Applications of Artificial Intelligence in Discovery and Development of Therapeutics for the Treatment of Alzheimer's Disease

Thesis submitted in partial fulfilment for the Award of Degree

Doctor of Philosophy

By

Ravi Singh, M.Pharm

Department of Pharmaceutical Engineering & Technology
Indian Institute of Technology
(Banaras Hindu University)
Varanasi-221005, India

Roll No. 19161001 **Year: 2024**

Department of Pharmaceutical Engineering & Technology
Indian Institute of Technology
(Banaras Hindu University)
Varanasi-221005

CERTIFICATE

It is certified that the work contained in the thesis titled "**Applications of Artificial Intelligence in Discovery and Development of Therapeutics for the Treatment of Alzheimer's Disease**" by **Mr. Ravi Singh** has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and SOTA for the award of Ph.D. Degree.

Dr. Ashok Kumar
(Supervisor)

अशोक कुमार / Ashok Kumar
वरिष्ठ वैज्ञानिक अधिकारी / Senior Scientific Officer
भैषजकीय अभियांत्रिकी एवं प्रौद्योगिकी विभाग /
Department of Pharmaceutical Engineering & Technology
भारतीय प्रौद्योगिकी संस्थान / INDIAN INSTITUTE OF TECHNOLOGY
(बनारस हिन्दू विश्वविद्यालय) / (BANARAS HINDU UNIVERSITY)
वाराणसी-२२१००५ / Varanasi-221005
E-mail:akmaurya.rs.phe@itbhu.ac.in

Date: 06/03/2024
Place: IIT (BHU), Varanasi

Department of Pharmaceutical Engineering & Technology
Indian Institute of Technology
(Banaras Hindu University)
Varanasi-221005

DECLARATION BY THE CANDIDATE

I, **Ravi Singh**, certify that the work embodied in this Ph.D. thesis is my own bona fide work and carried out by me under the supervision of **Dr. Ashok Kumar** from **July, 2019 to October, 2023** at the **Department of Pharmaceutical Engineering & Technology, Indian Institute of Technology (Banaras Hindu University), Varanasi.** The matter embodied in this Ph.D. thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credit to the research workers wherever their works have been cited in my work in this thesis. I further declare that, I have not wilfully copied any other's work, paragraphs, text, data, results, etc. reported in the journals, books, magazines, reports, dissertations, theses, etc., or available at websites and have not included them in this Ph.D. thesis and have not cited as my own work.

Date: 6/3/24
Place: IIT (BHU), Varanasi

Ravi Singh
Ravi Singh

CERTIFICATE BY THE SUPERVISOR(S) AND HEAD OF THE DEPARTMENT

It is certified that the above statement made by the student is correct to the best of our knowledge.

Dr. Ashok Kumar
(Supervisor)

अशोक कुमार/Ashok Kumar
वरिष्ठ वैज्ञानिक अधिकारी/Senior Scientific Officer
भैषजकीय अभियांत्रिकी एवं प्रौद्योगिकी विभाग/
Department of Pharmaceutical Engineering & Technology
भारतीय प्रौद्योगिकी संस्थान/INDIAN INSTITUTE OF TECHNOLOGY
(बनारस हिन्दू विश्वविद्यालय)/(BANARAS HINDU UNIVERSITY)
वाराणसी-२२१००५/Varanasi-221005
E-mail:akmaurya.rs.phe@itbhu.ac.in

S. Hemalatha 6/3/24
Prof. S. Hemalatha
(Head of the Department)

विभागाध्यक्ष/Head
भैषजकीय अभियांत्रिकी एवं प्रौद्योगिकी विभाग/
Department of Pharmaceutical Engineering & Technology
भारतीय प्रौद्योगिकी संस्थान/INDIAN INSTITUTE OF TECHNOLOGY
(बनारस हिन्दू विश्वविद्यालय)/(BANARAS HINDU UNIVERSITY)
वाराणसी-२२१००५/Varanasi-221005

Department of Pharmaceutical Engineering & Technology
Indian Institute of Technology
(Banaras Hindu University)
Varanasi-221005

CERTIFICATE

It is certified that the work contained in the thesis titled "**Applications of Artificial Intelligence in Discovery and Development of Therapeutics for the Treatment of Alzheimer's Disease**" by **Mr. Ravi Singh** has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and SOTA for the award of Ph.D. Degree.

Dr. Ashok Kumar
(Supervisor)

अशोक कुमार/Ashok Kumar
वरिष्ठ वैज्ञानिक अधिकारी/Senior Scientific Officer
भैषजकीय अभियांत्रिकी एवं प्रौद्योगिकी विभाग/
Department of Pharmaceutical Engineering & Technology
भारतीय प्रौद्योगिकी संस्थान/INDIAN INSTITUTE OF TECHNOLOGY
(बनारस हिन्दू विश्वविद्यालय)/(BANARAS HINDU UNIVERSITY)
वाराणसी-२२१००५/Varanasi-221005
E-mail:akmaurya.rs.phe@itbhu.ac.in

Date: 06/03/2024
Place: IIT (BHU), Varanasi

Acknowledgement

I am deeply grateful to Bharat Ratna Pandit Madan Mohan Malviya ji for his unwavering dedication in establishing the esteemed institution, Banaras Hindu University. His tireless efforts have created this sacred space of learning, and I feel privileged to be a part of it.

I extend my sincere gratitude to my supervisor, Dr. Ashok Kumar, for his unwavering support and guidance throughout my Ph.D. research journey. His extensive knowledge, vast research experience, and invaluable suggestions have been instrumental in shaping the trajectory of my work. His insightful feedback challenged me to refine my thinking and elevate the quality of my research. I am truly thankful for the opportunity to be his student, and for the enduring faith he has shown in me over the years. I find it a challenge to express adequately my appreciation for his role in shaping me into not only a better researcher but also a better person.

I am also thankful to Prof. S.K Singh for his constant guidance, support and motivation.

I extend my sincere thanks to Prof. S. Hemalatha, Head of the Department of Pharmaceutical Engineering & Technology, I.I.T. (B.H.U.), Varanasi, for her invaluable support in providing the necessary infrastructure facilities throughout the course of my work.

I would like to extend my sincere thanks to RPEC member Prof. S.K. Singh from the Department of Pharmaceutical Engineering & Technology), and Prof. R.S. Singh from the Department of Chemical Engineering & Technology. Their valuable input and constructive criticism have played a pivotal role in encouraging me to broaden the scope of my research from various perspectives.

I would like to thank Prof. Sairam Krishnamurthy and his lab members for their support in performing in-vivo experiments.

I extend my heartfelt thanks to all the esteemed faculty members of the department, including Prof. B. Mishra, Prof. K. Sairam, Prof. S. K. Srivastava, Prof. M.S. Muthu, Prof. A. Senthil Raja, Dr. Ruchi Chawla, Dr. Vinod Tiwari, Dr. A. N Sahu, Dr. S. K. Mishra, Dr. Prasanta Kumar Nayak, Dr. Gyan Prakash Modi, Dr. S. K. Jain, Dr. Ashish Agarwal, Dr. Deepak, Dr. Dinesh and Dr. Jairam Meena. Their unwavering support and encouragement have been instrumental during the progress of my research, and I am sincerely grateful for their kindness and guidance.

The support and the resources provided by PARAM Shivay Facility, Centre for Computing and Information Services (CCIS) and Central Instrument Facility (CIF), IIT(BHU), Varanasi, are gratefully acknowledged.

I pay sincere thanks to Mr Yashwant Singh, Mr Atul Gupta, Mr Anand and all the departmental non-teaching staff for their co-operation and help during my research work.

I am thankful to the Ministry of Education, GoI, for providing support in the form of a teaching assistantship.

I am immensely grateful to my mentors, Ankit Ganeshpurkar and Qadir Alam, for their invaluable guidance and unwavering support throughout the journey. Their wisdom, encouragement, and belief in my potential have been instrumental in shaping my path.

I want to express my heartfelt thanks to my amazing friends. Your constant encouragement, laughter, and support have played a crucial role in making this work a reality. Having friends like you by my side has made the journey not only memorable but also achievable. Your belief in me and our shared moments of joy have been the fuel that kept me going. I am truly grateful for the friendship we share. Thank you for being an integral part of this journey and for celebrating this achievement with me.

I would like to extend my deepest gratitude to my family. Without their unwavering support, encouragement, and love, achieving this would not have been possible. Their

belief in me has been my source of strength and motivation throughout this journey. I am truly blessed to have such an incredible family by my side. Thank you for being my pillars of support and for making this achievement a reality.

Date: 11th March 2024

Place: IIT(BHU) Varanasi ***Ravi Singh***

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Full Form
3D	Three dimensional
ACh	Acetylcholine
AChE	Acetylcholinesterase
AD	Alzheimer's disease
AI	Artificial intelligence
Aβ	Amyloid-β
APP	Amyloid precursor protein
AS	Anionic site
ATP	Adenosine triphosphate
AUC	Area under the curve
BChE	Butyrylcholinesterase
CaMKII	Ca^{+2}/calmodulin dependent protein kinase II
CAT	Catalase
ChAT	Choline acetyl transferase
ChE	Cholinesterase
CNN	Convolutional neural network
CNS	Central nervous system
CT	Computed tomography
DNP	Donepezil
EAAT2	Excitatory amino acid transporter 2
EC	Enzyme classification
FN	False negative
FP	False positive
GAFF	Generalised amber force field
GSK3	Glycogen synthase kinase-3
iGluRs	ligand-gated ionotropic glutamate receptors
JNK3	c-Jun N-terminal kinase 3
KNN	K-nearest neighbors
LBDD	Ligand based drug design
LDA	Linear discriminant analysis
LGA	Lamarckian Genetic Algorithm
LR	Logistic regression
MACCS	Molecular access system
MD	Molecular dynamics
MEKK	Mitogen-activated Protein/ERK Kinase Kinases
ML	Machine learning
MLP	Multi-layer perceptron
MRI	Magnetic resonance imaging
MSME	Mini-Mental state exam
Nct	Nicastrin
NFT	Neurofibrillary tangle
NMDA	N-methyl D-aspartate
PAC	Passive-aggressive classifier
PAINS	Pan-assay interference compounds
PAM	Positive allosteric modulator

PAMPA	Parallel artificial membrane assay
PAS	Peripheral anionic site
PBL	Porcine brain lipid
PLIP	Protein ligand interaction profiler
PS1	Presenilin 1
PS2	Presenilin 2
QDA	quadratic discriminant analysis
QSAR	Quantitative structure-activity relationship
R_f	Retention factor
RF	Random forest
Rg	Radius of gyration
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RO5	Lipinski rule of five
ROC	Receiver operation characteristic
ROS	Reactive oxygen species
RT	Room Temperature
SAR	Structure activity relationship
SASA	Solvent accessible surface area
SBDD	Structure-based drug design
SBVS	Structure based virtual screening
SCO	Scopolamine hydrobromide
SEM	Standard error of mean
SF	Scoring function
SOD	Superoxide dismutase
SVC	Support vector classifier
SVM	Support vector machine
SVR	Support vector regression
TI	Thermodynamic integration
TN	True negative
TP	True positive

Preface

Alzheimer's Disease (AD), a progressive neurodegenerative disorder characterized by inexorable cognitive decline and memory impairment, represents a significant and growing public health concern. While existing interventions primarily focused on mitigating cholinergic dysfunction offer symptomatic relief, they fail to address the underlying pathophysiological processes. Recognizing the limitations of current approaches, researchers are delving into the intricate molecular pathways implicated in AD pathogenesis, paving the way for a paradigm shift towards multi-target-directed therapeutic strategies.

C**hapter 1** initiates a comprehensive exploration of AD, encompassing its background, pathophysiology, and the current therapeutic landscape. Additionally, the various methodologies, such as Artificial intelligence (AI), that are involved in drug discovery and design are detailed.

Chapter 2: In this chapter, the objectives of the study and plan of work are mentioned.

Chapter 3: This chapter utilizes machine learning to predict potential BACE-1 inhibitors for AD. By analysing molecular properties and applying algorithms like SVM and Random Forest, the chapter identifies promising candidates and key structural features associated with BACE-1 inhibition, contributing to the development of targeted AD treatments.

Chapter 4: This section outlines the development and application of an XGBoost-based machine learning model to screen an in-house database for potential BACE-1 inhibitors, a key target in AD treatment. The model utilizes PubChem fingerprints to identify promising candidates, which are then subjected to experimental validation for their inhibition activity.

Chapter 5: This section explores a novel approach for predicting Blood-Brain-Barrier (BBB) permeability of molecules using Natural Language Processing (NLP) and Deep Learning. It utilizes the B3DB database and extracts features from molecules via SELFIES and N-gram tokenization, converting them into numerical vectors. These features are then fed into various Deep Learning models like ANN and LSTM to predict BBB permeability. The best-performing model achieved high accuracy, suggesting its potential for early screening of drugs targeting the central nervous system.

Chapter 6: This section describes developing and evaluating multi-target directed ligands (MTDLs) based on *N-benzylpiperidines* for AD treatment. Utilizing suitable approaches, promising *N-benzylpiperidine* scaffolds are identified. Synthesis of derivatives with predicted enhanced multi-target activity against BACE-1, cholinesterases, and amyloid aggregation is followed by *in-vitro* and, for promising candidates, *in-vivo* efficacy evaluation in AD animal models.

Chapter 7: This chapter outlines the summary and conclusions of the research work undertaken.

Chapter 8: The references used to carry out the research work are presented in the chapter.

An appendix of additional supporting information, spectral data of representative compounds, and a list of publications from the course of the Ph.D. are included.

Chapter-1
Introduction

1 Introduction

1.1 Neurodegenerative disorders

Neurodegenerative disorders impact nearly 30 million individuals, causing disability, cognitive impairment, and even death. These conditions are characterized by specific pathological changes in the brain and the degeneration of particular groups of neurons [1]. Among the most prevalent are Alzheimer's Disease (AD) and Parkinson's Disease (PD), which gradually lead to the dysfunction of brain neurons or the peripheral nervous system, resulting in cognitive disturbances. Additional neurodegenerative disorders include amyotrophic lateral sclerosis (ALS), Huntington's disease, frontotemporal dementia, and spinocerebellar ataxia. These disorders disrupt communication between nerve cells, affecting various functions such as locomotor abilities, speech, muscle coordination, cognitive functions, intelligence, behavior, and many more [2].

These conditions pose a significant threat to human health, and the risk of developing a neurodegenerative disease increases considerably with age. Given the current scenario, there is an urgent need to deepen our understanding of the exact causes of neurodegenerative disorders and explore novel methodologies for their treatment and prevention [3].

1.2 Alzheimer’s Disease

AD is a devastating neurodegenerative disorder that is characterized by progressive cognitive decline and is the leading cause of dementia and mortality. The prevalence of AD is increasing, with the rising average life-span resulting in a dramatic increase in age-related disorders such as dementia [4]. In addition to its impact on patients, AD also has profound effects on their families and society as a whole [5]. With increasing life expectancy and an ageing population, AD has become a growing socio-economic and medical problem in modern society [6]. This debilitating condition leads to a gradual loss

of memory, thinking skills, and the ability to carry out daily activities. It is estimated that around 50 million people worldwide currently suffer from AD and this number is projected to triple by 2050, if effective preventive and therapeutic strategies are not developed [7]. In today's rapidly changing world of healthcare and medical advancements, AD stands out as a significant challenge that requires urgent attention and research efforts. In recent years, there has been a growing understanding of the pathophysiology of AD and the underlying mechanisms that contribute to its progression. As researchers continue to uncover the complexities of this disease, there is a pressing need for accurate diagnosis, effective treatment options, and support for those affected by this condition.

1.3 Symptoms of Alzheimer's Disease

The symptoms of AD typically begin with mild memory loss and progress to more severe cognitive impairments over time. These cognitive impairments include difficulties with language, problem-solving, and decision-making. The neurodegenerative nature of AD also leads to behavioral and psychological changes, such as mood swings, agitation, and irritability. Furthermore, as the disease advances, individuals with Alzheimer's may experience disorientation, confusion about events, dates, and places, as well as suspicions about family, friends, and caregivers. As the disease progresses, individuals may also have difficulty speaking, writing, and walking.

1.4 Stages of Alzheimer's Disease

The gradual changes in the brain and behavior of AD patients have been summarized in a figure (**Figure 1.1**).

