
N	Negative Ionizable
nM	Nano molar
ns	Nano seconds
NFT's	Neuro Fibrillary Tangles
PD	Parkinson's disease
PDB	Protein Data Bank
PLS	Partial Least Square Analysis
ps	Pico seconds
QSAR	Quantitative Structure-Activity Relationship
P	Positive Ionizable

R	Ring Aromatic
RMSD	Root Mean Square Deviation
SAR	Structure-activity relationship
SBDD	Structure-Based Drug Design
ThT	Thioflavin T

Amino acids	**3 letter code**	**1 letter code**
Alanine	Ala	A
Arginine	Arg	R
Aspargine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoluecine	Ile	I
Luecine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Proteins are the essential biomolecules performing various roles and functions in the cells. They are polymers of amino acids joined by peptide bonds. The linear sequence of amino acids of any protein constitutes the primary structure. Specific sequences adopt basic secondary structural elements called α-helices and β-sheets (Parthasarathi, Raman, Subramanian, & Ramasami, 2007). Based on the arrangement of the polypeptide chain β–sheets can further be classified as parallel or anti-parallel β-sheet structures. These secondary structures, α-helices and β-sheets are stabilized by hydrogen bonds as shown in **Figure 1.1**. The arrangement of the secondary structure elements into compact well-folded 3-dimensional structures constitutes the tertiary structure of proteins which in some cases can associate to form the quaternary structures. The correctly folded 3-dimensional structures are responsible for function and are referred to as the native conformations of proteins. In addition to hydrogen bonds, interactions involving the side chains contribute to the structure and stability of proteins.

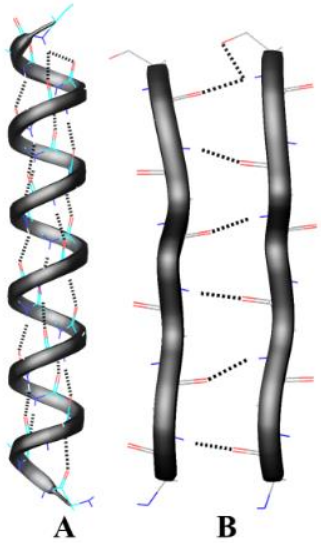


Figure 1.1 *Secondary structures of proteins. A and B represent α-helix and β-sheet structure respectively. The black dots indicate the pattern of hydrogen bonds stabilizing the α-helix and β-sheet.*

Protein folding to the native states is the process in which the sequences attain the correctly folded native, functional states. The process involves the search by the unfolded chain via intermediate states to the folded state via specific pathways and not by a random search. The process is accompanied by a gradual decrease in entropy of the folded state with the concomitant decrease in the folding free energy as depicted in **Figure 1.2.** Deviations in the pathways and process of folding can also lead to non-native states called the misfolded conformations. While, the correctly folded native states are responsible for function, the misfolded, non-native states are now believed to be responsible for various kinds of human diseases including, Alzheimer's, Parkinson's, cystic fibrosis, diabetes and cancer. Misfolding of proteins into highly ordered amyloids or aggregates involves a shift in the folding energy barriers between different states of proteins as shown in **Figure 1.2.** Misfolded proteins may have low energy conformations than the native state and thus be populated in the process (Hartl & Hayer-Hartl, 2009). Protein misfolding appears to be independent of the protein sequence and fold, a conclusion arrived from the general observation that proteins with different amino acid sequences and adopting different folds can form misfolded species(Chiti & Dobson, 2006).

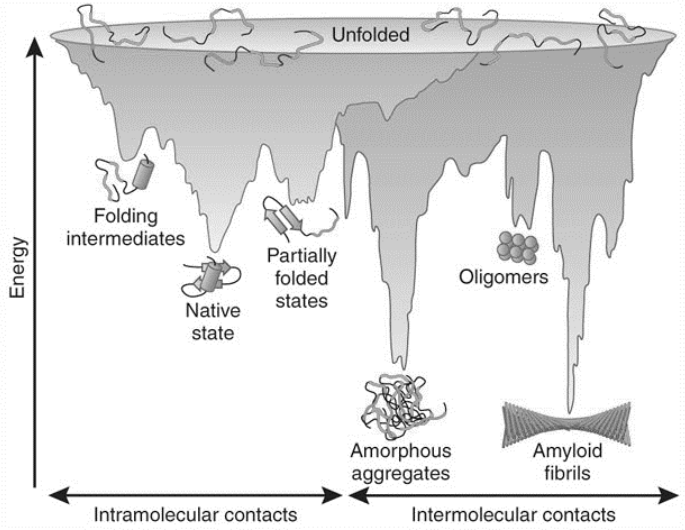


Figure 1.2 *Schematic representation of folding free energy landscape of proteins (Hartl & Hayer-Hartl, 2009).*

Role of misfolded proteins in human diseases

Protein misfolding is now recognized to be responsible for the progression of numerous human diseases (Chiti & Dobson, 2006; Valastyan & Lindquist, 2014). Misfolded proteins often result in the formation of highly ordered structures called amyloids/fibrils (Chiti & Dobson, 2006; Jahn & Radford, 2005; Meier & Bockmann, 2015). In some cases it is observed that it is responsible for the formation of higher order oligomers and aggregates. Irrespective of the primary sequence and fold, the proteins misfold into highly ordered β-sheet aggregates and/or amyloid plaques/fibrils. These non-native states are implicated in cellular toxicity and protein dysfunction in their associated diseases. The structure seen in amyloid fibrils is typically a β-sheet structure (Meier & Bockmann, 2015). Therefore, as discussed in the previous section, from the folding energy landscape of proteins, misfolding can be characteristically referred to as a shift from the typical 3D structure of the native protein into a highly ordered β-sheet conformation (T. P. J. Knowles, M. Vendruscolo, & C. M. Dobson, 2014; Landau et al., 2011). The diseases resulting from misfolded proteins or peptides are called conformational disorders or amyloidogenic diseases (T. P. J. Knowles et al., 2014).

Cells have protective mechanisms involving the degradation of the misfolded proteins. But, in some cases, these misfolded proteins escape these degradation mechanisms by rapidly aggregating into ordered β-sheet amyloid plaques (Eisenberg & Jucker, 2012; MB, 2006). These plaques are insoluble, cytotoxic and accumulate in various organs, disrupting the cells by interfering with the normal physiological functions which can result in organ failure (Rajasekhar, Chakrabarti, & Govindaraju, 2015).

The role of misfolding in human diseases is exemplified by the identification of numerous conformational disorders or amyloidogenic diseases. Depending on the type of protein or

peptide involved in misfolding, (SY & DE, 2014) Fabrizio Chiti and Christopher M. Dobson categorized the various depositions of misfolded proteins in various diseases in the review "Protein Misfolding, Functional Amyloid, and Human Disease" (Chiti & Dobson, 2006). **Table 1.1** listing various diseases, provides an indication of the role of protein misfolding in numerous diseases. As illustrated in **Table 1.1**, proteins with diversity in sequence, size and structures can be involved. For instance, the forty two residue amyloid β peptide is primarily responsible for Alzheimer's disease, a neurodegenerative condition leading to serious conditions.

Another protein that is involved in a neurodegenerative disorder is α-synuclein. α-synuclein, a natively disordered protein populates partially folded states which serve as precursor for the formation of amyloids in the brain leading to Parkinson's. Many neurodegenerative diseases including ALS, Huntington's etc., clearly indicate the devastating consequences of protein misfolding.

Disease	Aggregating protein or peptide	Number of residues	Native structure of protein
Alzheimer's disease	Amyloid β peptide	40 or 42	Natively unfolded
Spongiform encephalopathies	Prion protein	253	Natively unfolded (residues 1–120)and α-helical (residues 121–230)
Parkinson's disease	α-Synuclein	140	Natively unfolded
Dementia with Lewy bodies	α-Synuclein	140	Natively unfolded
Frontotemporal dementia with Parkinsonism	Tau	352–441	Natively unfolded
Amyotrophic lateral sclerosis	Superoxide dismutase 1	153	
Huntington's disease	Huntingtin with polyQ expansion	3144	Largely natively unfolded
Spinocerebellar ataxia	Ataxins with polyQ expansion	816	All-β, AXH domain (residues 562–694)
Spinal and bulbar muscular atrophy	Androgen receptor with polyQ expansion	919	All-α, nuclear receptor ligand-binding domain
Nonneuropathic localized diseasesType II diabetes	Amylin, also called islet amyloid polypeptide (IAPP)	37	Natively unfolded
Hereditary cerebral haemorrhage with amyloidosis	Mutants of amyloid β peptide	40 or 42	Natively unfolded
Aortic medial amyloidosis	Medin	50	
Cataract	γ-Crystallins		
Cutaneous lichen amyloidosis	Keratins		Unknown
Corneal amylodosis associated with trichiasis	Lactoferrin	692	α+β, periplasmic-binding protein like II
Pituitary prolactinoma	Prolactin	199	All-α, 4-helical cytokines

Table 1.1 *The list of misfolding diseases due to the formation of amyloids/aggregates. The Table is adapted from Protein Misfolding, Functional Amyloid and Human Disease by Fabrizio Chiti and Christopher M. Dobson published in The Annual Review of Biochemistry.2006, 75, 333-366. (Partial & not complete list)*

Misfolding and Alzheimer's disease

One of the vital causes in the appearance of several neurodegenerative diseases is misfolding of proteins. Neurodegenerative diseases like Alzheimer's, Amyotrophic Lateral Sclerosis, Huntington disease, Parkinson's disease, Spinocerebellar ataxia are due to the formation of amyloids in neuronal cells (Ow & Dunstan, 2014). The incidence of Central Nervous System (CNS) disorders is normally up to 10 fold higher after the age of 65 years and it increases exponentially every 5 years after the age of 70 (T. P. J. Knowles et al., 2014; Landau et al., 2011; O'Brien & Wong, 2011; Patterson et al., 2015). Alzheimer's disease (AD) is the most important of all the neurodegenerative diseases due to its high incidence and death rate. The disease is typically detected in elders >65 years but, studies indicating that presence of mutations in amyloid precursor protein (APP) gene can also lead to an early onset of AD affecting people in the age group of 30-60 years (O'Brien & Wong, 2011). In fact, it is reported that 44 million people around the world are diagnosed with AD and is the leading cause of death and disability after heart disease, breast cancer, stroke and HIV. The socio-economic burden is very high as 94% of people living with dementia in low and middle income countries and the global cost of Alzheimer's and dementia today is estimated to be $818 billion.

Alzheimer's disease progression is linked with the accumulation of insoluble aggregates of amyloid peptides in various regions of the brain (J & D, 1991). These peptides aggregate by forming hydrogen bonds in a β-sheet conformation leading to the formation of fibrils and plaques. These fibrils and plaques are very stable and cannot be easily cleared by the natural defensive mechanisms of the body.

In addition to the accumulation of amyloid peptides leading to neurodegeneration in AD, another pathway that was identified involves the abnormal deposition of Neuro Fibrillary

Tangles (NFTs) in the neurons of Alzheimer's patients (Killick et al., 2011). The slow and progressive accumulation of the fibrils will disrupt the normal functioning of brain by damaging the connections between neurons (Sheng, Sabatini, & Sudhof, 2012). The accumulation of amyloids in the brain will also trigger a cascade of pathways including increased reactive oxygen species and accelerating neuron death through apoptosis. These toxic changes in the brains will make the neurons non-functional by affecting the synaptic connections leading to progressive memory loss, also called as dementia, resulting in loss of cognitive functions due to cell death in key areas of the brain like the cerebral cortex and the hippocampus which are associated with memory.

As a result in the advancements in health care, early detection, prevention and treatment methods for numerous diseases and general quality of life, there is a substantial increase in life expectancy in both developed and developing countries (Ferri et al., 2005; Kalaria et al., 2008). Consequently, the number of people living beyond 65 years is significantly increasing along with the increasing risk for developing neurodegenerative diseases (Hindle, 2010; Vivar, 2015). Identification of effective treatment methods for Alzheimer's disease will reduce the unprecedented budget necessary for the management of these diseases both in India and throughout the world (Chandra et al., 1998; Mathuranath et al., 2012).

Mechanisms and modulation of Alzheimer's disease

Understanding the mechanisms underlying the progression of AD, proteins involved in the pathways and their modulation can lead to developing therapeutics for the disease (Han & Mook-Jung, 2014; Jia, Deng, & Qing, 2014). The major protein involved in the pathway of amyloid formation leading to Alzheimer's is believed to be the Amyloid Precursor Protein (APP). (O'Brien & Wong, 2011; Rabinovich-Nikitin & Solomon, 2014; Zhao et al., 2015) The

Amyloid Precursor Protein is an integral membrane protein with 365 to 770 amino acids, the variation in the size is a result of alternate splicing gene variants identified in humans (Yamada, Goto, & Sakaki, 1993). Although, the precise function of amyloid precursor protein is not known, however, it is known to be linked with synapse formation, iron transport and plasticity of neurons (O'Brien & Wong, 2011). Mutations in APP are strongly associated with Alzheimer's disease, especially in the amyloid-β region (Citron et al., 1992; Weggen & Beher, 2012). In Alzheimer's disease, the degradation product of Amyloid Precursor Protein or APP by α-secretase, β-secretase-1 and γ-secretase forming 40-42 residue amyloid-β peptides is known to be a major pathogenic species that often aggregate into oligomers called Amyloid plaques (De Strooper, Vassar, & Golde, 2010; Findeis, 2007).

Therefore, β-secretase-1 and γ-secretase play a very important role in the efforts for the discovery of remedies for Alzheimer's disease (De Strooper et al., 2010). **Figure 1.3** represents the cleavage patterns of the above 3 secretases involved in APP processing.

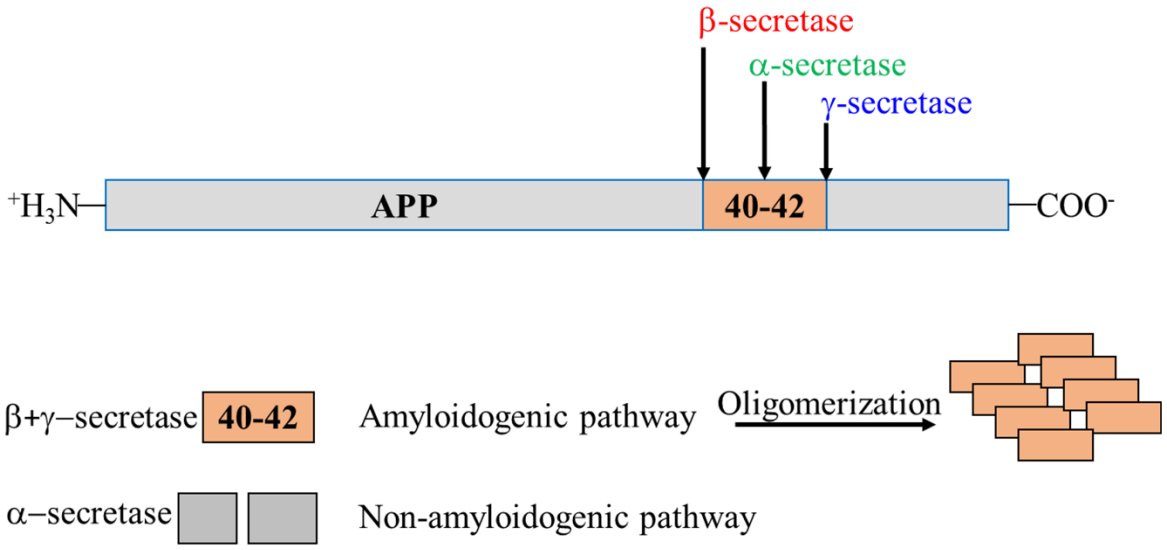


Figure 1.3 *Processing of APP by various secretases (α,β and γ) leading to the formation of Aβ.*

β-secretase1 (BACE-1) also called as beta-site APP cleaving enzyme 1. It is an aspartate protease known to be involved in the cleavage of Amyloid Precursor Protein (APP) (Robert Vassar, 2011),(Robert Vassar, 2012). The activity of BACE1 along with γ-secretase results in the formation of amyloid-β peptides from 40-42 amino acid residues in length (Kandalepas & Vassar, 2014),(Arun K. Ghosh, 2014) About 90% of the secreted amyloid-β peptides formed from processing of APP are Aβ-40, and 10% of secreted amyloid-β peptides are Aβ-42 and 43 (Martin Citron et al., 1996). These secreted peptides are cytotoxic and result in insoluble aggregates and eventually leads to amyloid formation (Knowles TP, Vendruscolo M, & CM, 2014; Rajasekhar et al., 2015; SY & DE, 2014). Also, BACE1 expression is known to be up-regulated in Alzheimer's patients and further increases the amyloid-β burden (X. Cheng et al., 2014; Robert Vassar, 2012).

The proteins β-secretase-1 and γ-secretase are of great value for the efforts in the discovery of new medicines for Alzheimer's disease. BACE1 is not involved in any major physiological pathways in the adult brain, inhibitors of this protein were considered to be relatively safe (Yan & Vassar, 2014).

BACE-1 is a 501 amino acid protein and an aspartate protease (Yan & Vassar, 2014). β-secretase has long been regarded as a therapeutic target for Alzheimer's disease in the development of inhibitor drugs for the reduction of amyloid-β (Robert Vassar, 2012; Yan & Vassar, 2014). Crystal structures of β-secretase inhibitor complexes have revealed significant information about the binding sites and nature of protein-ligand interactions, which in turn have prompted research groups to design novel inhibitors. APP is processed in two diverse pathways, the amyloidogenic and non-amylodogenic pathways (Lichtenthaler, 2010). The amyloidogenic

pathway comprises BACE1 to process APP along with γ-secretase, leading to the formation of amyloid-β.(Lichtenthaler, 2010) In the other pathway, APP is processed by α-secretase in place of β-secretase resulting in a fragmented form of the peptide and is not capable of forming amyloid (Bandyopadhyay S, Goldstein LE, Lahiri DK, & JT, 2007; Lichtenthaler, 2010).

Disease progression can be stopped/reversed via the activation of α-secretase or by the inhibition of BACE1 (Bandyopadhyay S et al., 2007; Lichtenthaler, 2010). γ-secretase is another important target in Alzheimer's disease as it is responsible for the cleavage on the c-terminal of the β-amyloid, whereas β-secretase is identified to cleave the APP on the n-terminal part of the APP, often called as the first cut for the amyloid-β peptide (Karran, Mercken, & De Strooper, 2011). Also, recently published studies detailed that γ-secretase is involved in the notch signaling pathway which is necessary for proper cellular functions, limiting its use as a viable candidate for drug design (Geling A, Steiner H, Willem M, Bally-Cuif L, & C., 2002). Amyloids formation in the brain can lead to the activation of several pathways leading to neurodegeneration through inflammation and apoptosis (Gonfloni, Maiani, Di Bartolomeo, Diederich, & Cesareni, 2012). Prevention of amyloid formation by the inhibition of β-secretase is one of the major goals towards the treatment for Alzheimer's disease (Chun-Jiang, Li, & Zhi-Ren, 2015). As BACE-1 knockout mice did not produce amyloids, confirming it as the most important protein's for intervention in Alzheimer's disease (Luo Y et al., 2001 ; Yan & Vassar, 2014). This proves that BACE1 has the most significant role in the amyloidogenic pathway (Yan & Vassar, 2014). As an aspartate protease, BACE-1 draws inspiration from HIV-protease as a success story for protease inhibitors (Arun K. Ghosh, 2014). But the search of BACE1 inhibitors is an on-going as there are currently no drugs as BACE1 inhibitors.

In addition to APP processing pathway, another important pathway leading to Alzheimer's disease is the formation of Neuro Fibrillary Tangles by the aggregation of abnormally phosphorylated 'Tau' protein (Ballatore C, Lee VM, & JQ, 2007). The presence of amyloid-β is also known to propagate Tau mediated cytotoxicity in Alzheimer's disease (Pooler et al., 2015). Tau normally functions as a microtubule stabilizing protein as it binds to microtubules in the neurons. Upon abnormal phosphorylation of Tau at Tyrosine residue 394, the Tau proteins dislodge from microtubules and aggregate to form NFT's- Neuro Fibrillary Tangles leading to neurodegeneration (Ballatore C et al., 2007). The hyper-phosphorylation of Tau is made by the phosphorylation mechanism of Abelson Tyrosine kinase protein (c-Abl). Thus, c-Abl's phosphorylation of the Tau protein leads to its aggregation and deposition as NFT's. These abnormal aggregates of Tau in the neurons are considered as another pathological phenotype of Alzheimer's disease which can be treated by the inhibition of Abelson's Tyrosine Kinase (Sydow et al., 2011). Thus c-Abl can serve as a new target for AD has been put forth very recently (Schlatterer, Acker, & Davies, 2011). c-Abl up-regulation also leads to abnormal cell cycle re-entry of the terminally differentiated neurons leading to cell death i.e., neuron-degeneration (Schlatterer, Acker, et al., 2011). Activation of c-Abl is also linked with the oxidative stress leading to apoptosis of nerve cells in the brain (Gonfloni et al., 2012; Tan et al., 2013). c-Abl is also involved in cell differentiation, cell division, and apoptosis and also in stress response (Schlatterer, Acker, et al., 2011; Schlatterer, Tremblay, Acker, & Davies, 2011). c-Abl is a crucial enzyme for the development of the nervous system, but it is not believed to be an active part of the adult brain (Schlatterer, Acker, et al., 2011).

Recent studies have shown that expression of c-Abl in the brain is directly linked to extensive apoptosis in the neurons (Schlatterer, Acker, et al., 2011). Recent experimental studies

confirmed that aberrant activation of c-Abl in the brain is associated with pathogenesis in various neurodegenerative diseases like Alzheimer's, Parkinson's and Picks disease (Eisenberg & Jucker, 2012). c-Abl and Tau are found co-precipitated in NFT's of neurons in brain sections from Alzheimer's cases. Interestingly, the presence of amyloid-β is also known to propagate Tau mediated cytotoxicity in Alzheimer's disease (Pooler et al., 2015). So it is clear that apoptosis of the brain cells could be prevented by treatment with Abelson tyrosine kinase inhibitors.

Prevention and disruption of amyloid fibrils

In addition to preventing the formation of amyloids leading to AD by inhibiting the enzymes/proteins responsible for generating the causative species, yet another approach for AD prevention is to inhibit amyloid formation and/or to disrupt the preformed amyloids. It has been long known that curcumin in Indian diet is believed to be the important ingredient implicated in the interference of amyloid formation (Mishra & Palanivelu, 2008; Ringman, Frautschy, Cole, Masterman, & Cummings, 2005). Either, by preventing the formation, or by disrupting/dissolving the amyloids, curcumin and other molecules such as polyphenols are being considered as small molecule inhibitors for AD prevention. Therefore, use of small molecules that are capable of destabilizing the preformed aggregates is an option for the treatment of AD and other numerous other amyloid diseases. Several natural compounds are known to target the preformed amyloid fibrils. The amyloid can itself be used as a target for the design of small molecule inhibitors. In fact, it is interesting to note that David Eisenberg and co-workers have proposed the idea of developing a common pharmacophore model of amyloid (Landau et al., 2011).

GAPS IN THE EXISTING RESEARCH

As discussed above, the 3 most important targets in Alzheimer's disease pathway are BACE1, c-Abl and amyloid-β. The inhibition of one or more of these target proteins can discourage the production of the misfolded proteins, prevent them from aggregation and also enhance the clearance of the preformed aggregates.

First generation BACE1 inhibitors developed were mostly peptide-based and displayed poor ADMET properties. Subsequent generations of inhibitors also displayed difficulties with the blood-brain barrier penetration. Even though recently, some of these β-secretase inhibitors advanced to clinical trials, so far, there are no approved drugs based on β-secretase inhibitory activity. Thus, identification of small molecules that can pass the blood brain barrier and inhibit BACE1 can be of great importance.

Abelson's Tyrosine Kinase is a well-known target in CML-Chronic Myeloid Leukemia. It is also recently confirmed to be a target (*J Alzheimers Dis*. 2011, 119-33) for Alzheimer's (also in Parkinson's disease). Inhibition of c-Abl in the brain can decrease the abnormal phosphorylation of TAU and prevent neuron apoptosis. Current drug molecules for CML are ineffective for AD due to lack of selectivity and their poor brain penetration ability. Identification of small molecule brain penetrating inhibitors with high selectivity and efficiency towards the target can be a successful strategy for the treatment of AD.

Another gap in the existing research is the identification of amyloid destabilizing compounds. Even though, several small molecule polyphenols, (Palmal, Maity, Singh, Basu, & Jana, 2014; Yang et al., 2005) natural compounds like curcumin have been identified to dissolve preformed (destabilize) amyloid fibrils from in-vitro assays, there are still no small molecule drugs

available so far. The mode of binding of these natural compounds to β-amyloid is poorly understood. Computational tools along with in-vitro assays can shed more light on the mechanism of destabilization of small molecules. MD simulation analysis of the fibril compound complexes can provide further insights into the binding mode as well as the process of destabilization in greater light.

OBJECTIVES

The following are the objectives proposed for the study:

1. Multiple crystal structures coupled with high active inhibitors for the identification of diverse inhibitors of β-secretase using e-pharmacophore and QSAR methodology.

2. Multiple e-pharmacophore strategy to identify inhibitors of Abelson's Tyrosine Kinase and selectivity assessment.

3. Computational methods for the identification of small molecules with fibril destabilizing properties and to understand fibril destabilizing mechanisms of amyloid-β.

Overview of the Thesis

This thesis primarily is focused on identifying potential modulators in the form of small molecules inhibitors for:

(i) the enzyme, β-secretase the most amenable target for AD. A recently implemented computational approach involving energy based pharmacophore modeling has been used. In addition, instead of using a single ligand bound structure for energy based modeling, multiple structures of β-secretase with bound ligands were considered. To the best of my knowledge, this is the first attempt wherein, energy based pharmacophore modeling using multiple β

secretase crystal structures was attempted. 3D based quantitative structure activity relationship (QSAR) approach was also used on a dataset of inhibitors with nano molar activities.

(ii) c-Abl, a new target implicated in AD. Energy based modeling using multiple crystal structures of c-Abl were employed in proposing lead molecules screened from CNS database. The stability of the proposed lead compounds were assessed by analyzing their docked conformations in comparison to the molecular dynamic trajectories for up to 10 ns. The final set of lead molecules is identified by applying selectivity criteria based on comparison of docking cores of these molecules against a group of closely related kinases.

(iii) disruption of Aβ-peptide amyloids. Small molecules and natural compounds as well as synthetic compounds were tested for their potential to prevent amyloid formation and or dissolving pre-formed amyloids.

The goal is to employ e-pharmacophore based computational approach together with MD simulations towards identifying potential inhibitors for two important enzymes implicated in AD, and screening for the efficacy of several small molecule using fluorescence and cellular toxicity measurements.

A chapter-wise elaboration of the thesis is presented below:

Chapter 1 deals with an e-pharmacophore based modeling of potential inhibitors for β-Secretase (BACE1), which is responsible for the pathogenesis of Alzheimer's disease. Given its role in the formation of amyloids leading to Alzheimer's disease, it has been a major therapeutic target for intervention. In the present work, we considered multiple crystal structures with bound inhibitors showing affinity in the range of 2–210 nM activity and optimize the pharmacophoric requirement based on the energy involved in binding termed as e-pharmacophore mapping. A

high throughput screening combined with molecular docking, ADMET predictions, logP values and in vitro assay led to the identification of 7 potential compounds showing inhibition at 10 μM from Fluorescence Resonance Energy Transfer (FRET) assays. These leads can be further developed as novel inhibitors for β-secretase. It is interesting to note that, a much wider diverse set of molecules could be achieved.

In Chapter 2 the strategy of ligand based pharmacophore modeling coupled with QSAR studies on BACE1 was carried out on a dataset of known inhibitors whose activities against β-secretase hovered in the nano molar range. The identified 5 feature pharmacophore model, AHHPR, was validated via three dimensional quantitative structure activity relationship as indicated by r^2, q^2 and Pearson R values of 0.9013, 0.7726 and 0.9041 respectively. For a dataset of compounds with nano molar activity, the important pharmacophore features present in the current model appear to be similar with those observed in the models resulting from much wider activity range of inhibitors. Virtual screening of the ChemBridge CNS-Set™, a dedicated database housing compounds with a better suitability for central nervous system based disorders resulted in the identification of eight prospective compounds. Molecular Dynamics Simulations of the top complexes revealed that two compounds have persistent interactions. The identified compounds by optimization can be potential therapeutics for Alzheimer's disease.

Chapter 3 is focused on the modeling studies identifying potential c-Abl inhibitors for AD. A very interesting aspect of this work is the target, c-Abl. The inherent complexity associated with the network of pathways leading to the progress of AD have lead researchers to discover additional targets that may be involved. Recent experimental findings have identified Abelson's Tyrosine Kinase (c-Abl), a non-receptor kinase involved in the abnormal phosphorylation of

tyrosine 394 of Tau protein as a new target for AD. In the present study, we employed energy optimized multiple pharmacophore modeling strategy from multiple c-Abl structures bound with ligands in the inactive ATP binding conformation. Virtual screening followed by docking of molecules from ChemBridge_CNS database, a dataset with small molecules biased towards increased blood brain penetration and bioavailability, resulted in the identification of ten best scoring molecules. MD simulations were further used to corroborate the choice from docking studies.

Selectivity is a major set-back for potential inhibitors. It is more so in the case of kinases which share a high degree of sequence and structural similarity. Therefore, in this chapter we have also attempted to assess the selectivity of the potential inhibitors in favor of c-Abl over other related kinases. Based on binding affinity analysis three out of the ten molecules appear to be more selective for c-Abl than other structurally related kinases. Given the implied role of c-Abl not only in AD but in Parkinson's disease the identified compounds may serve as leads to be developed as effective neurotherapeutics.

In Chapter 4 of the thesis, the efficacy of 15 small molecules to prevent the formation of amyloids, one of the strategies for AD therapy, was evaluated experimentally and supported by modeling. The widely used assay of measuring the increased fluorescence of Thioflavin T (ThT) when specifically bound to amyloids was used. Amyloid forming 1-42 residue Aβ peptide in AD was expressed, purified and was incubated with the molecules. ThT fluorescence assay was used to assess the amyloid destabilizing properties of small molecules. We have tested the efficacy of 15 compounds including natural compounds. Further, MD simulations were used to understand the binding pattern of these potential molecules to the structure of amyloid-β (2BEG) (Luhrs et al., 2005).

Chapter 2

Multiple e–Pharmacophore Modeling Combined with High–Throughput Virtual Screening and Docking to Identify Potential Inhibitors of β–Secretase (BACE1)

Introduction

β-secretase (BACE1) is an aspartate protease responsible for the building up of amyloid-β (Aβ) peptide is accountable for the pathogenesis of Alzheimer's disease (AD) (Arun K. Ghosh, Sandra Gemma, & Tang, 2008; Kandalepas & Vassar, 2012; Nelson et al., 2012; R. Vassar & Kandalepas, 2011; Venugopal C, Demos CM, Rao KS, Pappolla MA, & K., 2008). Although the two isoforms of β-secretase are proteases and cleave the APP, the role of BACE 2(Ahmed et al., 2010) in AD is not as clear as the role of its counterpart, BACE1 (Ahmed et al., 2010). BACE1 is highly expressed in brain and largely neuronal while BACE2 is found in peripheral tissues. The BACE2 homolog cleaves within the AB peptide. Interestingly, it was suggested that BACE2 can function as an alternative α-secretase-like enzyme cleaving with the Aβ peptide and can lower the formation of the full length Aβ peptide (Ahmed et al., 2010; R. Yan, 2001; Sun et al., 2005; R. Vassar, Kovacs, Yan, & Wong, 2009). Inhibiting β–secretase (BACE1, hereafter, β-secretase refers to BACE1) for therapeutic intervention of AD is advantageous as it prevents the production of Aβ and also causes minimal physiological consequences (Ohno et al., 2004). Therefore reduction in the production of the Aβ peptide could lead to normal brain function and for this reason β-secretase inhibition has been a therapeutic strategy towards the drug discovery for AD (A. K. Ghosh, Brindisi, & Tang, 2012). Since β-secretase is so important for the development of Alzheimer's disease, many laboratories have been working hard to find inhibitors to block its activity, hence we find many inhibitors developed and inhibitors bound

to β-secretase crystal structures are also available in the protein data bank (PDB) (Berman et al., 2000). However, as the β-secretase activity is mostly confined within the endosomes of brain neurons, inhibitors of the β-secretase pose a challenge of penetrating the blood–brain barrier(Zimmermann, Gardoni, & Di Luca, 2005) (BBB) and the neuronal membranes. As a result only a small number of β-secretase inhibitors have entered the early phase clinical trials (A. K. Ghosh et al., 2012). Contrary to the success in designing inhibitors of HIV protease which also is a member of aspartate protease family, the emergence of clinically effective small molecule candidates from ligand–based pharmacophore modeling and structure–based design routes to curtail AD has been slow. But the search continues in the fight against this debilitating disease (A. K. Ghosh, 2010).

The crystal structures of β–secretase confirm that the active site is a long cleft for substrate recognition, with two active site amino acids Asp-93and Asp-289(Dislich & Lichtenthaler, 2012; Kandalepas & Vassar, 2012; Luo & Yan, 2010) positioned at the site of the peptide bond hydrolysis.

Structural knowledge of β–secretase suggests that a preferable approach for the development of its inhibitors would be structure–based cycles(A. K. Ghosh, Kumaragurubaran, Hong, Koelsh, & Tang, 2008) rather than screening of existing chemical libraries. Ligand–based technologies, such as 2D fingerprint similarity searching, shape–based screening, and 3D-pharmacophore modeling are traditionally recognized as fast methods for screening large compound databases (J. Chen & Lai, 2006; Guner, 2002). Structure–based approaches, on the other hand, are generally more computationally expensive but can lead to structural insights and have been shown to yield more diverse actives (J. Chen & Lai, 2006; Guner, 2002). Recently, methods have emerged that attempt to capitalize on the speed of pharmacophore

screening coupled with structure–based information by developing pharmacophore hypotheses derived from protein–ligand complexes (J. Chen & Lai, 2006; Guner, 2002; Noeris K. Salam, 2009; Wolber & Langer, 2005). These methods show promise and have been used to discover novel leads (J. Chen & Lai, 2006; Guner, 2002; Noeris K. Salam, 2009; Wolber & Langer, 2005). Multiple pharmacophore modeling using various crystal structures had been successful in the past for the discovery of new leads (Z. Chen et al., 2010).

A pharmacophore model by definition (IUPAC) is an ensemble of steric and electronic features that are necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response. (Wermuth, C.G. et al. 1998) Pharmacophore can be modeled in two ways, a ligand based approach, by overlapping a set of active molecules by pulling out common chemical features essential for their activity and also using a structure-based approach, by exploring possible interaction points between the protein target and ligands based on crystal structure complexes.

In the recent years, more and more crystal structures of β-secretase in complex with small ligands and peptides have been reported, providing detailed information about the structural features β-secretase inhibitors. Based on the structural information, it is convenient now to carry out structure–based drug design for discovering novel β-secretase inhibitors as this target demonstrates multiple conformational states and active site reorganization depending on the ligand bound to it (Xu et al., 2012). To the best of our knowledge, no multiple pharmacophore–based screening has been reported for discovering β-secretase inhibitors.

Based on the above reports on advancement in the pharmacophore modeling approaches, we attempted to construct multiple energy–based pharmacophore (e-pharmacophore) models from

ten high resolution crystal structures of β-secretase bound with high affinity inhibitors with an activity range of 2–210 nM.

The method combines pharmacophore perception and database screening with protein–ligand energetic terms computed by the Glide XP (extra precision) scoring function to rank the importance of pharmacophore features as reported earlier (Eldridge, Murray, Auton, Paolini, & Mee, 1997; Friesner, Banks, Murphy, et al., 2004; Friesner et al., 2006b; T. A. Halgren et al., 2004; K. Loving, N. K. Salam, & W. Sherman, 2009). Validation of the pharmacophore models is carried out using enrichment factor calculations and identified that the pharmacophore models display efficiency in screening. All the ten pharmacophore models were confirmed effective, based on the analysis. The entire procedure used here is given in a flowchart in **Figure 2.1**. Remarkably, the hits retrieved by these effective pharmacophore models were diverse, demonstrating that different pharmacophore models may have different performances in database screening (Chen Z et al., 2010). Therefore, all these models were employed to screen the compound database Asinex for finding potent leads with structural diversity. Combining all the screened hits based on the ten pharmacophore models, followed by molecular docking and filtering based on interacting amino acids and blood brain penetration parameters, together with fluorescence based in-vitro assay, yielded seven compounds which showed inhibition at 10 μM.

These compounds could serve as potential leads for development as β-secretase inhibitors useful for the treatment of AD.

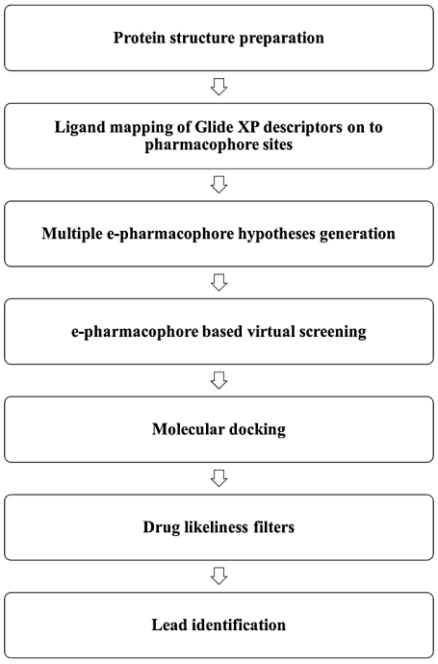


Figure 2.1 *The systematic representation of the work flow for the identification of β-secretase inhibitors.*

Materials and Methods

Protein Structure Preparation

Ten crystal structures of β-secretase bound with inhibitors whose affinity were measured to be in in the range of 2–200nM and having resolution < 1.9 Å were retrieved from PDB(Berman et al., 2002) and prepared using Protein Preparation Wizard which is part of the Maestro software package (Maestro, v9.2, Schrodinger, LLC, New York, NY). The PDB codes of the crystal structures with bound inhibitors employed in this work were 2wf1, 2vnn, 2qmf, 2qmd, 2qp8, 2wf4, 2vj7, 3lpj, 2vj9 and 2wf0 are presented in **Figure 2.2.** Bond orders and formal charges

were added for hetero groups, and hydrogen atoms were added to all atoms in the system. Water molecules in all structures were removed and the resulting structure was energy minimized.

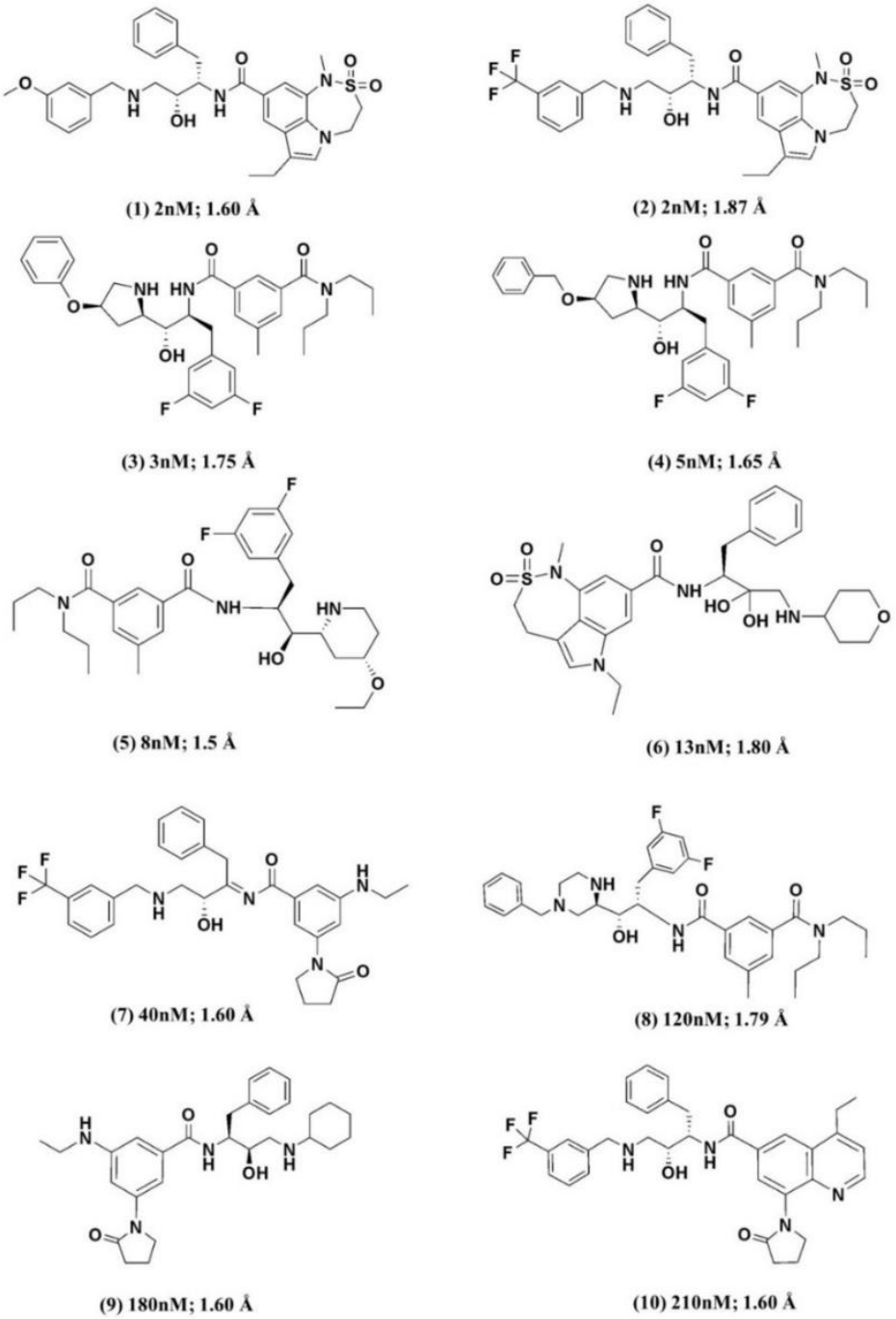


Figure 2.2 *Co-crystalized ligands obtained from β-secretase taken from the PDB structures (1) 2wf1, (2) 2vnn, (3) 2qmf, (4) 2qmd, (5) 2qp8, (6) 2wf4, (7) 2vj7, (8) 3lpj, (9) 2vj9 and (10) 2wf0. The values in μM and Å indicate the activity of the ligand and resolution of the PDB structure respectively*.