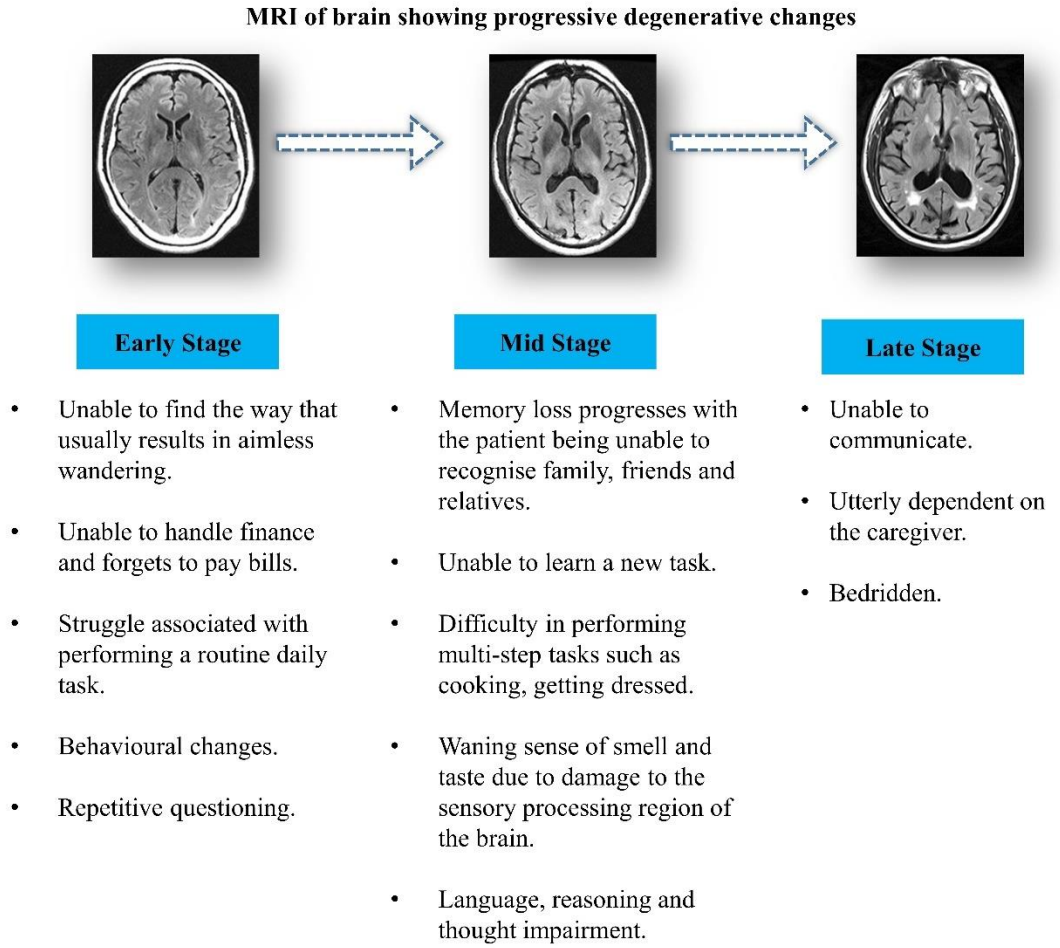


Figure 1.1 Stages of Alzheimer's disease

1.5 Diagnosis of Alzheimer's disease

Diagnosis of AD is a complex and challenging task due to various factors. These factors include the intricate pathogenesis of the disease, the high misdiagnosis rate, and the difficulty in identifying early-stage disease. AD features complex pathogenesis involving multiple pathophysiological processes, making diagnosing difficult [8]. In addition, clinical presentations of AD often overlap with other types of dementia, making it challenging to differentiate between them [9]. Moreover, the co-occurrence of multiple pathologies further complicates the diagnosis process. Furthermore, the availability of increasing biomarkers has opened up the possibility of diagnosing the early stages of AD before the onset of dementia [9].

The Alzheimer's Association provides a set of guidelines for the diagnosis of AD. The diagnostic process includes a physical examination and a neurological assessment to evaluate body reflexes, muscle coordination, strength, sensation, and eyeball movement. Additionally, mental examinations like the Mini-Cog test and the mini-mental state exam (MMSE) are conducted. The Mini-Cog test involves two tasks: firstly, the patient is asked to remember the names of three objects, which they are later prompted to repeat after a certain time. Secondly, the patient is asked to draw a clock with all 12 numbers and indicate a specified time [10].

The examination results may indicate the need for further investigation. The MMSE, a questionnaire-based assessment administered by a healthcare practitioner, evaluates various mental skills and has a maximum score of 30 points. Scores between 20 and 24 indicate mild dementia, 13 to 20 indicate moderate dementia and fewer than 12 indicate severe dementia. People with AD tend to experience an average decrease in their MMSE score by two to four points every year [11].

To diagnose other symptoms present in AD, depression and mood assessment tests are also conducted. Brain imaging using magnetic resonance imaging (MRI) or computed tomography (CT) is recommended to investigate the underlying cause of dementia. In some cases, dementia may be caused by factors other than AD, such as tumors, strokes, or trauma. For AD pathology diagnosis, specific PET ligands like Florbetapir, Florbetaben, and Flutemetamol are utilized [12-14]. Other tests involve estimating $A\beta_{1-42}$, hyperphosphorylated tau (p-tau), and total tau protein levels in cerebrospinal fluid [15].

1.6 Pathophysiology of Alzheimer's disease

In 1907, Dr. Alois Alzheimer published the results of an autopsy conducted on Auguste Deter, a 55-year-old woman who succumbed to a degenerative behavioral and cognitive disorder. The examination of Deter's brain revealed two distinct features: plaque composed of amyloid-β and neurofibrillary tangles (NFT) containing hyperphosphorylated tau protein. Alongside these classical hallmarks, AD involves intricate pathophysiological factors that are not fully understood. AD is associated with a multitude of pathways and targets, adding to the complexity of the condition [16].

Numerous hypotheses have been proposed to elucidate the multifactorial nature of this disorder, including the cholinergic hypothesis, Aβ hypothesis, tau hypothesis, and inflammation hypothesis. However, recent research has demonstrated that the widely accepted Aβ hypothesis, prevailing for the past two decades, fails to fully explain the intricate pathophysiology of this debilitating disease. Recent studies have shed light on the significance of Aβ oligomers in causing synaptic impairment, suggesting that they play a central role in disrupting brain function, among various other signals. Additionally, the formation of amyloid plaques, which typically occur later in life, appears to be a relatively delayed event in the progression of the disease [17].

According to the amyloid cascade hypothesis, the normal cleavage of APP involves α-secretase, while β- and γ-secretases lead to abnormal processing, resulting in an imbalance between the production and clearance of Aβ peptide [18]. This leads to the spontaneous aggregation of Aβ peptides into soluble oligomers, which then coalesce to form insoluble beta-sheet fibrils, eventually depositing as diffuse senile plaques [8]. Recent studies have highlighted the cooperative involvement of both neurons and associated astrocytes in producing $A\beta_{42}$ oligomers (**Figure 1.2**) [9]. These oligomers have been shown to induce oxidative damage, promote tau hyperphosphorylation, and

adversely affect synapses and mitochondria [6,7]. Additionally, $A\beta_{1-42}$ senile plaques are observed during later stages, attracting microglia and leading to the production and release of proinflammatory cytokines like IL-1β, TNF-α, and IFN-γ. This, in turn, stimulates nearby astrocytes and neurons to produce more $A\beta_{1-42}$ oligomers, further activating $A\beta_{1-42}$ production and dispersion [9]. Aβ oligomers are also responsible for the destruction of oligodendroglia (OLGs), resulting in oxidative stress due to their low reduced glutathione (GSH) content and high iron concentration, affecting their ability to scavenge oxygen radicals. $A\beta_{1-42}$ oligomers have an increased capability to damage cholesterol-rich membranes found in OLGs and myelin [19].

Previous studies have explored the pharmacology of Aβ receptors, with $A\beta_{1-42}$ monomers activating the neuroprotective signaling of insulin-like growth factor-1 receptor (IGF-1R). In contrast, $A\beta_{1-42}$ oligomers interact with a range of neurons' and astrocytes' membrane receptors, including RAGE, Frizzled receptor, insulin receptor, NMDA glutamate receptor, p75 neurotrophin receptor (p75NTR), a7 nicotinic ACh receptor (a7nAChR), ApoE receptors, formyl peptide receptor-like 1 (FPRL1/2), cellular prion protein (PrPc), and the calcium-sensing receptor (CaSR) [20].

The removal of Aβ oligomers from the brain involves various pathways, including proteolytic degradation by neprilysin and insulin-degrading enzyme (IDE), uptake by astrocytes and microglia, passive flow into the cerebrospinal fluid, and sequestration into the vascular compartment through the soluble form of the low-density lipoprotein receptor-related protein 1 (LRP1) [21]. Increased levels of NO observed in AD can decrease IDE enzymatic function, potentially leading to an increase in Aβ oligomer deposition in the brain and the development of AD [22]. Recent findings suggest a "contagion" like diffusion of $A\beta_{1-42}$ oligomers and hyperphosphorylated tau oligomers

via exocytosis or exosomes to closely associated target cells such as astrocytes and oligodendrocytes, turning them into producer cells of Aβ and tau oligomers [23].

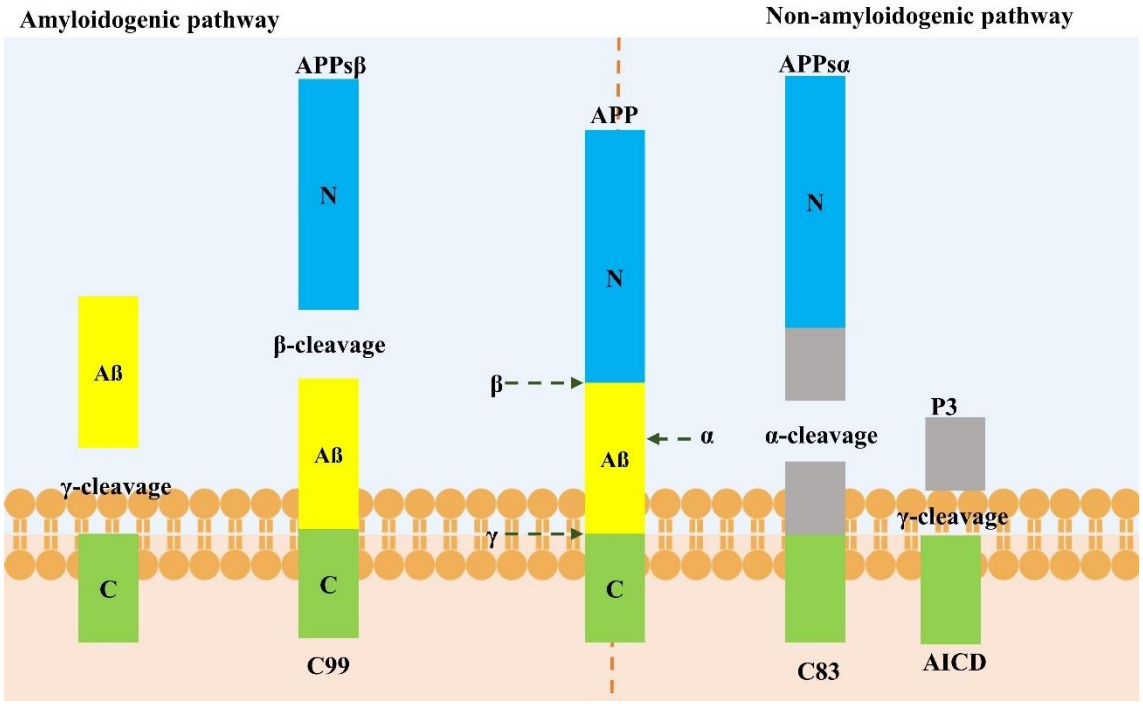


Figure 1.2 Metabolism of APP by secretases

Astrocytes have the ability to promote or reduce neurotransmitter release into enveloped synapses through the Ca^{2+} waves they respectively emit or absorb [12]. When neurons produce $A\beta_{1-42}$ beyond safe limits, toxic $A\beta_{1-42}$ oligomers overflow from neurons onto surrounding astrocytes, both cell types being equipped with $A\beta_{1-42}$ oligomer-binding receptors, accumulating or dispersing in the extracellular environment [9]. Due to the close physical and functional interconnections within the neurons' client group, the $A\beta_{1-42}$ oligomers released by neurons can directly bind to the a7nAChRs of partner astrocytes [23]. Signals from these receptors prompt the astrocytes to release the glutamate they have taken up from neuronal synapses. The discharged glutamate activates extrasynaptic NMDARs of the astrocytes' partner neurons, leading to Ca^{2+} surges, causing a cascade of events, including dysfunctional mitochondria emitting ROS, oxidative damage, caspase-3 activation, tau hyperphosphorylation, increased production of NO, ROS, and VEG-F,

resulting in the destruction of dendritic spines, neuronal synapses, and disruptions in astrocyte-neuron communication [24]. Studies have shown that a CaSR selective allosteric antagonist (calcilytic) NPS 2143 can specifically prevent the excess release of endogenous $A\beta_{42}$ from $A\beta_{25-35}$-exposed human astrocytes and neurons [25].

1.7 Management of Alzheimer's Disease

The only currently available treatments are symptomatic and aim to balance the disease's neurotransmitter imbalance. The use of three cholinesterase inhibitors (CIs) for the treatment of mild to moderate AD has been approved, i.e., Donepezil, Rivastigmine and Galantamine. Memantine, an NMDAR antagonist, is another treatment option for moderate-to-severe AD (**Figure 1.3**) [26, 27]. Additionally, amyloid-directed antibodies (eg, Aducanumab, Lecanemab) have been approved [28, 29]. More information about the FDA-approved drugs is summarized in **Table 1.1**. The behavioral signs of the illness are treated with antipsychotics and antidepressants at the same time.

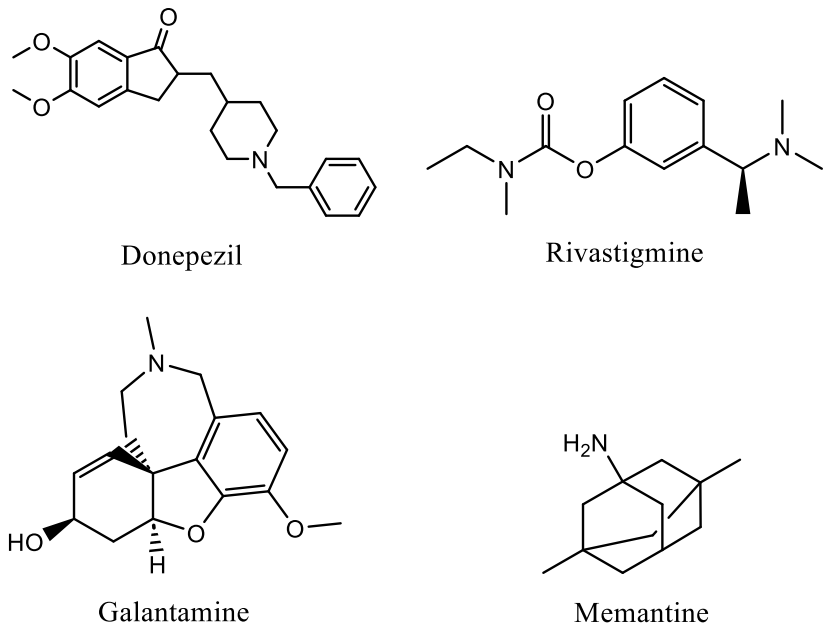


Figure 1.3 Chemical structures of FDA approved drug for the treatment of AD

Table 1.1 FDA-approved medications to treat AD

Drug	Mechanism of Action	Side effects	Administration
Donepezil	Cholinesterase inhibitor	Nausea, vomiting, diarrhea, insomnia, muscle cramps, fatigue, and weight loss.	Delivered orally once a day through a tablet that is either swallowed or dissolved in the mouth.
Rivastigmine	Cholinesterase inhibitor	Nausea, vomiting, diarrhea, weight loss, indigestion, decreased appetite, anorexia, and muscle weakness.	Delivered orally through a capsule twice a day or through a skin patch that is replaced once a day.
Galantamine	Cholinesterase inhibitor	Nausea, vomiting, diarrhea, decreased appetite, weight loss, dizziness, and headache.	Delivered orally through an extended-release capsule, tablet, or liquid. Extended-release capsule is taken once a day. Tablet and oral solution are each taken twice a day.
Memantine	NMDA antagonist	Dizziness, headache, diarrhea, constipation, and confusion.	Delivered orally through an extended-release capsule, tablet, or liquid. Extended-release capsule is taken once a day. Tablet and oral solutions are each taken once a day.
Memantine and Donepezil (manufactured combination)	Cholinesterase inhibitor and NMDA antagonist	Headache, nausea, vomiting, diarrhea, dizziness, anorexia, and ecchymosis (small bruising from leaking blood vessels).	Delivered orally through an extended-release capsule once a day.
Aducanumab	Immunotherapy-Removes abnormal beta-amyloid to help reduce the number of plaques in the brain.	Possible side effects include ARIA, headache, dizziness, falls, diarrhea and confusion.	Possible side effects include ARIA, headache, dizziness, falls, diarrhea, and confusion. Delivered through IV over one hour every four weeks.
Lecanemab	Immunotherapy-Removes abnormal beta-amyloid to help reduce the number of plaques in the brain.	Possible side effects include ARIA, headache, dizziness, falls, diarrhea, and confusion.	Delivered through IV over one hour every two weeks.

1.8 Artificial intelligence and Machine Learning in drug discovery

Machine learning (ML), a subfield of artificial intelligence (AI), studies the methods and mechanics of enabling machines to expertly carry out intelligent activities without being explicitly programmed for those tasks. In some recent jobs, including playing video games and picture identification, AI systems have come close to or surpassed human performance; nevertheless, these tasks have traditionally been extremely specific and focused. However, AI in its different forms is now successfully used in a wide range of fields and for difficult tasks, including robotics, voice translation, image analysis, and logistics, in addition to its continuous usage in molecular design [30].

Since the 1960s, medicinal chemistry has used AI to create molecules in a variety of ways and to variable degrees of success. Labeled training datasets are frequently used in supervised learning, which trains models. A common method for predicting characteristics, such as logP, solubility, and bioactivity for specific chemical compounds, is the quantitative structure-activity relationship (QSAR) approach. Contrarily, unsupervised learning, which does not rely on labels, is also well-liked in medicinal chemistry, where examples like principal components analysis, hierarchical clustering, and algorithms are frequently used to analyze and separate large molecular libraries into smaller collections of related compounds [31].

Since the development of atomistic theory, chemists have worked to predict the properties of compounds without synthesizing them. Although it is still difficult to define, Alexander Crum Brown argued in 1869 that a compound's physiological response is simply a result of its chemical makeup. After Hansch and Fujita's initial proposal of QSARs and their relationships in 1962, this field of study has continued to be active [31].

Figure 1.4 depicts a roadmap for discovering new drugs using AI. It starts with gathering data from various sources like databases, publications, and lab experiments. This data is

then processed and organized to feed into machine learning models. These models predict the properties of potential drugs, leading to the identification of promising candidates. The journey continues with optimizing these candidates, predicting their behavior in the body, and conducting virtual and laboratory tests. If successful, animal studies and, eventually, human clinical trials follow. Finally, after thorough evaluation, approved drugs reach patients. While still in its early stages, AI holds immense potential to revolutionize drug discovery by making it faster, more efficient, and ultimately, more successful in delivering life-saving solutions.

In particular, ClogP, which calculates the octanol/water partition coefficient, is a notable example of how the work on QSAR has advanced the practice of specific physicochemical property predictions [11]. Since the initial introduction of QSAR more than 50 years ago, there have been considerable increases in the number of modeling methodologies, molecular representations, and the volume of data and computing resources available. These datasets can now be used with techniques like deep learning that were previously inappropriate or unavailable due to the advancements in all of these domains. Predictive models can be created using the vast amounts of chemical structure data that are currently available, along with relevant, measurable endpoints. However, the availability of this data is still restricted, and even when it is, the quality is inconsistent. In this situation, it is anticipated that more advanced ML techniques will be able to handle this noisy data.

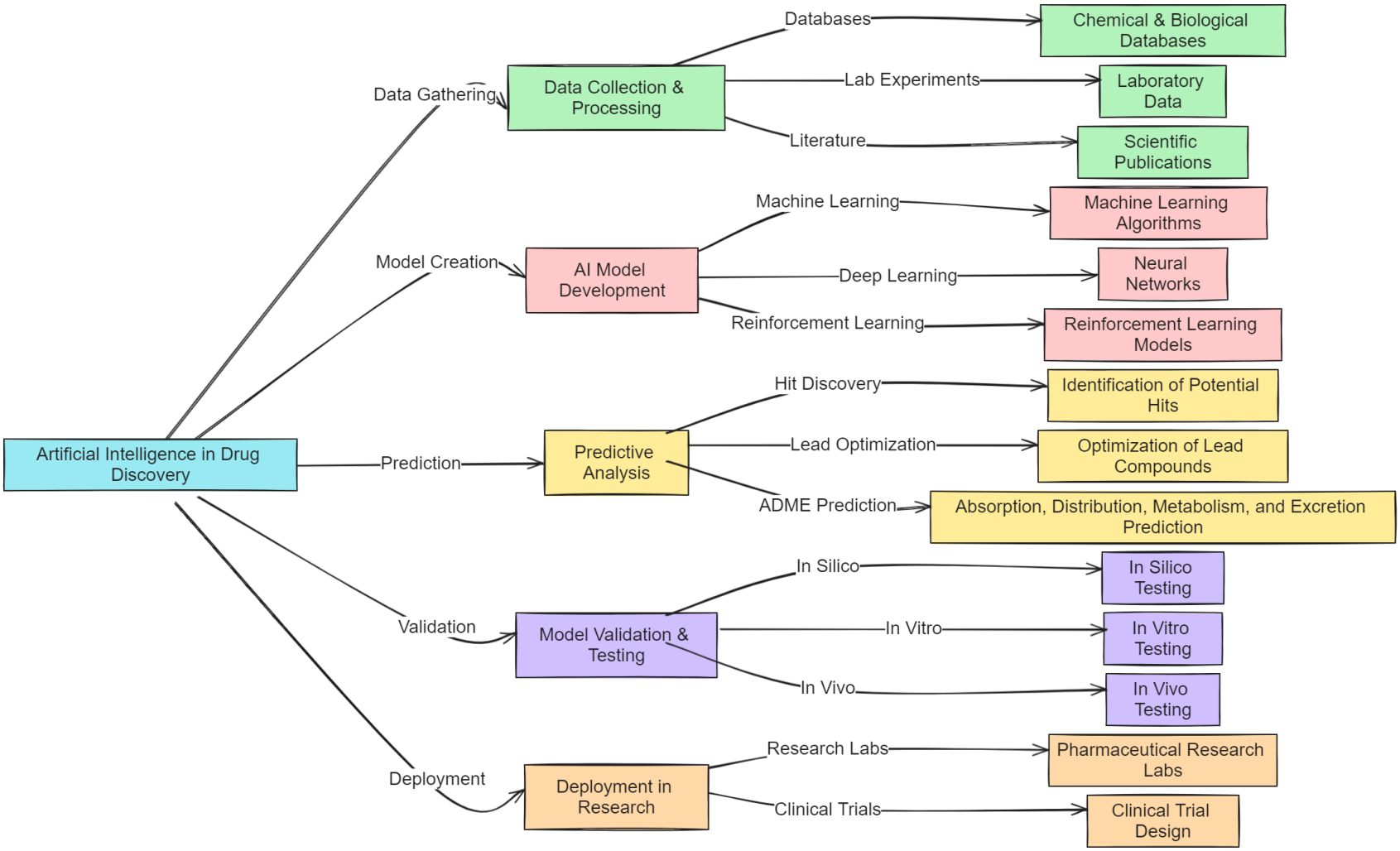


Figure 1.4 Flowchart summarizing the applications of AI in drug discovery

The Merck molecular activity challenge led to one of the earliest uses of deep learning for chemical property prediction, using multitask neural networks to predict many end points at once in addition to just one [32]. A very important field of research right now is deep learning in chemical property prediction [33].

Chapter 2

Objective and Plan of Work

2 Objective and Plan of Work

2.1 Objective

AD is a neurodegenerative disease which progresses into loss of the memory, cognition and thinking ability of a patient. The effects of disease lead to difficulty in performing activities of daily living and make the patient dependent on the caregiver. AI is rapidly transforming drug discovery, bringing speed and efficiency to the traditionally slow and expensive process. Here's how:

- Mining data mountains: AI excels at analyzing vast datasets of molecular structures, genetic information, and clinical trials, uncovering hidden patterns, and predicting drug properties [34].
- Designing new molecules: AI algorithms can design new drug candidates with desired properties, even suggesting novel scaffolds not found in nature [31].
- Repurposing existing drugs: AI can identify new uses for existing drugs, accelerating development and reducing costs [35].
- Optimizing clinical trials: AI can analyze patient data to design more efficient and targeted clinical trials, leading to faster drug approvals [36].

In the present research, efforts have been made to employ AI to find drugs for the treatment of AD.

2.2 Plan of work

2.2.1 Classification of BACE-1 inhibitors using machine learning

Plan of Work:

- Utilize existing datasets of BACE-1 inhibitors and their activity data.
- Extract various molecular descriptors such as Mordred descriptors and fingerprints.

- Apply and compare different machine learning algorithms, including Naïve Bayes, kNN, SVM, Random Forest, and XGBoost.
- Evaluate model performance using metrics like accuracy, precision, recall, and F1 score.
- Analyze the structural features of active and inactive compounds based on predicted classifications.

2.2.2 Machine learning-based screening of in-house database for BACE-1 inhibitors

Plan of Work:

- Build a robust ML model based on a suitable dataset of known BACE-1 inhibitors and their properties.
- Optimize the model for accurate prediction of inhibitor activity.
- Screen the in-house database using the developed model to identify promising candidates.
- Experimentally validate the activity of shortlisted compounds through assays like BACE-1 inhibitor activity evaluation.
- Perform docking studies to understand the binding interactions of identified hits.

2.2.3 Natural Language Processing (NLP) for deep learning-based prediction of BBB permeability

Plan of Work:

- Utilize public databases like B3DB containing molecules and their BBB permeability data.
- Generate SELFIES strings to represent molecular structures.

- Apply NLP techniques like word-level and N-gram tokenization to convert SELFIES into numerical vectors.
- Train and evaluate various deep learning models like ANN and LSTM using the extracted features.
- Optimize the model for accurate prediction of BBB permeability.

2.2.4 Design, synthesis and biological evaluation of *N-benzylpiperidines* as potential multitargeted ligands for the treatment of Alzheimer's disease

Plan of Work:

- Perform structure-based drug design (SBDD) to identify promising N-benzylpiperidine scaffolds.
- Synthesize N-benzylpiperidine derivatives with structural modifications predicted to enhance multi-target activity.
- Evaluate the *in-vitro* inhibitory activity of synthesized compounds against BACE-1, cholinesterases, and amyloid aggregation assays.
- Assess the *in-vivo* efficacy of promising candidates in relevant AD animal models.

Chapter 3
Classification of BACE-1 inhibitors by using machine learning methods

Summary:

The beta-site amyloid precursor protein cleaving enzyme 1 (BACE-1) is a transmembrane aspartyl-protease, that cleaves amyloid precursor protein (APP) at the β-site. The sequential proteolytic cleavage of APP, first by β-secretase and then by γ-secretase complex, leads to the production and release of amyloid-β peptide, a pathological hallmark of AD. BACE-1 inhibitors are reported to possess considerable potential in decreasing the level of amyloid-β in the brain and preventing the progression of AD. A classification study has been conducted on 3536 diverse BACE-1 inhibitors, obtained from the Binding DB database, by extracting two types of descriptors, that is, molecular property (Mordred) and fingerprints (Pubchem, MACCS and KRFP). Furthermore, based on the descriptors, various machine learning algorithms such as Naïve Bayesian (NB), nearest known neighbors (kNN), support vector machine (SVM), random forest (RF) and gradient-boosted algorithms (XGB) were applied to develop classification models. The performance of the models was evaluated using accuracy, precision, recall, and the F1 score of the test set. The best NB, kNN, SVM, RF and XGB classifiers had F1 scores of 0.74, 0.85, 0.86, 0.87 and 0.87, respectively. The diverse 3536 BACE-1 inhibitors were clustered into 11 subsets, and the structural features of each subset were evaluated. The important fragments present in active and inactive compounds were also identified. The model developed in the study would serve as a valuable tool for the design of BACE-1 inhibitors and the virtual screening of molecules to identify these.

3 Classification of BCAE-1 inhibitors using machine learning methods

3.1 Introduction

The β-site amyloid precursor protein cleaving enzyme 1 (BACE-1), also known as β-secretase, is mainly expressed in the neurons of the brain. It is a transmembrane aspartyl-protease responsible for producing toxic Aβ from amyloid precursor protein (APP), which is a pathological hallmark in AD [37]. The cascade proteolytic cleavage of APP is led by BACE-1 and followed by γ-secretase, which leads to the production and release of amyloid-β peptide [38]. The initial cleavage of APP by BACE-1 yields a membrane-bound C-terminal fragment, C-99, which is also a rate-limiting step in the formation of Aβ peptide. This membrane-bound fragment is further cleaved by γ-secretase to produce 39- to 43-amino acid Aβ peptide fragments [39].

BACE-1, a 501 amino acid long protease, was first cloned in 1999 by five research groups simultaneously [40]. Its gene is located on chromosome 1. BACE-1 predominantly cleaves the β site of APP at M596-D597, and also at Y606-E607, but to a lesser extent. It operates mainly in an acidic environment. The BACE-1 possesses hallmark features of eukaryotic aspartic proteases of the pepsin family with two active site motifs, i.e., asp-thr-gly-ser (residues 93-96) and asp-ser-gly-thr (residues 289-292) [41]. The fact that the protease accounted for most of the β-secretase activity in the brain, was supported by subsequent demonstrations in BACE-1 knockout mice that did not produce Aβ peptides in the brain [42].

The X-ray co-crystal structure of BACE-1 with peptidomimetic inhibitors showed important interactions and was a key step in the further development of BACE-1 inhibitors [43]. The R289 and hydrophobic pockets formed by the active site played an important role in substrate binding [44]. Efforts are being made to develop compounds that specifically inhibit BACE-1, in view of its strong *in-vivo* and *in-vitro* validation as a

major β-secretase enzyme in the brain. Many drug candidates reached the clinical trials but failed mainly due to safety concerns **(Table 3.1)**. The other factor of failure may be the very late initiation of the treatment to make any impact on Aβ production [45]. Further, the size of BACE-1 active site (twenty-eight amino acids) is also relatively large and having small molecules to occupy such an active site is a challenge [46].

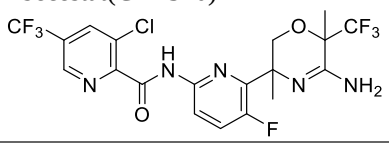
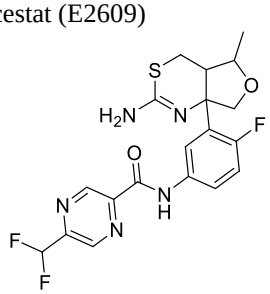
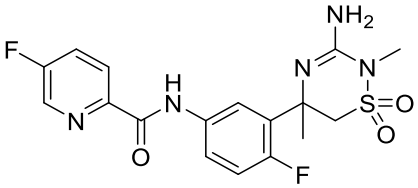
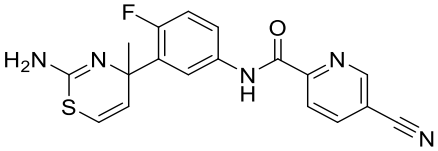
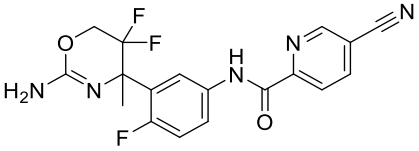
The blood-brain barrier permeability is another major challenge. Many of the BACE-1 inhibitors developed were prone to efflux by P-gp protein. Therefore, the process of drug entry into the brain is complex, even the drug has successfully crossed BBB [47]. Despite the challenges, various research groups have been able to design, develop, and synthesize selective, potent, and bioavailable inhibitors.