Hypothesis generation for energy–optimized structure–based pharmacophores (e–pharmacophores)

The energy optimized pharmacophore modeling method achieves the benefits of both ligand and structure-based approaches by generating an energetically optimized, structure-based pharmacophore based on energetic binding terms of Glide XP that can be used to screen millions of compounds. Glide energy grids were generated for each of the prepared complexes. The binding site was defined by a rectangular box surrounding the crystal ligand. Ligands were refined using the "Refine" option in Glide, with the option to output Glide XP descriptor information was employed (*Glide, version 5.7, Schrödinger, LLC, New York, NY, 2011*) (Friesner et al., 2006b). Default settings were used for the refinement and scoring. Starting with the refined crystal ligand, pharmacophore sites were automatically generated with Phase (Phase, v3.0, Schrodinger, LLC, New York, NY)(Dixon, Smondyrev, Knoll, et al., 2006; K Loving, N.K Salam, & W Sherman, 2009) using the default set of six chemical features: hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobe (H), negative ionizable (N), positive ionizable (P), and aromatic ring (R). Hydrogen bond acceptor sites were represented as vectors along the hydrogen bond axis in accordance with the hybridization of the acceptor atom. Hydrogen bond donors were represented as projected points, located at the corresponding hydrogen bond acceptor positions in the binding site. Projected points allow the possibility for structurally dissimilar active compounds to form hydrogen bonds to the same location, regardless of their point of origin and directionality. Each pharmacophore feature site is first assigned an energetic value equal to the sum of the Glide XP(Friesner et al., 2006b) contributions of the atoms comprising the site, allowing sites to be quantified and ranked on the basis of the energetic terms (Ligand mapping). Glide XP descriptors include terms for

hydrophobic enclosure, hydrophobically packed correlated hydrogen bonds, electrostatic rewards, $\pi...\pi$ stacking, π...cation, and other interactions (Friesner et al., 2006b). ChemScore,(Eldridge et al., 1997) hydrogen bonding and lipophilic atom pair interaction terms are included when the Glide XP terms for hydrogen binding and hydrophobic enclosure are zero.

Most of the hypotheses generated 5 feature models except 2qmd which had 6 features/sites and four of the crystal structures (2vj9, 2wf4, 2vj7 and 2wf0) yielded hypotheses with 4 features. Data indicating the number of sites present in the pharmacophore hypothesis with each of the ten crystal ligands along with the nature of features are represented in **Table 2.1** and **Figure 2.3**.

S. No	PDB code	Total no. of available sites	No. of optimized sites	Hypothesis
1	2wf1	6	5	RRADD
2	2vnn	6	5	RRADD
3	2qmf	9	5	RRAAD
4	2qmd	10	6	RRRAAD
5	2qp8	7	5	RAADD
6	2wf4	6	4	RRAD
7	2vj7	8	4	RRAD
8	3lpj	8	5	RRAAD
9	2vj9	7	4	RAAD
10	2wf0	9	4	RADD

Table 2.1 *Number of pharmacophore sites available for each ligand prior as well as post energy minimization and the final selection of optimized sites from all 10 PDB structures. A-represents an Acceptor group, D-represents a Donor group and R- represents a Ring aromatic group.*

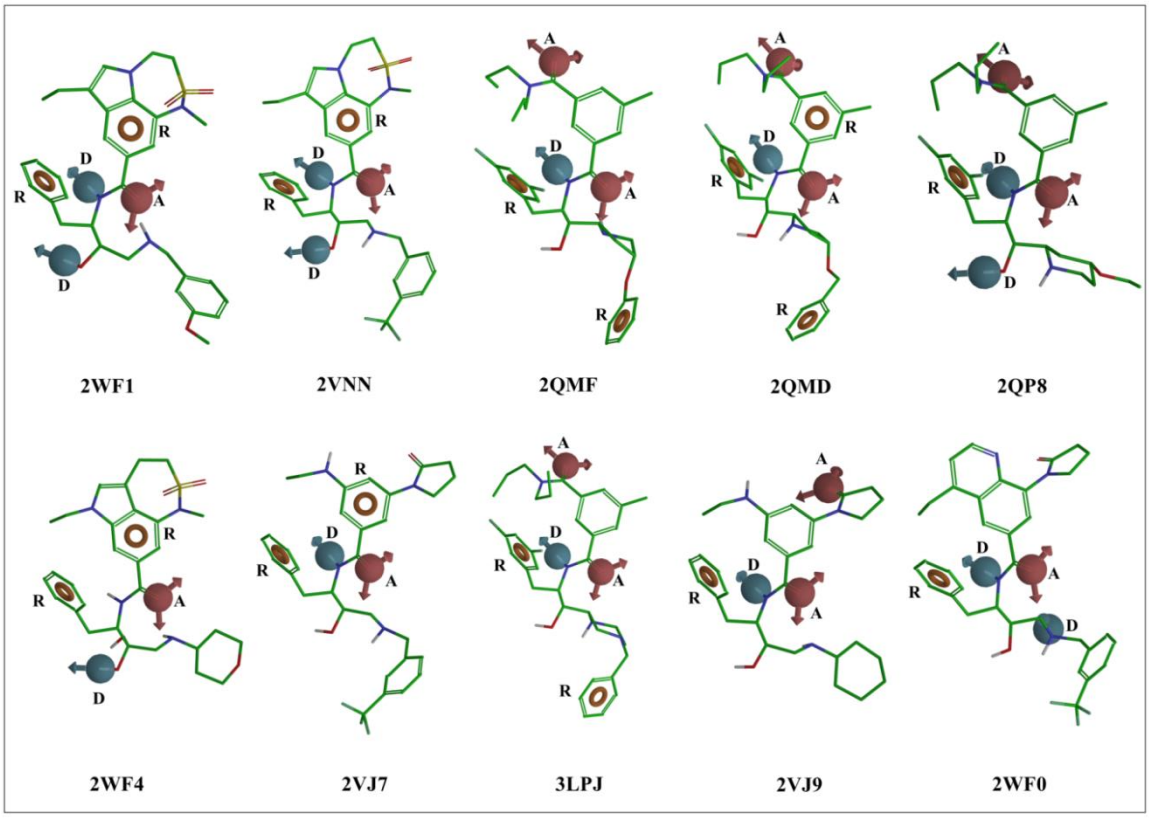


Figure 2.3 *Energy-optimized pharmacophore hypotheses generated from the 10 selected PDB structures. Brown sphere=hydrogen bond acceptor, open circle=aromatic ring, light blue=hydrogen bond donors. A-represents an acceptor group, D-represents a donor group and R-represents a ring aromatic group.*

Database Preparation

For validating the reliability of the 10 e–pharmacophore models, an initial data set of 1000 drug–like decoys, with an average molecular weight of 400 D (the "dl–400" data set), was employed. Ligand decoy sets were available for download (http://www.schrodinger.com/glide_decoy_set). The property distributions of these data sets have been well characterized (Friesner, Banks, R. B. Murphy, et al., 2004) and they have been employed in several retrospective virtual screens (Friesner, Banks, R. B. Murphy, et al., 2004; Friesner et al., 2006a). We supplemented the dl–400 data set with 30 known inhibitors for the target in the nano molar range.

The commercial chemical database Asinex over 500000 compounds was processed through redundancy checking and Lipinski filters to select compounds that have better drugability. Database molecules were prepared using LigPrep (*LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2011.*) with Epik (*Epik, version 2.2, Schrödinger, LLC, New York, NY, 2011.*). Conformational sampling was performed on all database molecules using the ConfGen search algorithm (Eldridge et al., 1997). ConfGen samples the conformations based on a heuristic search algorithm and energetic evaluations to efficiently explore diversity around rotatable bonds, flexible ring systems, and nitrogen inversions. We employed ConfGen with the OPLS_2005 force field and a duplicate pose elimination criterion of 1.0 Å rmsd (root mean square deviation) to remove redundant conformers. A distance–dependent dielectric solvation treatment was used to screen electrostatic interactions. The maximum relative energy difference of 10 kcal mol^{-1} was chosen to exclude high–energy structures. Using Phase, the database was indexed with the automatic creation of pharmacophore sites for each conformer to allow rapid database alignments and screening.

e–Pharmacophore Database Screening

For the e–pharmacophore approach, explicit matching was required for the most energetically favorable site (scoring better than –1.0 kcal mol^{-1}). Screening molecules required to match a minimum of 4 sites for hypotheses with 5 or more sites. Distance matching tolerance was set to 2.0 Å as a balance between stringent and loose–fitting matching alignment. Database hits were ranked in order of their Fitness score (Dixon, Smondyrev, Knoll, et al., 2006) a measure of how well the aligned ligand conformer matches the hypothesis based on rmsd site matching, vector alignments and volume terms. The fitness scoring function is an equally weighted composite of these three terms and ranges from 0 to 3, as implemented in the default database screening in

Phase (*Phase, version 3.3, Schrödinger, LLC, New York, NY, 2011) (Dixon, Smondyrev, Knoll, et al., 2006).*

Molecular Docking

Database ligands were docked into the binding sites of the 10 selected pdb structures with Glide 5.7 (*Glide, version 5.7, Schrödinger, LLC, New York, NY, 2011*), utilizing the high-throughput virtual screening (HTVS) scoring function to estimate protein–ligand binding affinities. Glide HTVS is faster and more tolerant to suboptimal fits than Glide XP, making it better for comparison in this work (Eldridge et al., 1997). The center of the Glide grid was defined by the position of the co-crystallized ligand. Default settings were used for both the grid generation and docking. Post docking minimization was implemented to optimize the ligand geometries. Compounds with best docking and Glide scores were then subjected to Glide XP screening. The generated grid information for each crystal ligand for docking is given in the **Table 2.2.**

S. No	PDB code	X-Centre	Y-Centre	Z-Centre
1	2wf1	29.772	1.610	34.455
2	2vnn	30.043	1.449	34.348
3	2qmf	23.960	12.383	23.052
4	2qmd	23.563	12.301	23.151
5	2qp8	24.076	12.810	23.223
6	2wf4	29.999	1.176	35.000
7	2vj7	29.904	1.829	34.201
8	3lpj	19.635	32.785	57.702
9	2vj9	23.563	12.301	23.151
10	2wf0	29.996	1.489	34.214

Table 2.2. *Grid information for the crystal structures used for the docking process.*

Enrichment Calculations

Enrichment Factor (EF) was employed for the fraction of known actives recovered when a fraction of database is screened. EF(X%) is the fraction of actives recovered after X% of decoy

database has been screened (T. A. Halgren et al., 2004). For this, we focused primarily on EF (1%), the enrichment in the top 1% of the decoys (K. Loving et al., 2009). A second enrichment metrics the Boltzmann–enhanced discrimination of receiver operating characteristic (BEDROC)(Truchon & Bayly, 2007) was also used as a way to ensure that the results and conclusions were significant. BEDROC is a generalization of the receiver operating characteristic (ROC)(Truchon & Bayly, 2007) that addresses the "early scoring problem" by Boltzmann weighting the hits based on how early they are retrieved. We used both α= 20.0 and α= 160.9 for the comparison. The value of α= 20.0 was suggested as a reasonable choice for virtual screening evaluations (K Loving et al., 2009). We have also found a value of α= 20.0 to be useful in virtual screening, which corresponds to 80% of the score being accounted in the top 1% of the database.

In-Vitro Enzyme Assay

A fluorescence resonance energy transfer (FRET) based assay was used to assess the inhibition of β-secretase by the top ranked compounds using commercially available KIT from Cayman, Michigan, USA. In the assay, the fluorescence intensity from the Swedish mutated APP peptide is an indication of the activity of β-secretase which cleaves the peptide. The fluorophore (EDANS) at the N-terminal of the peptide is quenched by the Dabcyl at the C-terminal end in the intact peptide. The compounds were purchased from the commercial databank, Asinex and dissolved in dimethyl sulfoxide (DMSO). Their potency was measured by monitoring the decrease in the fluorescence as compared to un-inhibited β-secretase activity using FLX800 Multi-Detection Microplate Reader using excitation wavelengths in the range 335-345 nm and emission wavelengths 485-510 nm. The assay was performed in triplicates on the selected

compounds at 10 μM to determine the percentage of inhibition. The kit protocol was followed with no modifications and following the suggested positive and negative controls.

Cell Viability Assay

The cell viability in the presence of the compounds identified to be inhibitors of BACE1 from the FRET assay was analyzed using MTT /cell viability assay (Mosmann, 1983). It is a reliable and simple method to evaluate the potential toxicity of drug like compounds. The cell lines were grown in T-25 flasks with 4-5 ml DMEM, 10% fetal calf serum, 50 U/ml penicillin, 50 mg/ml Streptomycin for 48 hrs. After >80% confluence is achieved, the cells were trypsinized/scraped and centrifuged. The cells were re-suspended in fresh DMEM and counted in a Neubauer counting chamber. Based on the density of cells, the cell suspension was appropriately diluted in DMEM at a density of 10000 cells per well i.e., per 100 μl. The cells were constantly stirred to ensure equal distribution while being transferred to the 96 well plates. After 24 hrs, the culture media is replaced with serum free media containing compounds at appropriate concentrations (100 and 50 μM). The compounds to be tested are initially dissolved in DMSO and later in the cell culture media to get a final concentration of 100 μM and 50 μM. The concentration of DMSO is kept below 5% to avoid cell death. The cells were maintained at 37^{o} C and 5% CO_2. MTT assay is performed by the reduction of tetrazolium salts by metabolically active cells forming Formazan crystals that are solubilized and quantified using spectrometry.

Results and Discussion

Protein preparation

The human protein β–secretase (BACE1) structures with bound inhibitors were retrieved from PDB and it was confirmed that the bound ligands showed interactions with the two important

catalytic residues, Asp93, and Asp289 (Kandalepas & Vassar, 2012). As our approach was structure based modeling, it is a prerequisite to have accurate starting structures. Here, we used a total of ten high resolution crystal structures with a resolution between 1.5 Å – 1.87 Å and activity of inhibitors (IC_{50}) ranging from 2 nM to 210 nM to enhance the chances of retrieving similarly active and diverse hits from the database screening. The prepared structures were energy minimized using the force field OPLS–2005 and are shown in **Figure 2.4.** The orientation of the ligand in the active site pocket with the interactions between the ligand atoms and protein atoms for all the 10 proteins is clearly seen. Some changes in the biding pattern of different ligands can be observed suggesting the flexibility of the active site of β-secretase.

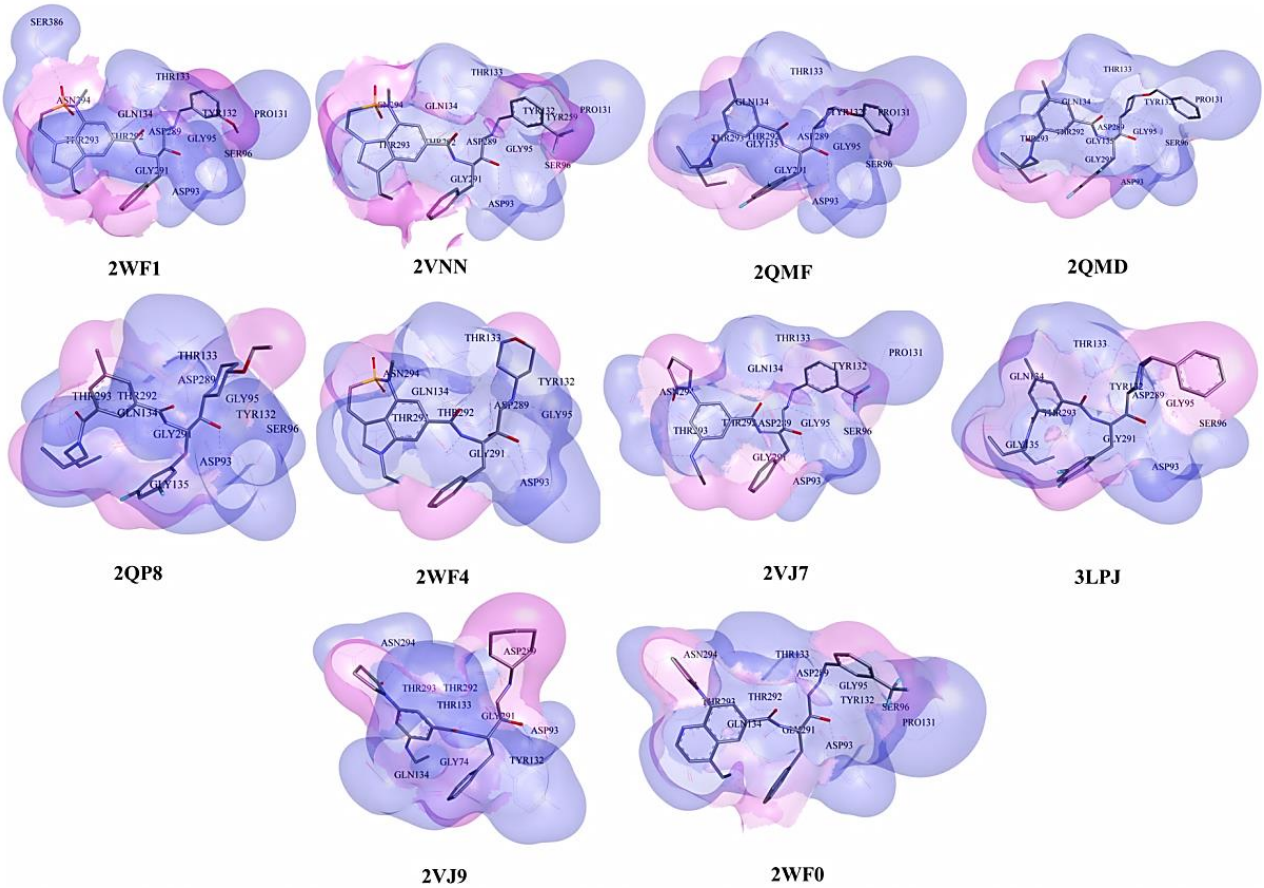


Figure 2.4 *Small changes in the binding pattern of the 10 crystal structures used indicating the flexibility of the active site pocket of β-secretase. Blue and magenta correspond to active site pocket of the protein and ligand surface respectively. *The figure is generated using the Discovery Studio 3.1.*

Pharmacophore Generation

The e–pharmacophore method which combines the aspects of structure–based and ligand–based techniques was explored for 10 crystal structures of β–secretase used in this study. The pharmacophore hypotheses were developed by mapping Glide XP(Friesner, Banks, Murphy, et al., 2004; Friesner, Banks, R. B. Murphy, et al., 2004; T. A. Halgren et al., 2004; K. Loving et al., 2009) energetic terms onto pharmacophore sites which are calculated based on the structural and energy information between the protein and the ligand. The initial number of pharmacophore sites was set up to ten for all the crystal structures. The total number of pharmacophore sites for each ligand prior to energy–based site selection and the optimized sites for hypothesis generation for the 10 crystal structures of β–secretase are provided in **Table 2.1 and Figure 2.3.** On an average, there were 7.6 sites per hypothesis, many of which do not even appear to be directly involved in protein–ligand interactions. Across all ten crystal systems, the above protocol was able to reduce the average number of sites per ligand by more than half, to 5.8, which made the hypotheses more practical for screening. **Figure 2.3** shows the final e–pharmacophores for each of the 10 crystal ligands studied.

These ten pharmacophore models showed three kinds of features: Ring Aromatic: R, Hydrogen Acceptor: A, Hydrogen Donor: D. The only pharmacophore hypothesis with six features was obtained from the crystal structure 2qmd which showed three ring aromatic (R) feature compared to other models (**Table 2.1**). The minimum featured pharmacophore model with four points were obtained with four crystal structures (2wf4, 2vj7, 2vj9, and 2wf0), among which 2wf4 and 2vj7 showed similar features, while 2vj9 and 2wf0 had an extra acceptor and an extra donor respectively in place of ring aromatic feature. Five other crystal structures (2wf1, 2vnn, 2qmf, 2qp8 and 3lpj) yielded five–point pharmacophore hypothesis in which 2wf1 and 2vnn

showed similar features of RRADD, while 2qmf and 3lpj showed similar features of RRAAD. All the ten pharmacophore models generated from ten crystal structures show three central features RAD and we calculated the distance mapping of these three features specifically **(Table 2.3)** and it was observed that the distances for all nine models except the 2wf4 pharmacophore model were within ± 0.05Å, ± 0.2 Å, and ± 0.36 Å for A to D, D to R and R to A respectively. Based on the most energetically favorable sites, we selected 4–6 features per hypothesis which were tested for performance for enriching the active compounds using a decoy set database.

S.No	PDB code	Central A to D distance	Central D to R distance	Central R to A distance	Maximum distance	Minimum distance
1	2wf1	3.158	3.849	4.779	7.382	3.158
2	2vnn	3.201	3.649	5.034	7.681	3.201
3	2qmf	3.173	3.857	4.762	13.650	3.173
4	2qmd	3.170	3.809	4.607	13.841	3.170
5	2qp8	3.172	3.853	4.638	9.498	3.172
6	2wf4	5.270	_a	5.157	7.467	3.694
7	2vj7	3.150	3.684	4.893	5.590	3.150
8	3lpj	3.170	3.671	5.038	13.933	3.170
9	2vj9	3.143	3.818	4.864	9.235	3.143
10	2wf0	3.147	3.723	4.850	7.966	3.147

'_a' indicates the absence of the central donor feature.

Table 2.3 *Distance comparison between the three central common features ADR. A-represents an Acceptor group, D-represents a Donor group and R- represents a Ring aromatic group.*

The pharmacophore model obtained from 2wf4 showed longer distances in central RAD region and hence a different pharmacophoric requirement with respect to central RAD distances had been observed. Based on this observation, analysis of differences in maximum distances in each pharmacophore models obtained from ten crystal structures were also studied as each crystal structures had differences in the active site grids. The distances are presented in **Table 2.3** and it was observed that the maximum distance between pharmacophoric points was high with

2qmd (13.84Å), 2qmf (13.65Å) and 3lpj (13.93Å), while 2vj7 pharmacophore model had the maximum distance of 5.59Å. These data are critical to assess the diversity among the pharmacophore models within the same protein target based on different ligands. The use of different pharmacophore models developed from different crystal structures can not only increase the chance of identifying active compounds but also diversity and also the flexibility of the active site can contribute to the changes in the energy pattern in various crystal structures.

Validation of constructed pharmacophore models

The enrichment results for all 10 targets using the e–pharmacophore method were compared for the enrichment factor (EF), BEDROC (α=20.0), based on recovery rate of actives against the ranked decoy database as in **Table 2.4**.

S. No	PDB code	EF 1%	BEDROC (α=160.9)	BEDROC (α=20.0)
1	2wf1	27	0.644	0.542
2	2vnn	30	0.641	0.518
3	2qmf	43	0.738	0.682
4	2qmd	27	0.679	0.563
5	2qp8	33	0.676	0.576
6	2wf4	37	0.681	0.610
7	2vj7	33	0.710	0.572
8	3lpj	30	0.518	0.564
9	2vj9	43	0.738	0.682
10	2wf0	27	0.610	0.568

Table 2.4 *Enrichment Factor at 1% of the decoy database and BEDROC at α=20.0 and 160.9.*

The average EF1% value from all the 10 pharmacophore models was $\geq$ 34 and is a good indication that this method can identify actives, while the average BEDROC values (α=20.0) were $\geq$ 0.51. The highest EF was observed with the pharmacophore model 2qmf which is a 5 feature hypothesis. Further, the recovery rate of the known actives from the constructed decoy

database versus the ranked database screened with 10 pharmacophore models were plotted and are depicted in the **Figure 2.5**.

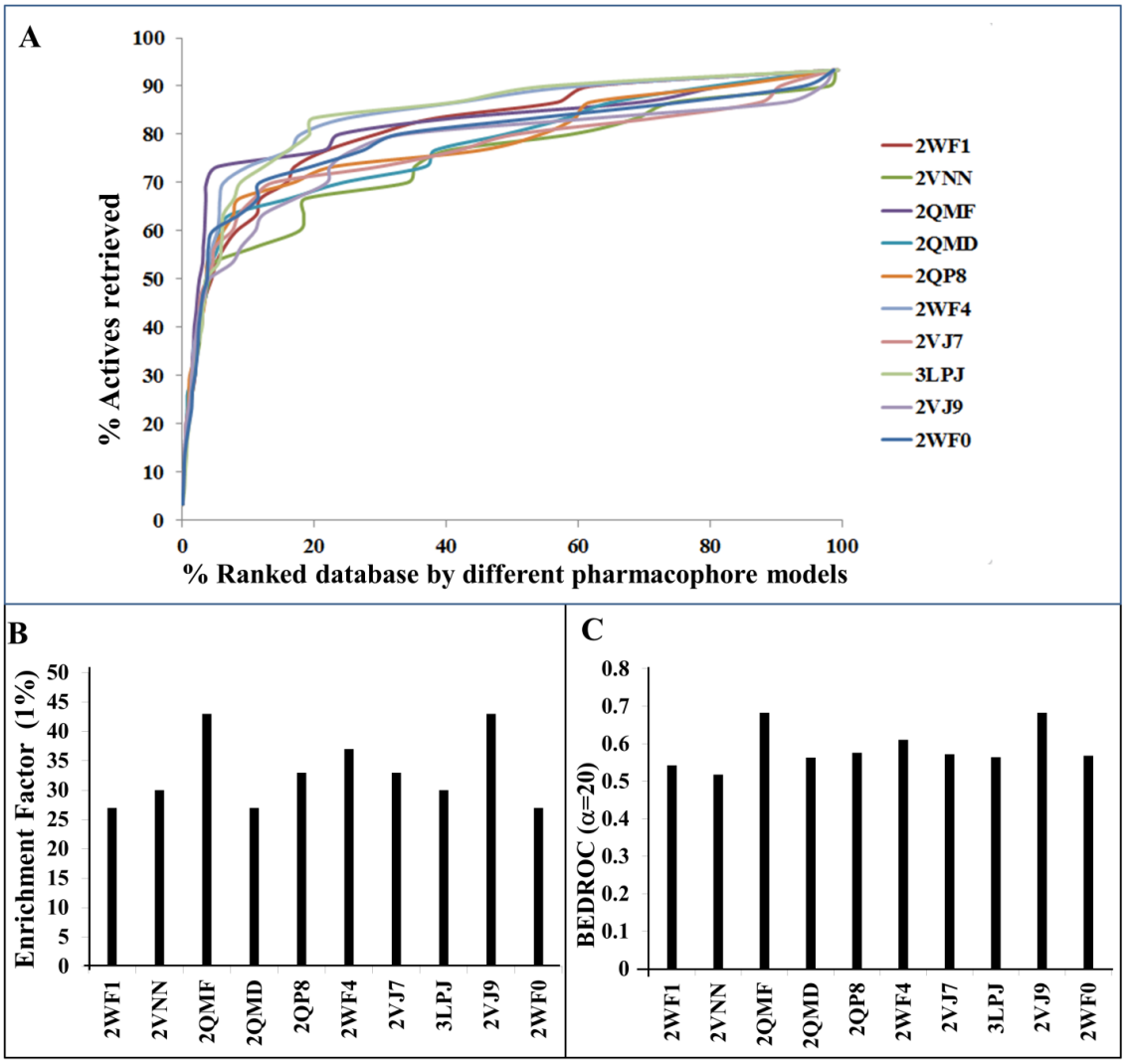


Figure 2.5 A. *The recovery rate of known actives from the constructed decoy database versus the ranked database.* ***B****. Enrichment values (EF1%) for the crystal structures used in docking and the respective PDB codes from which the model is generated.* ***C.*** *BEDROC (α=20.0) values of the crystal structures from Enrichment calculations.* ***D.*** *Correlation between EF1% and BEDROC values of the e-pharmacophore models.*

The result indicates that the pharmacophore models generated from 10 crystal structures of BACE1 could identify the actives, with a recovery rate of the known actives close to 80% in the total ranked decoy database. We also found a good correlation value of 0.9421 for EF1% versus BEDROC (α=20). Based on the validation results these multiple pharmacophore models

could be employed for thoroughly identifying potential hits. Therefore, the 10 pharmacophore models were all subjected to the virtual screening.

Multiple Pharmacophore models and Virtual Screening

A fit value is a measure of how well the ligand fits the pharmacophore. Therefore, the hits with a high fit value are probably very active. The number of the pharmacophore matches as hit compounds predicted by each of the 10 pharmacophore models from Asinex database are summarized in **Table 2.5**. It was obvious from the results that the pharmacophore model generated from 2qmd yielded the lowest number of matches which may be due to more features in the pharmacophore model (RRRAAD). For further analyzing the performance of different pharmacophore models, the extent of overlap in the hits retrieved from each pair of two different pharmacophore models were analyzed (**Figure 2.6**). The black–colored area represents the extent of overlap of the hits by the two corresponding pharmacophore models. Therefore the larger the black–colored area, the more similar the two corresponding pharmacophore models are. All of the pairs had less than 30% black–colored area indicating that the results from the 10 pharmacophore models were different from each other. In this study, the hits retrieved by the pharmacophore models with fit values above 1.5 were regarded as potential hits and were carried forward for further high throughput virtual screening. Noticeably three pharmacophore models (2wf4, 2vj7 and 2vj9) yielded hits of more than 10000. In contrast, the pharmacophore model from 2qmd was found very restrictive and retrieved very limited hits from Asinex with only 1218, demonstrating that different pharmacophore models may have quite different performance in screening a chemical database. Therefore, multiple pharmacophore models can be used to improve the overall screening efficacy.

S. No	PDB code	Phase Find matches hits	Phase compounds selected for docking	HTVS compounds	clusters	XP compounds	clusters	Aspartic acid interacting compounds[a]	clusters	Blood brain barrier penetration parameterizations	clusters	Nonpeptidic/ Small molecules	Peptidic compounds
1	2wf1	15005	2250	1753	289	909	80	53	10	33	8	27	6
2	2vnn	22535	3380	3151	421	3135	418	35	8	13	7	11	2
3	2qmf	25252	3787	3090	387	224	3	12	5	5	3	5	0
4	2qmd	1218	1218	974	85	554	104	3	1	3	2	3	0
5	2qp8	13408	2011	1461	109	1438	108	15	8	11	5	9	2
6	2wf4	73693	11053	6647	920	1058	152	14	1	9	7	5	4
7	2vj7	143985	21597	10000	645	1037	161	17	8	15	7	15	0
8	3lpj	4332	3124	1777	206	357	31	12	7	7	6	5	2
9	2vj9	135642	20346	10000	854	778	79	25	6	17	7	13	4
10	2wf0	61433	9214	9214	1224	1956	330	74	14	54	11	51	3

[a]compounds showing interactions with Asp93 and Asp289.

Table 2.5 *Screening results at various stages from pharmacophore based screening (PFM-Phase find matches screening, HTVS-High throughput virtual screening, XP-Extra Precision docking) with their respective clusters and the final set of compounds selected based on filters like catalytic aspartic acid interactions and blood brain barrier filter. Non-peptide and peptidic molecules are also shown. The numbers indicate compounds and clusters retrived at each stage.*

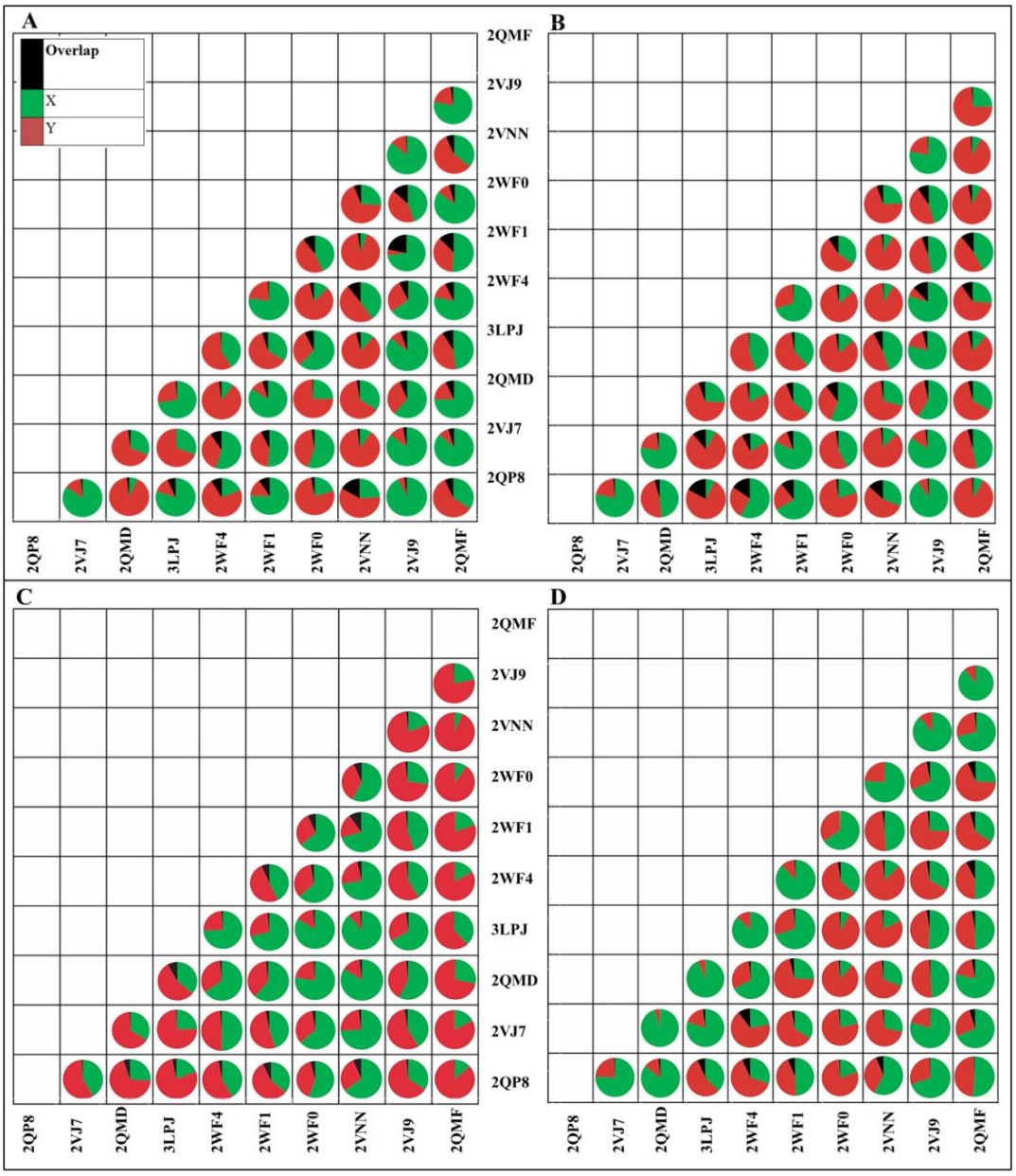


Figure 2.6 *Representation of the extent of overlapping between the compounds obtained from 10 pharmacophore models at each stage of screening.* ***A****-Pharmacophore based screening,* ***B****-HTVS docking,* ***C****-XP docking,* ***D****-Aspartic acid interaction and blood brain barrier penetration parameterization. Red and green colored regions represent the number of screened hit compounds by any two models while the part colored black represent the extent of overlap of the screened hit compounds.*

Docking study was carried out for the high fitness compounds retrieved from pharmacophore screening process using HTVS and Glide XP (**Table 2.5**). The pharmacophore based screening output was analyzed for the overlap between the pharmacophore models and presented in

Figure 2.**6A**. Similar findings were found with very few overlaps in the HTVS docking method **Figure 2.6B**, extra precision docking method **Figure 2.6C** and the aspartic acid interacting compounds along with the BBB parameterization filter **Figure 2.6D** are shown in black. Top compounds from HTVS results with a score ≥–6.0 kcal mol^{-1} and showing hydrogen bonds ≥ 2 were subjected to another round of docking by extra precision Glide (XP Glide). The XP Glide combines accurate, physics–based scoring terms and thorough sampling and the results showed scores ranging from –11.88 to –4.10 kcal mol^{-1}.

Structural diversity is an important index for the quality of the hits by an *in-silico* approach. Accordingly, we compared the structural diversity of the hits retrieved from the Asinex database. The compounds retrieved from each of the 10 crystal structure docking results were clustered using Canvas (*Canvas, version 1.4, Schrödinger, LLC, New York, NY, 2011*) a clustering algorithm from Schrodinger. The total number of clusters from each structure is represented in **Table 2.5**. The hit list in Glide XP was analyzed for structural diversity among the compounds generated (**Table 2.5**).

For further analyzing the performance of different pharmacophore models in the Glide XP screening which is based on the protein active site pocket grids employed, the extent of overlap of the hits retrieved from each pair of different pharmacophore models was studied (**Figure 2.6B**). Interestingly, the 2qmf pharmacophore model when overlapped with other 9 models, the number of common hits were less than 8 with 2vnn and 2vj9; 5 with 2wf0 and 2qp8; 2 with 2wf4 and 3lpj; 1 with 2qmd and none with 2vj7 and 2wf1. When we compare overlaps among the similar pharmacophore models, 2wf4 and 2vj7 (RRAD) showed 31 similar compounds, 2vnn and 2wf1 (RRADD) showed 437 similar compounds and with 2qmf and 3lpj (RRAAD) showed only 2 compounds. These differences could be attributed to the differences in the active

site pockets of the crystal ligand. This analysis gives an insight to the differences observed with virtual screening based on both ligand–based and structure–based hit identification methods. The XP screened compounds showed the interactions to Asp93 and Asp289, the two residues implicated in catalytic role of β-secretase **(Figure 2.7)**.

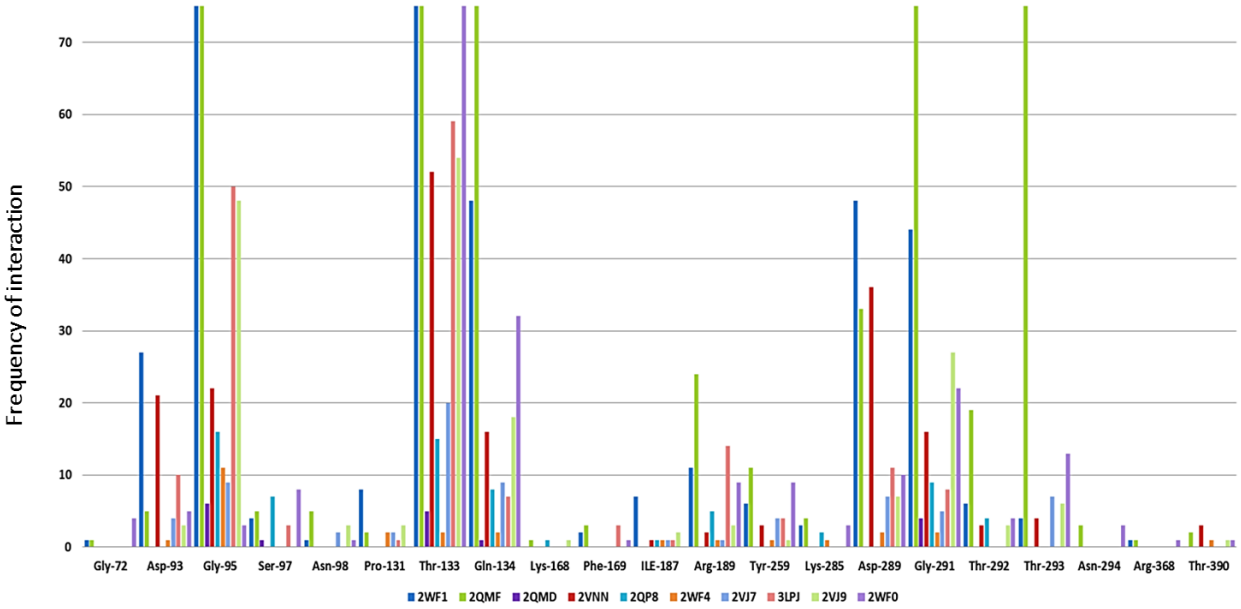


Figure 2.7 *The interaction profile of the high active compounds with the active site Asp93 and Asp289 with dock score < -8.0 kcal mol*$^{-1}$.