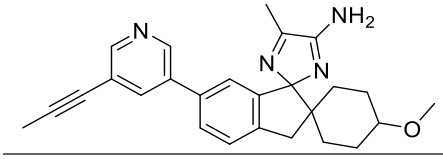
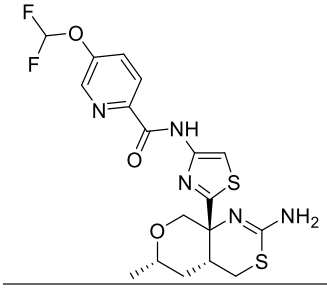
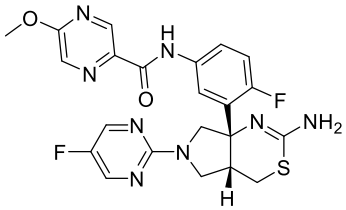
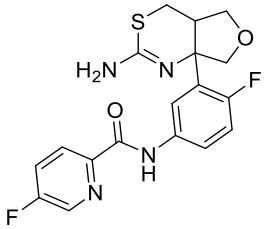
Several efforts were made to develop predictive models to classify BACE-1 inhibitors. Ponzoni *et al*. used a chemically diverse dataset with 215 molecules and calculated chemical descriptors to classify BACE-1 inhibitors. The best result was obtained with random forest algorithm (85% accuracy) [48]. Huang *et al.*, performed 3D-QSAR of BACE-1 inhibitors based on topomer CoMFA on a dataset of 125 compounds. The correlation coefficient of the fitting model was 0.966 [49]. Subramaniam *et al.* published a QSAR model with molecular descriptors that encompass molecular fingerprint, 1-dimensional, 2-dimensional, constitutional, physicochemical descriptors, topological descriptors and 3-dimensional molecular fields. The group performed both qualitative classification and quantitative regression involving linear, nonlinear and deep neural network methods. The best accuracy (81% accuracy) was observed for the model created on deep neural network using Constitutional, physicochemical and topological descriptors [50]. In another study by Kumar *et al*., a PLS-regression-based 2D-QSAR model, from 98 diverse compounds having defined BACE-1 enzyme inhibitory activity, was used to explore the essential structural requirements or molecular properties for

enzyme inhibition. The correlation coefficient of best model was found to be 0.826. In the same study, the authors performed a qualitative study using 3D QSAR pharmacophore model. The test set accuracy was found to be 73.43% [51].

In the present study, we used machine learning classification algorithms to build models on a very large dataset to identify active BACE-1 inhibitors. These models would serve as practical screening tools for designing and virtual screening of molecules to identify BACE-1 inhibitors.

Table 3.1 Clinical drug candidates as BACE-1 inhibitors

Name	**Sponsor**	**Phase**	**Clinical Trial accession number**	**Status**
Umibecestat(CNP520) 	Amgen, Novartis	Phase 2/3	NCT03131453	Discontinued in July, 2019
Elenbecestat (E2609) 	Biogen, Eisai	Phase 3	NCT03036280	Discontinued in September, 2019
Verubecestat (MK-8931) 	Merck	Phase 3	NCT01953601	Discontinued in April, 2018
Atabecestat (JNJ-54861911) 	Janssen, Shionogi Pharma	Phase 3	NCT02569398	Discontinued in May, 2018
RG7129 (RO5508887) 	Roche	Phase 1	NCT01664143	Discontinued in October 2013

Lanabecestat (AZD3293) 	AstraZeneca Eli Lilly & Co.	Phase 3	NCT02783573	Discontinued in June, 2018
PF-06751979 	Pfizer	Phase 1	NCT03126721	Completed in April, 2017
LY3202626 	Eli Lilly & Co.	Phase 2	NCT02791191	Discontinued in October 2018
LY2886721 	Eli Lilly & Co.	Phase 2	NCT01561430	Discontinued in June, 2013
BI 1181181 (VTP 37948)	Boehringer Ingelheim, Vitae Pharmaceuticals	Phase 1	NCT02044406	Completed in September, 2014. Discontinued in 2015.

3.2 Computational Method

3.2.1 Deriving experimental data for BACE-1 inhibitors

Molecular structures of BACE-1 inhibitors and their IC_{50} values were obtained from Binding DB by using the following criteria: (i) data of human BACE-1 inhibition assay (ii) data of IC_{50} assay, and (iii) filtration of compounds with multiple entries or without assay data [52]. The process resulted in the identification of a total of 4960 compounds, and their IC_{50} values ranged from 0.011 to 50,000 nM. The compounds having activity within the range of 50-500 nM, based on the differences in *in-vitro* studies, were further

filtered to avoid any bias (N=1424). The compounds having IC_{50} values less than 50 nM were marked as active (designated as 1), and those with IC_{50} values more than 500 nM were marked as inactive (designated as 0). This resulted in 1527 active and 2009 inactive compounds.

3.2.2 Ligand pre-treatment and preparation

The obtained dataset was further checked for any duplicity. The data from Binding DB was imported into the RDkit Python module. The molecules were standardized using the SaitizeMol module of RDkit. The standardization process involved removals of salts, mixtures, metal ions and correction of the geometry of the chemical structure. The correct protonation states were assigned at pH 7.4 using Open Babel.

3.2.3 Molecular descriptor calculation and feature selection

Mordred 1.2 was used to calculate molecular descriptors for all the ligands. It calculated a total 1826 descriptors, out of which 213 were 3-dimensional. In the present study, we used only 2D descriptors for the model development to avoid any complexity from 3D geometry considerations. The calculated descriptors were normalized using Sci-kit Learn Standard Scaler. Pearson correlation coefficients between the descriptors and log-transformed IC_{50} values were calculated to identify descriptors related to the activity. All the properties with a correlation coefficient of more than 0.25 and less than -0.25 were selected, resulting in 215 descriptors.

A correlation matrix among the descriptors was obtained to remove highly co-related features. One of the two highly co-related features i.e., Pearson correlation coefficient greater than 0.8 and lesser than -0.8 was also removed. This resulted in the identification of 43 features. The feature selection was based on the backward feature elimination using a linear regression model and the calculation of the P-value. Further, the threshold value was set to 0.05 [53]. Finally, there were 22 features left that were used for building

models, apart from molecular descriptors and fingerprint descriptors, viz. MACCS, PubChem fingerprint, and Klekota–Roth fingerprint (KRFP) were also calculated using the CDK descriptor calculator. All the features with zero variance were removed **(Table S2 of appendix**).

3.2.4 BACE-1 inhibitor training and test data sets

BACE-1 inhibitor dataset was split into training (70%), validation (15%) and test (15%) sets, by stratified *train_test_split* of python module *scikit learn* using *random state* as 1 into 2476 compounds in the training set and 530 compounds each in validation and test sets. Principal component analysis (PCA) was performed on 2D descriptors data to examine the chemical space and ensure uniform data distribution across training, validation and test sets. The number of components was kept to three (n_component=3).

3.2.5 Machine learning classification algorithms

Five machine learning algorithms viz. Naïve Bayesian (NB), Random Forest (RF), support vector machine (SVM), gradient boosting machine (XGB) and k- nearest neighbors (k-NN) were used to obtain classification models. All the machine learning classification algorithms were employed using Python library *scikit learn*-0.22.1 on Python 3.7. Grid search was performed to identify the optimal combination of parameters using *GridsearchCV* of Scikit Learn. Five-fold cross-validation (CV) accuracy was compared to determine the optimal combination of parameters.

3.2.5.1 Naïve Bayesian classifier

Gaussian type NB and Bernoulli type NB classifiers were used for Mordred 1.2 and fingerprint-based descriptors, respectively. The former is preferably used for continuous type of data, while the latter is used for Boolean or binary features. A grid search was performed to obtain the hyperparameter of the model. Gaussian type NB Models were

built for different values of *var_smoothing* ranging from 10^{-3} to 10^{-10} with a stepwise increment of 10^{-1}.

3.2.5.2 k-Nearest neighbours

Grid search was performed by using different combinations of parameters to optimize the performance of k-NN models. The parameter included:

- Number of neighbours- All the odd numbers between 3 to 20
- Distance metric- Euclidean, Manhattan and Minkowski distance

Thus, 27 models were built for each set of descriptors, and the best model was selected based on 5-fold CV accuracy.

3.2.5.3 Support vector classifier

Grid search was performed by using different combinations of parameters to optimize the performance of k-NN models. The parameters included:

- Regularization parameters (C)- 0.1, 1.0, 10, 100 and 1000
- Kernel coefficient (gamma)- 0.0001, 0.001, 0.01, 0.1 and 1.0
- Kernel- Poly, rbf and sigmoid

The combinations of the above parameters resulted in 75 models for each set of descriptors. The best model was selected on the basis of 5-fold CV accuracy.

3.2.5.4 Random forest classifier

Random forest (RF) is determined by three parameters, i.e., n estimator (number of trees in forest), max depth (maximum depth of the tree) and min sample split (minimum number of samples required to split an internal node). A grid search was performed to obtain parameters for maximum 5-fold CV accuracy by using the following parameters.

- n estimators 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1000
- Max depth ranges from 5 to 50 with a stepwise increment of 5
- Min sample split ranges from 2 to 10

3.2.5.5 Gradient boosting algorithm

The parameters of Gradient boosting algorithm (XGB) were similar to RF along with loss function. A grid search was performed to obtain hypertuned parameters and best model was selected on the basis of accuracy score. Since the XGB is computationally expensive, a small grid search was performed using the following parameters.

- n estimators 500, 600, 700, 800, 900 and 1000
- Max depth ranges from 5 to 10
- Min sample split ranges from 2 to 10

3.2.6 Validation of the performance of the models

The performance of the models was evaluated by using the following metrics: Accuracy (A), Cross validated accuracies (5-CV, 10-CV, and leave-one-out), true positive (TP), false positive (FP), true negative (TN), false positive (FP), false negative (FN), precision (PE), recall (RC), and F1 score.

$$A = \frac{TP + TN}{TP + FN + TN + FP} X\ 100$$

$$PE = \frac{TP}{TP + FP}$$

$$RC = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * PE * RC}{PE + RC}$$

3.2.7 Comparison of the performance of models created on 2D descriptors and fingerprints with 3D descriptors

3.2.7.1 Compound alignment and calculation of 3D descriptors

All the molecules were aligned to the co-crystallized ligand of BACE-1 in their bioactive conformation inside the binding site retrieved from the protein data bank (https://www.rcsb.org/). The bound conformation of compound 5HA (PDB ID-3tpp) had

a relatively better resolution (1.60 Å) and was used as a template for the alignment of molecular structures. The alignment was performed using the GetCrippenO3A module of RDkit. The number of conformers was set to 100, and the best score index for each molecule was selected from the multiple conformers. The 3D descriptors of the selected conformers were calculated using RDkit.

3.2.7.2 Evaluation of the performance of models build on 3D descriptors.

The models were trained and evaluated by using the methods discussed above in sections **3.2.5** and **3.2.6,** respectively.

3.3 Results and Discussion

3.3.1 Chemical space exploration and dataset distribution of BACE-1 inhibitors

The complete BACE-1 inhibitor dataset is represented in the figure with a pair plot of molecular property descriptors viz. Molecular weight, SlogP, TPSA, and number of rotatable bonds**.** Both active and inactive compounds were found to be well distributed in figure. There was no bias in the selection of active and inactive compounds for preparing the data sets, as was evident from the overlapping **(Figure 3.1)**.

3.3.2 Feature selection of calculated descriptors

The feature selection process contributed more to the prediction output. It was an important step to increase and improve the accuracy of the classification model. Mordred 1.2 calculated total 1826 molecular descriptors and building a classification model on such a large number of features would decrease the performance of a model. Pearson co-relation coefficient is a measure of relationship between the two variables. Since, the highly co-related features are not meaningful therefore, any one of features was selected. The co-relation of the activities and selected molecular properties are listed in table **(Figure 3.2).** The corresponding p-values indicated that all the co-relations were statistically significant and, could be used for building models in machine learning [54].

The final number of features used for building model has been summarised in table **(Table S1 of appendix)**.

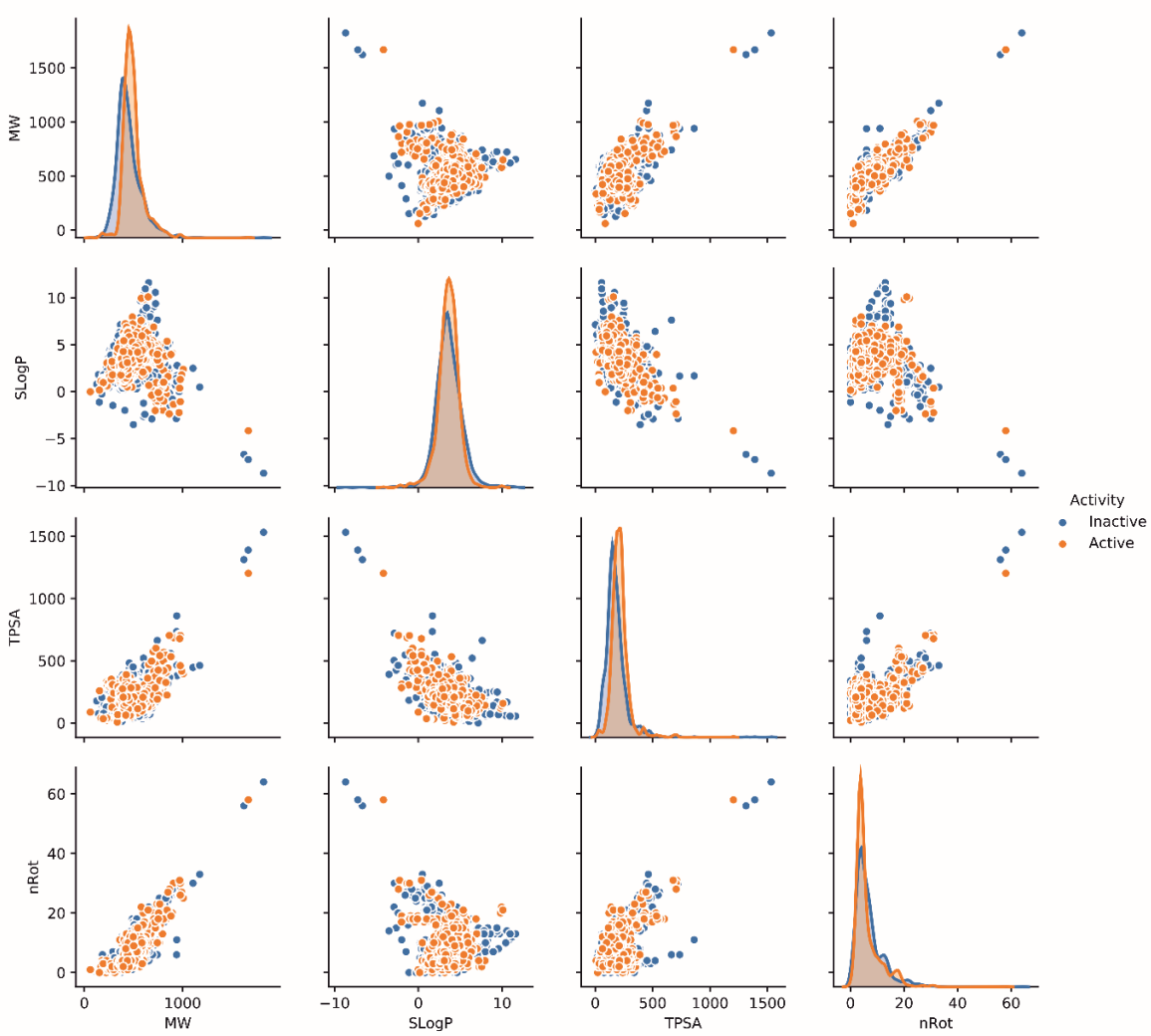


Figure 3.1 Pair plot of BACE-1 inhibitor dataset. The diagonal represents the frequency of distribution. (MW=Molecular weight, TPSA=Total Polar Surface Area, nRot=Number of rotatable bond)

3.3.3 BACE-1 inhibitor training, validation and test sets

The test-train split of Python toolkit, *scikit learn*, divided actives and in actives into test, validation and training sets in equal proportions **(Figure 3.3A)**. In an ideal splitting, the chemical space of training set should overlap to that of test set and validation set. The 3D scatter plot of PC1, PC2 and PC3 represented on X, Y and Z-axis respectively is depicted in figure **(Figure 3.3B)**. Three principal components obtained after performing PCA were

able to explain more than 50 % variance of the dataset. The scatter plot shows homogenous distribution (overlap) for the validation and test sets around the training set compounds **(Figure 3.3B)**.