Final shortlisting of possible lead compounds were based on visual inspection of the important amino acid residues involved in binding. BACE1 being an aspartate protease, catalytic function can be inhibited if the compounds bind to either or both of the two aspartic acid residues Asp93 and Asp289 based on the observation that all the 10 crystal ligands employed in this work did show interactions with these residues. The number of actual hits with a score <–8.0 kcal mol^{-1} and with Asp93 and Asp289 interactions and blood brain barrier penetration parameter, filtered from each of the pharmacophore model is presented in **Figure 2.8** and **Table 2.5** along with clustering data. It appeared that the highest number of hits was generated by top four e–pharmacophore models were in the following order 2wf0 (RADD) > 2wf1 (RRADD) >2vnn (RRADD) > 2vj9 (RAAD). Except for 2wf1 and 2vnn models, all other to pharmacophore

models were different. The lowest number of 3 hits was generated by the six feature model from 2qmd.

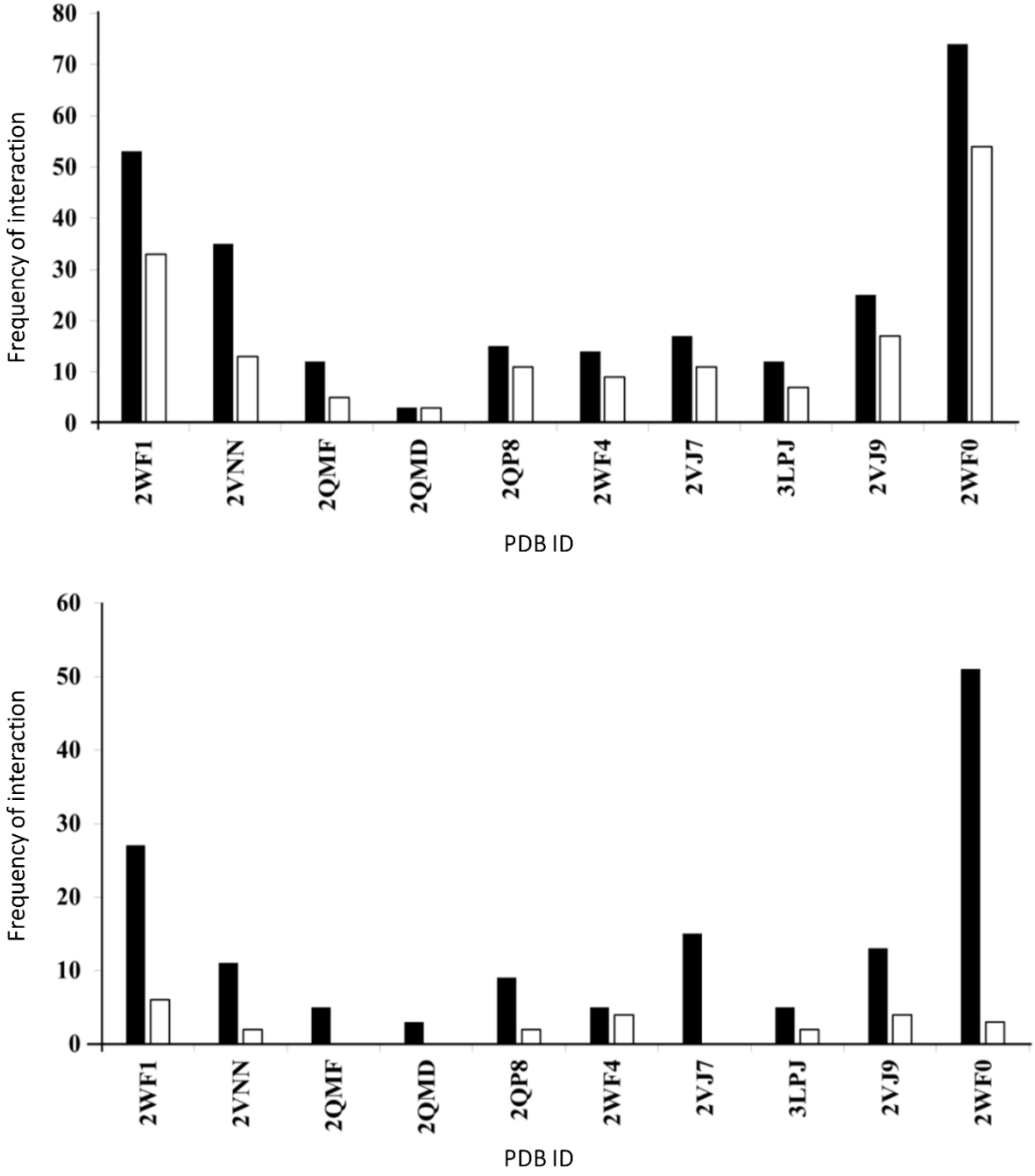


Figure 2.8A. *Asp93 and Asp289 interacting compounds before and after blood brain penetration parameterization analysis. The filled and unfilled bars represent the compounds before and after the blood brain barrier filter, respectively.* ***8B.*** *Peptidic and non-peptidic compounds short-listed at the end of the screening. The filled and unfilled bars represent non-peptidic and peptidic compounds respectively.*

Based on the previous reports on the drawback of BACE1 inhibitors being not efficiently transported to the brain, the final hits after inspection of the interacting amino acids were further scrutinized for BBB penetrating parameterization using ADMET descriptors module in Discovery Studio. Impressively, there are 7 of the e–pharmacophore models (2wf1, 2qmd, 2qp8, 2wf4, 2vj7, 2vj9 and 2wf0) in which the hit amounts retrieved was more than 60%, while the 2vnn e–pharmacophore model yielded the lowest hit amount of 37%. The final overlap of data after BBB parameterization filter is presented in **Figure 2.6D** and the data revealed that there was visible difference in the hit compound generated from each e–pharmacophore models as all of the pairs yielded similar hits less than 30%. Further analysis of the hit compounds from each e–pharmacophore was carried out on the basis of the compounds interactions with Asp 93/Asp 289 and their BBB penetration (LogP values) (**Figure 2.8A)**. Since all the crystal ligands employed in this work were peptidic, it was intriguing to observe that the non-peptidic hits outnumbered the peptidic hits (**Figure 2.8B**). The most promising hits from all the 10 structure–based design which were considered for in vitro assay are presented in **Figure 2.9.**

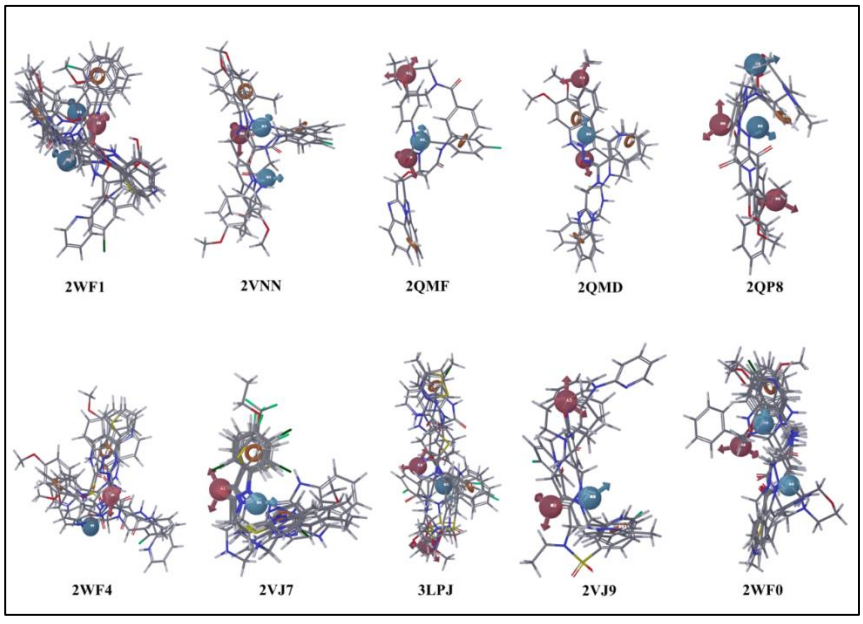


Figure 2.9 *Top 1% of the high scoring compounds aligned with their respective pharmacophore hypotheses.*

In-vitro enzymatic assay

Results from the FRET (Fluorescence Resonance Energy Transfer) based assay kit indicate that 7 compounds among 20 short-listed compounds appear to show inhibitory activity against β-secretase. The calculated percentage inhibition ranged from 68-91% at 10 μM inhibitor concentration. Structures of the 7 compounds with respective percent inhibition values are given in **Figure 2.10**. The highest active molecule showed a percentage inhibition of 91% and is derived from the pharmacophore model, 3lpj, which is a 5 point, RRAAD model.

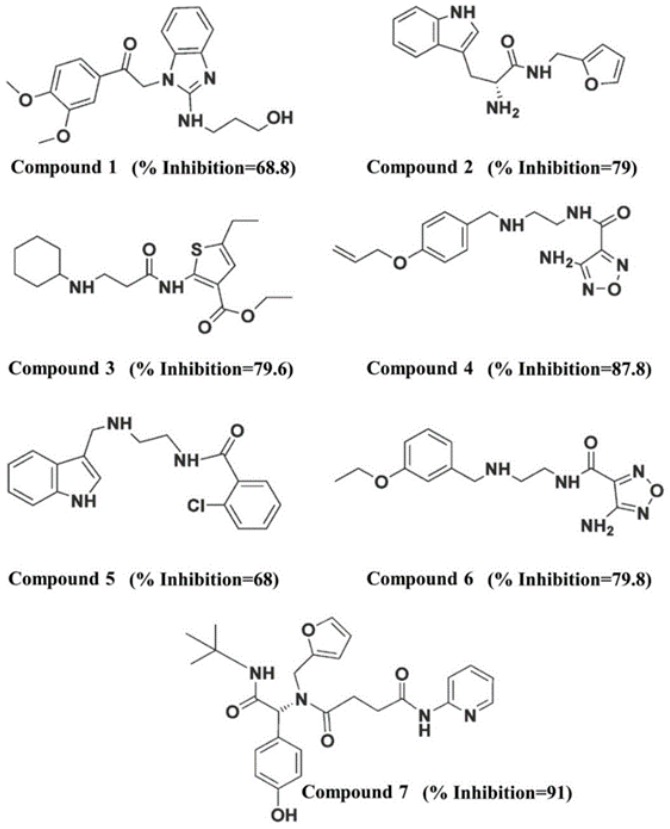


Figure 2.10 *The final set of compounds displaying the activity along with the percentage inhibition at a concentration of 10 μM.*

Cell viability assay

The 7 active compounds identified to be BACE1 inhibitors were tested for the effect of compounds on the viability of IMR-32 neuronal cells. The cells were treated with 50 and 100

μM compounds in triplicates and the results were compared with untreated cells as control. From the results/ analysis it appears that all compounds except compound 5 are non-toxic at 50 μM. In general, compounds 1, 2, 3 appears to be toxic compared to the others at 100 μM. It may be noted that, we have performed this study of toxicity on cells at druggable concentrations only, for the compounds which showed activity at 10 μM.

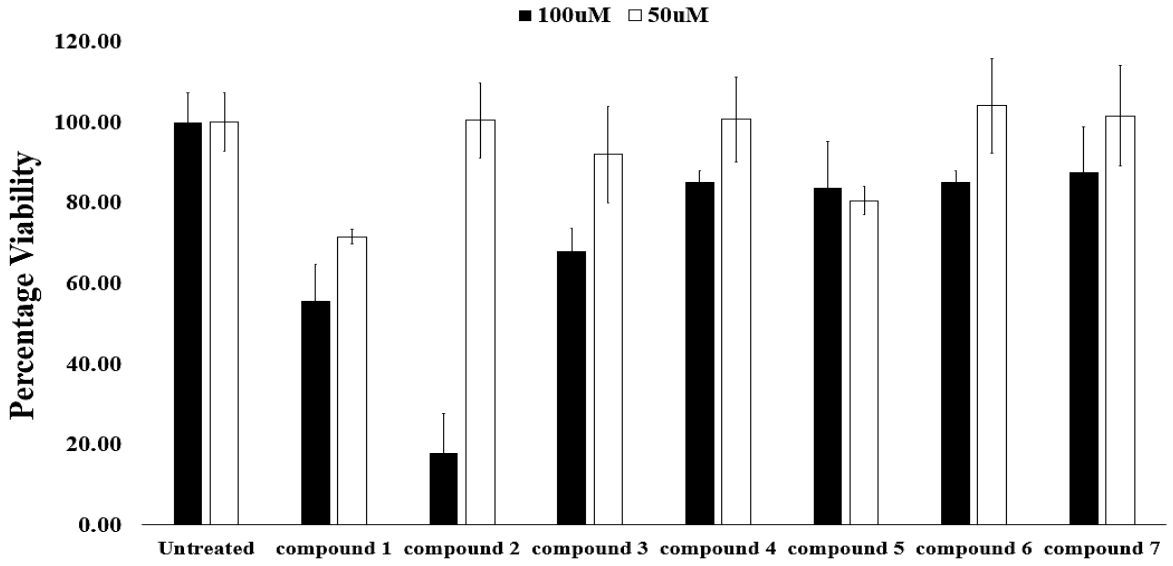


Figure 2.11 *Cell viability analysis of the 7 identified active compounds. The filled and unfilled bars represent the concentration of the compounds 100 μM and 50 μM respectively*.

CONCLUSIONS

The availability of the numerous crystal structures bound with inhibitors of BACE1 was explored using multiple pharmacophore models based on interaction energy and docking to yield diverse leads. Hence in this study a combination of ligand–based and structure–based modeling of e–pharmacophore, employed multiple crystal structures of BACE1with various filters that included pharmacophore fitness, docking score, hydrogen bonding, important aspartate interactions, logP and blood brain barrier penetrating ability had yielded strong

binding hits with requisite ADMET properties and structural diversity leading to identifying potential leads.

It is evident from the results that even for the same target the hits retrieved by different e–pharmacophore models were different. Systematic comparisons revealed that a ligand–receptor complex structure–based pharmacophore has advantages in efficiently identifying potent hits with structural diversity over simple ligand based pharmacophore search. The results also demonstrated that multiple pharmacophore models for the same target reflect different binding modes and should be used in the virtual screening.

Chapter 3

Pharmacophore based 3D-QSAR Modeling, Virtual Screening and Docking for Identification of Potential Inhibitors of β-secretase

Introduction

β-secretase (BACE1), a membrane associated aspartate protease is one of the most important protein targets for the treatment of Alzheimer's disease (A. K. Ghosh et al., 2012; Arun K. Ghosh et al., 2008; Kandalepas & Vassar, 2012; Kwak et al., 2011; Yan & Vassar, 2014). Inhibition of BACE1 can lead to cessation of the formation of β-amyloid (Aβ) peptide in Alzheimer's disease (AD) (Arun K. Ghosh et al., 2008; Kwak et al., 2011; Yan & Vassar, 2014). The development of β-secretase inhibitors being pursued for more than a decade has still not resulted in an effective therapy (A. K. Ghosh et al., 2012). However, steady progress in this field, has led to the development of inhibitors that show a wide range of activities, from nano to micro molar. Gene knock-out studies, highlight the advantage of the inhibition of β-secretase as a therapeutic intervention as it only results in minor physiological consequences (Luo Y et al., 2003; Roberds et al., 2001). Therefore, the development of inhibitors for β-secretase has been a valid therapeutic strategy towards the drug discovery efforts for AD.

Numerous studies resulted in the accumulation of large amount of experimental data in the form of identified potential inhibitors from several research groups (Kaller et al., 2012),(Cole DC, Cowling R, Lovering FE, & Wagner E, 2006; Cole et al., 2008; Dineen et al., 2012; Fobare WF, 2007; Huang H, O, & Jr, 2012 ; B. K. Malamas MS, Johnson M, Hui Y, Zhou P, Turner J, Hu Y, Wagner E, Fan & K, 2010 ; E. J. Malamas MS, Gunawan I, Barnes K, Hui Y, Johnson M, Robichaud A, Zhou & P, 2011 ; E. J. Malamas MS, Gunawan I, Turner J, Hu Y, Wagner E, Fan K, Chopra R, & Olland A, 2010 ; Niu et al., 2012; Nowak P, Jacobsen S, Robichaud AJ,

& JW., 2010; Rueeger H & Staufenbiel M, 2012; Tresadern G & Rombouts F, 2011; wahn BM, 2012 ; Weiss MM et al., 2012 ; Zuo et al., 2005). This repository of the existing inhibitors and their activity can be used for further understanding the common feature pharmacophore models for the identification of new lead compounds as well as optimization of the existing leads for enhanced activity and selectivity. Previous ligand based pharmacophore modeling studies on BACE1 identified the important features necessary for BACE1 inhibitors using data sets of compounds from a wide data range of activity, 0.01 μM to 140 μM (John, Thangapandian, Sakkiah, & Lee, 2011; Zuo et al., 2005). Modeling studies from these datasets resulted in the identification of pharmacophores with features DHHPR(John et al., 2011), AADDH(Niu et al., 2012) and AHH(Zuo et al., 2005) etc., containing hydrogen bond acceptor (A), hydrogen bond donor (D), positive ionizable (P), ring aromatic (R) and hydrophobic features (H) as predominant features (John et al., 2011; Niu et al., 2012).

In the current study we aspired to build and explore the common feature pharmacophore hypothesis from BACE1 inhibitors with an activity range of 1nm to 1000nM to understand the distinct contributing features for their high potency. Typically the selection of known inhibitors, for relating quantitatively the structure and activity, covers a range of at least one, two or three logarithmic units of activity (Ravichandran Veerasamy, 2011). The nano-molar range of chosen inhibitors falls within one order of magnitude. Thus, we have employed this approach to identify and explore the contributing features of inhibitors in the sub micro molar range (<1μM) with a dataset of reported BACE1 inhibitors. This order of magnitude of activity data may be useful in identifying the most critical features necessary for the identification of highly specific BACE1 inhibitors.

Further, using structural alignment docking followed by pharmacophore based QSAR generation, the field distributions of the various features in the pharmacophore models were analyzed using partial least square (PLS) factor analysis as a step towards validation of the identified pharmacophore model. The common features of the dataset compounds were aligned with the pharmacophore features and the resulting pharmacophore model was used for virtual screening of ChemBridge CNS-Set™ database. Based on the docking score and the interactions with the active site residues, potential BACE1 inhibitors were identified.

Materials and Methods

Dataset Pool

A pool of 58 reported β-secretase inhibitors (BACE1) with an activity reaching 1μM were considered. (**Figure 3.1**). The dataset compounds were downloaded from CHEMBL database. The reported IC_{50} values of the compounds are represented as the negative logarithm of IC_{50} (pIC_{50}) and shown in **Table 3.1**. Appropriate protonation states, addition of hydrogens and generation of conformers at neutral pH using OPLS_2005 force field was achieved using the Ligprep module. (LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2011.) The computations were carried out on an Intel Core 2 Duo E7400 2.80GHz processor and 2GB RAM running with the RHEL 5.2 operating system. PHASE 3.3 implemented in the Maestro 9.2 software package (Schrodinger, LLC) was used to generate pharmacophore and QSAR models.

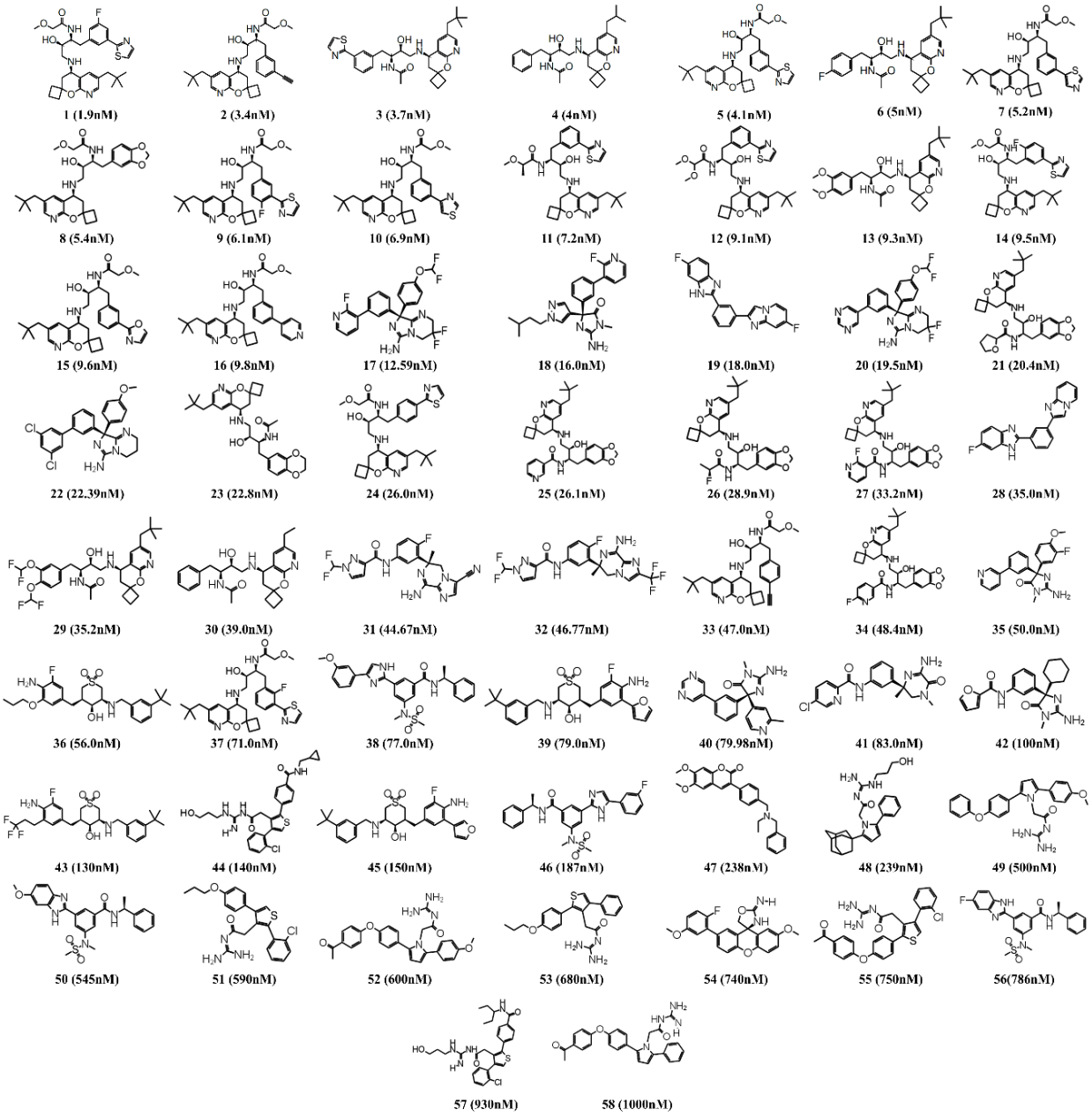


Figure 3.1 *The 2D structures of the dataset compounds indicating the experimentally reported inhibitory activity (in parentheses) against BACE1.*

Multiple Conformer Generation from the Dataset Pool

Multiple conformations with a maximum of 32 conformers per dataset compound were generated using Ligprep module (LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2011.) with Epik (Epik, version 2.2, Schrödinger, LLC, New York, NY, 2011) (Greenwood, Calkins, Sullivan, & Shelley, 2010). The dataset compounds were docked into the crystal structure, 2QMF,[27] for aligning the compounds within the active site pocket for proper alignment before producing the 3D descriptors needed for common feature pharmacophore development, only the aligned compounds are retained (**Figure 3.2**).

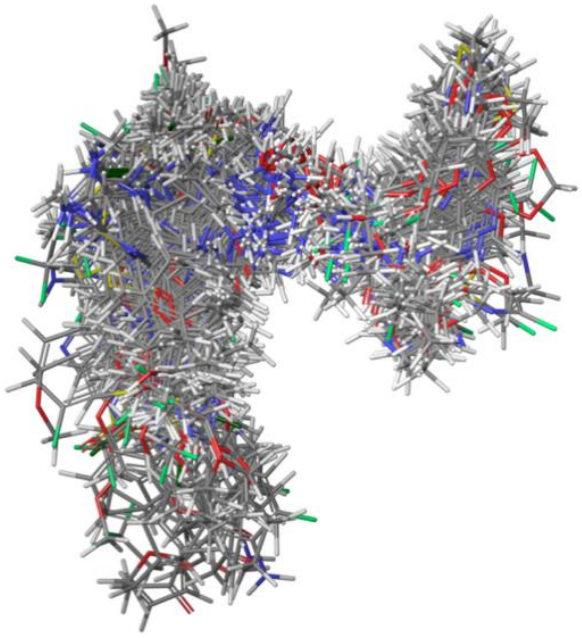


Figure 3.2 *Superimposition of the 58 compounds from the dataset used for generation of pharmacophore model.*

Pharmacophore Modeling and Validation

For alignment, the compounds from the dataset, were docked into the active site pocket of the crystal structure of BACE1 (2QMF) (Iserloh et al., 2008). Pharmacophore generation was performed using Phase module, (Phase, version 3.3, Schrödinger, LLC, New York, NY, 2011)

(Dixon, Smondyrev, Knoll, et al., 2006; Dixon, Smondyrev, & Rao, 2006). The dataset compounds were divided as high, moderate and low actives. Active compounds were used as representatives for the active group and low active compounds were selected as inactive group. Conformation generation protocol was used for the dataset compounds to identify diverse conformations. A default feature selection with acceptor (A), donor (D), hydrophobic (H), negative (N), positive (P) and aromatic ring (R) were used to create pharmacophore sites. This was followed by the clustering and rescoring of the models followed by the alignment of the dataset compounds to the pharmacophore models to obtain various statistical parameters to assess the pharmacophore hypotheses. Based on the parameters including (i) Survival score which measures the quality of alignment for the particular pharmacophore model, (ii) Site score is an indication of how closely the site points are superimposed in an alignment to the pharmacophore of the structures that contribute to the hypothesis. It is based on the root mean square deviation of the site points of a ligand from those of the reference ligand, (iii) Vector score which represents how well the vectors for acceptors, donors and aromatic rings are aligned in the structures that contribute to the hypothesis, (iv) Volume score, is the ratio of the common volume occupied by the matching conformer and the reference conformer, to the total volume (the volume occupied by both), (v) Energy value indicates relative conformational energy, (vi) Activity indicates the alignment of the pharmacophore to the most active ligand, the best of the pharmacophore models was validated further.

Database Screening and Docking

ChemBridge CNS-Set™ is a pool of more than 50,000 compounds with a higher prospect of oral bioavailability and blood-brain barrier penetration. Therefore this collection of compounds is more suitable for drug discovery efforts focused on diseases of the central nervous system

such as Alzheimer's and Parkinson's. Pharmacophore based virtual screening was performed by using Phase Find Matches together with QSAR information using Maestro software package, (Maestro, v9.2, Schrödinger, LLC, New York, NY) (Sastry GM, 2013). Docking was carried out by using Glide with the crystal structure, 2QMF. (Iserloh U & R, 2008) The positional co-ordinates of the crystal ligand were used to create the grid file necessary for docking of the database molecules. The docking studies were performed using Glide in two steps, a high throughput virtual screening (HTVS) and extra precision modes respectively.

Molecular Dynamic Simulations

MD simulations of the top three complexes were performed to understand the persistence of the interactions identified in the docking studies. As docking involves a rigid protein structure, it is necessary to evaluate the binding mode of the ligand in a flexible protein system. The docked protein-ligand complexes were subjected to 10 ns MD simulation analysis using Desmond software.

Each system comprising the protein/ligand complex (A, B and C) is generated individually using the system builder application. Each complex is solvated using SPC water molecules and neutralized in an Orthorhombic box with appropriate number of counter ions (Na^+). The box volume is then minimized. Using NPT ensemble at a temperature of 300 kelvin and 1 bar of pressure. The thermostat used is a Nose-Hoover chain method. Martyna-Tobias-kleinbarostat is used for pressure. Interaction cutoff radius for short range method is 10 A^o and a smooth particle mesh Ewald is used in the long range method. A 10 ns MD simulation is carried out with a recording interval of 1.2 ps. Root mean square deviations (rmsd), root mean square fluctuations (rmsf) and persistence of interactions are observed and compared between the complexes.

Results and Discussion

Nature of Dataset Pool

The compounds from the dataset as earlier indicated fall within an activity range of 1 nm to 1 μM. Their reported activity is from a uniform assessment of BACE1 inhibition measured from Fluorescence Resonance Energy Transfer (FRET) experiments. It may be noted that based on the alignment of the compounds on to the crystal structure, they show similar binding interactions to BACE1. Interestingly, of the approximately sixty dataset compounds a majority of them are hydroxyethylamine based (40%) (Dineen et al., 2012; Kaller et al., 2012; Rueeger H & Staufenbiel M, 2012; Weiss MM et al., 2012). Other compounds are acylguanidine,(Cole DC et al., 2006; Cole et al., 2008; Fobare WF, 2007) aminoimidazole,(wahn BM, 2012) aminohydantoin,(B. K. Malamas MS, Johnson M, Hui Y, Zhou P, Turner J, Hu Y, Wagner E, Fan & K, 2010 ; E. J. Malamas MS, Gunawan I, Barnes K, Hui Y, Johnson M, Robichaud A, Zhou & P, 2011 ; E. J. Malamas MS, Gunawan I, Turner J, Hu Y, Wagner E, Fan K, Chopra R, & Olland A, 2010 ; Nowak P et al., 2010) aminopiperazinone (Tresadern G & Rombouts F, 2011) and aminooxazoline derivatives (Huang H et al., 2012). Diversity and activity of the 58 dataset compounds was taken into consideration to divide them into training and test compounds.

Pharmacophore Hypothesis

A common feature pharmacophore hypothesis model to identify the important interactions between BACE1and the inhibitors was generated. The IC_{50} values (1.9 nM to 1000 nM) were converted into pIC_{50} and the pIC_{50} values range from 8.7 to 6.0 (**Table 3.1**). The compounds were divided into 3 groups based on the activity range from high active to low active. The high

active group with activity 8.7 to 8.1 were selected as representatives for common pharmacophore hypothesis generation. Default interaction features such as acceptor (A) donor (D) hydrophobic (H), positive ionizable (P), negative ionizable (N), ring aromatic (R), were used for generating favorable pharmacophore models. By using Phase common pharmacophore generation protocol, with default values, three best 5 feature pharmacophore hypotheses namely AHHPR, AHHRR and AAPRR were short-listed.

Ligand	IC_{50} (nM)	$-logIC_{50}$ (Actual)	Ligand	IC_{50} (nM)	$-logIC_{50}$ (Actual)
1	1.9	8.72	**30**	39	7.4
2	3.4	8.46	**31**	44.67	7.35
3	3.7	8.43	**32**	46.77	7.33
4	4	8.39	**33**	47	7.32
5	4.1	8.38	**34**	48	7.31
6	5	8.3	**35**	50	7.3
7	5.2	8.28	**36**	56	7.25
8	5.4	8.26	**37**	71	7.14
9	6.1	8.21	**38**	77	7.11
10	6.9	8.16	**39**	79	7.1
11	7.2	8.14	**40**	79.98	7.09
12	9.1	8.04	**41**	83	7.08
13	9.3	8.03	**42**	100	7
14	9.5	8.02	**43**	130	6.88
15	9.6	8.01	**44**	140	6.85
16	9.8	8	**45**	150	6.82
17	12.59	7.9	**46**	187	6.72
18	16	7.79	**47**	238	6.62
19	18	7.745	**48**	239	6.62
20	19.5	7.71	**49**	500	6.3
21	20.4	7.69	**50**	545	6.26
22	22.3	7.65	**51**	590	6.22
23	22.8	7.64	**52**	600	6.22
24	26	7.58	**53**	680	6.16
25	26.1	7.58	**54**	740	6.13
26	28.9	7.53	**55**	750	6.12
27	33.2	7.47	**56**	786	6.1
28	35	7.45	**57**	930	6.03
29	35.2	7.45	**58**	1000	6

Table 3.1 *The IC_{50} values in nano molar (nM) and the corresponding $-logIC_{50}$ values of the data set compounds.*

The common feature pharmacophore hypotheses were retained based on the parameters listed in **Table 3.2** and by also by taking into account highest number of matching pharmacophore sites and root mean square error values between conformers. Further, validation for the best model among the three hypotheses is based on the external decoy set method having 15 known BACE1 inhibitors and 1000 decoy molecules from the dl-400 decoy database (Schrodinger) (T. A. Halgren et al., 2004). Based on the Enrichment Factor (EF) value of 93.33% and Goodness of Hit value of 0.68, AHHPR was considered the best pharmacophore compared to the other two models. For validation purpose, the phase pharmacophore based screening is initiated with the combined active as well as decoys following which the enrichment factor values were generated using the formula [(Ha/4HtA)(3A + Ht)*(1 – (Ht – Ha)/(D –A))], wherein the number of molecules in the database is denoted by (D), total no. of actives in the database (A), total hits (Ht), active hits (Ha). The goodness of hit value 0.6 to 0.8 indicates highly acceptable model (John et al., 2011). In summary, the 5 feature model AHHPR was identified as the best pharmacophore hypothesis and is shown in **Figure 3.3** along with distances between the features. **Table 3.3** summarizes details of the number of actives, decoys, total hits, number of actives retrieved for enrichment validation of AHHPR.

Model	Survival active	Survival inactive	Site	Vector	Volume	Energy	Activity
AHHPR	3.851	2.939	0.95	0.999	0.897	9.077	8.721
AHHRR	3.838	2.781	0.97	0.987	0.882	7.821	8.721
AAPRR	3.833	3.723	0.95	0.999	0.884	9.077	8.721

Table 3.2 *Scores for the parameters used for the selection of the pharmacophore model.*

S.no	Parameters	Results
1	Number of molecules in the database (D)	1015
2	Total no. of actives in the database (A)	15
3	Total hits (Ht)	23
4	Active hits (Ha)	14
5	% Yield of actives	60.86
6	Enrichment Factor	93.33
7	False negatives	1
8	False Positives	9
9	Goodness of Hit	0.68

Table 3.3 *Statistical parameters of Good of Hit score validation for the pharmacophore model AHHPR. The formula for the calculation is [(Ha/4HtA) (3A + Ht) X (1 – (Ht – Ha)/(D –A))]. Goodness of hit score of 0.6-0.8 indicates a very good model.*

When compared, similarities may be observed between the pharmacophore (AHHPR) and the pharmacophore models identified earlier using different datasets. For instance the two hydrophobic features identified here along with P and R features are also present in DHHPR model proposed by Hualiang Jiang et.al. 2005 (Zuo et al., 2005). Although, not to a significant extent, the other earlier identified model, AADDH, shares A and H features with our current hypothesis,(John et al., 2011) with the exception of the absence of the donor feature. It may be noted that, although the model from our study may share common features, the distances and orientation in 3D space can be different.

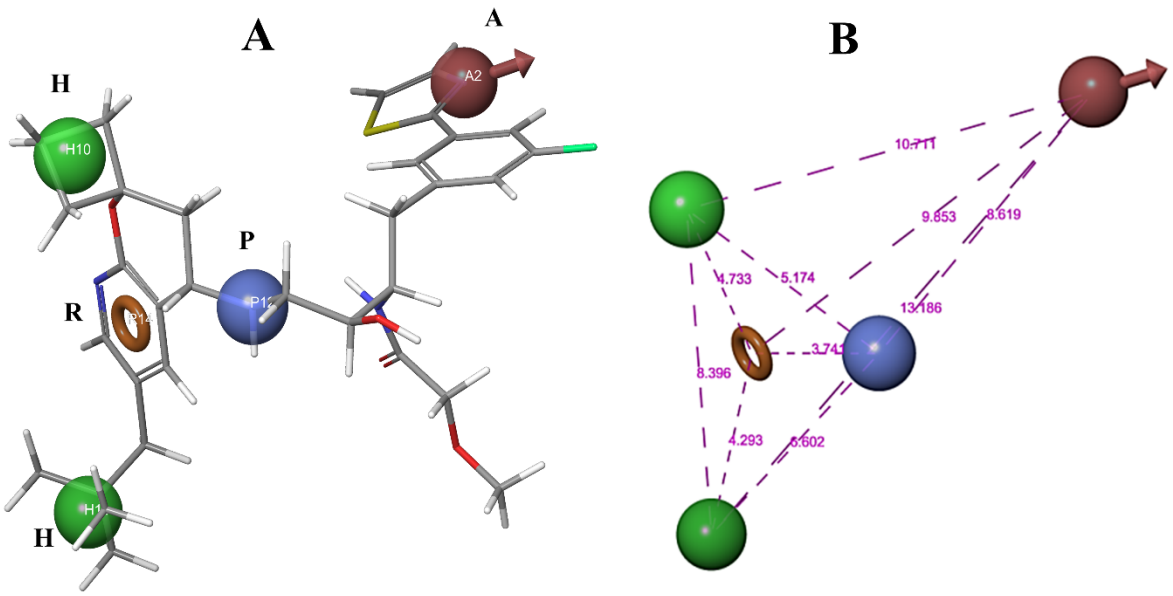


Figure 3.3 *(A) The pharmacophore model AHHPR aligned with the highest active compound in the dataset. (A) Acceptor Group, (H) Hydrophobic, (P) Positive ionizable and (R) Aromatic ring. (B) The inter-feature distances between the pharmacophore features AHHPR represented as pink lines with the distances in Angstroms.*

Activity Prediction of Test Set Compounds using 3D QSAR

The dataset of 58 compounds was divided as training and test set compounds from random seed optimization using the Build QSAR module of the Maestro software to assess the predictive ability of the pharmacophore hypothesis, AHHPR, in comparison to the other closely related models, AHHRR and AAPRR employing 3D QSAR approach. All dataset compounds were aligned to the three pharmacophore models including compounds that are not properly aligned (non-model ligands). Using the training set compounds, the Build QSAR Model algorithm predictability of the test set was assessed. The pharmacophore models show reasonably good correlation between actual and predicted activities (**Table 3.4**) as indicated by the r^2 and q^2 correlations. QSAR prediction is supported by higher F values indicating a more statistically significant regression along Pearson-R (Pearson correlation coefficient) values close to unity representing stability and reliability of the models.

S. No	ID	R-squared	F	Q-squared	Pearson-R
1	AHHPR	0.9013	109.6	0.7726	0.9041
2	AHHRR	0.8095	51	0.6729	0.8502
3	AAPRR	0.8854	92.7	0.6975	0.8548

Table 3.4 *Correlations observed for the three best pharmacophore models. r^2 represents the correlation of the training set, q^2 represents the correlation of the test set.*

The AHHPR model exhibited values for the training set and test set with a correlation of 0.90 and 0.77 respectively (**Figure 3.4, Table 3.4**). The test set compounds, in general, exhibited predicted activity close to the observed activity (**Table 3.5**). Three compounds could not be aligned and thus their predicted activity is not reported. The contour maps represent the field distribution in a series of equally spaced bins. As the model used here is a pharmacophore based QSAR model, the contour fields are aligned to the pharmacophore model. The contour maps also make it easy to interpret the distribution of the fields on the mapped pharmacophore model as well as understanding the type of feature involved. The negative fields are marked in red whereas the favorable fields are marked blue.