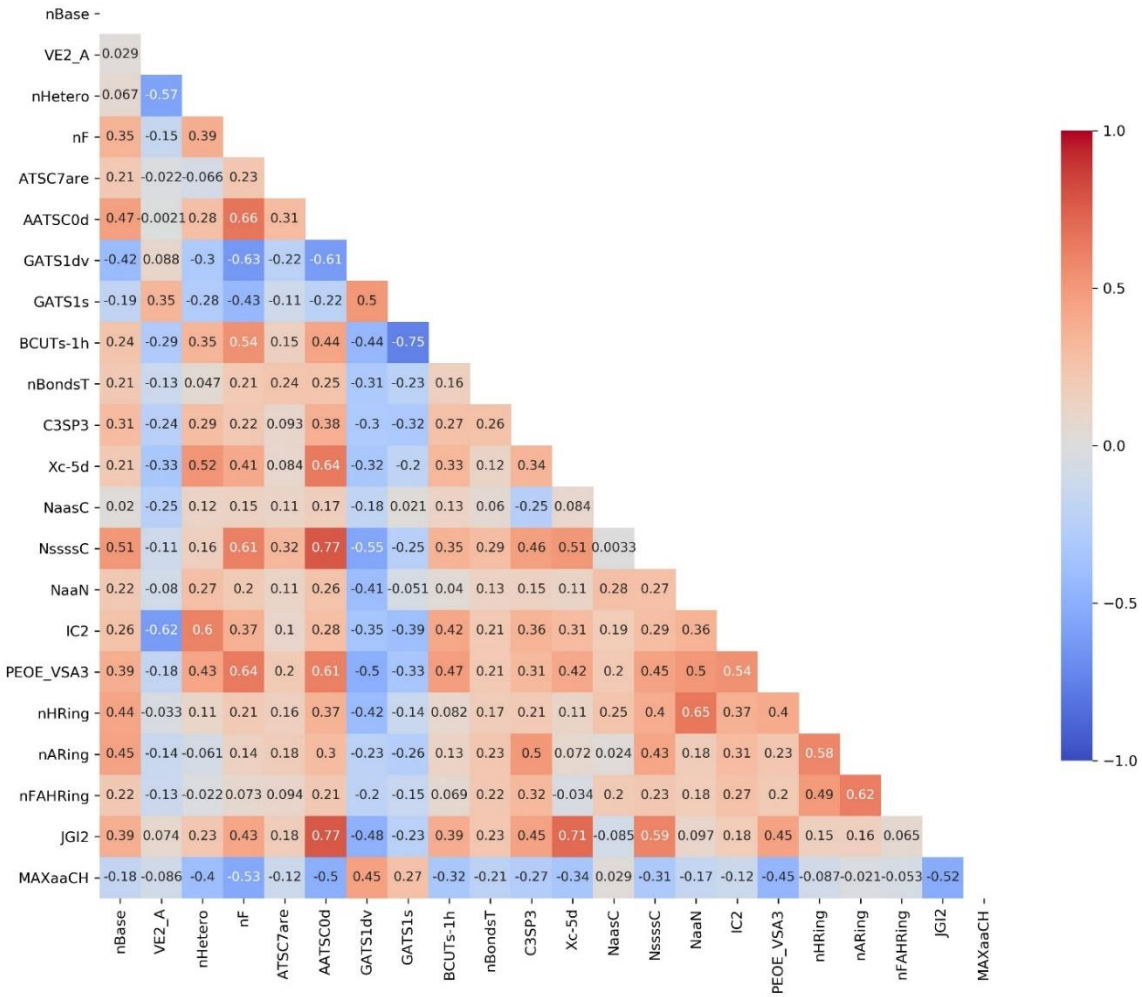


Figure 3.2 Pearson correlation coefficient of calculated descriptors after performing feature selection

Table 3.2 Pearson correlation coefficient between selected molecular properties and BACE-1 inhibition.

Descriptor code	Description	Pearson correlation coefficient*
nBase	Basic group count	-0.33
VE2_A	VE2 of adjacency matrix	0.30
nHetero	Number of hetero atoms	-0.27
nF	Number of F atoms	-0.33
ATSC7are	Centered moreau-broto autocorrelation of lag 7 weighted by allred-rocow EN	-0.25
AATSC0d	Averaged and centered moreau-broto autocorrelation of lag 0 weighted by sigma electrons	-0.37
GATS1dv	Geary coefficient of lag 1 weighted by valence electrons	0.30
GATS1s	Geary coefficient of lag 1 weighted by intrinsic state	0.26
BCUTs-1h	First highest eigenvalue of Burden matrix weighted by intrinsic state	-0.31
nBondsT	Number of triple bonds in non-kekulized structure	-0.29
C3SP3	SP3 carbon bound to 3 other carbons	-0.25
Xc-5d	5-ordered Chi cluster weighted by sigma electrons	-0.33
NaasC	Number of aasC	-0.25
NssssC	Number of ssssC	-0.40
NaaN	umber of aaN	-0.26
IC2	2-ordered neighborhood information content	-0.43
PEOE_VSA3	MOE Charge VSA Descriptor 3 (-0.25 <= x < -0.20)	-0.40
nHRing	Hetero ring count	-0.35
nARing	aromatic ring count	-0.27
nFAHRing	aromatic fused hetero ring count	-0.25
JGI2	2-ordered mean topological charge	0.33
MAXaaCH	max of aaCH	0.25

*P<0.05

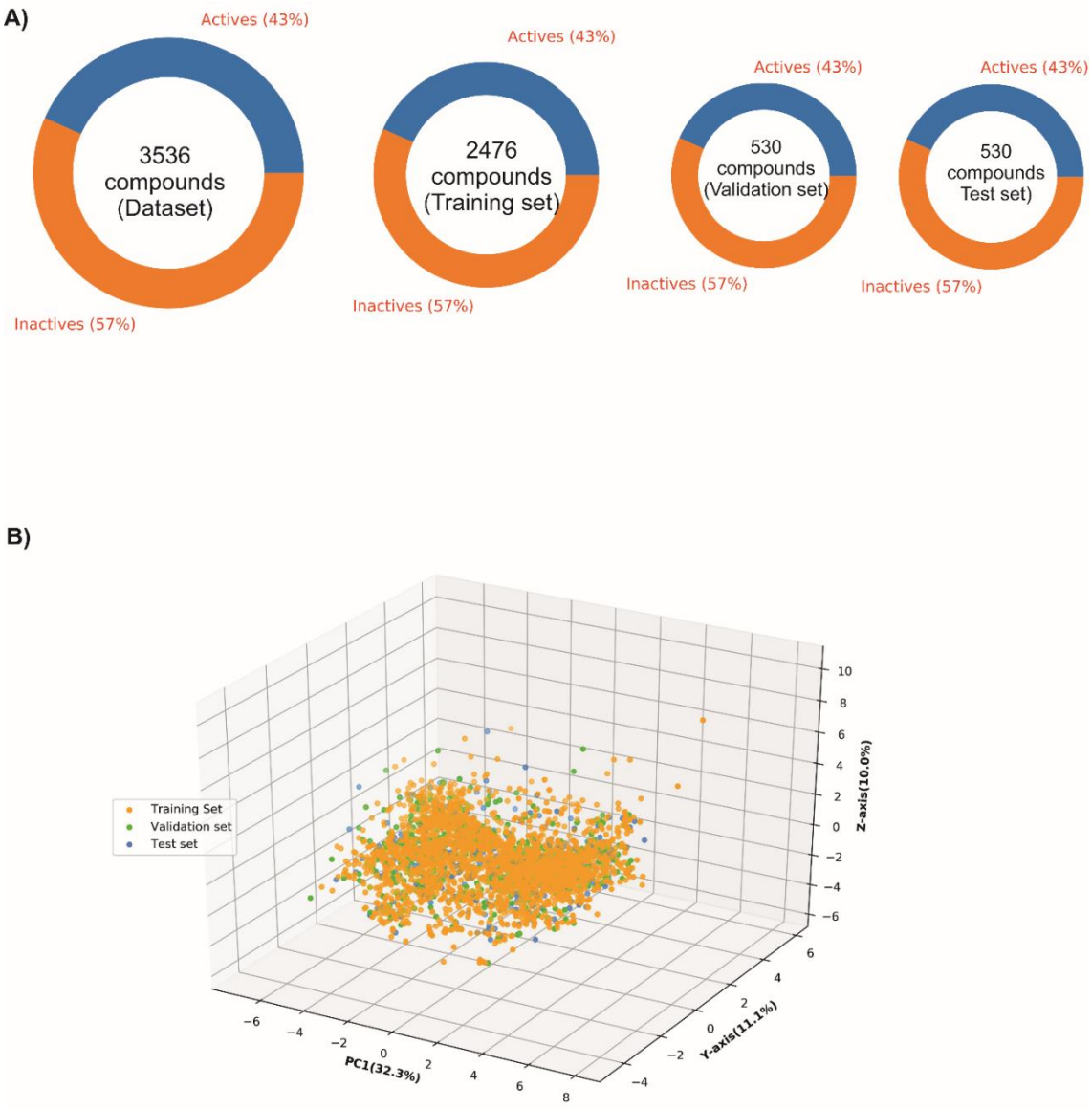


Figure 3.3 Data distribution for (A) complete BACE-1 dataset and (B) Chemical space of training, validation and test set as represented by PCA plot

3.3.4 Machine learning classification algorithms

3.3.4.1 Naïve Bayesian classifier

Naïve Bayesian classifier is based on Bayesian theorem, where it searches through each feature having ability to separate in an unbiased way. A molecule is represented in the form of 1-D array of molecular descriptors and structural fragments. The probability of a

molecule belonging to active or inactive class can be predicted by combining the-calculated probabilities of all descriptors using the following Bayesian theorem:

$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$

Where, A represents the activity class of a compound and B represents a molecular property; P(A) is prior probability of compound activity; P(B) is probability that a property is owned by a compound in sample data; P(B|A) is conditional probability of likelihood of the event B, when A has already occurred. On performing grid search, it was observed that the parameter *var_scaling* had no impact on the performance of the model. The four NB models, built on dataset from Pubchem fingerprints i.e., NB-4, showed better performance than the rest. The accuracies of training and validation sets were 74.82% and 74.51%. The result has been summarised in table **(Table 3.3).**

Table 3.3 Parameters and performance of classification model built by NB classifier

Model number	**Type**	**Var smoothing**	**Descriptors**	**Training set**			**Validation set**			
				Q	5-CV	10-CV	Q	PR	RC	F1
NB-1	Gaussian	10^{-9}	Mordred	74.26	75.47	75.25	75.14	0.70	0.72	0.71
NB-2	Bernoulli	10^{-9}	KRFP	73.81	72.72	74.33	74.38	0.68	0.77	0.72
NB-3	Bernoulli	10^{-9}	MACCS	70.78	70.62	70.54	70.62	0.64	0.73	0.68
NB-4	Bernoulli	10^{-9}	Pubchem	74.82	74.62	76.79	74.51	0.70	0.75	0.72

Q-Accuracy, CV-cross validation, PR- precision, RC- Recall

3.3.4.2 k-Nearest neighbours

The k-NN algorithm is one of the simplest machine learning algorithms. The data is represented in a high-dimensional feature space and labels from the closest nodes are assigned to query. Value k specifies the number of closest neighbours and is kept variable; too high or too low value may result in overfitting [55]. The optimum values of parameters have been summarized in **Table 3.4.** The dataset of KRFP descriptors i.e.,

KNN-2, showed best performance in kNN models, with training and validation set accuracies of 91.03% and 88.51%, respectively. The results of the models are summarized in **Table 3.4.**

Table 3.4 Parameters and performance of classification model built by kNN classifier

Model number	**N-Neighbours**	**Distance metrics**	**Descriptors**	**Training set**			**Validation set**			
				Q	5-CV	10-CV	Q	PR	RC	F1
KNN-1	5	Manhattan	Mordred	90.10	84.24	84.44	85.49	0.82	0.83	0.83
KNN-2	5	Manhattan	KRFP	91.03	86.78	87.31	88.51	0.87	0.86	0.86
KNN-3	7	Manhattan	MACCS	89.05	83.79	84.24	85.87	0.81	0.82	0.82
KNN-4	5	Manhattan	Pubchem	89.93	84.84	84.52	83.61	0.80	0.81	0.81

3.3.4.3 Support vector machines model

SVM model solves the classification problem by using a linear or non-linear kernel function to map data into high-dimensional space by finding an optimally separating hyperplane [56]. It can be determined by three parameters i.e., a penalty parameter (C), a kernel function and a kernel coefficient.

The performance of SVM model, built on Mordred 1.2 descriptors, was not as good as the corresponding models on fingerprint descriptors. The model raised on KRFP fingerprints SVM-2, showed best validation set accuracy of 91.33% having precision, recall and F1 scores of 0.91, 0.88 and 0.89, respectively. The other models built on Modred descriptors, MACCS and Pubchem fingerprints showed comparable performances. The parameters and performance of model having maximum accuracy for each dataset is included in **Table 3.5.**

Table 3.5 Parameters and performance of classification model built by SVM classifier

Model number	Parameters			Descriptors	Training set			Validation set			
	C	gamma	kernel		Q	5-CV	10-CV	Q	PR	RC	F1
SVM-1	10	0.1	Rbf	Mordred	96.68	86.50	87.11	88.88	0.87	0.86	0.86
SVM-2	100	0.1	Rbf	KRFP	95.95	89.13	89.33	91.33	0.91	0.88	0.89
SVM-3	1	0.1	Rbf	MACCS	93.73	86.90	87.02	90.01	0.91	0.85	0.88
SVM-4	10	0.01	Rbf	Pubchem	95.11	87.03	87.03	88.13	0.86	0.86	0.86

3.3.4.4 Random-forest classifier

Random forest is an ensemble learning method in which multiple decision trees are built on the training dataset. It is a not linear modelling algorithm. On each split node of tree, different subset of descriptor is evaluated for best split. Thus, a forest of such decision tree is created and the final prediction for a new sample is the average prediction of all the trees in a forest [57]. The parameters and performance of the model having maximum accuracy for each dataset is summarised in **Table 3.6.**

Table 3.6 Parameters and performance of classification model built by RF classifier

Model number	Parameters			Descriptors	Training set			Validation set			
	N estimator	Max depth	Min samples split		Q	5-CV	10-CV	Q	PR	RC	F1
RF-1	500	10	4	Mordred	94.98	85.29	85.73	86.62	0.84	0.84	0.84
RF-2	700	20	5	KRFP	95.67	88.00	88.44	89.83	0.91	0.84	0.87
RF-3	500	15	5	MACCS	94.86	86.94	87.39	89.26	0.89	0.84	0.87
RF-4	800	30	5	Pubchem	96.04	86.86	87.19	87.94	0.89	0.81	0.85

The performance of the random forest models, built on the Mordred dataset, was not as good as the model built on fingerprint descriptors. For the four random forest models,

RF-2 created on KRFP dataset showed maximum test set accuracy, precision, recall and F1 scores of 90.96%, 0.88, 0.90 and 0.89, respectively.

3.3.4.5 Gradient boosting algorithm

The ensemble method combined prediction from multiple decision trees to generate final predictions. In gradient boosting (XGB) machines, the learning procedure fits new models after each iteration, to provide a more accurate estimate of the response variable. The key idea behind the algorithm is to develop new base-learners, to be utterly correlated with the negative gradient of the loss function and to be associated with the whole ensemble [58]. The performance of model GB-4, built on the Pubchem dataset showed maximum validation set accuracy of 89.45%. The parameters and performance of the best model, corresponding to every dataset has been represented in **Table 3.7**.

Table 3.7 Parameters and performance of classification model built by XGB classifier

Model	**Parameters**			**Descriptors**	**Training set**			**Validation set**			
number	N estimator	Max depth	Min samples split		Q	5-CV	10-CV	Q	PR	RC	F1
GB-1	500	3	7	Mordred	98.98	85.74	85.73	86.25	0.85	0.81	0.85
GB-2	500	4	10	KRFP	97.85	85.73	85.73	89.07	0.89	0.84	0.86
GB-3	700	4	10	MACCS	93.17	86.22	86.26	87.75	0.88	0.82	0.85
GB-4	500	3	8	Pubchem	96.72	87.72	87.28	89.45	0.88	0.86	0.87

3.3.5 Prediction of the test set data

A test set was used to validate the performance of all the 20 models. The performance of models on test set is summarised in Table 3.8**.** All the 20 classification models showed overall accuracy between 71.69 - 89.62% on the test set. The model SVM2, built on KRFP fingerprint by SVM classifier, showed maximum accuracy, precision and recall score (0.87). The models showed precision score in the range of 0.65 – 0.90. Further, model RF2, built on KRFP fingerprint by RF classifier also showed precision score of 0.90.

3.3.6 Comparison of the performance of models build on 2D descriptors and fingerprints with 3D descriptors

The AUC of ROC plot for the best model of all the classification algorithms has been summarized in table (**Table S3 of appendix** and **Figure S1 of appendix)**. The best AUC of ROC was shown by the model build on RF classifier, RF-5 having score of 0.93, which was less than the models built on 2D descriptors and fingerprint datasets.

3.3.7 Structural diversity of BACE-1 inhibitors

The dataset of BACE-1 inhibitors consisted of diversity of structures such as benzopyran, imidazole, thiazolidines, pyrimidine etc. To evaluate the structural diversity between every two ligand pairs, we calculated Tanimoto coefficient-based distance matrix from their MACCS fingerprints. In general, two inhibitors were considered similar, if they have a Tanimoto coefficient of more than 0.70. It was observed that only 6.25% of the ligands had Tanimoto coefficient greater than 0.70, indicating that the dataset is structurally diverse. The frequency distribution of Tanimoto coefficient of each ligand pair is depicted in **Figure 3.4.**

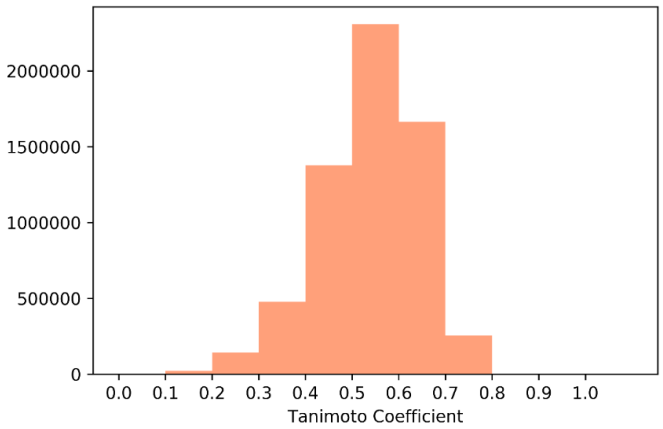


Figure 3.4 Frequency distribution histogram of Tanimoto similarity coefficient for pair of BACE-1 inhibitor based on MACCS fingerprints.

Table 3.8 Performance of 20 classification models on test set.

Model	TN	FP	FN	TP	Q	PR	RC	F1	AUC of ROC
NB-1	227	74	59	170	74.90	0.69	0.74	0.71	0.81
NB-2	219	82	47	182	75.66	0.68	0.79	0.73	0.85
NB-3	207	94	56	173	71.69	0.65	0.75	0.69	0.81
NB-4	231	70	52	177	76.98	0.71	0.77	0.74	0.82
KNN-1	263	38	31	198	86.98	0.83	0.86	0.85	0.92
KNN-2	269	32	36	193	87.16	0.85	0.84	0.85	0.93
KNN-3	256	45	31	198	85.66	0.81	0.86	0.83	0.93
KNN-4	264	37	42	187	85.09	0.83	0.81	0.82	0.92
SVM-1	271	30	37	192	87.35	0.86	0.83	0.85	0.91
SVM-2	280	21	34	195	89.62	0.90	0.85	0.87	0.95
SVM-3	276	25	37	192	88.30	0.88	0.83	0.86	0.93
SVM-4	274	27	31	198	89.05	0.88	0.86	0.87	0.94
RF-1	266	35	33	196	87.16	0.84	0.85	0.85	0.94
RF-2	281	20	37	192	89.24	0.90	0.83	0.87	0.94
RF-3	276	25	39	190	87.92	0.88	0.82	0.85	0.95
RF-4	279	22	34	195	89.43	0.89	0.85	0.87	0.95
GB-1	266	35	41	188	85.66	0.84	0.82	0.83	0.93
GB-2	275	26	35	194	88.49	0.88	0.84	0.86	0.94
GB-3	276	25	38	191	88.11	0.88	0.83	0.85	0.94
GB-4	275	26	30	199	89.43	0.88	0.86	0.87	0.94

3.3.8 k-Means clustering of BACE-1 inhibitors

In order to explore the structural features of 3536 inhibitors in the dataset, clustering was performed on the MACCS fingerprints data. The k-Means clustering is a method to classify data into k-groups, by minimizing distances within-group to the centroid. Further, most commonly used Euclidean distance metric was employed in the evaluation. The optimum value of k was determined by using the elbow method and silhouette score. For MACCS data with 11 clusters, the silhouette score was found to be only 0.14. Therefore, dimensionality reduction was carried out by using t-distributed stochastic neighbours embedding (t-SNE). The data was reduced to two-dimensions with the help of t-SNE, and was designated as t-SNE1 and t-SNE2. The silhouette score was found to be 0.52, after clustering on this dataset. The clusters and their centroids are included in **Figure 3.5.** Pairwise distances were calculated by using Euclidean distance metric, to visualize the structures closest to the centroid of cluster. The structure of compounds closest to the

centroids and activity class of all the 11 clusters have been represented in **Figure 3.6.** Further, subset 0 consisted of 187 compounds, with 63.64 % actives. The cluster mainly consisted of compounds with napthyridin and thiazine moieties. The subset 1 consisted of 322 compounds, with 31.06% of actives haveing oxopyrrolidin, pyrrolidine and sulphone moieties. The subset 2 had 453 compounds with similar numbers of active and inactive compounds. This cluster had a few azaspiro and pyridyl containing compounds too. Subset 3 consisted of compounds containing napthaquinone ring. This subset of 253 compounds, had only 8.3 % of actives. The compounds in cluster 4 had groups such as azabicyclo and fluorophenyl. This cluster had 424 compounds in which 59.67 % of compounds were active. In cluster 5, there were 209 compounds with 92.82 % actives. This cluster had the maximum percentage of active compounds among the clusters. Most of the compounds in this cluster possessed nitrile group, phenyl pyrrolidine ring and aza-

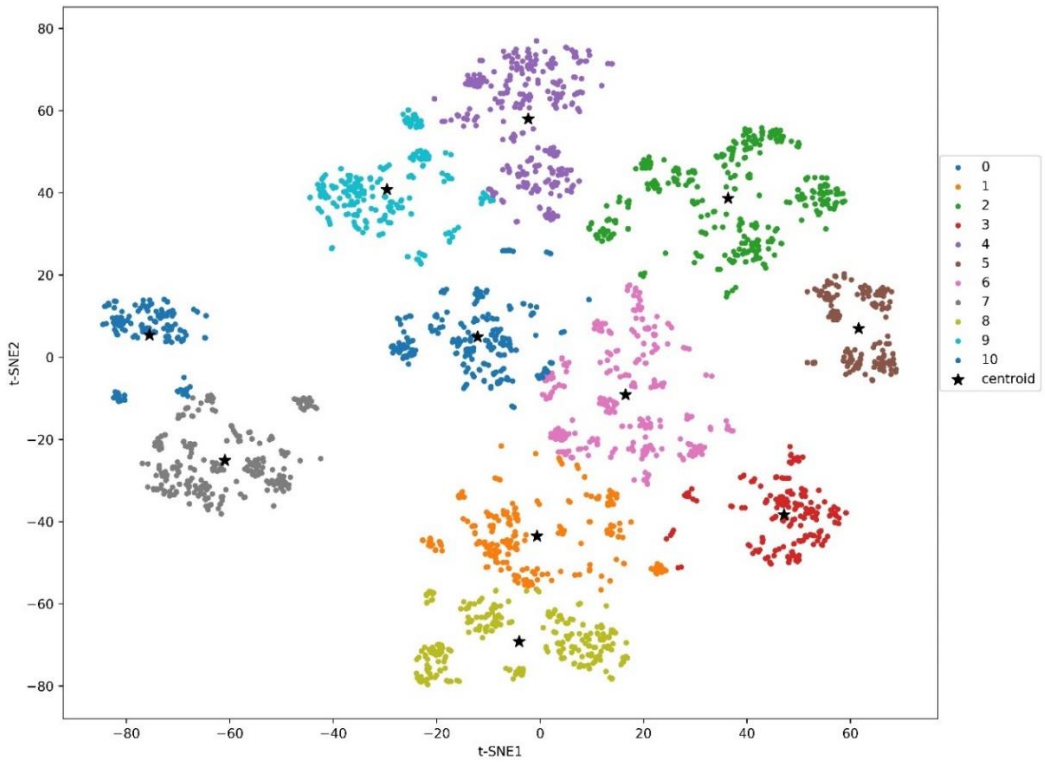


Figure 3.5 Clustering of compounds in 11-subsets along with their centroids (represented by ★). t-SNE1 and t-SNE2 are the two dimensions reduced from 166 dimensions of MACCS fingerprints.

bicyclo ring. The cluster 6 had 390 compounds with only 15.38 % of active compounds. The compounds consisted of groups such as aminopyridine, pyrrole ring and N-cycloalkylbenzamide. Cluster 7 was composed of compounds containing piperazinyl, isoxazolyl and sulfonyl groups. The cluster had 45.27 % of active compounds out of total 349 compounds. In cluster 8, there were 353 compounds in which 41.93 % of compounds were active. The compounds had fluorophenyl, hydroxy and acetamide groups. Cluster 9 had 287 compounds and 51.22 % of them were active. The percentage of active and inactive compounds in this subset was almost similar. The compounds in this subset consisted of oxazine ring and hydroxypicolinamide group. Finally, cluster 10 had 310 compounds in which 33.87% of compounds were active. The compounds in this subset had thiophene, imidazolyl and fluoropyridyl groups.

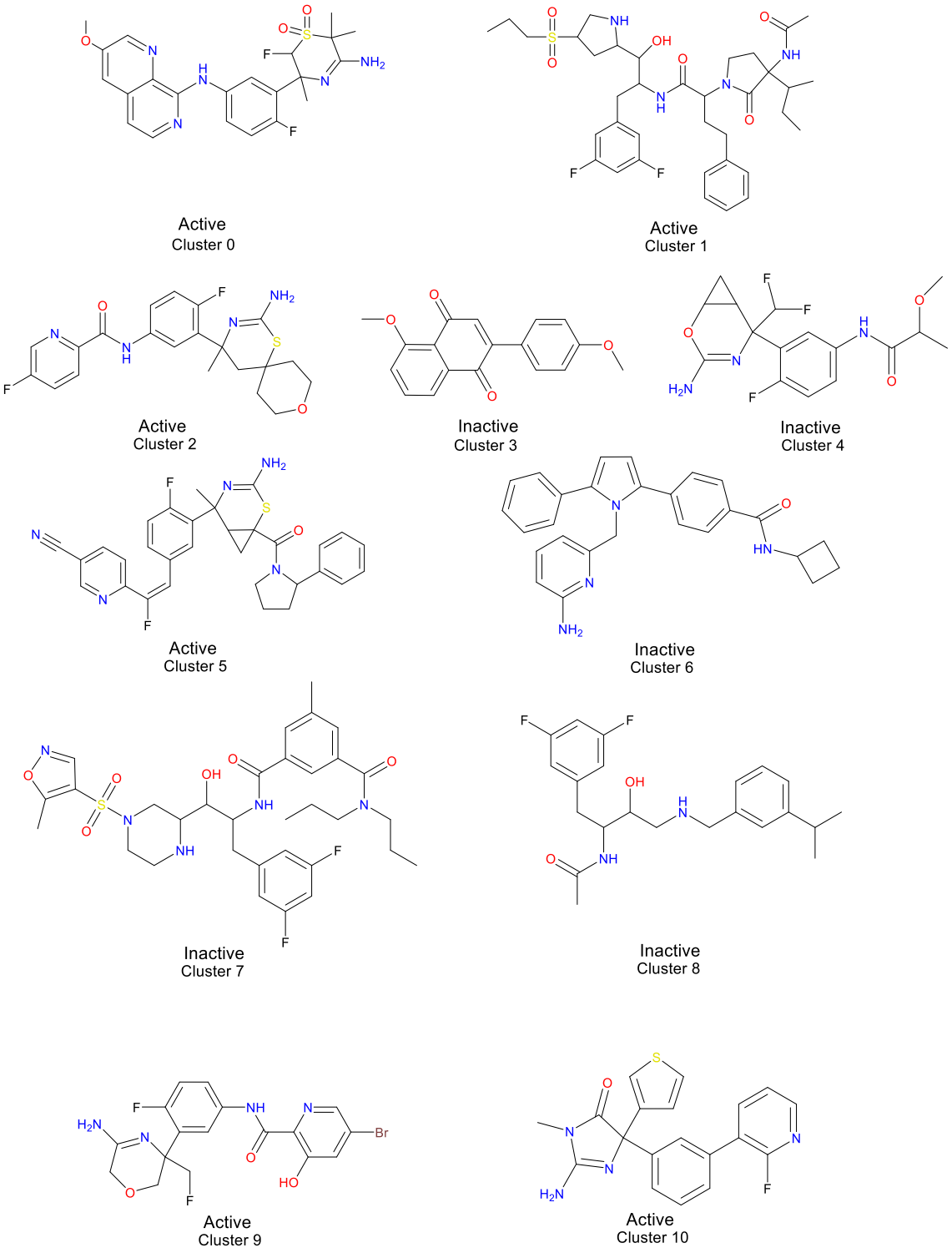


Figure 3.6 Central compounds and their corresponding activities in the eleven subsets

3.3.9 Comparison of performance of classification algorithms and descriptors

3.3.9.1 Molecular descriptors

Performance of the models built on molecular descriptor dataset was evaluated by using area under curve (AUC) of the receiver operating characteristic (ROC) plot. The values were found to be 0.81, 0.92, 0.91, 0.94 and 0.93 for NB, kNN, SVC, RF and GB algorithms, respectively **(Figure 3.7 A)**. The performance of random forest classifier was observed to be best considering the AUC and F1 score. The class specific feature importance and overall feature importance of top 10 properties have been summarized in **Table 3.9**.

3.3.9.2 KRFP fingerprints

The AUC of ROC plots for machine learning algorithms built on KRFP fingerprint dataset revealed that the performance of NB classifier (AUC=0.85) was not as good as other classification algorithms i.e., kNN, SVC, RF and GB. The AUC of ROC for kNN. Further, SVC, RF and GB was found to be 0.93, 0.95, 0.95 and 0.94, respectively **(Figure 3.7 B)**. The F1-score of RF classifier was also finest among the other classification algorithms. The top ten good and bad fingerprints obtained from random forest classifiers are included in **Figure 3.8.**

3.3.9.3 Pubchem fingerprints

In case of Pubchem fingerprints, the performance of kNN, SVC, RF and GB classifiers were almost similar, but the performance of NB classifier was less than the remaining classifiers. The AUC of ROC was found to be 0.82, 0.92, 0.94, 0.95 and 0.94 for NB, kNN, SVC, RF and GB, respectively **(Figure 3.7 D)**. On comparing the other accuracy metrics, RF classifier was found to be the best. The top ten good and bad fragments are represented in **Figure 3.9.**

3.3.9.4 MACCS fingerprint

In case of MACCS fingerprint, NB classifier showed poor result (AUC of ROC=0.81), but rest of the classifiers showed better result. AUC of ROC was found to be 0.93 for kNN, 0.94 for SVC, 0.95 for RF and 0.94 for XGB algorithm **(Figure 3.7 C)**. The top ten important features, along with their class specific feature importance were calculated. The fragments contributing to active and inactive BACE-1 inhibitors were identified. The active compounds were mainly found to contain a fluorine and an aromatic nitrogen atom **Table 3.10**.