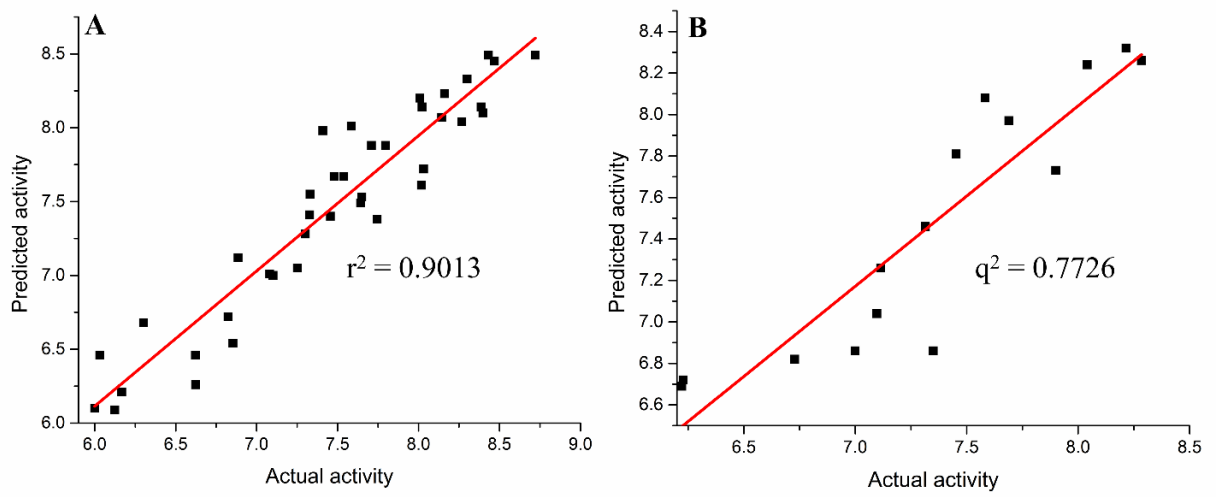


Figure 3.4 *Scatter plots between observed versus predicted activity using thirty nine compounds in the training set (A) and validated using fifteen compounds in the test set (B). The r^2 and q^2 values 0.90 and 0.77 indicate training and test set correlations respectively.*

The contour map giving the field distribution for a given feature in the highest actual active compound is shown in **Figure 3.5**. As can be seen from the figure, the role of hydrophobic feature is significant. This appears to be consistent with the results obtained from using a wider range of active dataset (John et al., 2011; Niu et al., 2012). The analysis of the contour maps based on the ligand based pharmacophore revealed that the hydrophobic groups in this pharmacophore model contribute to the most positive energy with a large field distribution which is well in agreement with the pharmacophore model generated by John et. al. 2011 (John et al., 2011).

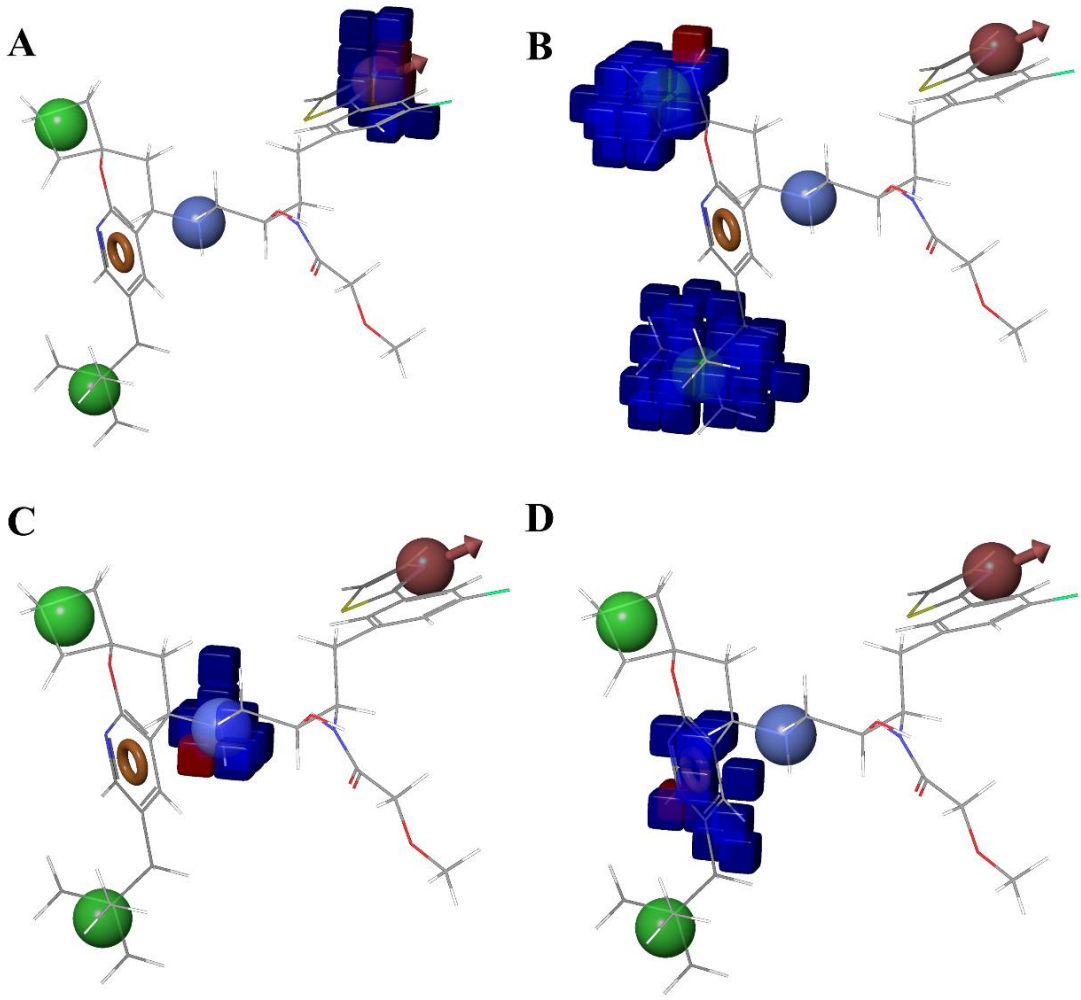


Figure 3.5 *QSAR contour map representation of the most active ligand with both positive (blue) and negative (red) field distribution of the respective feature, (A) Acceptor Group, (B) Hydrophobic, (C) Positive ionizable and (D) Aromatic ring.*

Ligand Name	QSAR Set	-logIC$_{50}$ (Actual)	-logIC$_{50}$ (Predicted)	Ligand Name	QSAR Set	-logIC$_{50}$ (Actual)	-logIC$_{50}$ (Predicted)
1	training	8.72	8.49	**28**	training	7.45	7.4
2	training	8.46	8.45	**29**	test	7.45	7.81
3	training	8.43	8.49	**30**	training	7.4	7.98
4	training	8.3	8.1	**31**	test	7.35	6.86
5	training	8.38	8.14	**32**	training	7.33	7.55
6	training	8.3	8.33	**33**	training	7.32	7.41
7	test	8.28	8.26	**34**	test	7.31	7.46
8	training	8.26	8.04	**35**	training	7.3	7.28
9	test	8.21	8.32	**36**	training	7.25	7.05
10	training	8.16	8.23	**38**	test	7.11	7.26
11	training	8.14	8.07	**39**	training	7.1	7
12	test	8.04	8.24	**40**	test	7.09	7.04
13	training	8.03	7.72	**41**	training	7.08	7.01
14	training	8.02	8.14	**42**	test	7	6.86
15	training	8.01	7.61	**43**	training	6.88	7.12
16	training	8	8.2	**44**	training	6.85	6.54
17	test	7.9	7.73	**45**	training	6.82	6.72
18	training	7.79	7.88	**46**	test	6.72	6.82
19	training	7.74	7.38	**47**	training	6.62	6.26
20	training	7.71	7.88	**48**	training	6.62	6.46
21	test	7.69	7.97	**49**	training	6.3	6.68
22	training	7.65	7.53	**51**	test	6.22	6.72
23	training	7.64	7.49	**52**	test	6.22	6.69
24	training	7.58	8.01	**53**	training	6.16	6.21
25	test	7.58	8.08	**55**	training	6.12	6.09
26	training	7.53	7.67	**57**	training	6.03	6.46
27	training	7.47	7.67	**58**	training	6	6.1

Table 3.5 *The dataset compounds divided into training and test sets with the actual activity converted to pIC$_{50}$. The QSAR predicted activity values (**-logIC$_{50}$ (Predicted)**) values were also indicated.*

Virtual screening

The crystal structure of BACE1 (2QMF)(Iserloh et al., 2008) is employed for docking as it has a high enrichment value compared to the other β-secretase crystal structures (2WF1, 2VNN,

2QMD, 2QP8, 2WF4, 2VJ7, 3LPJ, 2VJ9 and 2WF0) as identified by using thirty known BACE1 inhibitors and thousand decoy compounds as reported in our previous study (Palakurti, Sriram, Yogeeswari, & Vadrevu, 2013). Higher enrichment values are generally preferable as they significantly increase the retrieval of hit compounds as well as remove false positives during virtual screening. The pharmacophore model AHHPR was used to screen ChemBridge CNS-Set™ database using the Phase Find Matches screening module. The screening resulted in the identification of ~2500 hits. The fitness score resulting from the alignment of the database compounds reflects the degree of alignment of the compounds to the pharmacophore features. A cut-off >1, was used to shortlist the compounds for docking into BACE1 crystal structure (2QMF). 2QMF Grid is generated using the Grid generation protocol and optimized to produce the crystal pose with <1 A^{o} rmsd. The docking was performed in two steps, high through virtual screening (HTVS) followed by extra precision, (XP) screening. A binding affinity/docking score of >-8.0 (kcal/mole) along with more than two hydrogen bond interactions were used to identify the lead compounds. The eight compounds identified from the ChemBridge CNS-Set™ (Figure 3.6) with their respective docking scores and number of hydrogen bond interactions and stacking interactions are shown in **Figure 3.7** and **Table 3.6**.

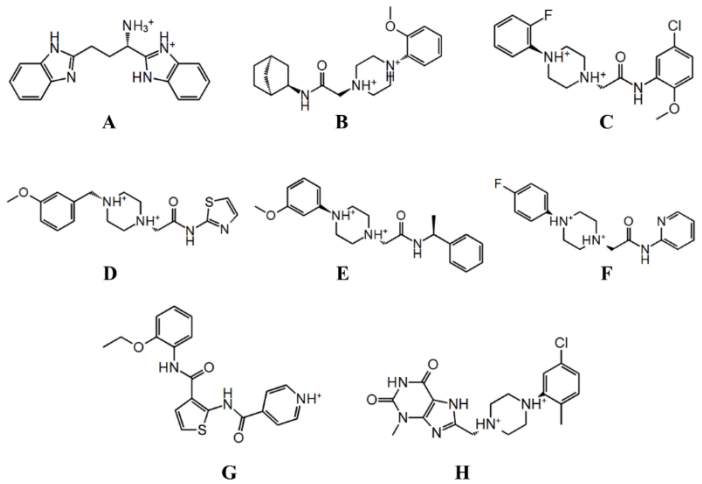


Figure 3.6 *2D structures of the potential lead compounds for BACE1 inhibition identified from ChemBridge CNS-Set™ database.*

Compound	Docking Score (2QMF) (kcal/mole)	Hydrogen bond interactions	Stacking interactions
A	-11.121	2	1
B	-10.190	4	0
C	-9.822	4	0
D	-9.340	4	1
E	-8.963	4	0
F	-8.932	5	2
G	-8.435	6	0
H	-8.317	3	0

Table 3.6 *Docking Score**,** number of hydrogen bonds, stacking interactions of the top 8 hit compounds (A-H) when docked into the crystal structure of BACE1 (2QMF) to a docking score > 8.0 kcal/mole for the corresponding compound.*

The eight compounds are involved in hydrogen bonding interactions with the active site residues Asp-32 (Asp-93,), Gly-95, Thr-133, Asp-228(Asp-289), Gly-291. The protein ligand interactions along with the interacting residues and hydrophobic contacts are represented in **Figure 3.7** using PoseView (Stierand & Rarey, 2010). As seen from the **Figure 3.7**, the tyrosine 132 residue is involved in hydrophobic contacts. For the final set of 8 lead compounds it may be noted that the predicted activity range as observed from the Phase Find Matches module taking into account the QSAR aspect is between 10-200 nM. The lead compounds identified represent piperazin, thiazole and benzodiazole based derivatives. Interestingly, a couple of compounds are similar to the existing BACE1 inhibitors. For instance, compound A resembles an inhibitor that was identified from our previous e-pharmocophore based modeling and discovery of potential BACE1 inhibitors (Palakurti et al., 2013). However, minor differences in their structure and functional groups can be noticed.

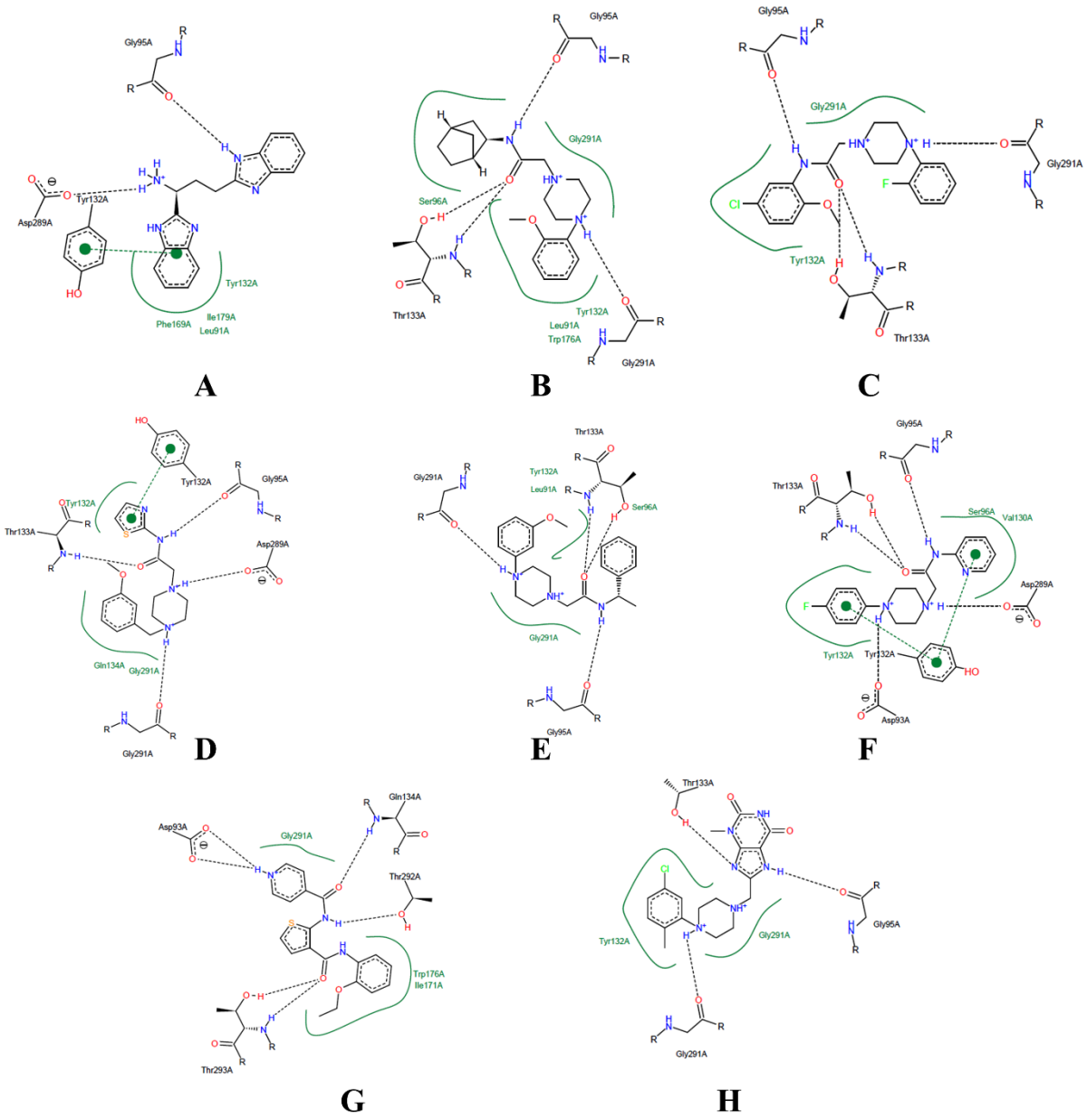


Figure 3.7 *Protein-Ligand interactions showing the hydrogen bonding and stacking interactions for the top A-H compounds. The dotted lines represent the hydrogen bonds, the green dotted lines represent stacking interactions and the solid green lines represent hydrophobic contacts.*

Molecular Dynamic Simulation Analysis

Analysis of the docking results of the identified compounds from QSAR screening reveals that piperazine moiety is predominant in the compounds (6/10 compounds) identified involving interactions with Gly-291 and Asp-289.

Docking and MD Simulation Analysis of 2QMF/Compound A complex

The best docking pose of Compound A reveals that the nitrogen group of the benzodiazole ring is involved in hydrogen bonding with Gly-95. The amino group between the two benzodiazole rings is involved in hydrogen bonding with the catalytic site residue Asp-289. The benzene ring of the benzodiazole is helpful in the stacking interactions with Tyr-132.

The interactions and flexibility of the 2QMF/Compound A complex is subjected to a 10 ns all atom simulation. The rmsd of the complex is stable throughout the simulation (**Figure 3.8)**. The rmsf of the complex is also compared to the rmsf of the protein in the absence of any ligands to identify the local fluctuations in the protein. It may be observed that 2QMF/Compound A complex does not show any major fluctuations when bound to compound A (**Figure 3.9**). The 2QMF/Compound A complex remained stable with minor fluctuations for residues in the loop regions; the interactions with Asp-289 and Tyr-132 being retained. Also persistent hydrogen bonding interactions with Asp-93 and Tyr-259 are also observed (**Figure 3.10 a, b**).

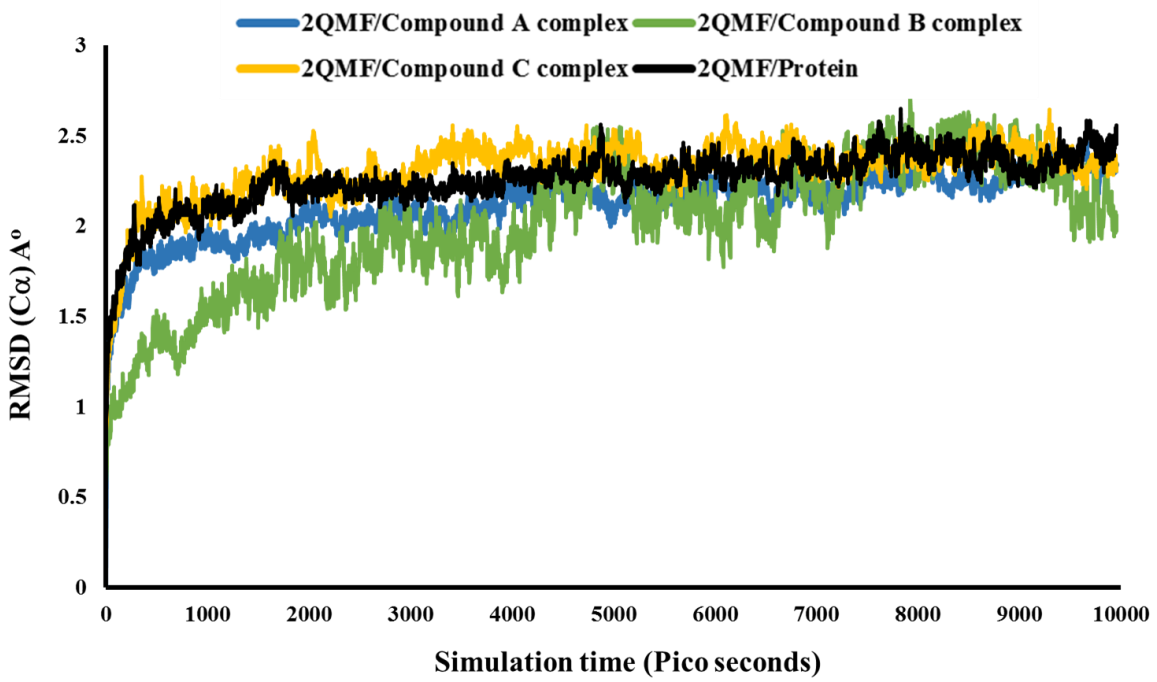


Figure 3.8 *RMSD plots of the protein, protein-ligand complexes*

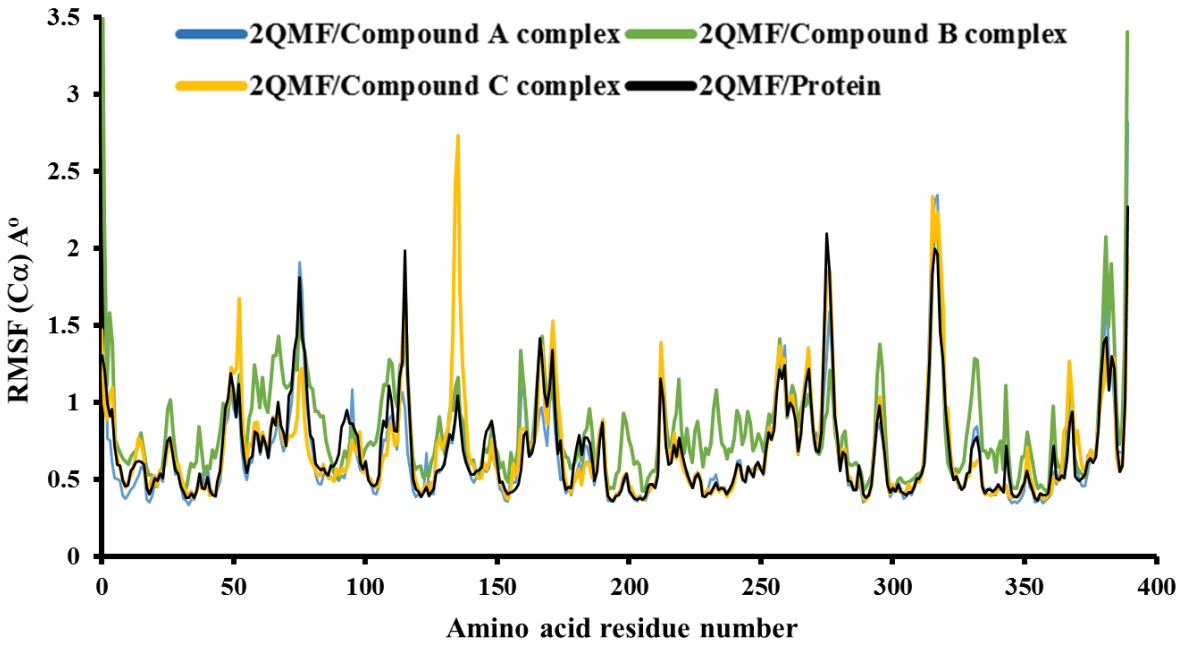


Figure 3.9 *RMSF plots of the complexes of the protein and protein-ligand complexes.*

Docking and MD Simulation Analysis of 2QMF/Compound B complex

The best docking pose for compound B reveals that the nitrogen of the piperazine derivative is hydrogen bonded with Gly-291 and the oxygen atom of the acetamide group is hydrogen bonded with Thr-133 and the nitrogen is hydrogen bonded with Gly-95.

The rmsd of the 2QMF/Compound B complex reveals insights into the persistence of interactions and stability of the complex. From the **Figure 3.8**, it appears that the complex is stable with no major as reflected by the root mean square deviations for the entire length of simulation, although not as stable as 2QMF/Compound A complex. The persistent interactions involving Gly-95, Tyr-132, and Thr-133 are retained throughout the length of the simulation. Further, compound B also involved in hydrophobic interactions with Leu-91 and hydrogen bonding with Asp-93 (**Figure 3.10 a, b**). All the interactions that are stabilizing protein compound A complex, are also observed in 2QMF/compound B complex.

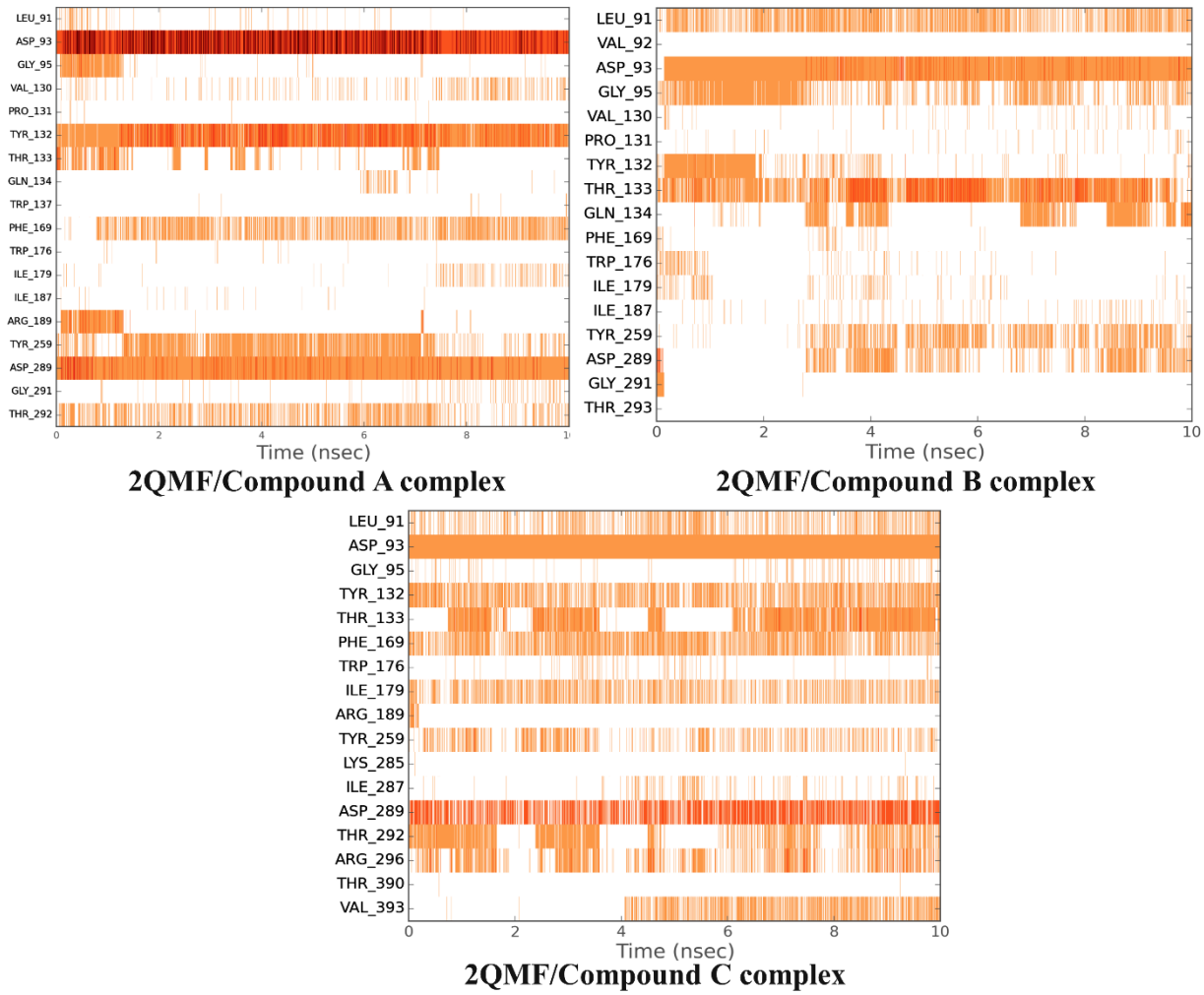


Figure 3.10a *Analysis of the interactions with β-secretase (2QMF) and their persistence in 3 complexes (A, B and C).*

Docking and MD Simulation Analysis of 2QMF/Compound C complex

The oxygen atom of the acetamide group of the compound 6 forms 2 hydrogen bonds with Threonine-133, the nitrogen atom is hydrogen bonded with Glycine-95. Also, the nitrogen atom of the Piperazine is hydrogen bonding to Glycine 291 (**Figure 3.7**).

The rmsd of the 2QMF-compound C complex is stable throughout the simulation (**Figure 3.8**). It may also be noted that compound C is involved in more persistent interactions than A and B complexes (**Figure 3.10 a, b**). The hydrophobic contributions by the residues Leu-91, Tyr-132, Phe-169, Ile-179, Val-393 are persistent throughout in the case of compound C. The amino

acids Gly-95, Thr-133, Thr-292, Arg-296 are also involved in water mediated hydrogen bonding interactions. Interestingly, the complex also displays Ionic interactions with Asp-289. Although the number of persistent hydrogen bond interactions are less in this complex, the hydrophobic contacts might be contributing to the stability of this complex.

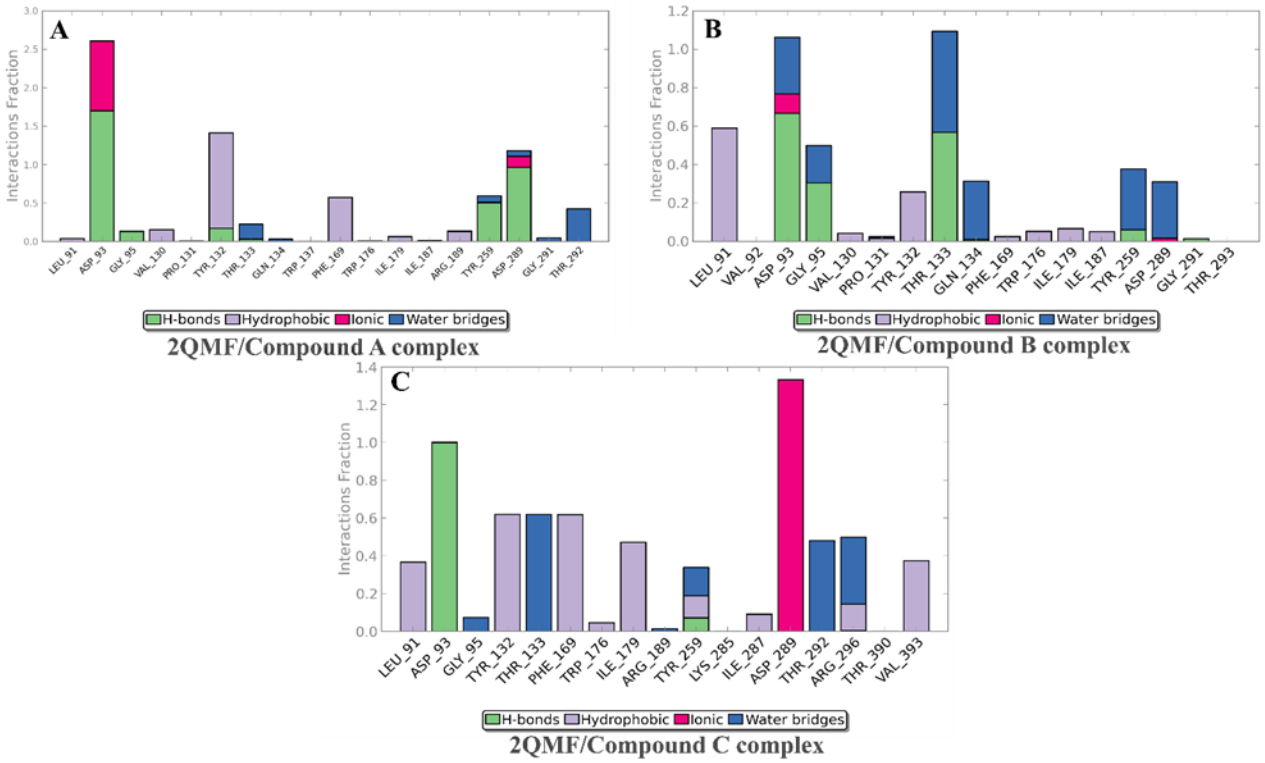


Figure 3.10b: *Plots indicating different type of interactions in the three protein-ligand complexes (A, B and C) during the simulation run.*

Overall all the complexes A, and C appear to be comparably stable and displayed persistent interactions with the active site residues. Eventhough complex B displayed similar interactions to the other two complexes (A and C), the persistance of these interactions.

Conclusions

In the current study, common feature pharmacophore model was developed using a dataset of BACE1 inhibitors showing activity values in nano molar range. It appears that the ring aromatic feature and the hydrophobic features play an important role, consistent with the previously reported models.[21-23] Although, a high active dataset in nano molar range consisting of ~60 compounds was chosen the important pharmacophore features present in the current model appear to be similar with those observed in the models resulting from much wider activity range of inhibitors. This suggests that the hydrophobic and ring aromatic features are distinctly responsible for the high activity of the reported potent inhibitors. Therefore, the current pharmacophore model can be useful for the identification, design and development of potent BACE1 inhibitors and with further optimization can be potential brain penetrating inhibitors for BACE1.

Chapter 4

Identification of Potential Inhibitors for Abelson Tyrosine Kinase, for Alzheimer's disease, using Multiple E-Pharmacophore Modeling, Virtual Screening and Molecular Dynamics

Introduction

Alzheimer's disease is the most common neurodegenerative disease affecting more than 40 million currently and it is estimated that by 2050 the number of people affected world-wide to exceed 100 million (Brookmeyer, Johnson, Ziegler-Graham, & Arrighi, 2007; K & L, 2014). Unfortunately, despite efforts to combat the disease, an effective therapeutic intervention is lacking for this fourth most common cause of death in the world (J & D, 1991). Given the enormous socio-economic burden on the society and the lethality factor, there is an urgent need to develop therapies for AD, in particular, and neurodegenerative diseases in general. It is believed that abnormal deposition of amyloid arising from aggregation of proteins involved in different pathways is responsible for not only Alzheimer's, but other neurodegenerative diseases including Parkinson's (Eisenberg & Jucker, 2012; Imam et al., 2011; Karran et al., 2011; Mahul-Mellier et al., 2014; O'Brien & Wong, 2011).

Current drug discovery approaches for AD are primarily focused on inhibiting β/γ secretase enzymes, for preventing the production of causative Aβ amyloid forming peptide (Iqbal, Liu, Gong, & Grundke-Iqbal, 2010). In fact, in addition to enzyme inhibition based approaches, strategies to explore chemical moieties to inhibit the formation, prevention/clearance of Aβ and Tau amyloids is gaining momentum (Cohen et al., 2009; Eisenberg & Jucker, 2012; Landau et

al., 2011). However, the inability of the molecules to cross blood brain barrier remains a general challenge concerning the development of neurotherapeutics (Alavijeh, Chishty, Qaiser, & Palmer, 2005; Pardridge, 2009). Specific to Alzheimer's disease, a bigger concern is the inherent complexity associated with the network of pathways leading to the progress of the disease and involving additional targets. Therefore a combinatorial approach wherein multiple targets are modulated to block the formation of amyloids will be an effective way to prevent AD (Lane, Shineman, Steele, Lee, & Fillit, 2012; Zhang, 2005).

Very recently Abelson tyrosine kinase (c-Abl) has been implicated in neurodegenerative diseases including Alzheimer's and Parkinson's disease. Experimental data unambiguously indicate its role in Alzheimer's and Parkinson's diseases via abnormal phosphorylation of tyrosine 394 of Tau protein and tyrosine 39 of α-synuclein respectively (Cancino et al., 2011; Derkinderen et al., 2005; Mahul-Mellier et al., 2014; Schlatterer, Tremblay, et al., 2011). Therefore, preventing the phosphorylation of specific tyrosine residues in Tau protein and α-synuclein can lead to effective treatment of not only AD but also Parkinson's disease. c-Abl thus serves as an alternate target for AD. In this context, Qingzhang Zhu et.al., targeted c-Abl with known kinase inhibitors for the identification of potential therapeutics for AD (Zhu, Chen, Wu, Jin, & Ruan, 2014). In fact, Parkinson's disease mouse models displayed protective effects on the neurons upon administration of c-Abl inhibitors indicating the feasibility and usefulness of brain penetrable c-Abl inhibitors for the treatment of neurodegenerative diseases (Karuppagounder et al., 2014; Zhu et al., 2014).

The cellular form of the Abelson tyrosine kinase, a relative of the c-Src family of tyrosine kinases is normally regulated by an auto inhibitory mechanism and participates in diverse cellular functions like cell cycle regulation, apoptosis etc. (Schlatterer, Acker, et al., 2011;

Shaul & Ben-Yehoyada, 2005)(Derkinderen et al., 2005; Jing, Caltagarone, & Bowser, 2009) Disruption of the regulation of c-Abl leads to chronic myelogenous leukemia. Thus, c-Abl has served as a target leading to the successful development of a therapeutic intervention for leukemia. Currently, Imatinib (Gleevec) is administered as a drug for treating Chronic Myeloid Leukemia (CML) (Nagar, 2007). Although Imatinib also inhibits c-Abl of non-cancer cells, it may be noted that normal cells have additional redundant kinases which allow them to continue their function even if one kinase is inhibited.

Unfortunately, the currently used drug molecule for treating leukemia by inhibiting c-Abl, suffers from the inability to penetrate the blood brain barrier and therefore, ineffective for treating AD (Pardridge, 2009; Senior, 2003). This necessitates the search for inhibitors that can effectively cross into the brain to inhibit c-Abl. In the present study we have used energy optimized pharmacophore modeling strategies for identifying potential inhibitors targeting the ATP-binding site of c-Abl. Further, in order to enhance the binding affinity of the compound towards the target, the structural features necessary for e-pharmacophore generation, we employed multiple crystal structures of c-Abl in the inactive aspartate-phenylalanine-glycine (DFG)-out conformation having bound inhibitors at the ATP binding site. Only crystal structures in the DFG-out conformation of c-Abl are used since this geometry provides an additional hydrophobic binding site (Okram et al., 2006; Reddy & Aggarwal, 2012). It is noteworthy that, Imatinib, binds to the inactive conformation of kinases. Keeping in view the limitation on the requirement of lipophilic molecules, a focused library of compound database, ChemBridge_CNS, consisting of molecules with enhanced probability to cross over the blood brain barrier has been used for virtual screening to identify potential inhibitors for c-Abl. Ten potential inhibitors were identified of which it appears that three compounds may be selective

for Abelson tyrosine kinase as against other structurally related kinases. To the best of our knowledge, for the first time, multiple e-pharmacophore modeling approach has been employed for the identification of potential c-Abl inhibitors for AD.

Materials and methods

Protein Preparation

Four c-Abl crystal structures in the inactive conformation with bound inhibitors and activities in the range of 8 nM to 33 μM and a resolution of 2.2 to 2.4 A^o were retrieved from PDB (Berman et al., 2000). The proteins were prepared using Protein Preparation Wizard, which is part of the Maestro software package (*Maestro, v9.2, Schrödinger, LLC, New York, NY*). The crystal structures of c-Abl along with the reported IC_{50} values of the bound inhibitors 2HIW/8nM, 2E2B/11nM, 2HYY/170nM and 3CS9/33uM used in the study are represented in the **Fig. 4.1a, 1b** (Cowan-Jacob et al., 2007; Horio et al., 2007; Weisberg et al., 2005). Bond orders and formal charges were added for hetero groups, and hydrogen atoms were added to all atoms in the system. Water molecules from the four crystal structures were removed followed by the energy minimization of the resulting structure.