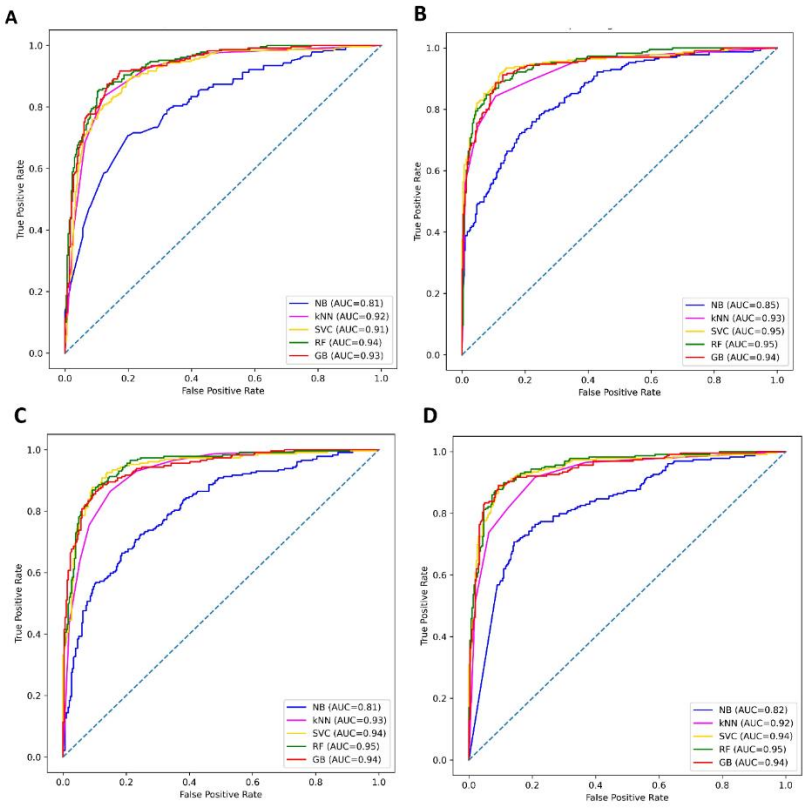


Figure 3.7 Receiver operating characteristics (ROC) curves of models build on (A) Mordred, (B) KRFP fingerprints, (C) MACCS fingerprints and (D) Pubchem fingerprints datasets

Table 3.9 Top ten features from Mordred dataset with their feature importance values from RF classifier

S. No.	Descriptor code	Feature importance (Active class)	Feature importance (Inactive class	Feature importance (Overall)
1	BCUTs-1h	1.826	1.750	0.128
2	nHetero	0.295	0.382	0.104
3	nBase	0.525	0.581	0.089
4	GATS1dv	0.077	0.073	0.081
5	nF	0.080	0.007	0.081
6	ATSC7are	0.052	0.115	0.080
7	GATS1s	0.052	0.047	0.072
8	NaasC	0.138	0.059	0.067
9	Xc-5d	0.029	0.021	0.066
10	C3SP3	0.036	0.028	0.062

Table 3.10 Top ten features from MACCS dataset with their feature importance values from RF classifier

S.No.	MACCS Key	Description	Feature Importance (Active Class)	Feature importance (Inactive class)	Overall feature importance
1	MACCSFP42	F	0.03292	0.015881	0.041926
2	MACCSFP65	C%N	0.031471	0.018215	0.040214
3	MACCSFP112	AA(A)(A)A	0.030114	0.019467	0.032893
4	MACCSFP66	CC(C)(C)A	0.02617	0.015269	0.032021
5	MACCSFP144	Anot%A%Anot%A	0.017212	0.013341	0.022599
6	MACCSFP92	OC(N)C	0.019617	0.014014	0.022523
7	MACCSFP154	C=O	0.016109	0.013349	0.018649
8	MACCSFP19	7M RING	0.004582	0.001327	0.017893
9	MACCSFP134	X (HALOGEN)	0.014445	0.00942	0.017506
10	MACCSFP95	NAAO	0.012847	0.009494	0.015532

(F-fluorine, %- aromatic query bond, A- Any valid periodic table element symbol, =- double bond)

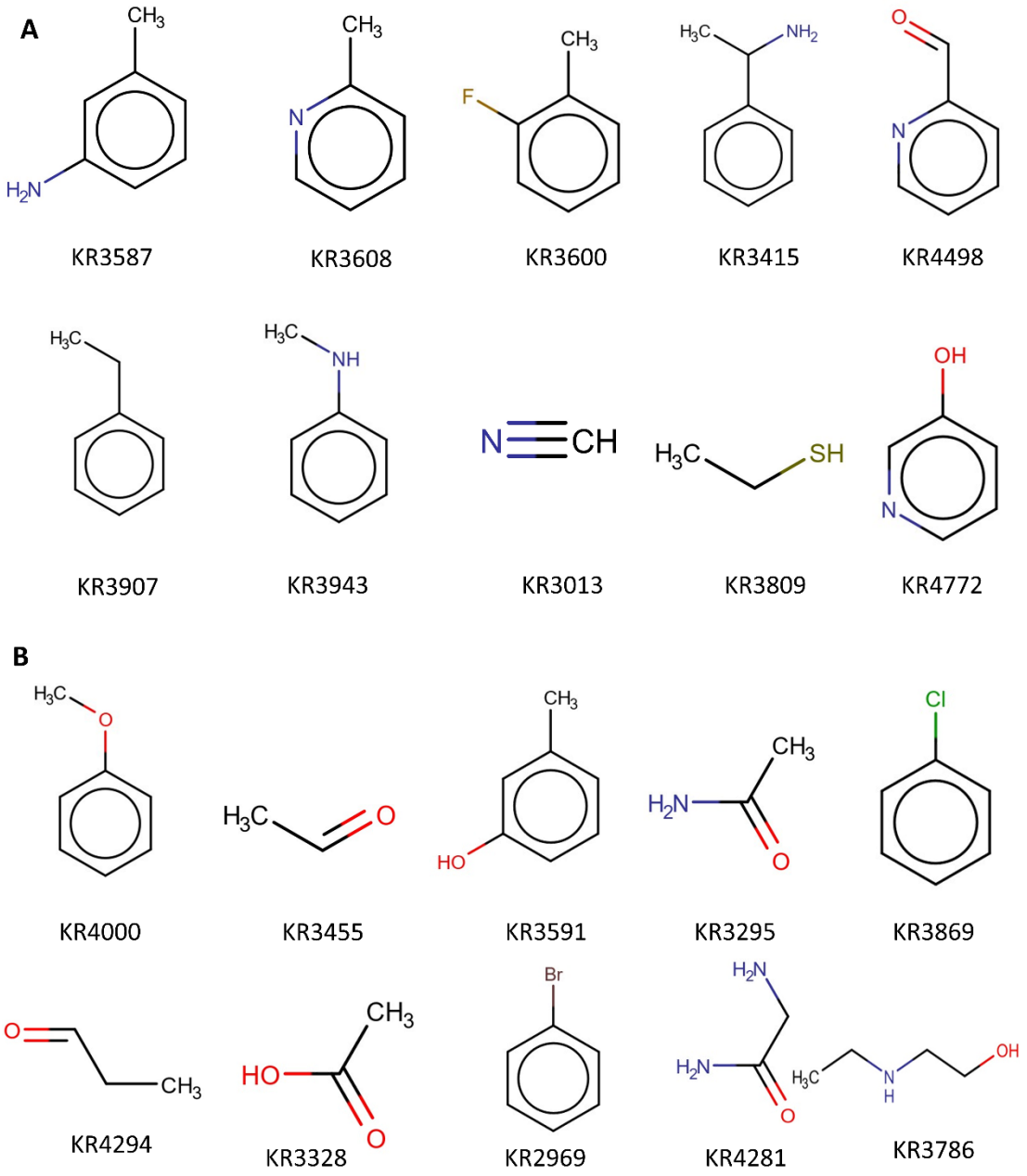


Figure 3.8 KRFP fingerprints (A) beneficial and (B) adversely affecting the BACE-1 inhibitory activity obtained from RF classifier

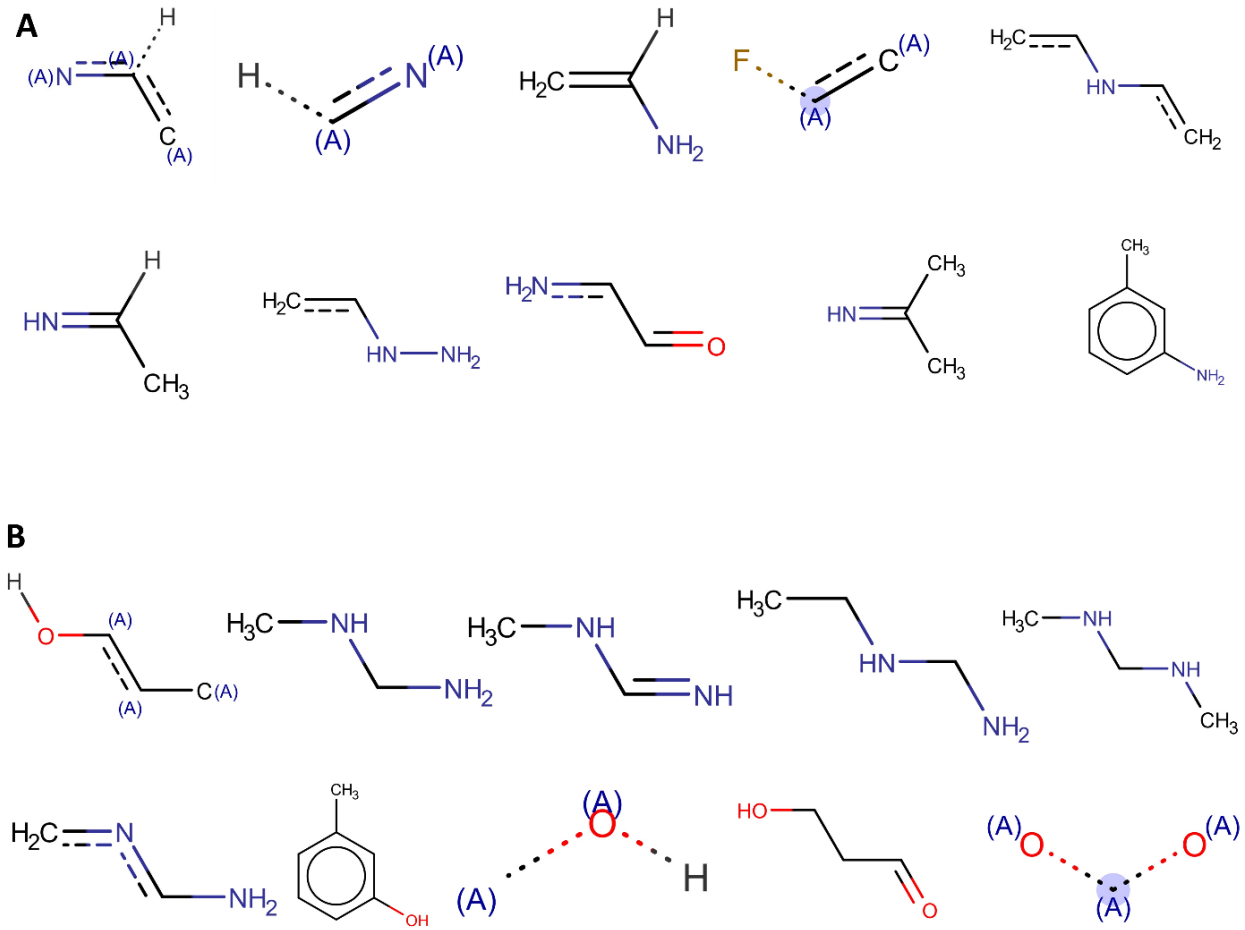


Figure 3.9 Pubchem fingerprints (A) beneficial and (B) adversely affecting the BACE-1 inhibitory activity obtained from RF classifier

3.3.10 Defining applicability domain of models

The applicability domain of a model specifies the limit within which the prediction is considered reliable. This restricts the applicability of a model to predict the compounds, which are similar to the training set used in model building. Similarity measurement was used to define the applicability domain. The applicability domain of the model was defined using the following equation:

$$AD = \gamma + \sigma Z$$

At first, the average pairwise Euclidean distance between all the training dataset based on Mordred descriptors was calculated. Then, the pairs with distances lower than the average

were formulated and γ and σ were calculated as the average and standard deviation, respectively. *Z*, an empirical cut-off value, was set to 0.5 [59]. The applicability domain of our model was found to be 9.74.

3.4 Conclusion

In the study, we developed twenty classification models, based on 3536 BACE-1 inhibitors. The molecular structures were quite diverse according to the Tanimoto coefficients. The dataset was divided into training, validation and test sets in the ratio of 70:15:15. Two types of descriptors i.e., molecular properties (Mordred 1.2) and chemical fingerprints (KRFP, Pubchem and MACCS) were calculated. Five machine learning algorithms were used to develop classification models viz. NB, SVM, kNN, RF and XGB. The performance of NB classifier was not up to the mark in comparison to the other classifiers, as was evident from the F1 score and AUC of ROC of all the classification algorithms.

The classification-based machine learning models developed in the study, showed better accuracy of 89.62% than the models developed by Ponzoni *et. al.* The accuracy of the best model in the earlier study, built by using RF algorithm, was 85% [59].The contribution of various fingerprints was analysed through the feature importance of *scikit learn*. The fragments such as aromatic amine, pyridine ring, fluoropyridine and nitrile groups were important for the active BACE-1 inhibitors. The model and the dataset used for its development can be accessed at https://github.com/ravisingh15/BACE-1_inhibitor. The broad chemical diversity of the compounds used in the dataset and the structural details would be certainly helpful in designing active BACE-1 inhibitors. The classification model developed could also be used as a virtual screening tool for the identification of BACE-1 inhibitors.

Chapter 4
Machine Learning based screening of *in-house* database to identify BACE-1 inhibitors

Summary:

The β-site APP cleaving enzyme-1 (BACE-1) is one of the key targets for novel drugs to treat Alzheimer's disease (AD). The BACE-1 plays a key role in the amyloidogenic process, leading to the production of amyloid-β (Aβ) plaques in the brain. In the present work, we have developed an ML model based on the sulfonamides dataset. The best ML model was built using the XGBoost algorithm on PubChem fingerprints. The model had an accuracy, precision, recall and F1 score of 0.89, 0.88, 0.99 and 0.93, respectively, on the validation set. The same model was used to screen the database of previously synthesized and reported in-house compounds. The screening resulted in the identification of two hits, i.e., compound 28 and compound 37. Both the compounds were screened for their BACE-1 inhibitor activity. The IC_{50} value of compound 28 was found to be 0.431 ± 0.006 μM, and compound 37 showed an IC_{50} value of 0.272 ± 0.019 μM. The docking study revealed that compound 37 also showed interactions with the catalytic dyad of BACE-1, i.e., Asp32 and Asp228.

4 Machine Learning based screening of in-house database to identify BACE-1 inhibitors

4.1 Introduction

BACE-1 is an aspartyl protease of the pepsin family, discovered in 1999. BACE-1initiates the production of Aβ, which represents the rate-limiting enzyme in the amyloidogenic pathway. BACE-1 cleaves the Aβ precursor protein (APP) to its membrane-bound C-terminus fragment C99 (CTF) and soluble APPβ fragment. The BACE-1 is essential for the generation of all monomeric units of Aβ, including $A\beta_{42}$, which plays a crucial role in the pathogenesis of AD. The concentrations and activity rates of BACE-1 are actively increased in AD brains and body fluids. Therefore, BACE-1 emerged as a primary drug target for decreasing the production of Aβ in the AD brain [60]. BACE-1 is a type-1 transmembrane protein that is different from other peptidases of the same family. The catalytic domains of BACE have two significant motifs of the sequence DTGS and DSGT that together forms the active site of the enzyme [61]. BACE-1 consists of metal binding sites; it has a copper-binding site in its cytosolic domain [62]. The crystal structure of BACE-1 reveals that its proteolytic pocket is relatively large and is less hydrophobic; therefore, it becomes challenging for developing small-molecule inhibitors using high-throughput virtual screening [63, 64].

4.1.1 Sulphonamides as BACE-1 Inhibitors in Human Clinical Trials

Non-peptide BACE-1 inhibitors such as sulphonamides had some success in pre-clinical studies as some of the drugs were seen in various phases of clinical trials as well.

BACE-1 inhibitor MK-8931 (Verubecestat) entered the Phase-3 of clinical trial conducted in mild to moderate Alzheimer's patients and was terminated as it failed to show efficacy over the placebo. MK-8931 reduced the levels of $A\beta_{40}$ in healthy participants, whereas it showed a decrease in cognitive performance compared to the placebo [65].