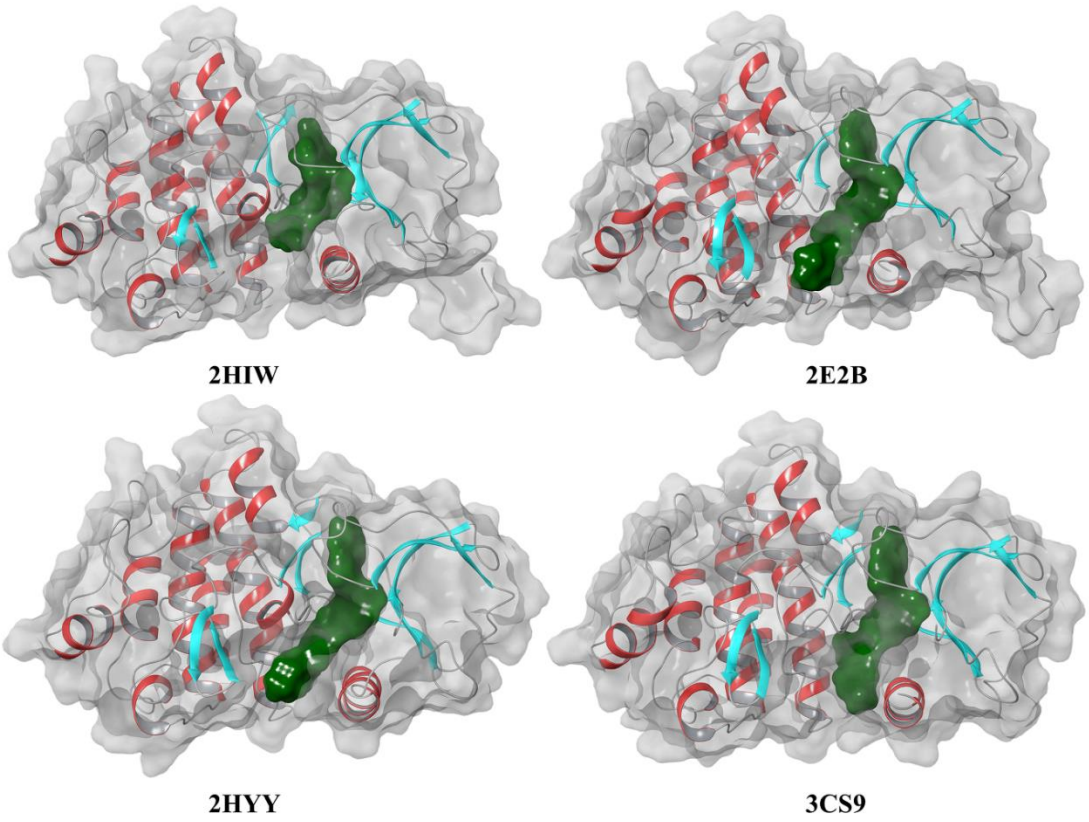


Figure 4.1a *Crystal structures of Abelson's Tyrosine Kinase with bound inhibitors in inactive form. Ligands are represented as a dark green surface in the ATP binding site. PDB codes of the respective crystal structures are given below as 2HIW, 2E2B, 2HYY and 3CS9.*

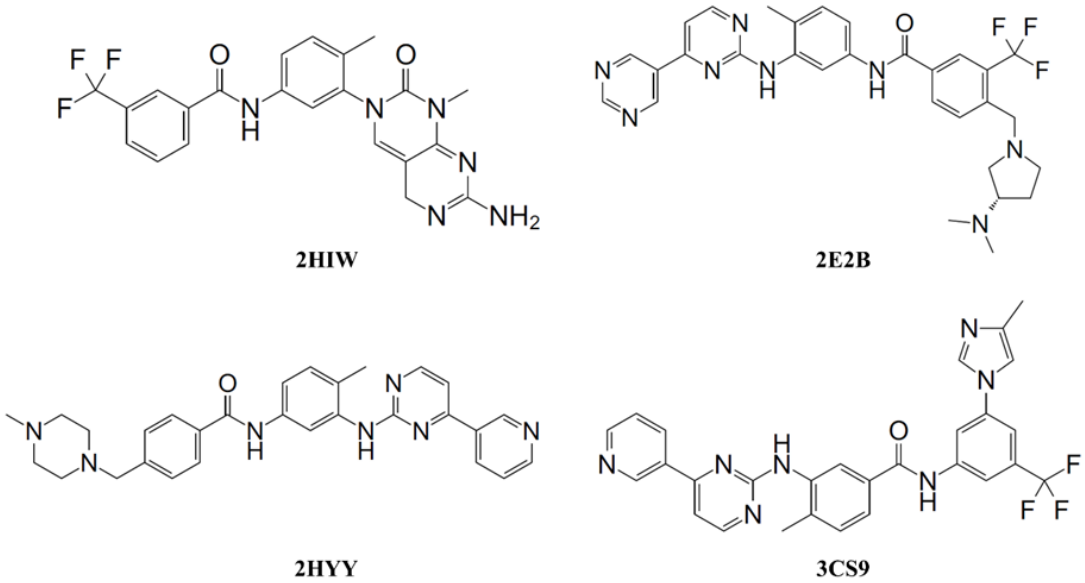


Figure 4.1b *Co-crystalized ligands bound to Abelson's Tyrosine Kinase. 7MP (2HIW), Bafetinib (2E2B), Imatinib (2HYY) and Nilotinib (3CS9) represent the PDB codes of the crystal structures.*

e-Pharmacophore models generation and Validation

The procedure to generate energy based pharmacophores is adapted as reported earlier (K. Loving et al., 2009; Palakurti et al., 2013). Glide energy grids were generated for each of the prepared complex. The binding site was defined by a rectangular box surrounding the crystal ligand. Ligands were refined using the "Refine" option in Glide, with the option to output Glide XP descriptor information. (Glide, version 5.7, Schrödinger, LLC, New York, NY, 2011) (Friesner, Banks, R. B. Murphy, et al., 2004; Friesner et al., 2006a; T. A. Halgren et al., 2004). Hypothesis selection was based on the default set of 6 chemical features: Hydrogen bond Acceptor (A), Hydrogen bond donor (D), Hydrophobe (H), Aromatic ring (R), Positive ionizable (P), Negative ionizable (N). Following the general strategy, each pharmacophore feature is first assigned an energetic value equal to the sum of the Glide XP (K Loving et al., 2009; Salam, Nuti, & Sherman, 2009) contributions of the atoms comprising the site, allowing sites to be quantified and ranked on the basis of the energetic terms. Glide XP descriptors include terms for hydrophobic enclosure, hydrogen bonds, electrostatic contributions, π-π stacking, π-cation, and other interactions. Standard parameters are used in the study unless specified otherwise. The features with favorable interactions are mapped and retained whereas the features with unfavorable interactions were omitted and not considered for the procedure.

The pharmacophore models generated were validated using DUD database (A Database of Useful Decoys-Enhanced). The database has 295 of actives and 11180 decoys for c-Abl and can be used to assess the performance of the pharmacophore models in database screening. The validation of the pharmacophore models is based on the percentage of yields, enrichment factor, false positives, false negatives and goodness of hit score. The validated models were used for virtual screening followed by docking.

Database Preparation

The ChemBridge_CNS database with a biased set of 50000 compounds having an enhanced probability of crossing the blood brain barrier was prepared using LigPrep (*LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2011*) with Epik (*Epik, version 2.2, Schrödinger, LLC, New York, NY, 2011*) (Greenwood et al., 2010). Conformational sampling was performed on all the database compounds using the ConfGen, a search algorithm which samples the conformations and performs energetic evaluations to efficiently explore diversity around rotatable bonds (Watts et al., 2010). OPLS_2005 force field and a duplicate pose elimination criterion of 1.0 Å rmsd (root mean square deviation) to remove redundant conformers was used in the ConfGen sampling procedure. To screen electrostatic interactions a distance–dependent dielectric solvation treatment was used. Further, exclusion of the high–energy structures was based on the maximum relative energy difference of 10 kcal/mol (Watts et al., 2010).

Virtual Screening and Molecular Docking

e-pharmacophore based virtual screening was employed for the best scoring pharmacophore hypotheses. Screening molecules required to match all the pharmacophore sites. Distance matching tolerance was set to 2.0 Å as a balance between stringent and loose–fitting matching alignment. From the database screening, the hits are ranked in order of their Fitness scores. The fitness scoring function is an equally weighted composite of these three terms and ranges from 0 to 3, as implemented in the default database screening in Phase. *(Phase, version 3.3, Schrödinger, LLC, New York, NY, 2011*) (Dixon, Smondyrev, Knoll, et al., 2006; Dixon, Smondyrev, & Rao, 2006).

The shortlisted molecules from virtual screening based on high fitness were docked into the four selected c-Abl crystal structures. For this, purpose Grid files are generated by using the centroid of the bound ligand of the respective crystal structures. The grids generated were verified to reproduce the crystal ligand bound conformation with an rmsd of <1. Initial rapid screening of the docked compounds was performed using Glide HTVS screening function derived from protein-ligand binding affinities. Subsequently, the compounds with favorable docking scores resulting from high throughput screening step were subjected to Glide XP (extra precision) screening to calculate the precise docking score. The compounds showing Glide XP scores (*Glide, version 5.7, Schrödinger, LLC, New York, NY, 2011*) (Friesner, Banks, R. B. Murphy, et al., 2004; Friesner et al., 2006a; T. A. Halgren et al., 2004). with a cut off value of >-9.0 kcal/mole were selected as potential inhibitors.

Molecular dynamic simulations

To understand the binding interactions between the docked ligand poses and the protein, there is a need to incorporate protein flexibility into the system. Glide docking procedure treated the protein as rigid, whereas the ligand is flexible. As the scores contributing to Glide docking can only be related to interaction energies between atoms. We considered MD simulations which provides an assessment of protein ligand stability as well as the persistence of the interactions contributing in their binding.

The initial model system for the simulations were generated using the Glide XP docking file having the protein-ligand complex pose. The ligand protein complexes A-C were processed individually using the system builder in which, each crystal ligand complex is used for solvation with water molecules (SPC) in an Orthorhombic box. The box volume is calculated and

minimized. The system is neutralized using appropriate number of counter ions for individual protein ligand systems.

The system thus generated earlier is used to run MD simulations. A default relaxation protocol is used before the simulation. Using NPT ensemble at a temperature of 300 kelvin and 1 bar of pressure. The thermostat used is a Nose-Hoover chain method. Martyna-Tobias-klein barostat is used for pressure. Interaction cutoff radius for short range method is 10 A^o and a smooth particle mesh Ewald is used in the long range method. A 10 ns MD simulation is carried out with a recording interval of 1.2 ps. The simulation is analyzed for the RMSD, RMSF and the interaction patterns of the complexes.

Relative selectivity of c-Abl inhibitors

The potential inhibitors identified by the virtual screening were ascertained for relative selectivity towards c-Abl in comparison to other structurally similar kinases. The crystal structures of c-Src (2OIQ), c-Kit (1T46) and Syk (1XBB) bound with Imatinib,(Atwell et al., 2004; Lin, Meng, Jiang, & Roux, 2013; Seeliger et al., 2007) were used to generate Grid files necessary for docking using Glide program in Maestro. The compounds identified from virtual screening are docked into the crystal structures of the kinases c-Src, c-Kit and Syk (Atwell et al., 2004; Lin et al., 2013; Seeliger et al., 2007).

Results and Discussion

Protein Preparation

The prepared crystal structures of c-Abl from the Protein Preparation Wizard retained the interactions between the protein and bound ligand as confirmed by the analysis of the protein ligand interactions before and after the protein preparation. The prepared crystal structures with bound ligands are shown in the **Fig. 4.1a**. Also, the active site pocket of each crystal structure

of c-Abl is occupied by a different ligand; thus the four structures provide the necessary diversity anticipated from a multiple crystal structure model (Chen Z et al., 2010; Palakurti et al., 2013). Although the ligand in the case of 2HIW has the highest activity, it may be observed that the molecule is relatively smaller than the other bound ligands. This increased activity may be attributed to two hydrogen bonds (Anilic NH_2 of the pyridine diazin-2-one to backbone carbonyl oxygen of Met-336, and pyridine N-atom at position 3 of pyridine diazin-2-one to the backbone NH of Met-336) made by the 2HIW ligand with Met-336 in the crystal structure complex. The presence of 2 hydrogen bonding interactions contributed by 2HIW ligand distinguishes it from the other three ligands.

Pharmacophore generation and Validation

The e-pharmacophore models generated from the four DFG-out crystal structures of c-Abl employed in this study are listed in **Table 4.1** with the number of sites in each model and the type of features. The e-pharmacophore hypotheses are found with three kinds of features, R, A, H; ring aromatic, acceptor and hydrophobic (**Table 4.1**). The minimum featured pharmacophore model, resulting from 2HIW is a 4 feature model, RRRH, with 3 ring aromatic features and a hydrophobic group. 2E2B and 2HYY structures generated 5 feature models with 4 ring aromatic groups and an acceptor group (ARRRR), while, the crystal structure 3CS9 resulted in a 6 feature model with 4 ring aromatic groups and 2 acceptor groups (AARRRR). The respective features of the energy optimized pharmacophore hypothesis identified are given in **Table 4.1** and **Fig 4.2a.**

S. No	PDB code	Number of Sites	Features
1	2HIW	4	RRRH
2	2E2B	5	ARRRR
3	2HYY	5	ARRRR
4	3CS9	6	AARRRR

Table 4.1 *Number of optimized pharmacophore features derived for each ligand from the respective PDB structures. A-acceptor group; R-ring aromatic group and H-hydrophobic group.*

The four e-pharmacophore models generated from the crystal structures possesses a common feature of having three adjacent aromatic rings (RRR). The distances between these three aromatic rings in all the four pharmacophore models appeared to be similar as seen in (**Fig. 4.2b**) with a conserved V-like shape (Mucs, Bryce, & Bonnet, 2011). The observed three aromatic ring features in the pharmacophore models is consistent with the interactions of three aromatic amino acids, Phe317, Tyr253 and Phe382 of c-Abl with the bound inhibitor molecules. In fact, calculations have indicated that interactions of the aromatic amino acids in c-Abl with its bound ligand contribute to the additional stability of the protein-ligand complex (YZ & PY, 2008).

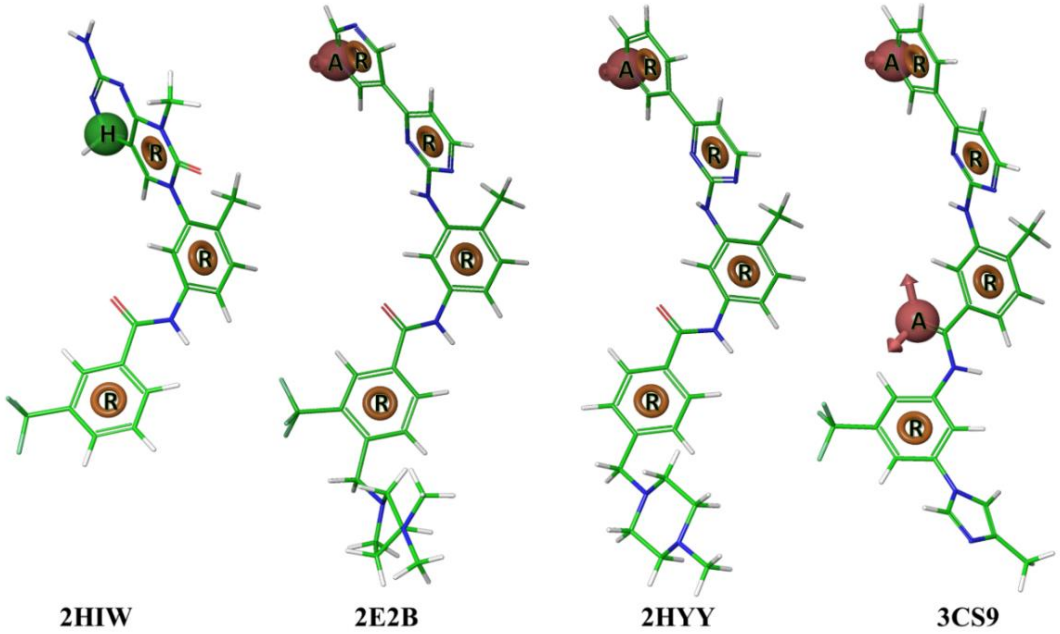


Figure 4.2a *Energy-optimized pharmacophore hypotheses generated from the 4 selected PDB structures. Brown sphere: hydrogen bond acceptor (A), open circle: aromatic ring (R) and Green sphere: Hydrophobic group (H).*

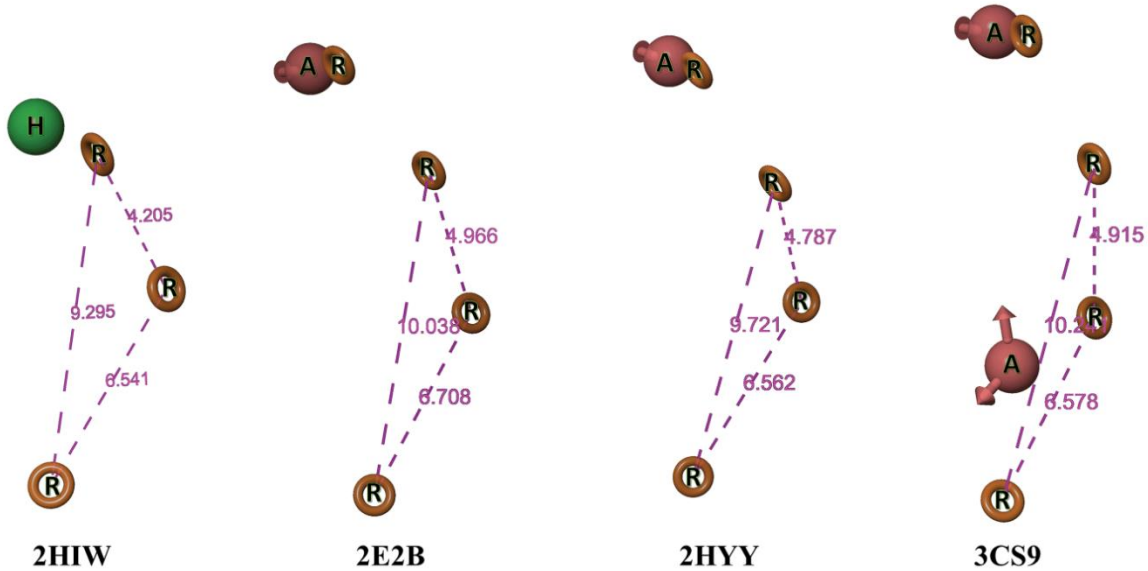


Figure 4.2b *Distance comparison of the three central common features RRR from all the 4 e-pharmacophore models obtained from 2HIW, 2E2B, 2HYY and 3CS9. Brown sphere: hydrogen bond acceptor (A), open circle: aromatic ring (R) and Green sphere: Hydrophobic group (H).*

Table 4.2 lists statistical parameters of the Goodness of Hit score validation for the pharmocophore models. The recovery rate from the decoy set together with a Goodness of Hit score of ≥ 0.6 for the pharmacophore models 2E2B, 2HYY and 3CS9 suggests a robust screening performance of the pharmacophore models. The three pharmacophore models could identify active compounds close to 75% in most cases, with very low percentage of false negatives. The validation indicated that 2HIW displays very low scores. This could probably be resulting from the lower number of pharmacophore features identified for 2HIW, a 4 feature model. Therefore, the three models, 2HYY, 2E2B and 3CS9 were used for virtual screening followed by docking.

Parameters	**2HIW**	**2HYY**	**2E2B**	**3CS9**
Number of molecules in the database (D)	11180	11180	11180	11180
Total no. of actives in the database (A)	295	295	295	295
Total hits (Ht)	1118	41	43	44
Active hits (Ha)	45	32	34	32
% Yield of actives [(Ha/Ht) X 100]	4	78	79	72
Enrichment Factor [(Ha X D)/(Ht X A)]	1.52	29.57	29.96	27.56
False negatives [A - Ha]	250	263	261	263
False Positives [Ht - Ha]	1073	9	9	12
Goodness of Hit	0.06	0.61	0.62	0.57

Table 4.2 *Parameters for the Goodness of Hit and enrichment factor estimation of the e-pharmacophore models.*

Compound	Docking Score (kcal/mole)	Hydrogen bond interactions	Stacking interactions	PDB structures
A	-12.61	2	2	2HIW, *2E2B*
B	-11.66	3	3	2E2B, 3CS9, *2HYY*
C	-11.60	3	1	2E2B, *3CS9*, 2HYY
D	-11.11	3	1	2HIW
E	-10.95	3	1	2HIW
F	-10.95	3	0	2HIW
G	-10.39	2	2	2HIW
H	-10.29	2	3	*2E2B*, 3CS9
I	-10.07	2	1	3CS9, 2HYY
J	-9.37	2	1	2E2B

Table 4.3 *Docking Score**,** number of hydrogen bonds, stacking interactions of the top 10 hit compounds (A-J) when docked into the crystal structures of 2HIW, 2E2B, 3CS9 and 2HYY. Underlined PDB structures represent the most favorable docking score for the particular compound. Structures (PDB ids) included in italics correspond to a docking score > 9.0 kcal/mole for the corresponding compound.*

Virtual screening and docking

The screening for the top 3 individual pharmacophore models was achieved using Phase Find Matches module in Maestro with a feature match criteria containing all the features in the top three validated e-pharmacophore models, 2E2B, 2HYY, 3CS9 (**Fig. 4.2a**). The pharmacophore screening is followed by docking of the compounds displaying a fitness score of >1.5. It may be noted that the 2HIW model though showed poor performance for validation, is actually the one which shows the best experimentally measured activity. Therefore, 2HIW was also included for docking the compounds displaying a fitness score of > 1.5 resulting from the three validated phamacophore models. The ATP binding pocket of c-Abl was the site for docking the compounds, considering the success of kinase drugs targeted to bind the ATP binding site in general. Further ATP binding pocket contains predominantly hydrophobic residues; valine (3) isoleucine (2), phenylalanine (2), leucine (2), methionine (2) and alanine (2). Polar residues

including one each of glutamine, lysine, threonine and aspartic acid also form a part of the active site pocket.

The most favorable compounds identified from virtual screening and docking (**Figure 4.3**) along with the hydrogen bond and stacking interactions are shown in **Figures 4.4**. It is interesting to note that, of the top 10 compounds the highest number of compounds were retrieved based on the crystal structure, 2HIW followed by 2E2B and 2HYY. The least number of hits were obtained from 3CS9.

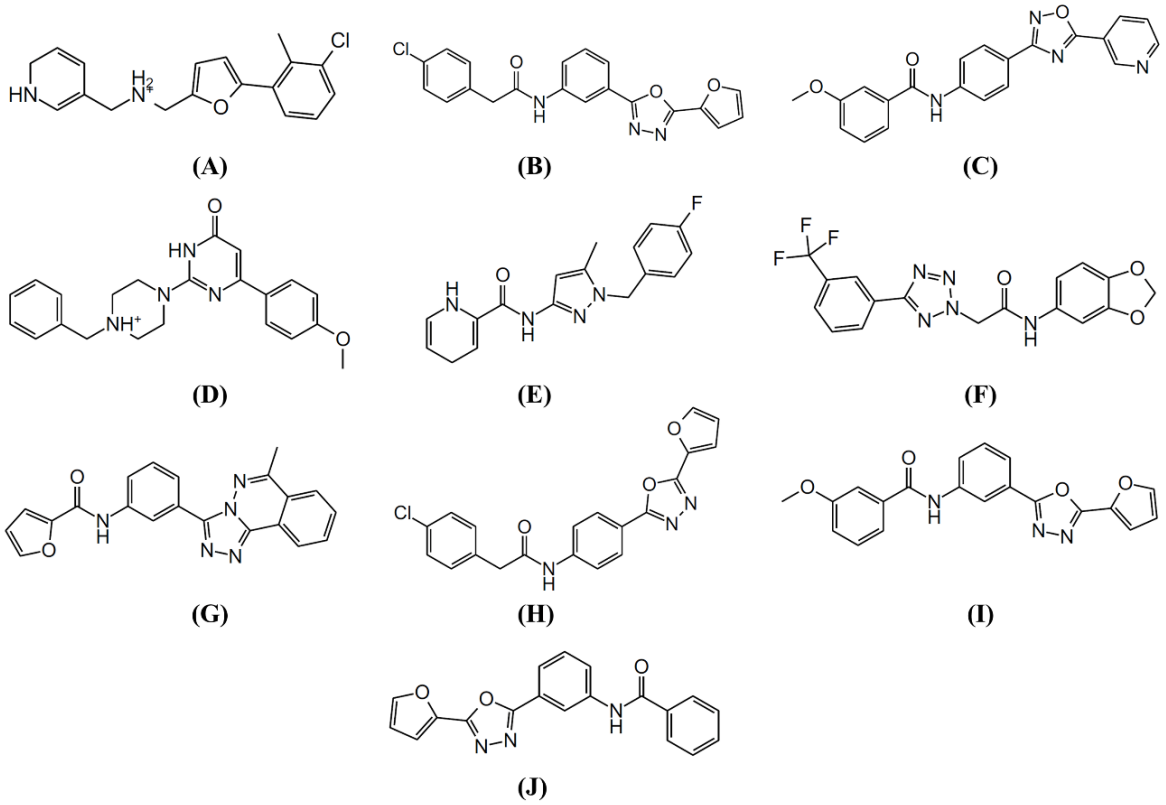


Figure 4.3 *2D structures of the compounds A-J, identified based on the high docking score from the virtual screening protocol.*

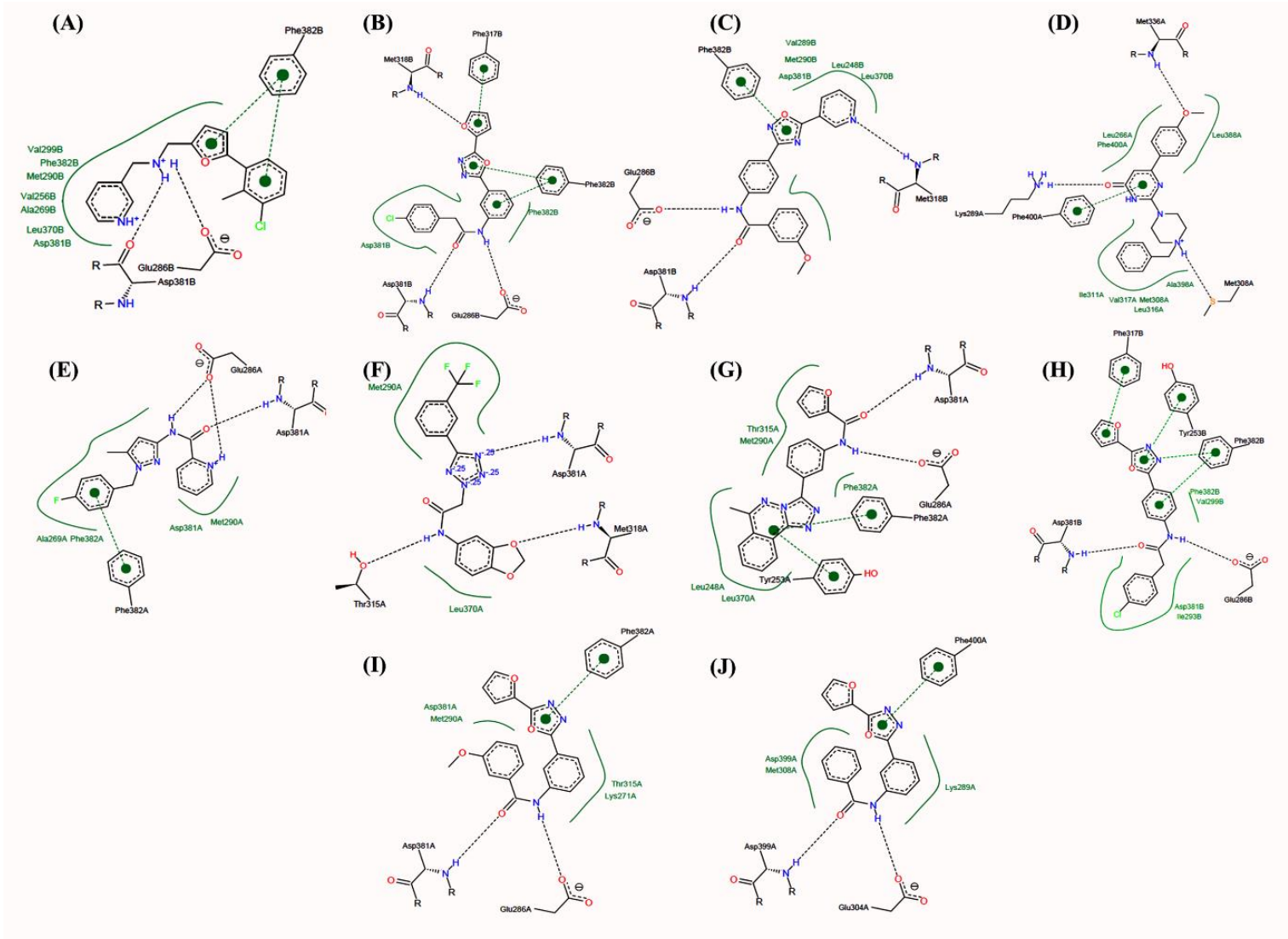


Figure 4.4 *Protein-Ligand interactions: Hydrogen bonding and stacking interactions are presented for the best scoring compounds A-J.*

The top 10 compounds were docked into all four crystal structures of c-Abl to identify the compounds that show a favorable docking score when bound to more than one crystal structure. Compounds A, B, C and D display docking score of more than > -11 kcal/mole for more than one crystal structure (**Table 4.3**). The summary of docking score, along with number of hydrogen bonds and stacking interactions that the compounds make in the respective crystal structures is shown in **Table 4.3**. Based on this summary, it is proposed that compounds A-D may possess a higher inhibitory activity against c-Abl. The hit compounds and their binding poses along with all the crystal ligands are represented using PoseView (Stierand & Rarey, 2010) (**Fig.4.4**). It may be observed from the protein-ligand interactions that the hit compounds align very closely with the bound ligand in the active site pocket of c-Abl (**Fig.4.5**). Hydrogen bonding and π-π stacking interactions are observed between the active site amino acids that consists of glutamine 286, aspartic acid 381, methionine 318 (methionine 336 in the case of 2HIW), threonine 315, tyrosine 253, phenylalanine 382.

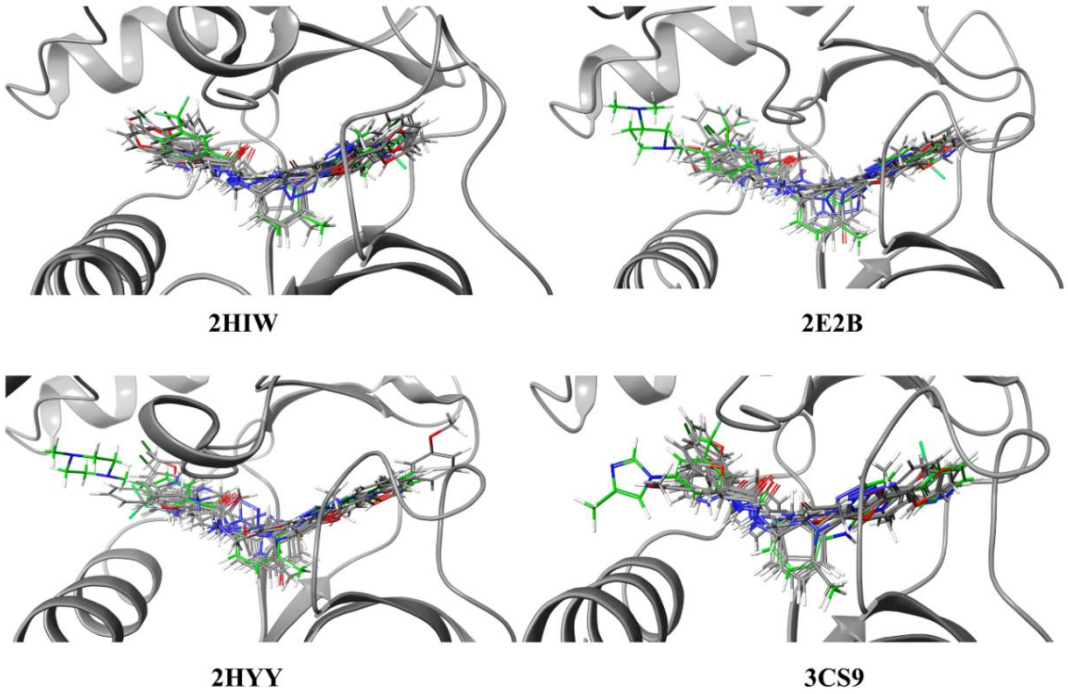


Figure 4.5 *Binding poses of the superimposed hit molecules (grey sticks) in the four crystal structures along with the respective bound crystal ligands in green stick form.*

Analysis of Docking and Molecular dynamics simulations

We performed MD simulations which provide an assessment of protein ligand stability as well as the persistence of the interactions contributing to their binding. Molecular dynamics simulations for the docked protein ligand complexes (c-Abl with 3 top compounds A, B and C) were carried out under identical conditions as described earlier. The simulations of the complexes were compared to the simulation performed in the absence of bound ligands to the protein to assess the stability of the protein ligand complex. The results from the 10 ns simulation were analyzed for changes in root mean square deviation (rmsd), interactions with the active site residues, nature of amino acids involved in the interactions and the root mean square fluctuations (rmsf) of the protein backbone C_α atoms. Deviations in the presence of bound ligands to protein will indicate the changes in comparison to unbound protein.

Docking and MD Simulation Analysis of C-Abl-Compound A complex

The docking pose and the interactions of Compound A with C-Abl (2E2B) reveals that the 3-chloro-2-methylphenyl and the 5-methylfuran ring are involved in stacking interactions with the amino acid Phe-382 **(Figure 4.3)**. Both the hydrogens in the NH_2 group at the center of the compound involved in hydrogen bonding with Glu-286 as well as Asp-381.

The interaction analysis of the MD simulation of the docked structure complex (2E2B/Compound-A) reveals that the RMSD of the complex is stable throughout the 10 ns simulation run. The active site residues Glu-286, Asp-381 retained the most stable hydrogen bonding interactions as indicated by the 50% and 82% retention of the interactions for Glu-286, Asp-381 respectively. Apart from this, an extra hydrogen bonding interaction with His-361 (37%) was also identified (**Figure 4.6a, b)**.

The amino acids Lys-271, Met-290, Val-299, Ala-380 and also His-361 are involved in strong hydrophobic contacts with the compound-A indicating that the compound is strongly bound deep in the hydrophobic pocket of the protein. Thus, ionic interactions between Glu-286 and Asp-381, water bridges between His-361 and Asp-381 may likely be contributing to the overall stability of the complex. The persistence of interactions leading to c-Abl/Compound A complex in the MD simulations is consistent with the selection of Compound A based on docking scores.

Docking and MD simulation Analysis of 2E2B/Compound B

The interaction analysis of the docking pose of c-Abl Compound b complex reveals that the Phenyl-1,3,4-oxadiazole is involved in the stacking interactions with Phe-382. N-phenyl acetamide is involved with the hydrogen bonding interactions with Glu-286 and Asp-381. The furan ring is also involved with the hydrogen bonding interactions with Met-318.

The RMSD of the protein-Compound B complex indicates that the resulting complex is not very stable. RMSD fluctuations are observed throughout the simulations (**Figure 4.6**). Further, indication in decreased stability of the complex is reflected from the interaction analysis of the ligand Compound B complex with the active site residues of the protein (**Figure 4.7**). For instance, interactions from Lys 271, His 36, A380 do not make stabilizing contacts with Compound B which are otherwise observed for compounds (**Figure 4.7a, b)**. However, it retains the interactions with active residues Asp 381 and Glu 286 and few other active site residues thus leading to a comparable docking score as that of Compounds A and C (**Table 4.3**). The amino acid Glu-382 shows 98 % hydrogen bonding, where as Phe-382 shows 51% persistant stacking interactions in the simulation. The residues Phe-382, Phe-317, Val-299 and Leu-248 show strong hydrophobic interactions with the Compound B (**Figure 4.7a, b**).

Docking and MD simulation Analysis of 2E2B/Compound C

The nitrogen atom of the pyridine ring is involved in the hydrogen bonding with Met-318, nitrogen and oxygen atoms of the benzamide group in Compound C are involved in hydrogen bonding with Glu-286 and Asp-381 respectively. The oxadiazole ring is involved with stacking interactions with Phe-382.

The RMSD of the c-Abl Compound C complex reveals that the complex is very stable throughout the simulation (**Figure 4.6**) and is involved in many persistant interactions as observed for Compounds A and B (**Figure 4.7a, b**). The hydrogen bonding interaction of the ligand with Met-318 that is absent in the two other cases appears persistent throughout in the case of Compound C. The amino acids Asp-381, Val-299, Glu-286 and Val-379 are also involved water mediated hydrogen bonding. Interestingly, the complex displayed stronger hydrophobic interactions with Phe-382, Ala-380, Leu-370, Met-290, Ile-293, Ala-269, Tyr-253 and Leu-248. The stability of this complex can be mostly contributed to the hydrophobic contacts.

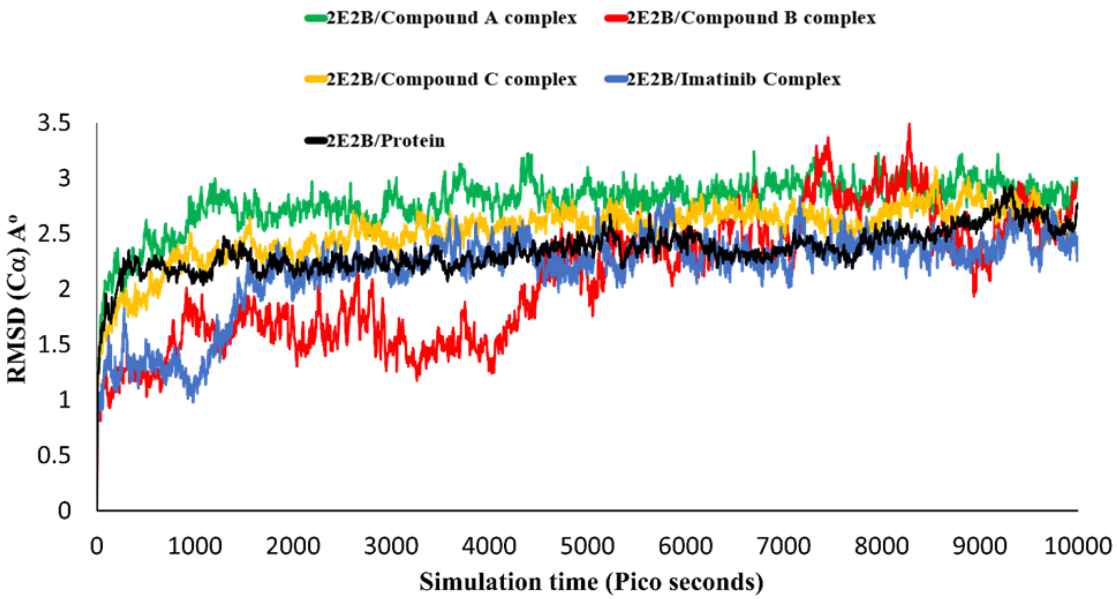


Figure 4.6 *RMSD plots of the protein, protein-ligand complexes.*

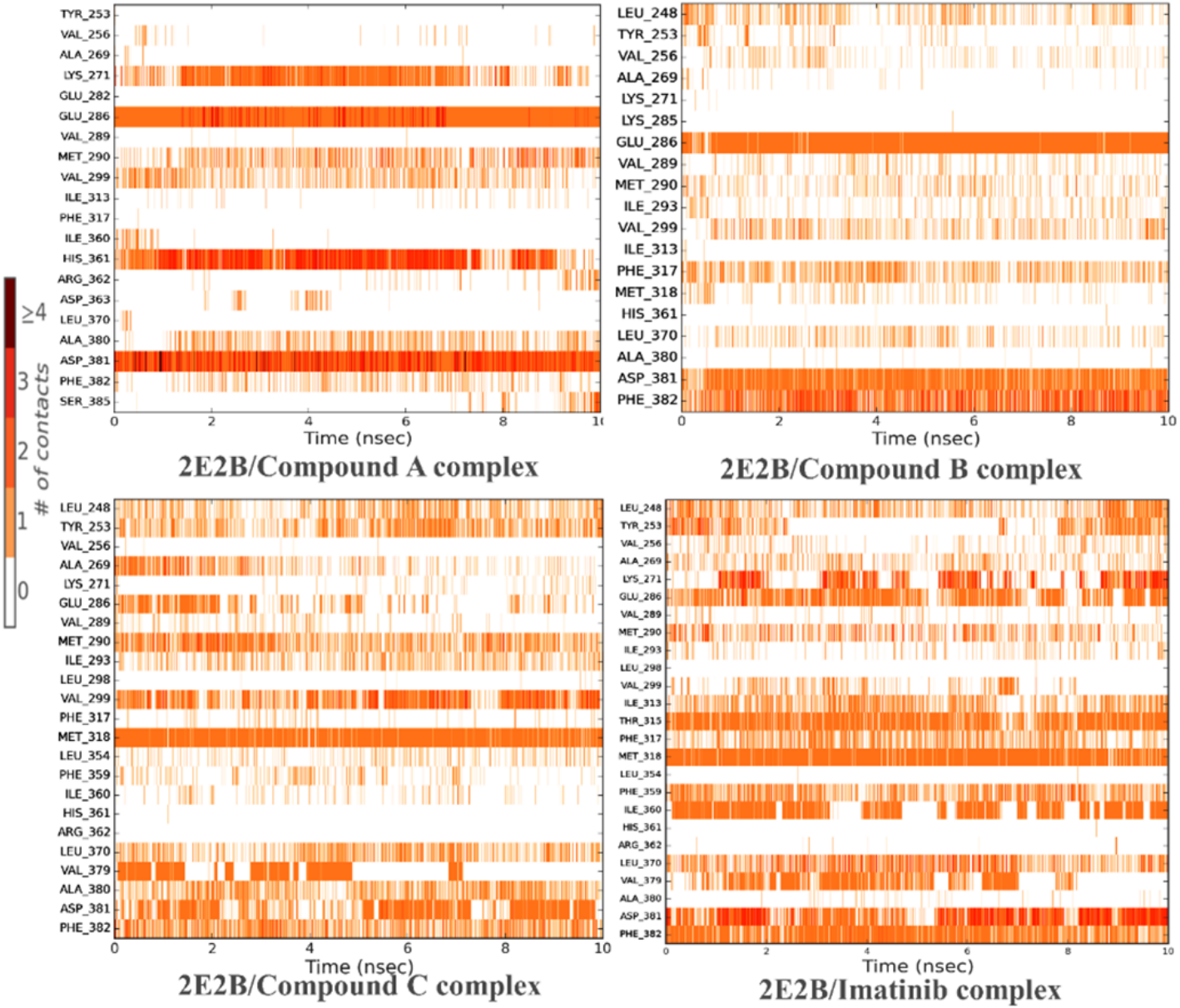


Figure 4.7a: *Analysis of the interactions with c-Abl and their persistence in 3 complexes (A, B and C).*

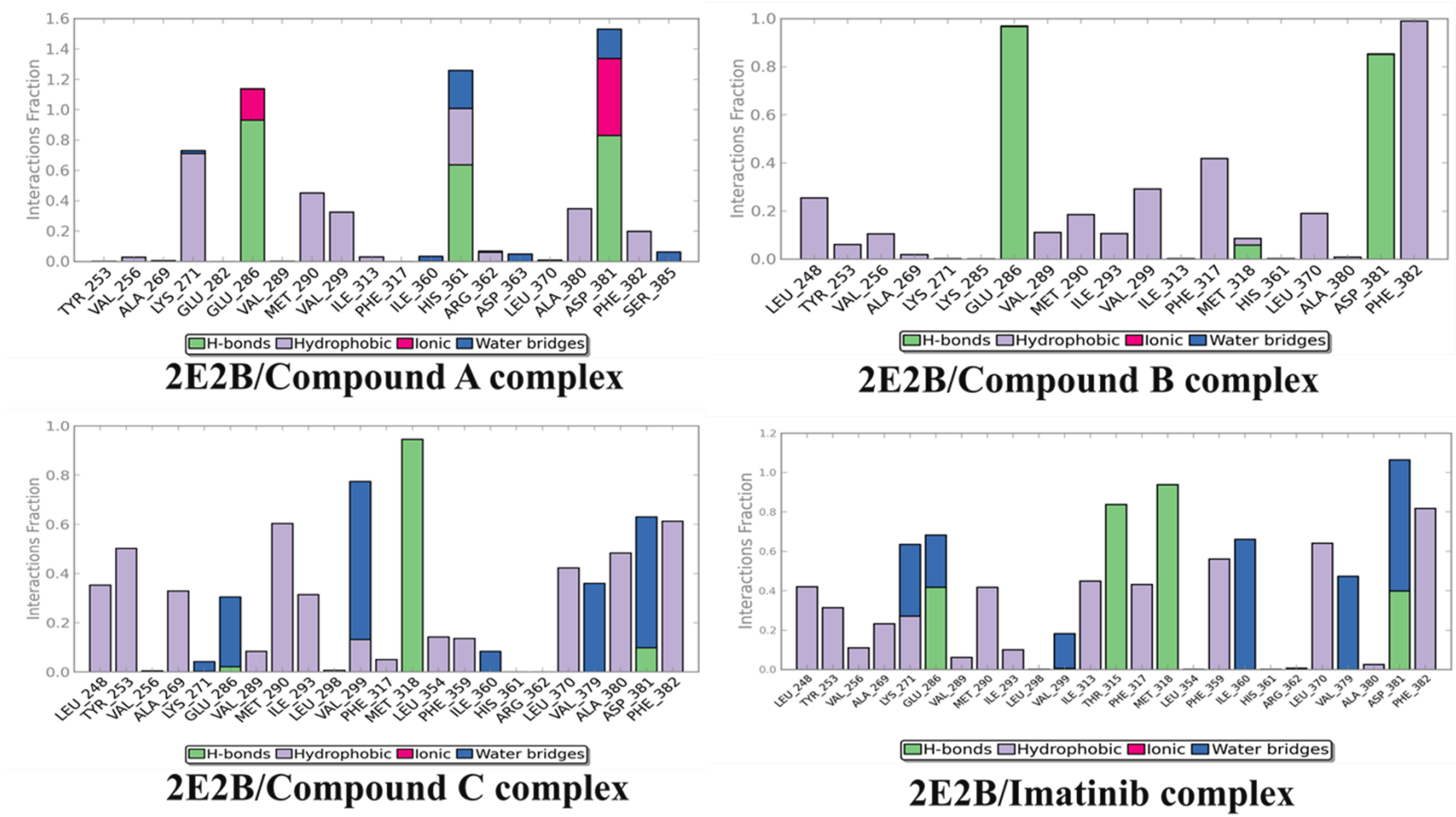


Figure 4.7b *Plots indicating different type of interactions in the three protein-ligand complexes (A, B and C) during the simulation run.*

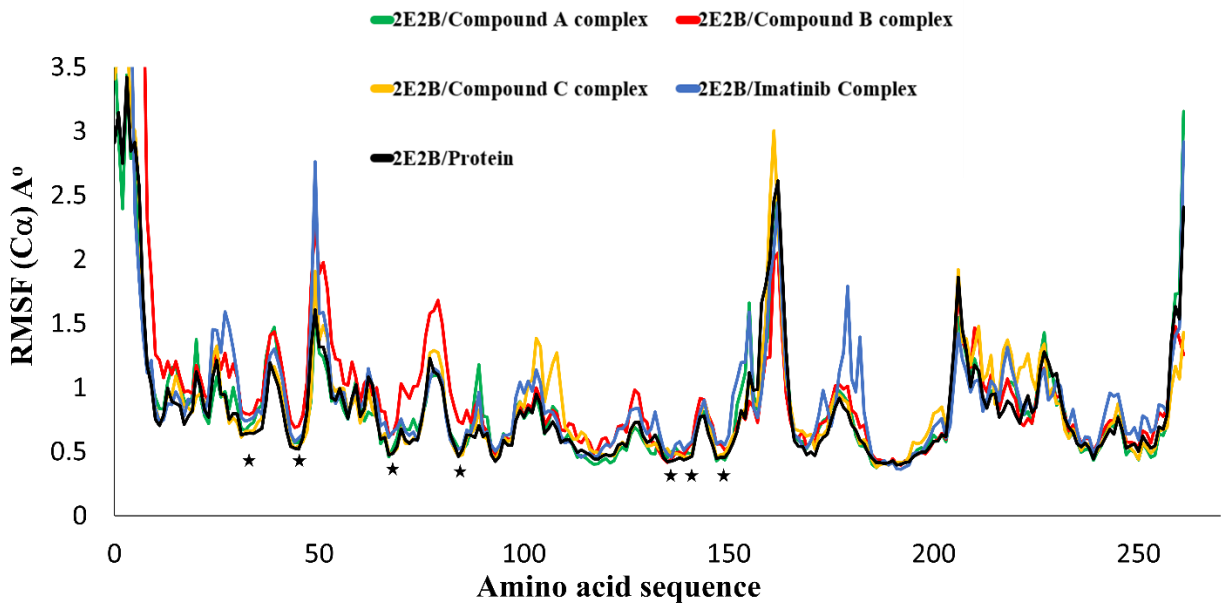


Figure 4.8 *RMSF plots of the complexes of the protein and protein-ligand complexes. The asterisks indicate the location of active site residue in the amino acid sequence of c-Abl.*

Overall, the complexes from Compounds A and C appear to be stable and show perisistant interactions with the protien active site residues. The Cα atom root mean square fluctuations (rmsf) of the protein, protein-ligand complexes (A, B and C) are shown in **Figure 4.8**. A decreased stability of the protein-Compound B complex is reflected in the increased Cα atom fluctations throught the length of the sequence, espcially, at the active site residues, indicating that binding of Compound B may lead to more dynamic fluctations.

Selectivity of potential inhibitors: c-Abl *versus* related kinases

The ability of a compound to bind specifically to a target protein in preference to other possible targets indicates its 'selectivity'. Lack of selectivity, may result in the binding of the potential inhibitor to other proteins in addition to the target protein. Therefore, addressing the concerns regarding specificity either in the design strategy or as a post assessment step following the virtual screening and identification of potential inhibitors will provide insights into possible

cross-reactions. It is even more vital in the case of targets such as kinases which display significant structural similarities (Anastassiadis, Deacon, Devarajan, Ma, & Peterson, 2011). In fact, measured activities for a set of commercially available kinase inhibitors for a large group of recombinant kinases revealed wide spectrum of promiscuity between them (Anastassiadis et al., 2011). The bias for particular kinase(s) can be assesses at an earlier stage if the selectivity of a compound is known. Experimental determination of kinase selectivity of a large set of potential compounds can add to expenses typically involved in assays etc., and therefore the assessment by analyzing docking score, docking affinity could be the inexpensive alternative.

One of the most clinically successful drugs against chronic myeloid leukemia is Imatinib which inhibits c-Abl (Nagar, 2007). However, it is found that, it also shows inhibition against the activity of other kinases, specifically Src kinase, Spleen tyrosine kinase (Syk), Tyrosine protein kinase Kit (c-Kit) (Atwell et al., 2004; Lin et al., 2013; Seeliger et al., 2007). Multiple sequence alignments of the four kinases represented in **Fig. 4.9** indicate that the sequence identity between c-Abl and c-Src, Syk and c-Kit kinase is 49%, 40%, 37% respectively. However, structures of the four kinases are very identical with Cα atom RMSD of ~1 A^o (**Fig. 4.10**).The observed specificity of Imatinib is attributed to the positioning of P and c-loops in c-Abl resulting in favorable π–π stacking interactions with Tyr253. Therefore, both sequence and structure of kinases are important in defining the specificity. In fact, experimental data indicate that kinases closely related by sequence and structure are inhibited by molecules possessing similar features (Anastassiadis et al., 2011; Dodson CA et al., 2010). Therefore, assessing selectivity for a specific target as against homologs and other proteins close in sequence or

structure will be benefitting before proceeding with the experimental validation of all the hit molecules resulting from virtual screening.

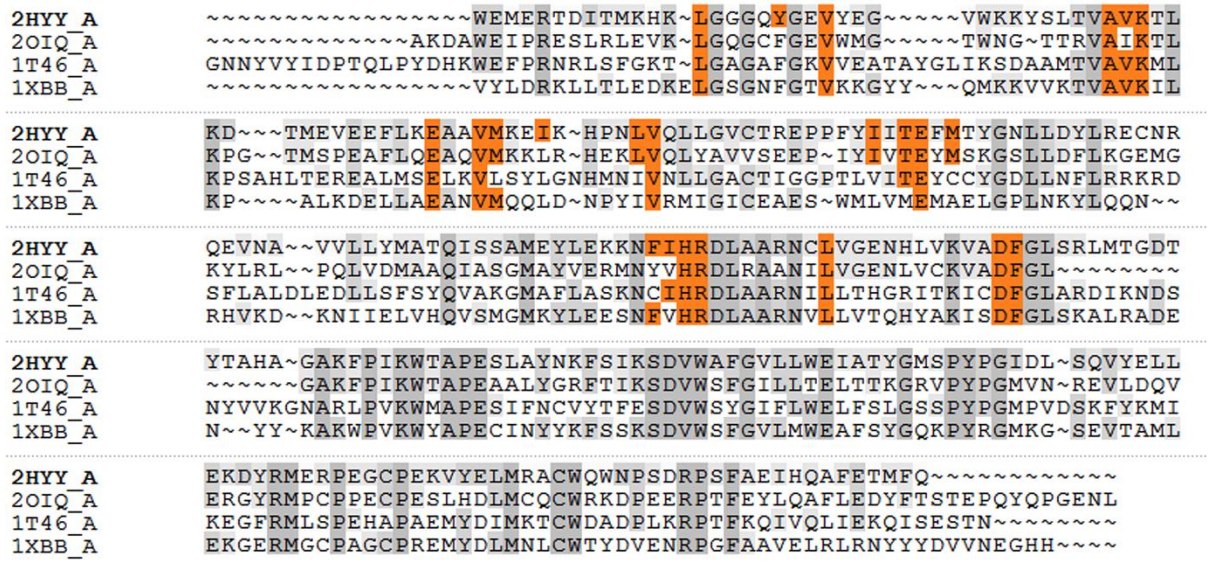


Figure 4.9 *(2HYY), c-Src (2OIQ), c-Kit (1T46) and Syk (1XBB). The conserved residues are in grey and the residues in the active site pocket of c-Abl (<5A°) are in orange. The figure is generated using multiple sequence viewer in Maestro software.*

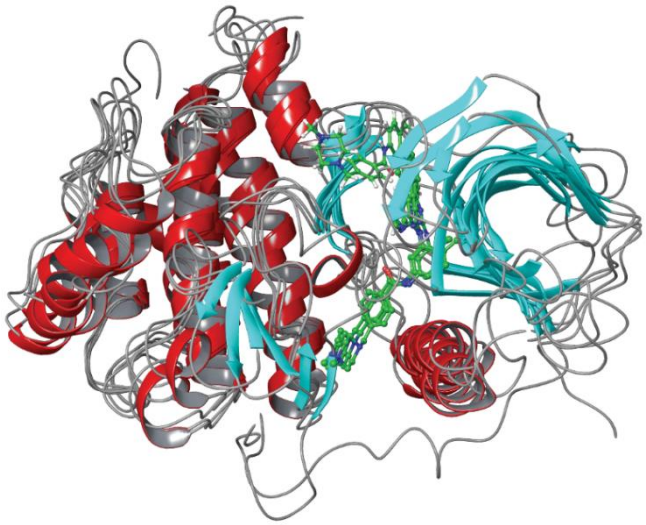


Figure 4.10 *Superimposition of the four kinases c-Abl, c-Src, c-Kit and Syk along with the bound inhibitor Imatinib.*

Comparative docking score analysis has been be applied to assess qualitatively the selectivity of a pre-selected set of shortlisted inhibitors (Ozbuyukkaya, Ozkirimli Olmez, & Ulgen, 2013; Szelag, Czerwoniec, Wesoly, & Bluyssen, 2015). In fact, selectivity for Yersinia outer protein

E (YopE) as against its homologs was verified by comparing the docking scores. The results identified for YopE showed lower binding score for the homologs of YopE indicating selectivity for YopE (Ozbuyukkaya et al., 2013). Similarly, in another study, docking results clearly pointed out specificity of natural compounds for human STAT1 and STAT3 (Szelag et al., 2015). Therefore, to address the selectivity of a particular hit compound towards a target, relating the calculated binding score with observed activity can be attempted by taking into account the known inhibitors Imatinib as positive controls. As a first step in that direction, we adopted the strategy of relating the docking score with known activity of Imatinib towards c-Abl and the other three structurally similar kinases. It may be noted that activity of Imatinb towards c-Abl (0.17 μM), Src kinase (100 μM), Syk (5-10 μM) and c-Kit (0.4μM) is experimentally observed (**Table 4.4**). When the experimentally measured activity (IC_{50}) of Imatinib was related to the calculated docking score when bound to the four kinases (**Table 4.4**), it may be observed that, qualitatively, a score of $\sim \geq 9.0$ kcal/mole implies certain degree of activity.

Based on the observed relation between the calculated binding score and experimentally measured activity, we attempted a similar analysis to assess selectivity of the identified compounds for c-Abl. The compounds identified using the virtual screening and docking results (**Fig. 4.3**), were docked into the crystal structure of c-Src (2OIQ), c-Kit (1T46) and Syk (1XBB). The four Imatinib bound crystal structures were prepared and their binding site grids were generated, as described in the protein preparation section. Glide XP docking was performed on all the kinases and the docking score differences as well as the interactions of hit compounds between these proteins was assessed.

Table 4.4 and **Fig.4.11(C)**, provide the calculated binding scores of Imatinib bound to c-Abl, along with the hit compounds (A-J) obtained from this study. Compounds A-J show a dock score of >-9.0 kcal/mole for c-Abl whereas, in comparison, the docking score is <-9.0 kcal/mole for the three other kinases, with the exception of Compound C and F showing a score slightly higher than -9.0 kcal/mole for c-Src kinase and c-Kit, respectively. In fact, Compounds A-D show affinity very similar to the binding score of Imatinib to c-Abl.

Kinases → / **Compounds ↓**	**c-Abl**	**c-Src/2OIQ**	**Syk/1XBB**	**c-Kit/1T46**
Imatinib/2HYY	-13.70	-11.07	-8.91	-13.51
Bafetinib/2E2B	-14.31	-14.37	-14.01	-13.44
Nilotinib/3CS9	-14.80	-12.46	-10.20	-13.92
7MP/2HIW	-14.43	-5.01	-4.09	-6.81
A	-12.61	-6.45	-4.74	-8.12
B	-11.66	-7.55	-4.16	-9.14
C	-11.60	-9.14	-4.86	-10.13
D	-11.11	-7.67	-6.21	-6.71
E	-10.95	-8.31	-4.96	-8.60
F	-10.95	-8.76	-5.33	-9.20
G	-10.39	-7.60	-5.33	-7.63
H	-10.29	-7.33	-4.67	-8.74
I	-10.07	-7.21	-5.19	-8.93
J	-9.37	-6.78	-5.21	-8.44

Table 4.4 *Comparison of the binding scores (kcal/mole) of Imatinib/2HYY,Bafetinib/2E2B, Nilotinib/3CS9, 7MP/2HIWand the identified hit compounds for c-Abl.c-Src/2OIQ, c-Kit/1T46 and Syk/1XBB represent Imatinib bound to different Kinases.*

For a compound to be active against c-Abl the difference in docking score between the hit Compounds (A-J) and Imatinib should be minimal, whereas to be specific for c-Abl, it needs to be well separated from docking scores when bound to other kinases. From **Fig. 4.11(C)**, it can be seen that in comparison Imatinib, Compounds A and D and to some extent B clearly show a

possibility of being both active and more selective for c-Abl. This observation stems from the fact that the difference in the binding score of Imatinib between c-Abl compared to c-Kit and c-Src is <3 kcal/mole whereas Compound A and D has a binding score of >4 kcal/mole compared to the structurally similar kinases.

Further, we have also considered other compounds which showed inhibitory activity towards c-Abl. **Table 4.3** lists the docking scores of Bafetinib (2E2B), Nioltinib (3CS9), 7MP (2HIW) (32-34) bound to c-Abl along with the docking scores of these molecules bound to Src kinase (2OIQ), Syk (1XBB), c-Kit (1T46).[46-48] Based on similar arguments presented above, it is indicative that the hit Compounds A, B and D appear to be relatively more specific for c-Abl than 7MP **(Fig. 4.11(A)),** Bafetinib (**Fig. 4.11(B))** and Nilotinib (**Fig. 4.11(D**)) as well. Experimental verification will provide clues to decipher the selectivity of selected set of compounds for c-Abl and help in further optimization.

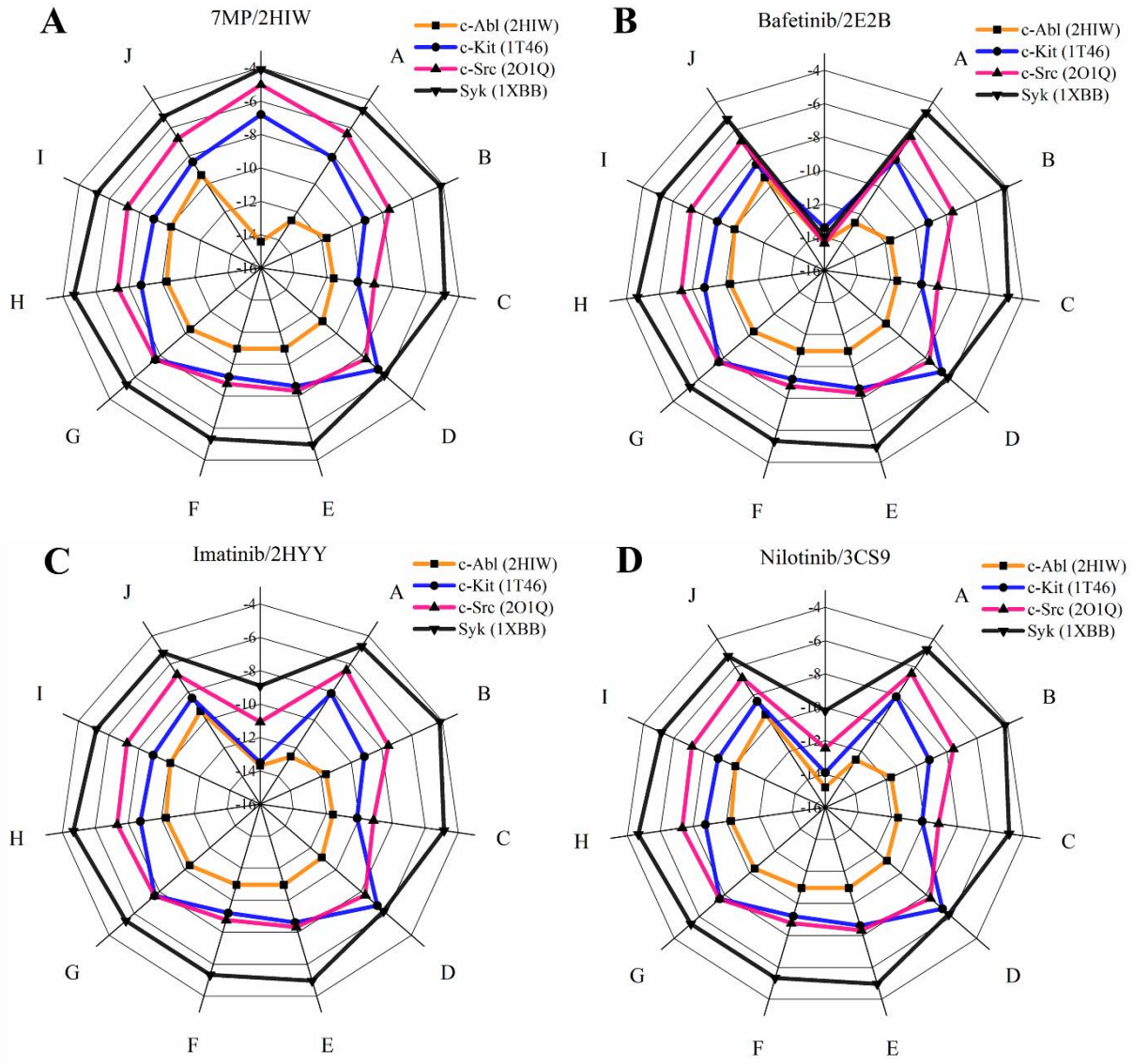


Figure 4.11 *Comparison of the docking scores of (A). 7MP(2HIW), (B). Bafetinib(2E2B), (C). Imatinib(2HYY) and (D). Nilotinib(3CS9) with the top 10 identified compounds docked into c-Abl and, also the kinases, c-Kit, c-Src, Syk. The numbers represent the docking score (kcal/mole). The figure is generated using Origin (OriginLab, Northampton, MA).*

Conclusion

Given the implicated role of c-Abl in AD and other neurodegenerative diseases and the lack of an effective treatment for Alzheimer's, it is important to explore virtual screening of a large number of molecules for their potency to curtail the activity of targets such as c-Abl. In this chapter we have employed e-pharmacophore modeling procedure to identify potential inhibitors against Abelson Tyrosine kinase activity. Multiple energy based pharmacophore modeling, virtual screening and docking coupled with selectivity analysis with other structurally similar kinases has helped us to propose three potential inhibitors against c-Abl activity. The molecules identified from virtual screening may be experimentally validated for activity and as well as selectivity for c-Abl. Availability of a successful and selective inhibitor may lead to exploring new strategies such as nano based approaches to deliver the molecules to the brain effectively.

Chapter 5

Small Molecules as Potential Amyloid Disruptors: ThT Fluorescence, Docking and Molecular Dynamics Studies

Introduction

Highly ordered amyloids are a consequence of proteins that misfold and are linked to a variety of human diseases (Iram & Naeem, 2014; T. P. Knowles, M. Vendruscolo, & C. M. Dobson, 2014; Ross & Poirier, 2004). They are formed by a wide variety of proteins differing in sequence and structure (Cardinale, Chiesa, & Sierks, 2014; Chiti & Dobson, 2006; T. P. Knowles et al., 2014). Another property associated with amyloids is their generic structure of highly ordered β-sheet structures stabilized by a network hydrogen bonding interactions between the strands of the β sheet structure (Chiti & Dobson, 2006; Jahn & Radford, 2005; Meier & Bockmann, 2015). Amyloid formation begins with a protofibril unit that can laterally grow into fully matured amyloid fibril (Eisenberg & Jucker, 2012; MB, 2006). Protein misfolding responsible for diseases such as Alzheimer's, Parkinson's, cystic fibrosis, diabetes, cancer etc., share a common amyloid structure (Cardinale et al., 2014; Chiti & Dobson, 2006; T. P. Knowles et al., 2014). Traditionally, proteins that are involved in the disease causing pathways are targeted for developing modulator molecules that can either inhibit or activate the target protein leading to the development of drug molecules. Complimentary to the common approach of inhibiting proteins, small molecules can be discovered that can inhibit amyloid formation and or disrupting the pre-formed amyloids. Efforts are in progress to discover naturally occurring and synthetic small molecules that can inhibit the formation of amyloid fibrils (Bu, Rao, & Wang, 2015; Doig & Derreumaux, 2015; Ho et al., 2015; Kitagawa et al.,

2015; Palmal et al., 2014; Ramshini, mohammad-zadeh, & Ebrahim-Habibi, 2015; Shariatizi, Meratan, Ghasemi, & Nemat-Gorgani, 2015; Wang & Raleigh, 2014; Wiesehan et al., 2008; Xie et al., 2015). At present there is no report of approved therapies to inhibit amyloid formation. However, many molecules have shown to be inhibiting and or disrupting amyloids. Thus discovery of small molecule inhibitors holds a great promise towards a therapeutic approach for protein misfolding diseases. In this context, it becomes necessary to elucidate the mode of interactions between the molecules and the amyloid in order to better understand the mechanism involved in their disintegration.

Molecular dynamics (MD) simulations together with docking studies play an important role in the identification of lead molecules that can potentially destabilize the ordered structures leading to fragmenting/disintegration of amyloids into less toxic amorphous deposits (B. Cheng et al., 2013). MD simulations are providing insights into the mode of interactions and underlying mechanisms at atomic detail (Baweja, Balamurugan, Subramanian, & Dhawan, 2015; Berhanu & Masunov, 2015; Hou et al., 2015; Kuang, Murugan, Tu, Nordberg, & Agren, 2015; Kumar et al., 2015; Tran & Ha-Duong, 2015; Ye, Wang, Jiang, Yu, & Chen, 2013).

As discussed in earlier chapters, the Aβ peptide is the main causative agent in the progression of Alzheimer's disease. Resulting from the cleavage of APP by β-secretase, the highly aggregation prone, 42 residue peptide spontaneously forms amyloid plaques causing toxicity to the neurons in the brain (B. Cheng et al., 2013).

Small molecules possessing anti-amyloidogenic property are being explored as potential inhibitors to treat not only AD, but other amyloid associated diseases. Natural compounds such as curcumin, quercetin, piceid, ferulic acid and other polyphenols are known to destabilize pre-

formed amyloids (Palmal et al., 2014; Yang et al., 2005). In pursuit of identifying potential amyloid inhibitors, we attempted to test the efficacy of small molecules to disintegrate the Aβ amyloids. The basis of selecting the compounds was based on two factors; (i) structural similarity to the molecules known to destabilize amyloids and (ii) the identified potential inhibitors of β-secretase from our studies.

To determine the efficacy to disintegrate preformed aggregates of Aβ-42 amyloid peptide, we employed fluorescence based studies with MD simulations to understand the interactions of a few small molecules that have shown a potential to destabilize the fibrils.

Materials and methods

Preparation of small molecules for Docking Studies

Small molecules that were tested for the potential to inhibit β–secretase were shortlisted. Out of the 15 compounds, five (Compounds 3, 8, 11, 13 and 15 in **Figure 5.1**) showed inhibitory activity against β-secretase in our earlier studies (Chapter 2) (Palakurti et al., 2013). Some of the shortlisted compounds also resemble the compounds previously identified to be amyloid disruptors. Natural compounds including Curcumin, Quercetin, Piceid, etc are also used to explore their binding mode with Aβ fibril. The structures of the 15 compounds along with the natural compounds are shown in **Figure 5.1** LigPrep module is used to generate 3D conformations of these compounds *(LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2011.)* Multiple conformations were generated for the selected molecules using the ConfGen search algorithm (Eldridge et al., 1997). **Figure 5.1** displays the 2D structures of the selected compounds, 1-15.

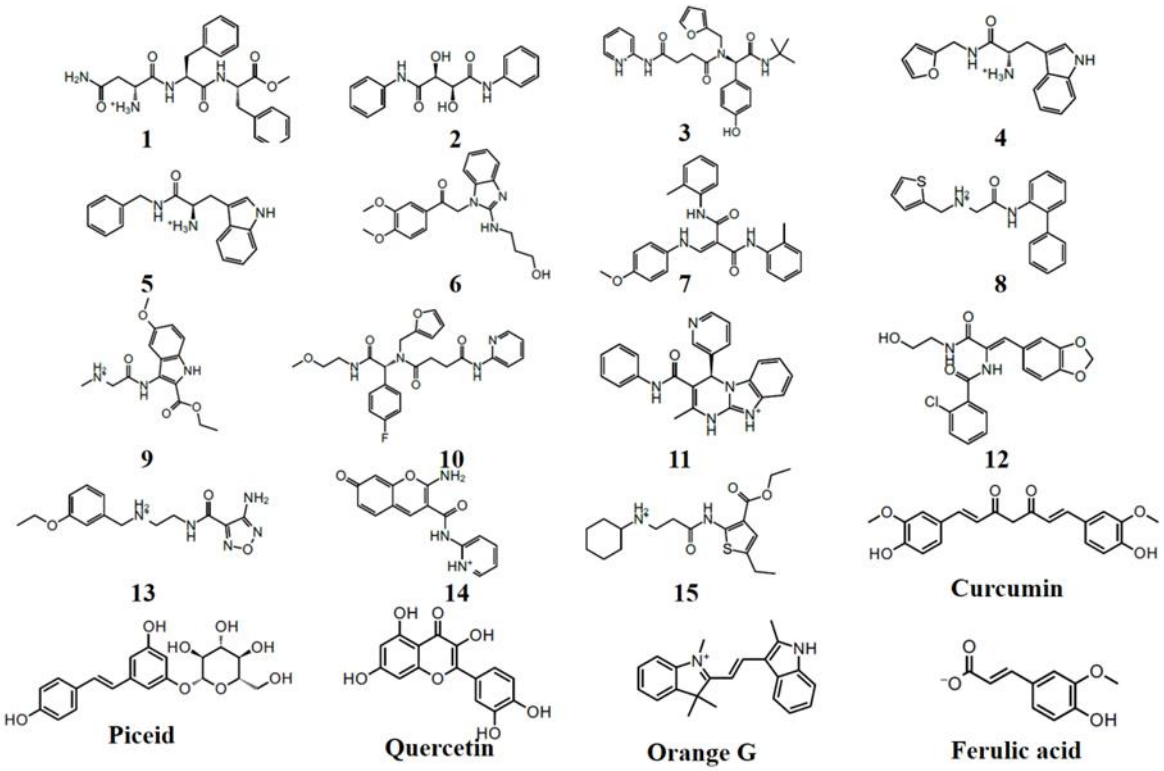


Figure 5.1 *The 2D structures of the shortlisted 15 compounds. Structures of the natural compounds are also included.*

Grid generation

The nuclear magnetic resonance (NMR) structure of Aβ-42 fibril, PDB code: 2BEG, (Luhrs et al., 2005) (**Figure 5.2**) was employed for the generation of the Grid file using Glide grid generation algorithm (Eldridge et al., 1997; Friesner, Banks, Murphy, et al., 2004; Friesner et al., 2006b; T. A. Halgren et al., 2004; K. Loving et al., 2009). The entire surface of the 2BEG structure is used for grid generation as the precise binding mode for these compounds is not known. The hydrogen bonding pattern of the β-sheets is represented as dotted lines while the side chains are shown in grey. The generated grid file is used for docking of the prepared compounds.

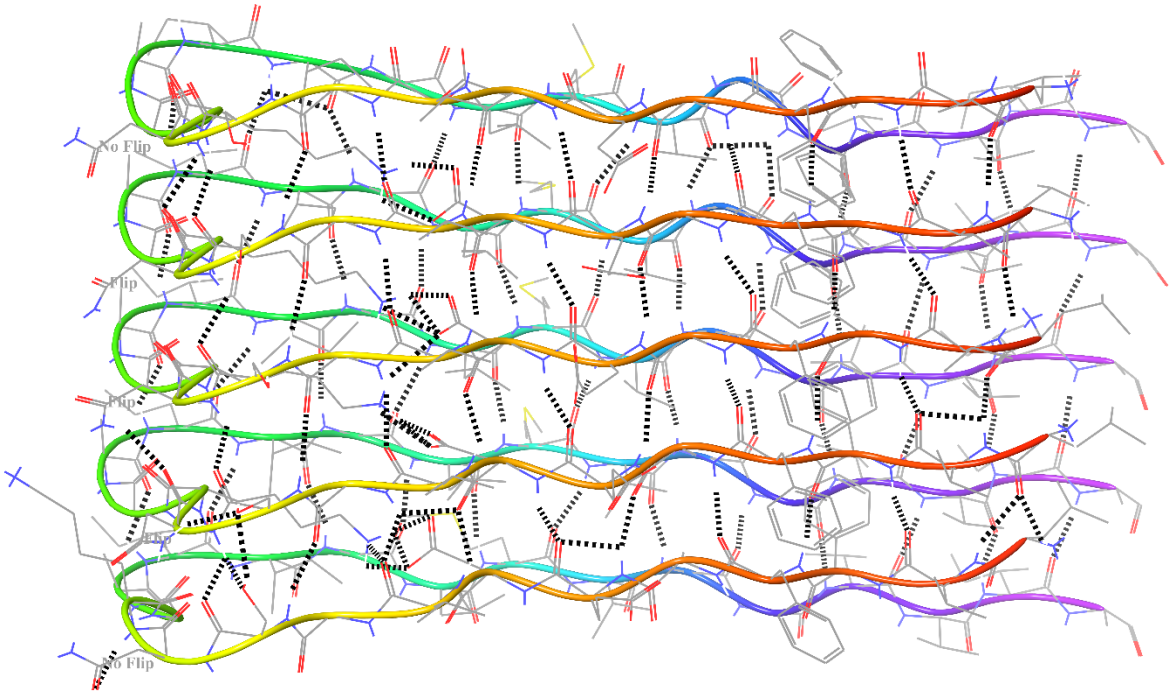


Figure 5.2 *3D structure of amyloid fibril PDB code: 2BEG, the black dotted lines represent the hydrogen bonding pattern in the fibril. Side chains are shown in grey.*

Docking Studies

Docking studies were performed using Glide module using extra precision mode. (*Glide, version 5.7, Schrödinger, LLC, New York, NY, 2011*) (Friesner et al., 2006b). The ligand flexibility as well as conformation generation is allowed for higher sampling during the docking. The 15 compounds (from Asinex database) as well as the natural compounds prepared with ligprep were assessed for their binding interactions.

ThT fluorescence assay

ThT fluorescence assay can be used to monitor the formation of amyloids as it specifically binds to the highly ordered, extensively hydrogen bonded β-sheet structure typical for amyloid. Fluorescence intensity is increased when ThT binds to the β amyloids while it decreases when the amyloids disintegrate. Therefore, increase or decrease of the ThT fluorescence is an indication of increase or decrease in the quantity of β-sheet structure.

Expression and purification of Aβ-42 peptide

Aβ-42 gene cloned into pRSETB vector is obtained from NIMHANS research centre (generous gift from Dr. Sharada Subramanian)The recombinant plasmid is transformed into competent BL21 (DE3) pLysS cells and plated on to LB ampicillin plates (50 μg). A single colony is inoculated into LB media until it reached an OD of 0.6. The protein expression was induced by the addition of IPTG of 1 mM concentration. After induction the cells were grown for 2 hours. The cells were then harvested by centrifugation and suspended in 5 ml lysis buffer (50 mM Tris-Cl, pH-8.0 containing 300 mM NaCl, 5 mM β-Mercaptoethanol and 2 mM PMSF). The cells were subjected to sonication on ice, followed by centrifugation at 17400 g for 20 min. The soluble fraction was discarded and insoluble fraction is dissolved in 8M urea containing Tris-Cl buffer followed by another round of centrifugation at 10000 g for 30 min. The supernatant obtained was loaded into Ni^{2+} CAM affinity resin pre-equilibrated with lysis buffer containing 10 mM imidazole. The recombinant peptide was eluted by 400 mM Imidazole buffer. Fractions of 1 ml were collected and run on 15% SDS to identify the presence of the recombinant peptide in the fractions. Followed by the pooling of the recombinant protein, dialysis is performed using a 3.5 kDa membrane to remove Imidazole using 20 mM Tris-Cl buffer (pH-8.0). This step is followed by Enterokinase digestion to release the Aβ 1-42 peptide. The cleaved peptide is purified by using Ni^{+2} affinity column. The purified Aβ-42 was concentrated and lyophilized and used for the destabilization assay. The purification method followed was as reported by in Sharada et. al 2007 (Subramanian S, 2007). **Figure 5.3** shows the 15% SDS PAGE gel with molecular weight markers, Enterokinase cleavage step and the purified amyloid-β 1-42.

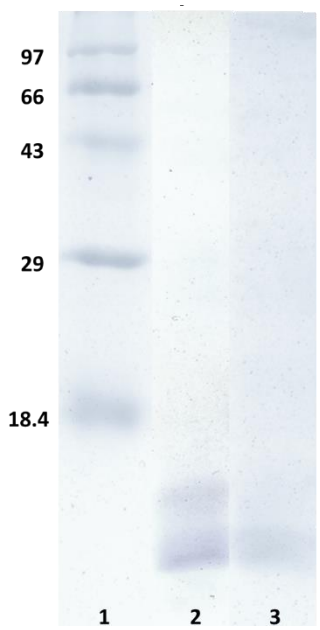


Figure 5.3 *15% SDS gel showing 1. Molecular weight markers, 2. Enterokinase cleavage showing the bands corresponding to cleaved and uncleaved peptide, 3 Purified Aβ-1-42(4 kDa).*

ThT fluorescence assay-(Amyloid Destabilization)

The compounds that showed a relatively high docking score (obtained from Asinex) were tested for their efficacy by ThT fluorescence assay to disrupt the amyloids formed by Aβ-42 peptide. Their amyloid destabilizing ability was compared to that of Ferulic acid, a natural compound known to destabilize Aβ amyloids taken as a positive control.

Aβ peptide is dissolved in 50 mM phosphate buffer, pH 7.5, and 100 mM NaCl. The stock solution is incubated at a temperature of 4^{o} C with intermediate agitation to enhance the formation of amyloids until an equilibration state is attained (Maya Mathew, 2013). ThT fluorescence of the fAβ (fibrillar Aβ) stock solution was measured for the formation of amyloid by the Aβ peptide indicating the formation fibrillar Aβ (fAβ). The fAβ is used for the destabilization assay to monitor decrease in ThT fluorescence intensity as a function of time. The assay is performed in duplicates and changes in the ThT fluorescence intensity in the presence of various compounds are monitored for a period of 6 hours.

The reaction mixture containing 25 μM of fAβ, 50 mM Phosphate buffer, pH-7.5 and 100 mM Nacl, 5 μM ThT. The reaction tubes were then incubated at 37^{o} C for 6 hours with hourly measurements of ThT fluorescence obtained by taking 50 μl aliquots of reaction mixture. The reaction tubes were not agitated during the incubation period. The measurements were done using SpectraMax® M Series Multi-Mode Microplate Reader. For fluorescence measurements, λmax wavelength values of 445 and 490 nm were used for excitation and emission of ThT (Maskevich et al., 2007).

Molecular Dynamic Simulation Analysis

The fibril binding mode of the compounds 3 was analyzed in greater detail using MD simulations. A 10 ns simulation of the docked conformation of the compounds in the amyloid-β-fibril is performed. The NMR structure, 2BEG, contains five Aβ-42 peptides in a β-sheet structure as shown in **Figure-5.1**.

Based on the docking and binding mode analysis of the complexes, MD simulation is performed using the same methodology as described in Chapter 3, 4. The stability of the interactions of compound 3 was ascertained using trajectory analysis.

Results and discussion

Molecular Docking

The compounds 1-15 were docked into the amyloid structure based on the gird file generated earlier. The docking results indicate that the compounds 1-6 display a docking score of >-4.0 kcal/mole. The natural compounds Curcumin, Piceid, Quercitin displayed higher docking scores -5.5 to -4.80 kcal/mol while Orange G and Ferulic acid showed -3.7 to -3.0 kcal/mol respectively, (**Table 5.1**) indicating the similarities of the docking scores for 1-6 and the natural

compounds. Most of these compounds bind at a similar position of the amyloid structure (**Figure 5.4a**) except compound 5, compound 12 and Quercitin, which are bound at a different position (**Figure-5.4b)**.

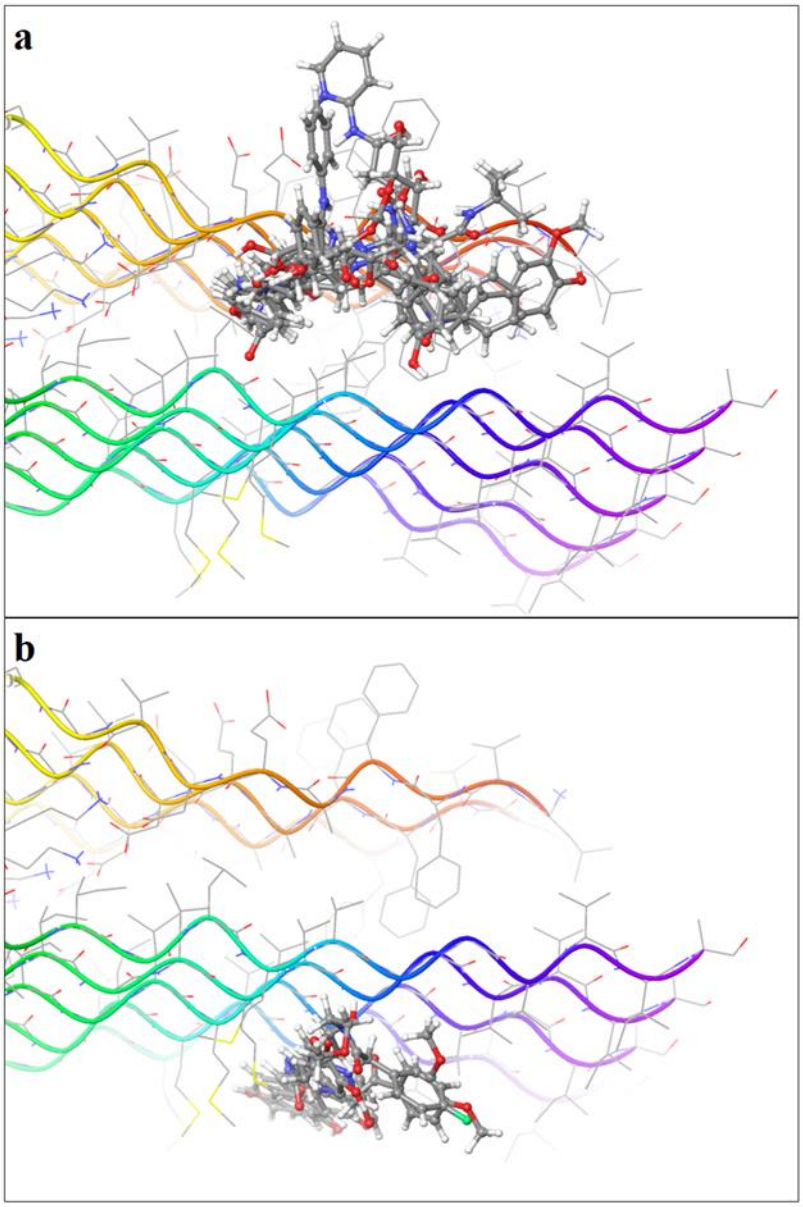
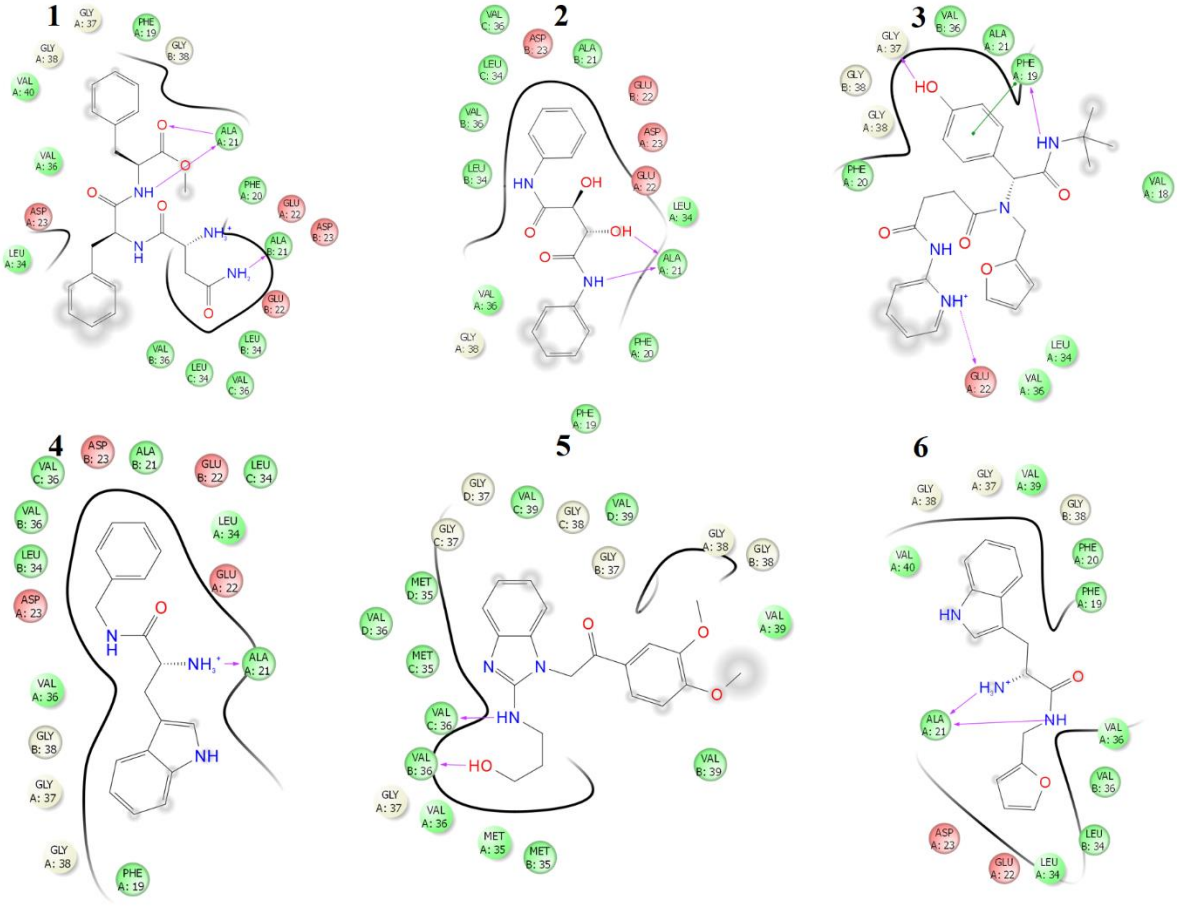


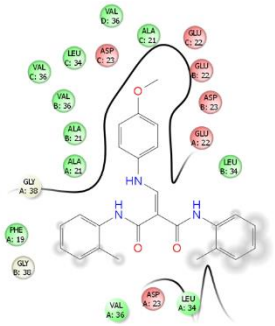
Figure 5.4 *The different binding mode of molecules to the amyloid fibril.*

Further we have analyzed the interactions of the compounds bound to the amyloids. Curcumin shows a single hydrogen bonding interaction with Leu-17, while Ferulic acid and Quercitin showed no hydrogen bonding interactions. Compound 1, 2, 4 showed hydrogen bonding to

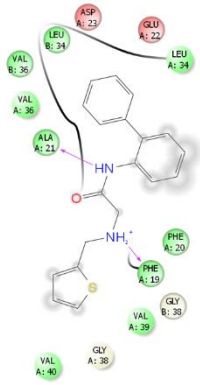
Alanine-21. The compound 3 displayed hydrogen bonding interactions with Glu-22 and Gly-37. Pi-Pi interaction with Phe-19 is also observed for this compound. Compound 5 hydrogen bonded with Val-36. Compound 6 hydrogen bonded with Ala-21 (**Figure 5.5).**



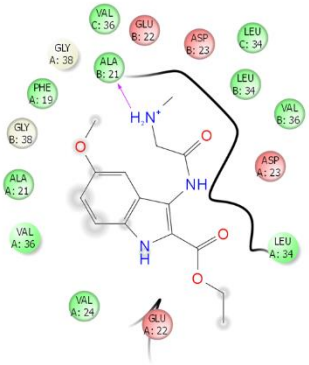
7



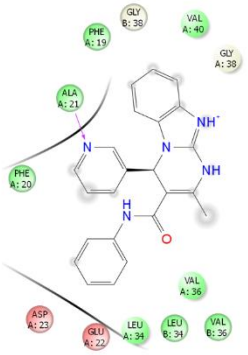
8



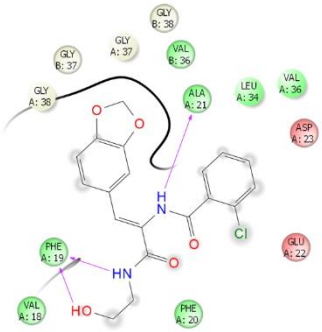
9



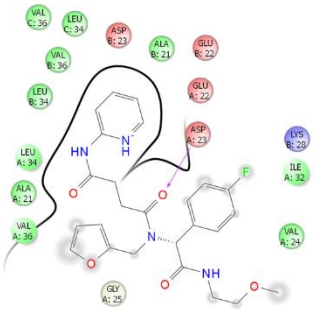
10



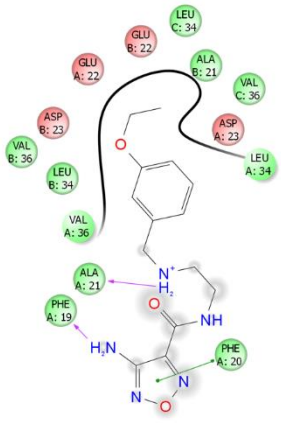
11



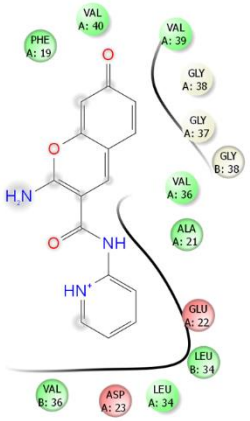
12



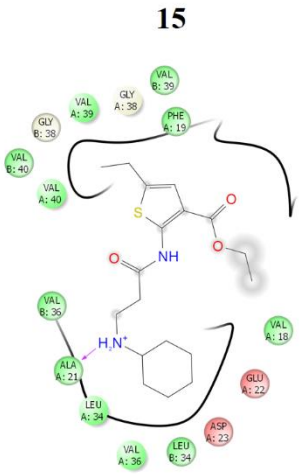
13



14



15



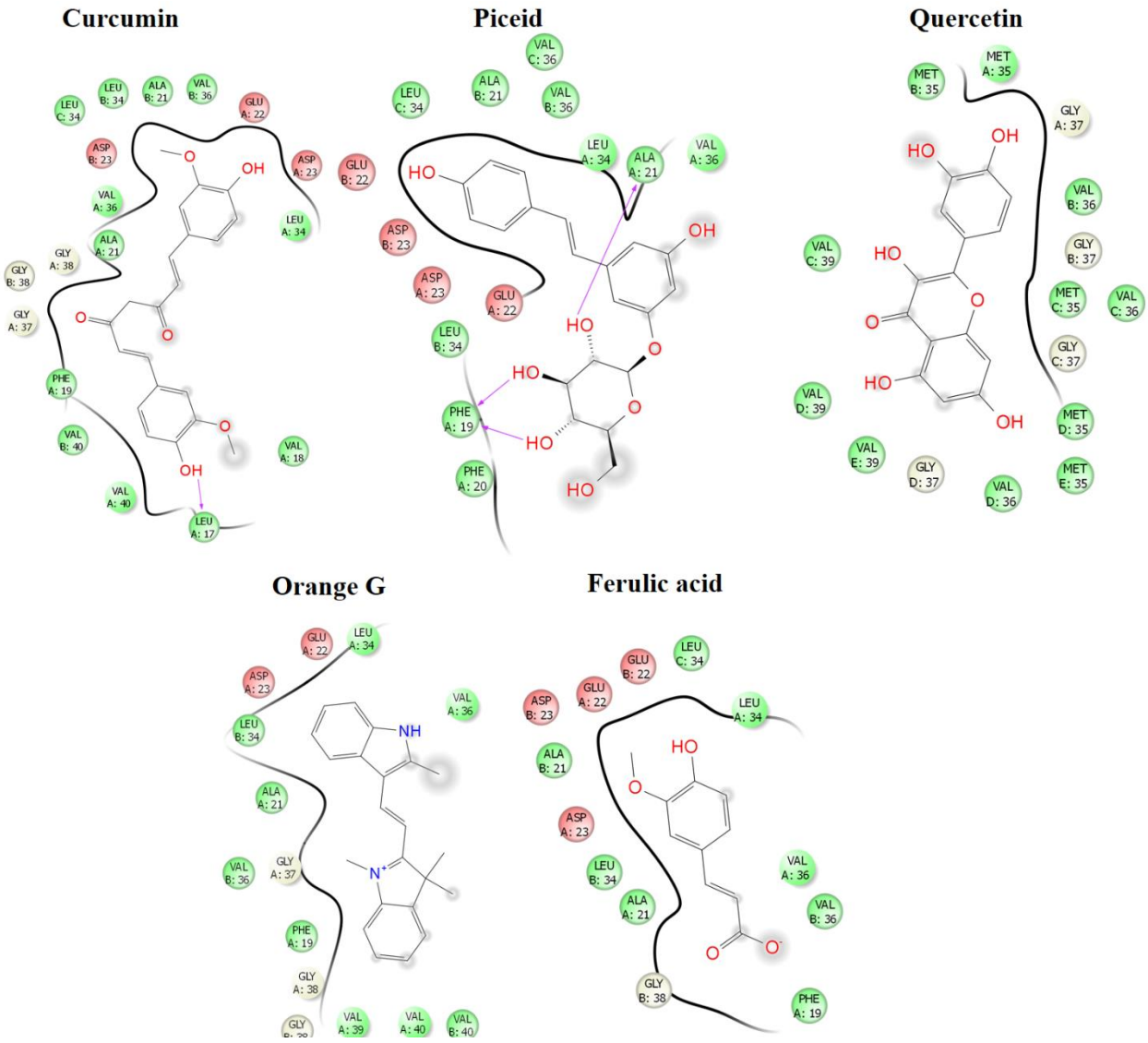


Figure 5.5 *Analysis of the interactions of the dataset compounds bound to amyloid fibrils.*

Compound	Docking Score	Interacting residues (via hydrogen bonding and stacking)
Curcumin	-5.503	Leu-17
Piceid	-5.474	Ala-21, Phe-19
Quercetin	-4.873	-
Orange G	-3.714	-
Ferulic acid	-3.046	-
1	-5.571	Ala-21
2	-5.091	Ala-21
3	-4.869	Gly-37, Glu-22, Phe-19
4	-4.335	Ala-21
5	-4.531	Val-36
6	-4.435	Ala-21
7	-3.768	-
8	-3.411	Phe-19, Ala-21
9	-3.253	Ala-21
10	-2.996	Ala-21
11	-3.097	Phe-19, Ala-21
12	-3.084	Val-36
13	-2.983	Phe-19, Phe-20
14	-2.846	-
15	-2.81	Ala-21

'-' denotes the absence of any interactions

Table 5.1 *Summary of docking scores and interacting amino acids of the amyloid peptide with the compounds.*

ThT fluorescence assay studies

The ThT fluorescence of the reaction mixture was recorded periodically for a period of 6 hours to identify decreased fluorescence intensity due to destabilization. 50 μM Ferulic Acid is used as positive control for the assay as it is a known amyloid destabilizing compound. Compounds 3 and 10 clearly showed decreased ThT fluorescence (**Figure 5.6**) indicating the destabilization of amyloid fibrils, though not as efficient as Ferulic acid.

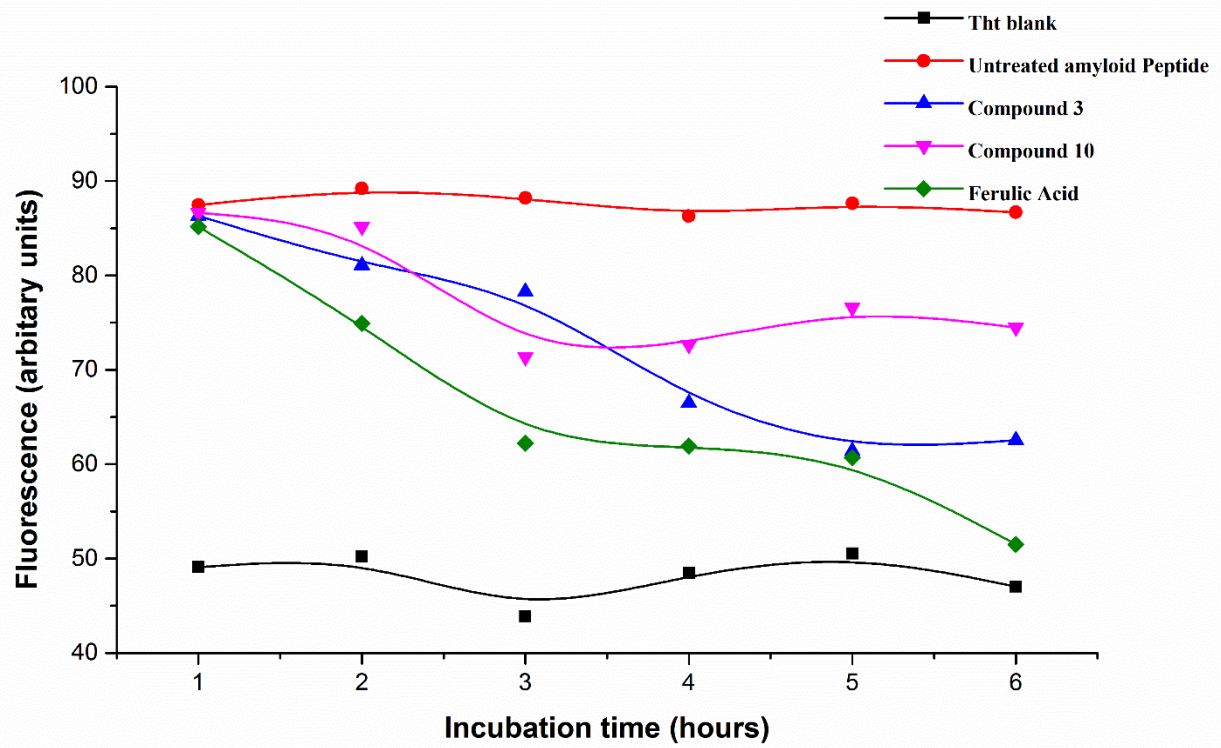


Figure 5.6 *Destabilizing effect on amyloid-β peptide. ThT fluorescence intensity monitored as a function of incubation time of peptide-compound mixture.*

Binding mode analysis of the lead compounds from MD simulations

Figure 5.7 shows the mode of interactions of compounds A and B with the amyloid structure before simulation. Compound 3 makes three hydrogen bonding interactions with Glu-22, Phe-19 and Gly-37 and compound 10 forms a single hydrogen bond with Ala-21.

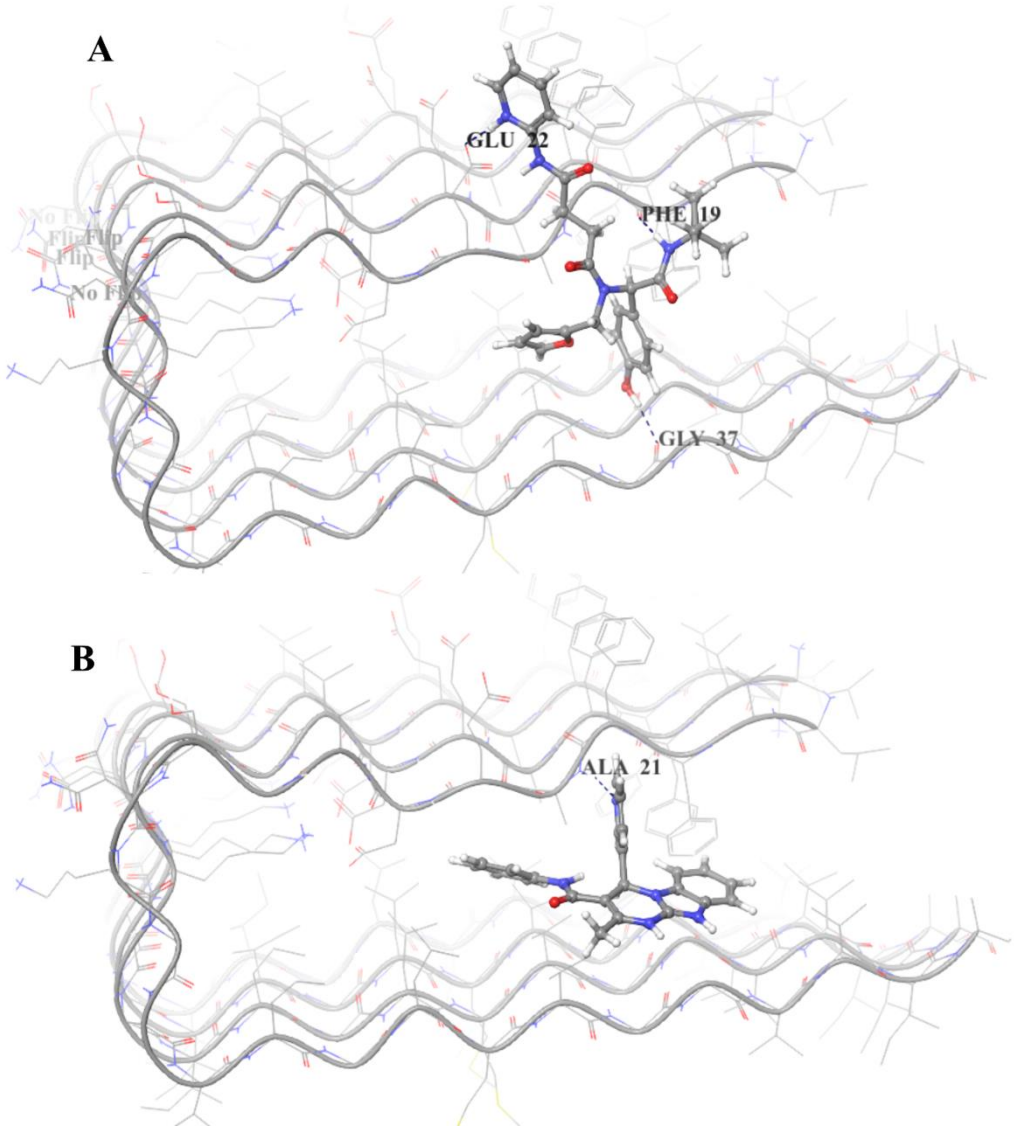


Figure 5.7 *Binding mode of compound 3 (A) and compound 10 (B) with the PDB structure 2BEG at the beginning of the simulation.. The hydrogen bonds are displayed as dark green dotted lines.*

Analysis of the trajectories from the simulation indicate that the compound 3 is retaining the interactions with the 3 amino acids that are observed in the docking studies. The RMSD of the fibril alone and the fibril-compound complex remained stable after 2 ns till the end of 10 ns.

Slight deviations in the RMSD when compared to the control simulation is corresponding to the change in distances between the peptide chains in the fibril (**Figure 5.8)**. Disruption of the hydrogen bonding between the individual fibrils from 5 ns onwards can be observed indicating that the binding results in destabilization. Even though the RMSD displayed minor fluctuations in the trajectory, the visual inspection (**Figure 5.9**) revealed that the complex was undergoing a significant extent of destabilization due to loss of hydrogen bond network (**Figure 5.10**). The hydrogen bonding interactions of the compound 3 with the residues of the amyloid peptide are persistent throughout the length of the 10 ns simulation (**Figure 5.11**).

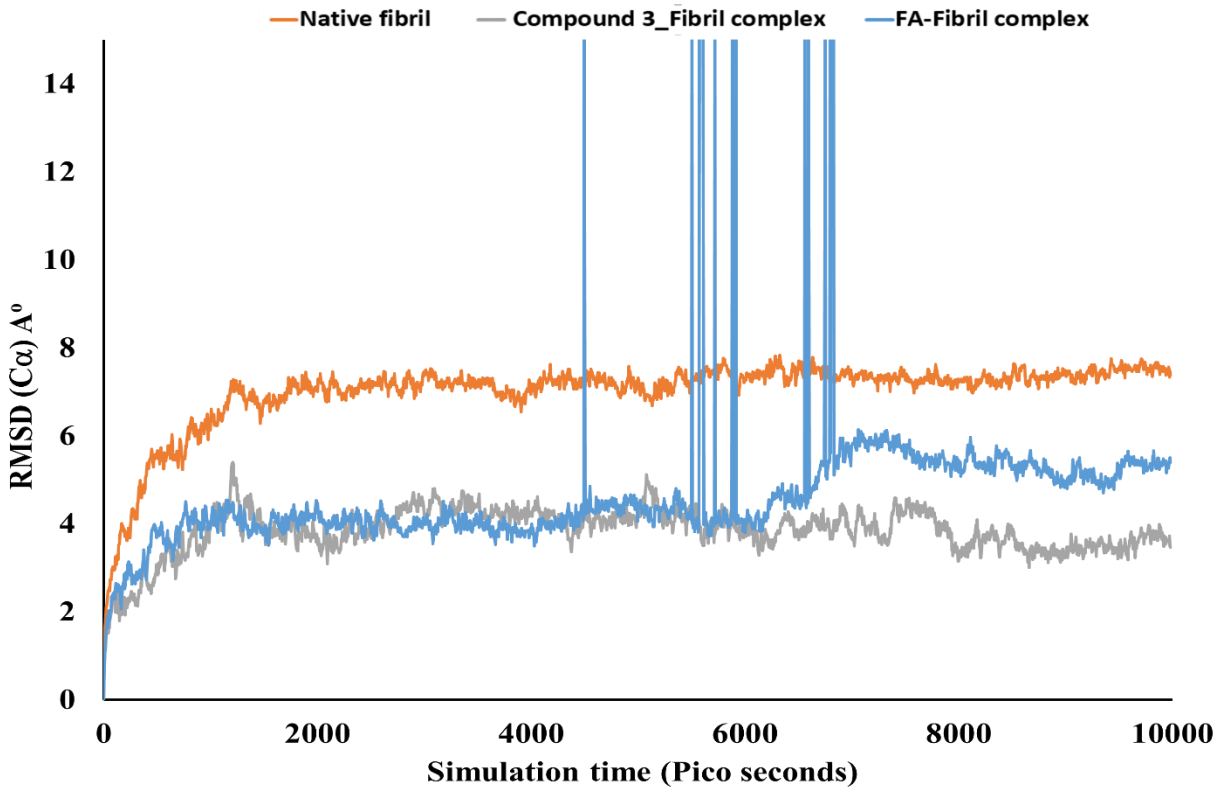


Figure 5.8 *The RMSD plot for the compound 3 complex (grey), Ferulic acid (blue) and the control simulation in the absence of ligand (orange).*

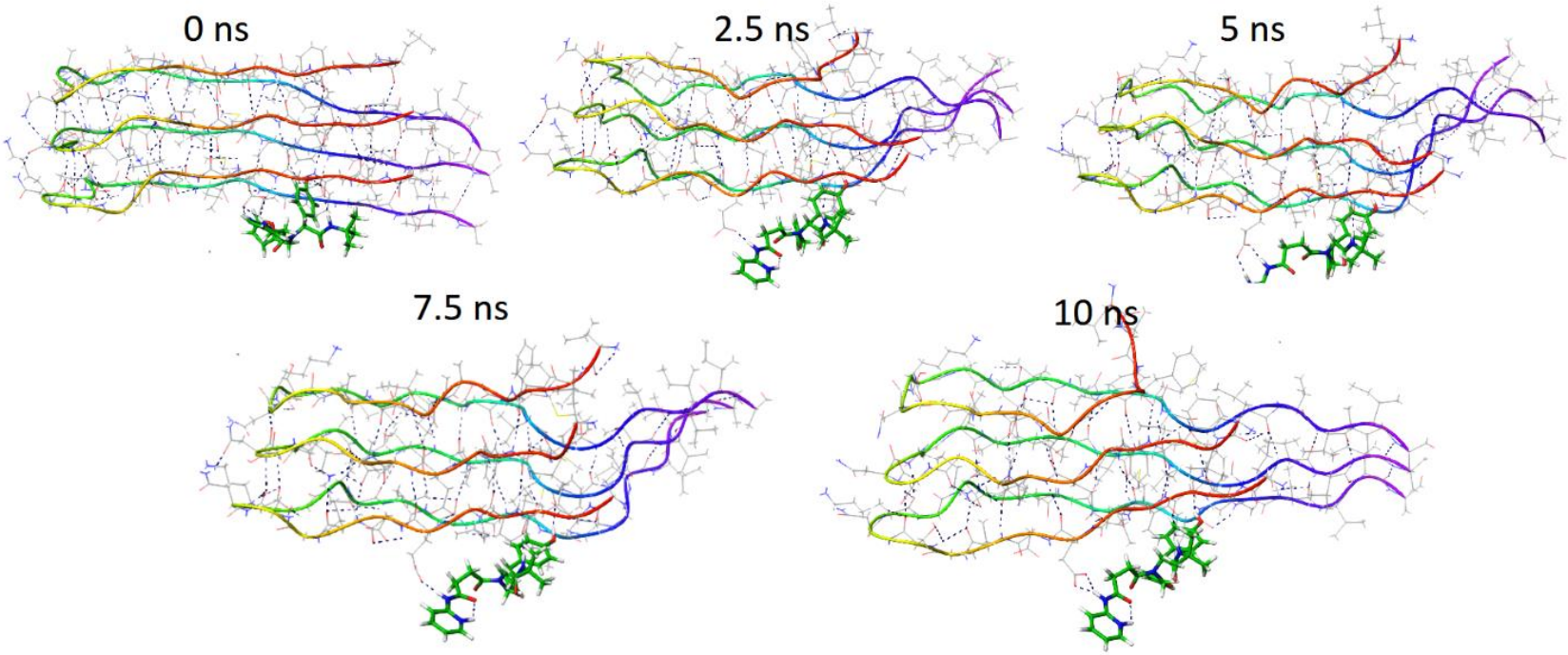
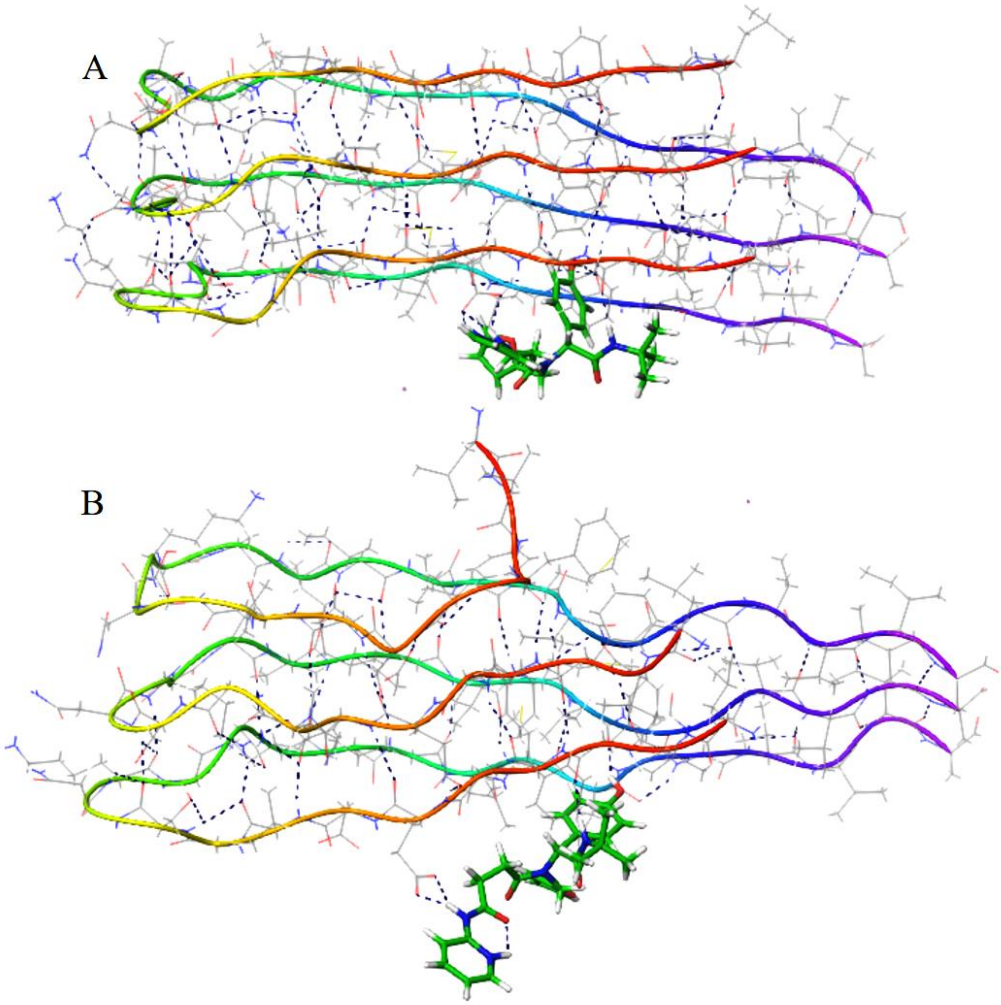


Figure 5.9 *Snapshots of different times of the simulation. The compound 3 is displayed in green sticks, the black dotted lines indicate the hydrogen bonding network of the fibril as well as the hydrogen bonding interactions of the compound 3 to the fibril.*

The destabilization of the amyloids in the simulations from binding of compound 3 is consistent from the results of the ThT fluorescence. Perhaps, simulation for longer time scales may conclusively reveal the destabilizing process as a result of ligand binding. However the study provided an insight into the nature of the molecular interactions of the fibril with compound 3.



Figures 5.10 *Interaction analysis of compound 3 with amyloid fibril before and after the 10 ns simulation. Black dotted lines indicate the hydrogen bonding network of the amyloid fibril and also the interactions with the ligand.*

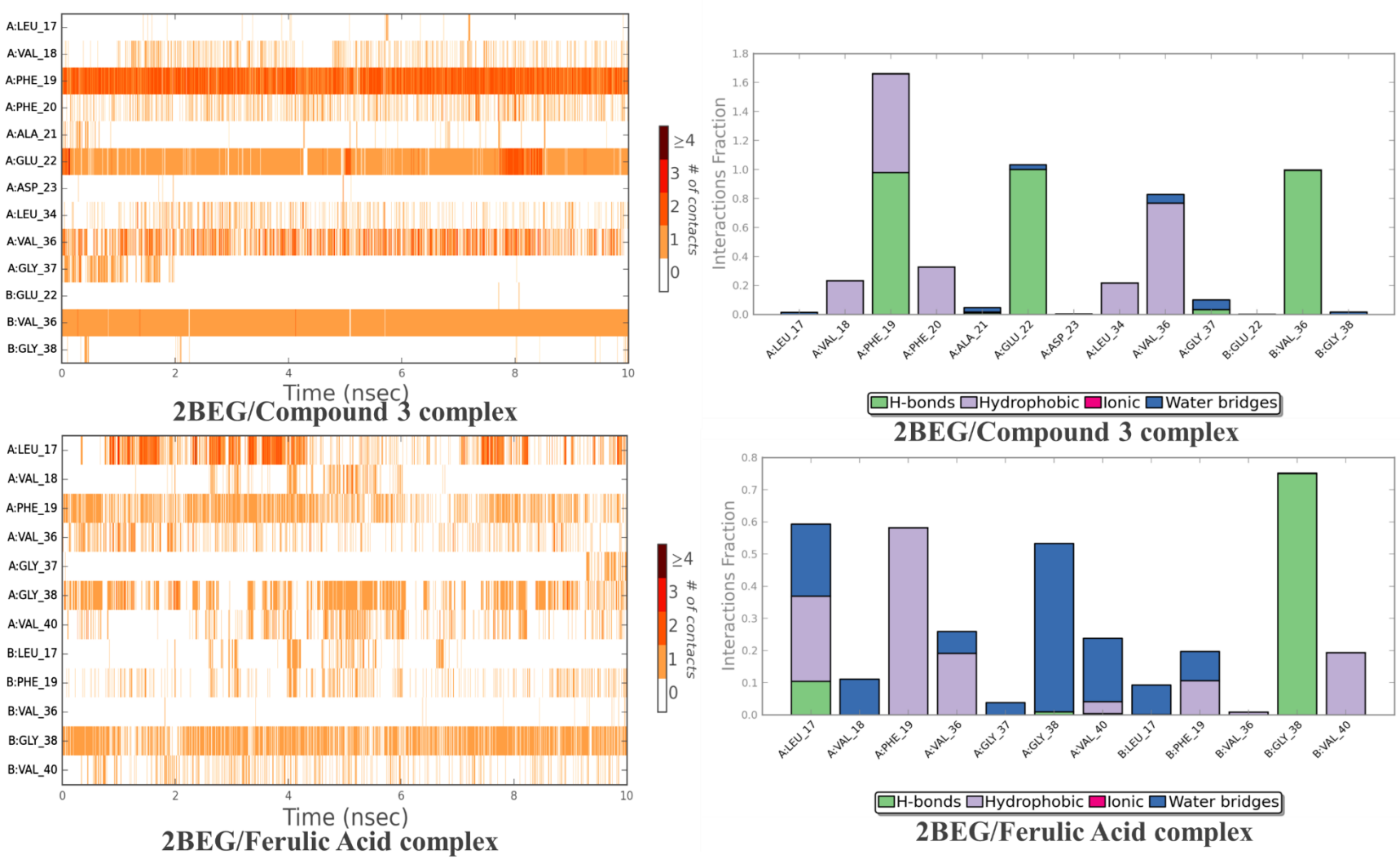


Figure: 5.11 *Interactions Persistence and the interactions fraction of compound 3 and Ferulic acid with amyloid fibril during the 10 ns simulation.*

Conclusion

In this study, we have explored the possible interactions between amyloid fibril binding and destabilizing compounds. The docking based analysis of binding to the amyloid fibril provided insights into the mode of interactions of known amyloid destabilizing compounds like Ferulic acid and curcumin as well as the test compounds employed in this study. ThT fluorescence analysis of these compounds in the presence of fibrils provided the validation of the computational methods as 2 compounds lead to destabilization of the preformed amyloid fibrils. Surprisingly, the compound 10 with a dock score of -3.0 kcal/mol also showed positive result in ThT fluorescence assays. Perhaps MD simulations can shed light on the dynamics and destabilizing effect of compound 10. It is interesting to note that one of the lead compounds (compound 3) was identified as an active inhibitor of β-secretase (Chapter 2) indicating the possibility of developing this and compounds of this nature as dual inhibitors for β-secretase and for Aβ-42 as well.

Chapter 6

CONCLUSIONS AND FUTURE PERSPECTIVES

CONCLUSIONS

In summary, in this thesis, we have attempted to identify potential leads that could be developed as drugs for AD. One of the approaches followed, was to target the enzymes involved in the pathway(s) leading to AD. A second approach was to identify molecules that possess amyloid destabilizing efficacy. A combination of modeling and experimental studies provided potential leads in addition to providing the important interactions that can be considered for lead optimization.

E-pharmacophore modeling using multiple crystal structures combined with high-throughput virtual screening resulted in the identification of lead compounds showing activity measured from *in-vitro* assay. Detailed analysis of the performance of screening and docking by each pharmacophore model helped in understanding the importance of using multiple crystal structures for arriving with a diverse set of inhibitors for β-secretase 1, a key target for AD (Palakurti et al, 2013).

Further, we have also explored using 3D quantitative structure activity relationship (3D QSAR) approach, a commonly used approach which provides binding affinity predictions and the interactions required for activity. In our studies, we have considered a dataset of compounds with a narrow activity range to explore the distinct and structurally important features responsible for increased activity of known inhibitors. The resulting QSAR properties of the

ligand based pharmacophore was used to screen databases to identify active compounds with distinct chemical structures *(manuscript under review)*.

In addition to considering an important target, β-secretase, efforts to identify potential leads were directed towards identifying potential inhibitors for c-Abl kinase, a very recently identified target for AD. Abnormal phosphorylation of Tau protein is implicated in AD and c-Abl is involved in this pathway. For the first time, we have used multiple e-pharmacophore based modeling and developed a pharmacophore model for virtual screening of molecules from CNS database, a dedicated database housing molecules with increased probability to penetrate the blood brain barrier. The short-listed promising candidates from docking energies were further subjected to molecular dynamics simulations to ascertain their binding energies, persistence of protein-ligand interactions, etc. Also we have made an attempt to address the challenges associated with selectivity for kinases comparing docking score with the observed activity of the existing inhibitors for c-Abl.

Disruption of amyloids is now considered as one of the alternate therapies for AD and other amyloid based diseases. Towards, this, we chose to identify compounds which can disrupt/destabilize amyloid fibrils. About 15 compounds were tested for their efficacy to disrupt the Aβ 1-42 amyloid fibrils using Fluorescence based assay for amyloid disruption. Aβ 1-42 peptide was over-expressed and purified and induced to form amyloid fibrils. ThT fluorescence experiments lead to the identification of two potential fibril destabilizing compounds. Analyzing the mode of binding and interactions involved from docking and molecular dynamics provided a structural rationale for the experimental observation.

Overall the thesis attempts to identify potential leads for Alzheimer's disease by not only targeting the relevant enzymes involved in the disease progression, but by also, identifying small molecules that can destabilize amyloid fibrils.

FUTURE PERSPECTIVES

There is yet no medical intervention for AD, and many research groups and multinational pharma industries are making efforts to discover drug for treating AD. The identification of potential leads will help in optimizing the leads to potential therapies for AD.

Identification of potential leads against c-Abl was successful leading to development of drugs to treat leukemia, the same is not true in the case of treating AD. The present set of identified leads from CNS database can optimized for activity, and efficacy for reaching the brain.

An interesting possibility that can be further explored is the possibility of developing small molecules to treat AD by targeting inhibition of enzymes as well as amyloids responsible for AD. Development of molecules with dual inhibitory action, against enzymes(s) and amyloids may prove to be a better therapeutic option.

Limitations

We have used a combination of new and existing strategies and arrived at the resulting compounds. Even though these compounds are tested for their inhibitory activity against β-secretase, but, their actual blood brain penetration ability still needs to be tested experimentally. Also the c-Abl inhibitor selectivity is still an important aspect to address all though the docking based selectivity score indicates selectivity of the identified inhibitors, it needs to be validated experimentally.

References

Parthasarathi, R., et al., *Bader's electron density analysis of hydrogen bonding in secondary structural elements of protein.* J Phys Chem A, 2007. **111**(30): p. 7141-8.

Hartl, F.U. and M. Hayer-Hartl, *Converging concepts of protein folding in vitro and in vivo.* Nat Struct Mol Biol, 2009. **16**(6): p. 574-81.

Chiti, F. and C.M. Dobson, *Protein Misfolding, Functional Amyloid, and Human Disease.* Annu Rev Biochem, 2006. **75**: p. 333-366.

Valastyan, J.S. and S. Lindquist, *Mechanisms of protein-folding diseases at a glance.* Dis Model Mech, 2014. **7**(1): p. 9-14.

Jahn, T.R. and S.E. Radford, *The Yin and Yang of protein folding.* FEBS J, 2005. **272**(23): p. 5962-5970.

Meier, B.H. and A. Bockmann, *The structure of fibrils from 'misfolded' proteins.* Curr Opin Struct Biol, 2015. **30**: p. 43-9.

Landau, M., et al., *Towards a pharmacophore for amyloid.* PLoS Biol, 2011. **9**(6): p. e1001080.

Knowles, T.P.J., M. Vendruscolo, and C.M. Dobson, *The amyloid state and its association with protein misfolding diseases.* Nat Rev Mol Cell Biol, 2014. **15**(6): p. 384-396.

Eisenberg, D. and M. Jucker, *The amyloid state of proteins in human diseases.* Cell, 2012. **148**(6): p. 1188-203.

Pepys, MB., *Amyloidosis.* Annu Rev Med, 2006. **57**: p. 223-241.

Rajasekhar, K., M. Chakrabarti, and T. Govindaraju, *Function and toxicity of amyloid beta and recent therapeutic interventions targeting amyloid beta in Alzheimer's disease.* Chem Commun (Camb), 2015. **51**(70): p. 13434-50.

SY, O. and D. DE, *A brief overview of amyloids and Alzheimer's disease.* Protein Sci, 2014. **23**(10): p. 1315-1331.

Ow, S.Y. and D.E. Dunstan, *A brief overview of amyloids and Alzheimer's disease.* Protein Sci, 2014. **23**(10): p. 1315-1331.

Patterson, B.W., et al., *Age and amyloid effects on human central nervous system amyloid-beta kinetics.* Ann Neurol, 2015. **78**(3): p. 439-53.

O'Brien, R.J. and P.C. Wong, *Amyloid precursor protein processing and Alzheimer's disease.* Annu Rev Neurosci, 2011. **34**: p. 185-204.

J, H. and A. D, *Amyloid deposition as the central event in the aetiology of Alzheimer's disease.* Trends Pharmacol Sci, 1991. **12**(10): p. 383-8.

Killick, R., et al., *p73: a multifunctional protein in neurobiology.* Mol Neurobiol, 2011. **43**(2): p. 139-46.

Sheng, M., B.L. Sabatini, and T.C. Sudhof, *Synapses and Alzheimer's disease.* Cold Spring Harb Perspect Biol, 2012. **4**(5).

Ferri, C.P., et al., *Global prevalence of dementia: a Delphi consensus study.* Lancet, 2005. **366**(9503): p. 2112-7.

Kalaria, R.N., et al., *Alzheimer's disease and vascular dementia in developing countries: prevalence, management, and risk factors.* Lancet Neurol, 2008. **7**(9): p. 812-26.

Vivar, C., *Adult Hippocampal Neurogenesis, Aging and Neurodegenerative Diseases: Possible Strategies to Prevent Cognitive Impairment.* Curr Top Med Chem, 2015. **15**(21): p. 2175-92.

Hindle, J.V., *Ageing, neurodegeneration and Parkinson's disease.* Age Ageing, 2010. **39**(2): p. 156-61.

Mathuranath, P.S., et al., *Incidence of Alzheimer's disease in India: a 10 years follow-up study.* Neurol India, 2012. **60**(6): p. 625-30.

Chandra, V., et al., *Prevalence of Alzheimer's disease and other dementias in rural India: the Indo-US study.* Neurology, 1998. **51**(4): p. 1000-8.

Jia, Q., Y. Deng, and H. Qing, *Potential therapeutic strategies for Alzheimer's disease targeting or beyond beta-amyloid: insights from clinical trials.* Biomed Res Int, 2014. **2014**: p. 837157.

Han, S.H. and I. Mook-Jung, *Diverse molecular targets for therapeutic strategies in Alzheimer's disease.* J Korean Med Sci, 2014. **29**(7): p. 893-902.

Zhao, Y., et al., *Beta-Amyloid Precursor Protein (betaAPP) Processing in Alzheimer's Disease (AD) and Age-Related Macular Degeneration (AMD).* Mol Neurobiol, 2015. **52**(1): p. 533-44.

Rabinovich-Nikitin, I. and B. Solomon, *Inhibition of amyloid precursor protein processing leads to downregulation of apoptotic genes in Alzheimer's disease animal models.* Neurodegener Dis, 2014. **13**(2-3): p. 107-9.

Yamada, T., I. Goto, and Y. Sakaki, *Neuron-specific splicing of the Alzheimer amyloid precursor protein gene in a mini-gene system.* Biochem Biophys Res Commun, 1993. **195**(1): p. 442-8.

Citron, M., et al., *Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production.* Nature, 1992. **360**(6405): p. 672-4.

Weggen, S. and D. Beher, *Molecular consequences of amyloid precursor protein and presenilin mutations causing autosomal-dominant Alzheimer's disease.* Alzheimers Res Ther, 2012. **4**(2): p. 9.

Findeis, M.A., *The role of amyloid beta peptide 42 in Alzheimer's disease.* Pharmacol Ther, 2007. **116**(2): p. 266-86.

De Strooper, B., R. Vassar, and T. Golde, *The secretases: enzymes with therapeutic potential in Alzheimer disease.* Nat Rev Neurol, 2010. **6**(2): p. 99-107.

Robert Vassar, P.C.K., *The β-secretase enzyme BACE1 as a therapeutic target for Alzheimer's disease.* Alzheimer's Research & Therapy, 2011. **2**(20).

Vassar, R., *BACE1, the Alzheimer's beta-secretase enzyme, in health and disease.* Molecular Neurodegeneration, 2012. **7**(Suppl 1): p. L3.

Kandalepas, P.C. and R. Vassar, *The normal and pathologic roles of the Alzheimer's beta-secretase, BACE1.* Curr Alzheimer Res, 2014. **11**(5): p. 441-9.

Arun K. Ghosh, H.L.O., *BACE1 (b-secretase) inhibitors for the treatment of Alzheimer's disease.* Chem Soc Rev, 2014. **43**: p. 6765--6813.

Martin Citron, et al., *Evidence that the 42- and 40-amino acid forms of amyloid β protein are generated from the β-amyloid precursor protein by different protease activities.* Proc Natl Acad Sci U S A., 1996. **93**: p. 13170-13175.

Knowles TP, Vendruscolo M, and D. CM, *The amyloid state and its association with protein misfolding diseases.* Nat Rev Mol Cell Biol, 2014. **15**(6): p. 384-396.

Cheng, X., et al., *High activities of BACE1 in brains with mild cognitive impairment.* Am J Pathol, 2014. **184**(1): p. 141-7.

Yan, R. and R. Vassar, *Targeting the β secretase BACE1 for Alzheimer's disease therapy.* The Lancet Neurology, 2014. **13**(3): p. 319-329.

Lichtenthaler, S.F., *Alpha-secretase in Alzheimer's disease: molecular identity, regulation and therapeutic potential.* J Neurochem, 2010. **116**(1): p. 10-21.

Bandyopadhyay S, et al., *Role of the APP non-amyloidogenic signaling pathway and targeting alpha-secretase as an alternative drug target for treatment of Alzheimer's disease.* Curr Med Chem, 2007. **14**(27): p. 2848-2864.

Karran, E., M. Mercken, and B. De Strooper, *The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics.* Nat Rev Drug Discov, 2011. **10**(9): p. 698-712.

Geling A, et al., *A gamma-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish.* EMBO Rep, 2002. **3**(7): p. 688-694.

Gonfloni, S., et al., *Oxidative Stress, DNA Damage, and c-Abl Signaling: At the Crossroad in Neurodegenerative Diseases?* Int J Cell Biol, 2012. **2012**: p. 683097.

Chun-Jiang, Y., W. Li, and Z. Zhi-Ren, *BACE1 RNA interference Improves Spatial Memory and Attenuates Abeta Burden in Streptozotocin induced tau hyperphosphorylated rat model.* Int J Neurosci, 2015.

Luo Y, et al., *Mice deficient in BACE1, the Alzheimer's beta-secretase, have normal phenotype and abolished beta-amyloid generation.* Nat Neurosci, 2001 **4**(3): p. 231-232.

Ballatore C, Lee VM, and T. JQ, *Tau-mediated neurodegeneration in Alzheimer's disease and related disorders.* Nat Rev Neurosci, 2007. **8**(9): p. 663-672.

Pooler, A.M., et al., *Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease.* Acta Neuropathol Commun, 2015. **3**: p. 14.

Sydow, A., et al., *Tau-induced defects in synaptic plasticity, learning, and memory are reversible in transgenic mice after switching off the toxic Tau mutant.* J Neurosci, 2011. **31**(7): p. 2511-25.

Schlatterer, S.D., C.M. Acker, and P. Davies, *c-Abl in neurodegenerative disease.* J Mol Neurosci, 2011. **45**(3): p. 445-52.

Tan, J.L., et al., *Mild oxidative stress induces redistribution of BACE1 in non-apoptotic conditions and promotes the amyloidogenic processing of Alzheimer's disease amyloid precursor protein.* PLoS One, 2013. **8**(4): p. e61246.

Schlatterer, S.D., et al., *Neuronal c-Abl overexpression leads to neuronal loss and neuroinflammation in the mouse forebrain.* J Alzheimers Dis, 2011. **25**(1): p. 119-33.

Ringman, J.M., et al., *A potential role of the curry spice curcumin in Alzheimer's disease.* Curr Alzheimer Res, 2005. **2**(2): p. 131-6.

Mishra, S. and K. Palanivelu, *The effect of curcumin (turmeric) on Alzheimer's disease: An overview.* Ann Indian Acad Neurol, 2008. **11**(1): p. 13-9.

Yang, F., et al., *Curcumin inhibits formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid in vivo.* J Biol Chem, 2005. **280**(7): p. 5892-901.

Palmal, S., et al., *Inhibition of amyloid fibril growth and dissolution of amyloid fibrils by curcumin-gold nanoparticles.* Chemistry, 2014. **20**(20): p. 6184-91.

Luhrs, T., et al., *3D structure of Alzheimer's amyloid-beta(1-42) fibrils.* Proc Natl Acad Sci U S A, 2005. **102**(48): p. 17342-7.

Ghosh, A.K., Sandra Gemma, and J. Tang, *β-Secretase as a Therapeutic Target for Alzheimer's Disease.* Neurotherapeutics, 2008. **5**(3): p. 399-408.

Vassar, R. and P.C. Kandalepas, *The beta-secretase enzyme BACE1 as a therapeutic target for Alzheimer's disease.* Alzheimers Res Ther, 2011. **3**(3): p. 20.

Kandalepas, P.C. and R. Vassar, *Identification and biology of beta-secretase.* J Neurochem, 2012. **120 Suppl 1**: p. 55-61.

Venugopal C, et al., *Beta-secretase Structure, Function, and Evolution.* CNS Neurol Disord Drug Targets, 2008. **7**(3): p. 278-294.

Nelson, P.T., et al., *Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature.* J Neuropathol Exp Neurol, 2012. **71**(5): p. 362-81.

Ahmed, R.R., et al., *BACE1 and BACE2 enzymatic activities in Alzheimer's disease.* J Neurochem, 2010. **112**(4): p. 1045-53.

Sun, X., et al., *Distinct transcriptional regulation and function of the human BACE2 and BACE1 genes.* FASEB J, 2005. **19**(7): p. 739-49.

Vassar, R., et al., *The beta-secretase enzyme BACE in health and Alzheimer's disease: regulation, cell biology, function, and therapeutic potential.* J Neurosci, 2009. **29**(41): p. 12787-94.

R. Yan, J.B.M., M. E. Shuck, M. J. Bienkowski, *BACE2 functions as an alternative alpha-secretase in cells.* J. Biol. Chem, 2001. **276**: p. 34019-34027.

Ohno, M., et al., *BACE1 deficiency rescues memory deficits and cholinergic dysfunction in a mouse model of Alzheimer's disease.* Neuron, 2004. **41**(1): p. 27-33.

Ghosh, A.K., M. Brindisi, and J. Tang, *Developing beta-secretase inhibitors for treatment of Alzheimer's disease.* J Neurochem, 2012. **120 Suppl 1**: p. 71-83.

Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.

Zimmermann, M., F. Gardoni, and M. Di Luca, *Molecular rationale for the pharmacological treatment of Alzheimer's disease.* Drugs Aging, 2005. **22 Suppl 1**: p. 27-37.

A. K. Ghosh, R.M., H. Kubinyi, *Aspartic Acid Proteases as Therapeutic Targets*. Aspartic Acid Proteases as Therapeutic Targets, ed. A.K. Ghosh2010: John Wiley & Sons.

Dislich, B. and S.F. Lichtenthaler, *The Membrane-Bound Aspartyl Protease BACE1: Molecular and Functional Properties in Alzheimer's Disease and Beyond.* Front Physiol, 2012. **3**: p. 8.

Luo, X. and R. Yan, *Inhibition of BACE1 for therapeutic use in Alzheimer's disease.* Int J Clin Exp Pathol, 2010. **3**(6): p. 618-28.

Ghosh, A.K., et al., *Memapsin 2 (beta-secretase) inhibitors: drug development.* Curr Alzheimer Res, 2008. **5**(2): p. 121-31.

Guner, O.F., *History and evolution of the pharmacophore concept in computer-aided drug design.* Curr Top Med Chem, 2002. **2**(12): p. 1321-32.

Chen, J. and L. Lai, *Pocket v.2: further developments on receptor-based pharmacophore modeling.* J Chem Inf Model, 2006. **46**(6): p. 2684-91.

Noeris K. Salam, R.N.a.W.S., *Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis.* J. Chem. Inf. Model., 2009. **49**(10): p. 2356-2368.

Wolber, G. and T. Langer, *LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters.* J Chem Inf Model, 2005. **45**(1): p. 160-9.

Chen, Z., et al., *Multiple pharmacophore models combined with molecular docking: a reliable way for efficiently identifying novel PDE4 inhibitors with high structural diversity.* J Chem Inf Model, 2010. **50**(4): p. 615-25.

Xu, Y., et al., *Flexibility of the flap in the active site of BACE1 as revealed by crystal structures and molecular dynamics simulations.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 1): p. 13-25.

Friesner, R.A., et al., *Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.* J Med Chem, 2006. **49**(21): p. 6177-96.

Friesner, R.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* J Med Chem, 2004. **47**(7): p. 1739-49.

Halgren, T.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* J Med Chem, 2004. **47**(7): p. 1750-9.

Eldridge, M.D., et al., *Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes.* J Comput Aided Mol Des, 1997. **11**(5): p. 425-45.

Loving, K., N.K. Salam, and W. Sherman, *Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation.* J Comput Aided Mol Des, 2009. **23**(8): p. 541-54.

Chen Z, et al., *Multiple Pharmacophore Models Combined with Molecular Docking: A Reliable Way for Efficiently Identifying Novel PDE4 Inhibitors with High Structural Diversity.* J. Chem. Inf. Model, 2010. **50**(4): p. 615–625.

Berman, H.M., et al., *The Protein Data Bank.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 6 No 1): p. 899-907.

Dixon, S.L., et al., *PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results.* J Comput Aided Mol Des, 2006. **20**(10-11): p. 647-71.

Loving, K., N.K. Salam, and W. Sherman, *Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation.* J. Comput. Aided Mol. Des, 2009. **23**: p. 541–554.

Friesner, R.A., et al., *Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy.* J. Med. Chem., 2004. **47**: p. 1739–1749.

Friesner, R.A., et al., *Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes.* J. Med. Chem, 2006. **49**(6177-6196).

Halgren, T.A., et al., *Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening.* J. Med. Chem, 2004. **47**: p. 1750–1759.

Truchon, J.F. and C.I. Bayly, *Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem.* J Chem Inf Model, 2007. **47**(2): p. 488-508.

Mosmann, T., *Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays.* J Immunol Methods, 1983. **65**(1-2): p. 55-63.

Kwak, Y.D., et al., *Differential regulation of BACE1 expression by oxidative and nitrosative signals.* Mol Neurodegener, 2011. **6**: p. 17.

Luo Y, et al., *BACE1 (beta-secretase) knockout mice do not acquire compensatory gene expression changes or develop neural lesions over time.* Neurobiol Dis, 2003. **14**(1): p. 81-88.

Roberds, S.L., et al., *BACE knockout mice are healthy despite lacking the primary beta-secretase activity in brain: implications for Alzheimer's disease therapeutics.* Hum Mol Genet, 2001. **10**(12): p. 1317-24.

Kaller, M.R., et al., *A Potent and Orally Efficacious, Hydroxyethylamine-Based Inhibitor of beta-Secretase.* ACS Med Chem Lett, 2012. **3**(11): p. 886-91.

Dineen, T.A., et al., *Design and synthesis of potent, orally efficacious hydroxyethylamine derived beta-site amyloid precursor protein cleaving enzyme (BACE1) inhibitors.* J Med Chem, 2012. **55**(21): p. 9025-44.

Weiss MM, et al., *Design and preparation of a potent series of hydroxyethylamine containing β-secretase inhibitors that demonstrate robust reduction of central β-amyloid.* J Med Chem, 2012 **55**(21): p. 9009-9024.

Rueeger H, L.R., Rogel O, Rondeau JM, Möbitz H, Machauer R, Jacobson L, and D.S. Staufenbiel M, Neumann U., *Discovery of cyclic sulfonehydroxyethylamines as potent and selective β-site APP-cleaving enzyme 1 (BACE1) inhibitors: structure-based design and in vivo reduction of amyloid β-peptides.* Med Chem, 2012. **55**(7): p. 3364-3386.

Cole, D.C., et al., *Acylguanidine inhibitors of beta-secretase: optimization of the pyrrole ring substituents extending into the S1 and S3 substrate binding pockets.* Bioorg Med Chem Lett, 2008. **18**(3): p. 1063-6.

Cole DC, M.E., Stock JR, Condon JS, Jennings LD, Aulabaugh A, Chopra R,, et al., *Acylguanidines as small-molecule beta-secretase inhibitors.* J Med Chem, 2006. **49**(21): p. 6158-6161.

Fobare WF, S.W., Robichaud AJ, Malamas MS, Manas E, Turner J, Hu Y, Wagner E, Chopra R, Cowling R, Jin G, Bard J., *Thiophene substituted acylguanidines as BACE1 inhibitors.* Bioorg Med Chem Lett, 2007. **17**(19): p. 5353-5356.

wahn BM, H.J., Kihlström J, Kolmodin K, Lindström J, Plobeck N, Rotticci D, Sehgelmeble F, Sundström M, Berg Sv, Fälting J, Georgievska B, Gustavsson S, Neelissen J, Ek M,

Olsson LL, Berg S. , *Aminoimidazoles as BACE-1 inhibitors: the challenge to achieve in vivo brain efficacy.* Bioorg Med Chem Lett., 2012 **22**(5): p. 1854-1859.

Malamas MS, E.J., Gunawan I, Turner J, Hu Y, Wagner E, Fan K, Chopra R, and B.J. Olland A, Jacobsen S, Magolda RL, Pangalos M, Robichaud AJ, *Design and synthesis of 5,5'-disubstituted aminohydantoins as potent and selective human beta-secretase (BACE1) inhibitors.* J Med Chem, 2010 **53**(3): p. 1146-1158.

Nowak P, C.D., Aulabaugh A, Bard J, Chopra R, Cowling R, Fan KY, Hu B,, et al., *Discovery and initial optimization of 5,5'-disubstituted aminohydantoins as potent beta-secretase (BACE1) inhibitors.* Bioorg Med Chem Lett, 2010. **20**(2): p. 632-635.

Malamas MS, B.K., Johnson M, Hui Y, Zhou P, Turner J, Hu Y, Wagner E, Fan and C.R. K, Olland A, Bard J, Pangalos M, Reinhart P, Robichaud AJ, *Di-substituted pyridinyl aminohydantoins as potent and highly selective human beta-secretase (BACE1) inhibitors.* Bioorg Med Chem, 2010 **18**(2): p. 630-639.

Malamas MS, E.J., Gunawan I, Barnes K, Hui Y, Johnson M, Robichaud A, Zhou and Y.Y. P, Solvibile W, Turner J, Fan KY, Chopra R, Bard J, Pangalos MN, *New pyrazolyl and thienyl aminohydantoins as potent BACE1 inhibitors: exploring the S2' region.* Bioorg Med Chem Lett, 2011 **21**(18): p. 5164-5170.

Tresadern G, D.F., Delgado O, Gijsen H, Macdonald GJ, Moechars D, and A.R. Rombouts F, Spurlino J, Van Gool M, Vega JA, Trabanco AA, *Rational design and synthesis of aminopiperazinones as β-secretase (BACE) inhibitors.* Bioorg Med Chem Lett, 2011. **21**(24): p. 7255-7560.

Huang H, L.D., Cheng AC, Whittington DA, Patel VF, Chen K, Dineen TA, Epstein, G.R. O, Hickman D, Kiang YH, Louie S, Luo Y, Wahl RC, Wen PH, Wood S,, and F.R. Jr, *Structure- and property-based design of aminooxazoline xanthenes as selective, orally efficacious, and CNS penetrable BACE inhibitors for the treatment of Alzheimer's disease.* J Med Chem, 2012 **55**(21): p. 9156-9169.

Zuo, Z., et al., *Molecular docking and 3D-QSAR studies on the binding mechanism of statine-based peptidomimetics with beta-secretase.* Bioorg Med Chem, 2005. **13**(6): p. 2121-31.

Niu, Y., et al., *The discovery of novel beta-secretase inhibitors: pharmacophore modeling, virtual screening, and docking studies.* Chem Biol Drug Des, 2012. **79**(6): p. 972-80.

John, S., et al., *Potent BACE-1 inhibitor design using pharmacophore modeling, in silico screening and molecular docking studies.* BMC Bioinformatics, 2011. **12 Suppl 1**: p. S28.

Ravichandran Veerasamy, H.R., Abhishek Jain, Shalini Sivadasan, *Validation of QSAR Models - Strategies and Importance.* International Journal of Drug Design and Discovery, 2011. **2**(3): p. 511-519.

Greenwood, J.R., et al., *Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution.* J Comput Aided Mol Des, 2010. **24**(6-7): p. 591-604.

Iserloh, U., et al., *Potent pyrrolidine- and piperidine-based BACE-1 inhibitors.* Bioorg Med Chem Lett, 2008. **18**(1): p. 414-7.

Dixon, S.L., A.M. Smondyrev, and S.N. Rao, *PHASE: a novel approach to pharmacophore modeling and 3D database searching.* Chem Biol Drug Des, 2006. **67**(5): p. 370-2.

Sastry GM, A.M., Day T, Annabhimoju R, Sherman W., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* J Comput Aided Mol Des, 2013. **27**(3): p. 221-234.

Iserloh U, W.Y., Cumming JN, Pan J, Wang LY, Stamford AW, Kennedy ME, Kuvelkar and C.X. R, Parker EM, Strickland C, Voigt J., *Potent pyrrolidine- and piperidine-based BACE-1 inhibitors.* Bioorg Med Chem Lett, 2008 **18**(1): p. 414-417.

Palakurti, R., et al., *Multiple e-Pharmacophore Modeling Combined with High-Throughput Virtual Screening and Docking to Identify Potential Inhibitors of β-Secretase(BACE1).* Mol. Inform., 2013. **32**(4): p. 385-398.

Stierand, K. and M. Rarey, *Drawing the PDB: Protein-Ligand Complexes in Two Dimensions.* ACS Med Chem Lett, 2010. **1**(9): p. 540-5.

K, F. and B. L, *2014 Alzheimer's disease facts and figures.* Alzheimers Dement, 2014. **10**(2): p. e47-92.

Brookmeyer, R., et al., *Forecasting the global burden of Alzheimer's disease.* Alzheimers Dement, 2007. **3**(3): p. 186-91.

Imam, S.Z., et al., *Novel regulation of parkin function through c-Abl-mediated tyrosine phosphorylation: implications for Parkinson's disease.* J Neurosci, 2011. **31**(1): p. 157-63.

Mahul-Mellier, A.L., et al., *c-Abl phosphorylates alpha-synuclein and regulates its degradation: implication for alpha-synuclein clearance and contribution to the pathogenesis of Parkinson's disease.* Hum Mol Genet, 2014. **23**(11): p. 2858-79.

Iqbal, K., et al., *Tau in Alzheimer disease and related tauopathies.* Curr Alzheimer Res, 2010. **7**(8): p. 656-64.

Cohen, A.D., et al., *Anti-Amyloid Effects of Small Molecule Abeta-Binding Agents in PS1/APP Mice.* Lett Drug Des Discov, 2009. **6**(6): p. 437.

Pardridge, W.M., *Alzheimer's disease drug development and the problem of the blood-brain barrier.* Alzheimers Dement, 2009. **5**(5): p. 427-32.

Alavijeh, M.S., et al., *Drug metabolism and pharmacokinetics, the blood-brain barrier, and central nervous system drug discovery.* NeuroRx, 2005. **2**(4): p. 554-71.

Zhang, H.Y., *One-compound-multiple-targets strategy to combat Alzheimer's disease.* FEBS Lett, 2005. **579**(24): p. 5260-4.

Lane, R.F., et al., *Beyond amyloid: the future of therapeutics for Alzheimer's disease.* Adv Pharmacol, 2012. **64**: p. 213-71.

Derkinderen, P., et al., *Tyrosine 394 is phosphorylated in Alzheimer's paired helical filament tau and in fetal tau with c-Abl as the candidate tyrosine kinase.* J Neurosci, 2005. **25**(28): p. 6584-93.

Cancino, G.I., et al., *c-Abl tyrosine kinase modulates tau pathology and Cdk5 phosphorylation in AD transgenic mice.* Neurobiol Aging, 2011. **32**(7): p. 1249-61.

Zhu, Q., et al., *Repurposing of Kinase Inhibitors to Target c-Abl as Potential Therapeutics for Alzheimer's Disease.* Journal of Pharmaceutical Innovation, 2014. **9**(4): p. 331-340.

Karuppagounder, S.S., et al., *The c-Abl inhibitor, nilotinib, protects dopaminergic neurons in a preclinical animal model of Parkinson's disease.* Sci Rep, 2014. **4**: p. 4874.

Shaul, Y. and M. Ben-Yehoyada, *Role of c-Abl in the DNA damage stress response.* Cell Res, 2005. **15**(1): p. 33-5.

Jing, Z., J. Caltagarone, and R. Bowser, *Altered subcellular distribution of c-Abl in Alzheimer's disease.* J Alzheimers Dis, 2009. **17**(2): p. 409-22.

Nagar, B., *c-Abl Tyrosine Kinase and Inhibition by the Cancer Drug Imatinib (Gleevec/STI-571)1–4.* J. Nutr, 2007. **137**: p. 1518-1523.

Senior, K., *Gleevec does not cross blood-brain barrier.* Lancet Oncol, 2003. **4**(4): p. 198.

Okram, B., et al., *A general strategy for creating "inactive-conformation" abl inhibitors.* Chem Biol, 2006. **13**(7): p. 779-86.

Reddy, E.P. and A.K. Aggarwal, *The ins and outs of bcr-abl inhibition.* Genes Cancer, 2012. **3**(5-6): p. 447-54.

Horio, T., et al., *Structural factors contributing to the Abl/Lyn dual inhibitory activity of 3-substituted benzamide derivatives.* Bioorg Med Chem Lett, 2007. **17**(10): p. 2712-7.

Cowan-Jacob, S.W., et al., *Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 1): p. 80-93.

Weisberg, E., et al., *Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl.* Cancer Cell, 2005. **7**(2): p. 129-41.

Salam, N.K., R. Nuti, and W. Sherman, *Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis.* J. Chem. Inf. Model, 2009. **49**: p. 2356–2368.

Watts, K.S., et al., *ConfGen: a conformational search method for efficient generation of bioactive conformers.* J Chem Inf Model, 2010. **50**(4): p. 534-46.

Atwell, S., et al., *A novel mode of Gleevec binding is revealed by the structure of spleen tyrosine kinase.* J Biol Chem, 2004. **279**(53): p. 55827-32.

Seeliger, M.A., et al., *c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty.* Structure, 2007. **15**(3): p. 299-311.

Lin, Y.-L., et al., *Explaining why Gleevec is a specific and potent inhibitor of Abl kinase.* Proc Natl Acad Sci U S A, 2013. **110**(5): p. 1664-1669.

Mucs, D., R.A. Bryce, and P. Bonnet, *Application of shape-based and pharmacophore-based in silico screens for identification of Type II protein kinase inhibitors.* J Comput Aided Mol Des, 2011. **25**(6): p. 569-81.

YZ, X. and C. PY, *ONIOM DFT/PM3 calculation on the interaction between STI-571 and abelson tyrosine kinase.* J Mol Model, 2008. **14**(11): p. 1083-1086.

Anastassiadis, T., et al., *Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity.* Nat Biotechnol, 2011. **29**(11): p. 1039-45.

Dodson CA, et al., *Crystal structure of an Aurora-A mutant that mimics Aurora-B bound to MLN8054: insights into selectivity and drug design.* Biochem J., 2010. **427**(1): p. 19-28.

Ozbuyukkaya, G., E. Ozkirimli Olmez, and K.O. Ulgen, *Discovery of YopE Inhibitors by Pharmacophore-Based Virtual Screening and Docking.* ISRN Bioinformatics, 2013. **2013**: p. 1-12.

Szelag, M., et al., *Identification of STAT1 and STAT3 specific inhibitors using comparative virtual screening and docking validation.* PLoS One, 2015. **10**(2): p. e0116688.

Ross, C.A. and M.A. Poirier, *Protein aggregation and neurodegenerative disease.* Nat Med, 2004. **10 Suppl**: p. S10-7.

Knowles, T.P., M. Vendruscolo, and C.M. Dobson, *The amyloid state and its association with protein misfolding diseases.* Nat Rev Mol Cell Biol, 2014. **15**(6): p. 384-96.

Iram, A. and A. Naeem, *Protein folding, misfolding, aggregation and their implications in human diseases: discovering therapeutic ways to amyloid-associated diseases.* Cell Biochem Biophys, 2014. **70**(1): p. 51-61.

Cardinale, A., R. Chiesa, and M. Sierks, *Protein misfolding and neurodegenerative diseases.* Int J Cell Biol, 2014. **2014**: p. 217371.

Bu, X.L., P.P. Rao, and Y.J. Wang, *Anti-amyloid Aggregation Activity of Natural Compounds: Implications for Alzheimer's Drug Discovery.* Mol Neurobiol, 2015.

Kitagawa, K., et al., *Inhibition of insulin amyloid fibril formation by cyclodextrins.* Amyloid, 2015. **22**(3): p. 181-6.

Ramshini, H., M. mohammad-zadeh, and A. Ebrahim-Habibi, *Inhibition of amyloid fibril formation and cytotoxicity by a chemical analog of Curcumin as a stable inhibitor.* Int J Biol Macromol, 2015. **78**: p. 396-404.

Wiesehan, K., et al., *Inhibition of cytotoxicity and amyloid fibril formation by a D-amino acid peptide that specifically binds to Alzheimer's disease amyloid peptide.* Protein Eng Des Sel, 2008. **21**(4): p. 241-6.

Xie, H., et al., *Inhibition of beta-amyloid peptide self-assembly and cytotoxicity by poly(LVFF-co-beta-amino ester).* J Pept Sci, 2015. **21**(7): p. 608-14.

Shariatizi, S., et al., *Inhibition of amyloid fibrillation and cytotoxicity of lysozyme fibrillation products by polyphenols.* Int J Biol Macromol, 2015. **80**: p. 95-106.

Ho, S.L., et al., *Inhibition of beta-amyloid Aggregation By Albiflorin, Aloeemodin And Neohesperidin And Their Neuroprotective Effect On Primary Hippocampal Cells Against beta-amyloid Induced Toxicity.* Curr Alzheimer Res, 2015. **12**(5): p. 424-33.

Doig, A.J. and P. Derreumaux, *Inhibition of protein aggregation and amyloid formation by small molecules.* Curr Opin Struct Biol, 2015. **30**: p. 50-6.

Wang, H. and D.P. Raleigh, *General amyloid inhibitors? A critical examination of the inhibition of IAPP amyloid formation by inositol stereoisomers.* PLoS One, 2014. **9**(9): p. e104023.

Cheng, B., et al., *Inhibiting toxic aggregation of amyloidogenic proteins: a therapeutic strategy for protein misfolding diseases.* Biochim Biophys Acta, 2013. **1830**(10): p. 4860-71.

Tran, L. and T. Ha-Duong, *Exploring the Alzheimer amyloid-beta peptide conformational ensemble: A review of molecular dynamics approaches.* Peptides, 2015. **69**: p. 86-91.

Kumar, A., et al., *Molecular insight into amyloid oligomer destabilizing mechanism of flavonoid derivative 2-(4' benzyloxyphenyl)-3-hydroxy-chromen-4-one through docking and molecular dynamics simulations.* J Biomol Struct Dyn, 2015: p. 1-12.

Kuang, G., et al., *Investigation of the Binding Profiles of AZD2184 and Thioflavin T with Amyloid-beta(1-42) Fibril by Molecular Docking and Molecular Dynamics Methods.* J Phys Chem B, 2015. **119**(35): p. 11560-7.

Hou, Y., et al., *Evaluation of beta-Amyloid Peptides Fibrillation Induced by Nanomaterials Based on Molecular Dynamics and Surface Plasmon Resonance.* J Nanosci Nanotechnol, 2015. **15**(2): p. 1110-6.

Berhanu, W.M. and A.E. Masunov, *Atomistic mechanism of polyphenol amyloid aggregation inhibitors: molecular dynamics study of Curcumin, Exifone, and Myricetin interaction with the segment of tau peptide oligomer.* J Biomol Struct Dyn, 2015. **33**(7): p. 1399-411.

Baweja, L., et al., *Effect of graphene oxide on the conformational transitions of amyloid beta peptide: A molecular dynamics simulation study.* J Mol Graph Model, 2015. **61**: p. 175-85.

Ye, W., et al., *Molecular dynamics simulations of amyloid fibrils: an in silico approach.* Acta Biochim Biophys Sin (Shanghai), 2013. **45**(6): p. 503-8.

Subramanian S, S.A., *Expression, purification and characterization of a synthetic gene encoding human amyloid beta (Abeta1-42) in Escherichia coli.* Indian J Biochem Biophys, 2007. **44**(2): p. 71-75.

Maya Mathew, B.K.C.S., Sarada Subramanian, I*dentification of small molecule inhibitors against amyloid β (aβ) oligomerization and toxicity from nootropic ayurvedic herbal extracts.* IJPSR, 2013. **4**(12): p. 4685-4691.

Maskevich, A.A., et al., *Spectral properties of thioflavin T in solvents with different dielectric properties and in a fibril-incorporated form.* J Proteome Res, 2007. **6**(4): p. 1392-401.

List of Publications in Peer-reviewed Journals and Conferences

Palakurti, R., Sriram, D., Yogeeswari, P. and Vadrevu, R. (2013), Multiple e-Pharmacophore Modeling Combined with High-Throughput Virtual Screening and Docking to Identify Potential Inhibitors of β-Secretase (BACE1). Mol. Inf., 32: 385–398. doi: 10.1002/minf.201200169.

Palakurti, R and Vadrevu, R, Pharmacophore based 3D-QSAR Modeling, Virtual Screening and Docking for Identification of Potential Inhibitors of β-secretase. (*Under review)*

Palakurti, R and Vadrevu, R, Identification of Potential Inhibitors For Abelson Tyrosine Kinase, for Alzheimer's Disease, Using Multiple E-Pharmacophore Modeling, Virtual Screening and Dynamics. (*Under review)*

Palakurti, R and Vadrevu, R. Small Molecules for Disruption of Amyloids in Alzheimer's disease. (*Manuscript under preparation*)

Palakurti, R and Vadrevu, R. Identification of Small Molecule Disruptors of Aβ1-42 Fibrils in Alzheimer's Disease: Insights into the Mechanism of Destabilization. *14th FAOBMB Congress & 84th Annual meeting of SBC (I)*, 2015.

Abelson Tyrosine Kinase, a new Enzyme target for Alzheimer's disease: Exploring multiple e-pharmoacophore modeling, virtual screening, selectivity assessment for potential inhibitors. **Ravichand Palakurti**, Ramakrishna Vadrevu (2015). **Protein Science,** 23, S1**.**

P. Ravichand, Ramakrishna Vadrevu, P. Yogeswari, D. Sriram (2012). Structure based pharmacophore and virtual screening of b-secretase inhibitiors. International Conference on Advances in Biological Sciences, Kannur University, India.

K. Raja Sekhar Varma, Nabeel Ahmed, **P. Ravichand**, Ramakrishna Vadrevu (2012). Exploring sequence and structural patterns of the b/a loops in TIM Barrel Proteins: Lessons for Loop grafting. International Conference on Advances in Biological Sciences, Kannur University, India.

Biography of Prof. Ramakrishna Vadrevu

Prof. Ramakrishna Vadrevu is currently working as an Associate Professor, Department of Biological Sciences. After obtaining Ph.D. from Indian Institute of Technology, Bombay, he joined as a Post Doc Fellow at Pennsylvania State University and later as a Faculty member in University of Massachusetts Medical School. He joined BITS Pilani-Hyderabad Campus in the year 2008.

His research interests include Protein Design & Engineering: Self-assembly and Bio nano-materials; NMR spectroscopy / Biophysical Approaches to understand protein Folding / mis-Folding and Dynamics in vitro & in vivo and Drug discovery. He is a member of Protein Society and authored more than 20 publications in peer reviewed international journals and chaired in some international conferences. He has served as a reviewer for research proposals submitted to funding agencies, manuscript submitted to journals and theses.

His administrative contributions include serving as the Head, Department of Biology (2012-2014), Chairperson DRC-Departmental Research Committee (2012-2014), General Secretary, Technology Business Incubator Society (2012 onwards).

Biography of Mr. P. Ravichand

Mr. P. Ravichand is a Ph. D student in the Department of Biological Sciences. He obtained M. Sc. in Biotechnology from Bangalore University in the year 2009 and joined GVK biosciences in January 2010, in the interest of learning computational methods for drug discovery. Mr. Ravichand was selected for the Ph.D. program in August 2010 at BITS, Pilani-Hyderabad campus at the Department of Biological Sciences under the Supervision of Dr. Ramakrishna Vadrevu.

He is well trained in experimental lab work and modeling studies. During his work he gained proficiency in protein expression in E.coli, protein purification using affinity and anion exchange chromatography etc. He also gained expertise in growing human cell lines for the cell based assays for his thesis. Currently his research interests focus on Alzheimer's disease therapies.