Phase I clinical trial study of SUVN-502 (Masupirdine) revealed that it is well-tolerated by healthy young and old adult participants. Phase II clinical trial (NCT02580305) for SUVN-502 in mild to moderate AD patients in combination with donepezil and memantine was completed but failed to show significant benefits.

Phase I clinical trial of SAM-760 was completed and well tolerated in healthy subjects and AD patients. Further, Phase II was terminated as it failed to show significant benefits [66].

Bertini *et al*. developed a series of substituted aryl sulphonamides **(I, Figure 4.1)** as BACE-1 inhibitors where the highest potency of a compound was found to be 1.6 µM [67]. Kang *et al* synthesized a series of sulphonamide chalcones (**II, Figure 4.1**) as dual inhibitor of BACE-1 and acetylcholinesterase. The compounds showed activity in the micromolar range; the best activity was 0.62 µM. Li *et al*. identified some sulphonamide derivatives via virtual screening as BACE-1 and PPARγ inhibitors (**III, Figure 4.1**). The IC_{50} value of one of the identified hits was found to be 1.24 µM. Zou *et al*. developed a series of pyrazole and sulphonamide-based BACE-1 inhibitors with potent activity. The best compound showed an IC_{50} value of 0.036 µM **(IV**, **Figure 4.1).**

Over the last decade, several research has been done on the therapeutic potential of BACE-1 inhibition. However, despite the fact that inhibitors effectively reduce Aβ levels, clinical trials still fail to show benefits in cognitive function when given to patients with mild-to-moderate AD. This raises concerns about the true value of these putative anti-AD medications as well as the design of the clinical trials. Recent research indicates that starting BACE-1 inhibitor therapy as soon as possible is the best course of action. A critical problem that may help to explain some of the prior failures is the best time to begin using BACE-1 inhibitors [68]. Furthermore, recent studies report multitarget approaches focused on BACE-1, whose ligands are synthesized as small molecules that can be used to alter both BACE-1 and other AD-related targets through synergistic pathways due to the complex nature of AD.

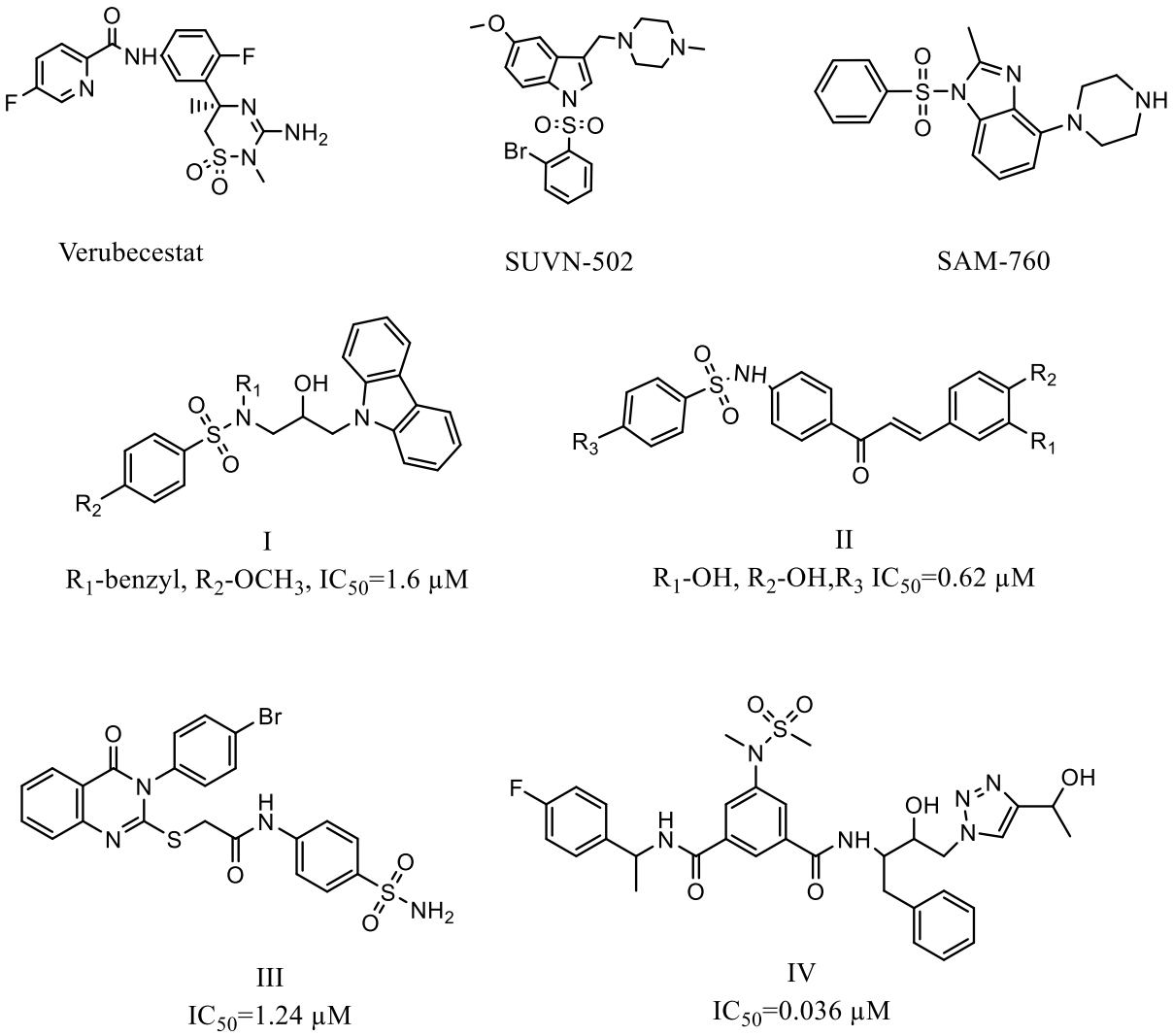


Figure 4.1 Sulphonamides as BACE-1 inhibitors in clinical and pre-clinical studies.

4.1.2 Machine Learning in drug discovery

ML techniques have been increasing and widely adopted in the early stages of drug discovery processes. ML is the branch of AI that focuses on developing and applying computer algorithms that use raw and unprocessed data to perform a specific task [69]. In the field of drug discovery, the applications of ML are growing enormously among a large number of pharmaceutical companies. The goal is to minimize the need for animal testing and primarily use high-throughput screening techniques to reduce the work and assist medication disclosure [70]. ML is classified into four groups based on the methodologies as: Supervised, Semi-

supervised, Unsupervised and Reinforcement learning. These techniques increase decision-making, QSAR analyses, hit discoveries and *de novo* drug designs more accurately. In the ML methodology of drug discovery, there are the following steps in the experimental setup: 1) data collection; 2) generation of descriptors; 3) searching best subset of variables; 4) model training; and 5) model validation [69].

4.1.3 Machine learning algorithms

4.1.3.1 Random Forest

RF is a supervised learning method which is composed by the combination of tree predictors such that each tree depends on the values of a random vector independently and with the same layout for each tree in the forest [71]. Each tree in Random Forest is transverse in a particular way:

1. Giving a training dataset *N*, *n* random samples with repetition taken as training set (Bootstrap).
2. For each node of the tree, *M* input variables are determined, where $m << M$ and the value of *m* remains constant. The node used are the randomly chosen variables.
3. Every tree is generated to its maximum expansion.

4.1.3.2 XGBoost classifier

XGBoost stands for Extreme Gradient Boosting, is an efficient and scalable machine learning Classifier model based on the Gradient Boosting Machine (GBM), providing parallel tree boosting and enhancing performance by using subsampling ratio, learning rate, and maximum tree depth to avoid overfitting. XGBoost defines additional features such as handling missing data with nodes, default directions and specifying efficiently splitting thresholds during split node [72]. XGBoost produces comparable and better predictive accuracy and supports the inherent ability to handle highly diverse and complex descriptors [73].

4.1.3.3 LightGBM

LightGBM is another scalable and flexible GBM approach that shows comparable performance with the other existing boosting tools by learning efficiency and accuracy with lower consumption of memory [74]. LightGBM is a fast, high-performance tree-based learning algorithm, used for both classification and regression tasks. It can reduce the cost of the gain for each split-up in training. In LightGBM, the tree grows vertically and leaf-wise, while most decision-tree learning algorithms grow horizontally and level-wise [75].

In the present work, we have collected a dataset of sulphonamides as BACE-1 inhibitors and then developed and validated an ML model to classify the BACE-1 inhibitors and used this model to screen our *in-house* library of sulphonamides. The identified hits were then screened for BACE-1 activity using an *in-vitro* assay.

4.2 Materials and Methods

4.2.1 Dataset Collection

The dataset for BACE-1 was obtained from BindingDB (https://www.bindingdb.org/), a public web-accessible database [76]. Only the compounds containing the sulphonamide group were selected further. The KNIME analytical tool was used to filter the compounds with multiple entries and IC_{50} values. The compounds having IC_{50} values less than 500 nM were marked as active (recognized as 1), while compounds with IC_{50} more than 500 nM were marked as inactive (recognized as 0). Hence, total of 327 actives and 194 inactive compounds were obtained [77].

4.2.2 Fingerprint Descriptors

KNIME analytical tools were used to generate the fingerprints descriptors for the BACE-1 dataset using Fingerprints and Fingerprints expander nodes. The five fingerprint descriptors viz. MACCS, Estate, PubChem, ECFP4, and ECFP6 were obtained.

4.2.3 Data Splitting

The dataset of BACE-1 inhibitors were split into training (80%), validation (10%), and test (10%) sets by train_test_split by using the scikit learn python module having a random state of 2529. Training dataset was used for model development and other two subsets (i.e., test and validation) were used to evaluate training model performance against new data.

4.2.4 Machine learning classification algorithms

Random Forest (RF), gradient boosting machine (XGBoost), and LightGBM machine learning algorithms were used for classification models using Python library *Scikit learn*. Grid search using *GridsearchCV* was performed to identify the optimal combination of values for the hyperparameters.

4.2.4.1 Random-Forest Classifier

Three different parameter combinations were used to determine the RF, that is the number of trees in the random forest (n_estimators), maximum depth of the tree (max_depth), and minimum number of samples required to split an internal node (min_samples_leaf) (*scikit-learn 1.2.2*). A grid search was performed to obtain the maximum accuracy using following parameters:

- n estimators- 50, 100, 200, 300, 400, and 500
- Maximum depth ranges from 5 to 50 with an increment of 5.
- Minimum sample split ranges from 2 to 10.

4.2.4.2 XG Boost Classifier

XGBoost or Extreme Gradient boosting classifier can work well in smaller datasets (*XGBoost 1.7.5*). A grid search was performed to tune hyperparameters, and based on accuracy score the best model was selected. XGBoost provides large range of hyperparameters such as:

- Maximum depth of a tree (max_depth)- 5,7,9,11,13, and 15.

- Learning rate ranges from 0.01 to 0.1.
- Gamma- 0, 0.25, and 1.
- Lambda (reg_lambda) ranges from 0 to 15.
- scale_pos_weight used for imbalanced classes having values 3,5,7,9, and 11.
- Subsample is the ratio of training instances having a value of 0.8.
- colsample_bytree is the subsample ratio of the column having value of 0.5.
- Tree construction algorithm (tree_method) used 'gpu_hist'.

4.2.4.3 LightGBM Classifier

LightGBM works on a histogram-based algorithm that results in faster and more accurate results compared to XGBoost (LightGBM 3.2.2). The most critical hyperparameters used by the LightGBM are:

- 'num_leaves': 10-50,
- 'reg_alpha': [0.1, 0.5],
- 'lambda_l1': [0-5],
- 'lambda_l2': [0, 1],
- 'min_data_in_leaf': 30-100,
- 'learning_rate': 0.9-0.001

4.2.4.4 Performance Evaluation of models

The methods of performance evaluation of models have been mentioned in the section **3.2.6.**

4.2.5 Screening Database preparation

The *in-house* database of previously synthesized and reported sulphonamides in our lab were prepared using DataWarrior V5.5.0 [78-80]. The database consists of 129 reported sulphonamide derivatives. The database was screened with the best model to identify the hits [81].

4.2.6 BACE-1 inhibition assay

The identified hits were evaluated for their BACE-1 inhibition potential using fluorescence resonance energy transfer (FRET) based BACE-1 fluorescence assay kit (Catalog No. CS0010, Sigma-Aldrich). The kit consists of fluorescent assay buffer, stop solution, substrate (7-Methoxycumarin-4-acetyl [Asn670, Lue671]-Amyloid βA4 Precursor Protein 770 Fragment 667-676-(2,4 dinitrophenyl) Lys-Arg-Arg amide trifluoroacetate salt) and BACE-1 enzyme. Different concentrations of test compounds were prepared. The fluorescence intensity was measured immediately after the addition of BACE-1 enzyme with the wavelength of excitation and emission was set at 320 nm and 405 nm, respectively. All the measurements were performed in triplicate. The percentage inhibition was calculated using the following formulae: [(Io–Ii)/Io] × 100, where Io and Ii are the fluorescence intensities obtained in the absence and presence of an inhibitor, respectively and the IC_{50} values were calculated using linear regression graph (GraphPad Prism 5.1, GraphPad Software Inc.).

4.2.7 Docking study

The docking study was performed to study the binding pose and interaction of the identified hits with the BACE-1 protein.

4.2.7.1 Grid generation and validation

The amino acid residues involved in the protein-ligand interactions of the selected protein (PDB ID-6EQM) were identified by using BIOVIA Discovery studio visualizer. The identified residues were used to construct a grid box around the active site as the reference points. The Autogrid 4.0 was used to calculate grid maps of interaction energies having various atom types present in the ligand (A, C, HD, N, NA, S, OA, Br, Cl and I). The grid size was set to xyz points at 60×60×54, having a grid spacing of 0.336 Å, and the grid centers were placed at the coordinates X: 28.936, Y: 79.442, Z: 18.584, respectively. Further, the obtained grid was validated by redocking ligand (BUH) and the Root Mean Square Deviation (RMSD) value was

calculated between experimentally obtained co-crystalized ligand and docked pose using Maestro. The RMSD was found to be 0.389 Å. Precision docking was performed using AutoDock 4.2 by engaging Lamarckian Genetic Algorithm (LGA) with the genetic algorithm runs kept at 100.

4.3 Result and Discussion

4.3.1 Machine learning models

The training data set had total of 521 compounds out of which 416 were taken for training set and remaining compounds were equally divided into test set and validation set using stratified splitting**.**

4.3.1.1 Random Forest classifier

Random forest is an ensemble of decision trees. The **Table 4.1** and **Table 4.2** summarizes the performance of Random Forest classifiers build using different fingerprints on the training and test set. The summary of hyperparameters of all the models have been summarized in **Table S4 of appendix).** The result indicates that the model build using PubChem fingerprints had the best F1 score of 0.91. The model had an accuracy, precision and recall score of 0.86, 0.84 and 0.98 respectively. The model build using Estate fingerprint showed recall score of 1.0 on training and test set but the precision score was low.

4.3.1.2 XGBoost Classifier

It is an ensemble of several weak classifier that uses a gradient boosting framework. The hyperparameters of the best model for every descriptor has been summarized in **Table S5** of **appendix.** The **Table 4.1** summarizes the performance of XGBoost classifier on training and test set. The accuracy and F1 score of models build using XGBoost classifier was better than that of RF classifier when evaluated on test set. The models built using PubChem fingerprint showed the best F1 score on the test set.

4.3.1.3 LightGBM classifier

It is also a boosting algorithm based on decision tree. It is considered to be fast and less computational memory intensive. The hyperparameters corresponding to each fingerprint for LightGBM model has been summarized in **Table S6 of appendix.** The model build using PubChem showed the best accuracy and F1 score on the test set i.e., 0.87 and 0.92 respectively. The model performed better than the other two algorithms. The summary of performance of models has been given is the **Table 4.1**.

4.3.2 Performance of ML models on validation set

In order to check the robustness of machine learning models it is necessary to evaluate the models on independent validation set. The performance of every algorithm on different type of fingerprint has been summarized in **Table 4.2.** The result indicates that the model build using XGB classifier with PubChem fingerprints showed the best performance on the external validation set with accuracy, precision, recall and F1 score of 0.89, 0.88, 0.99 and 0.93, respectively. The best model was selected and feature importance were calculated. The top 20 features and their importance has been summarized in a figure **(Figure S2 of appendix)** and the top ten fragments have been shown in **Figure 4.2**.

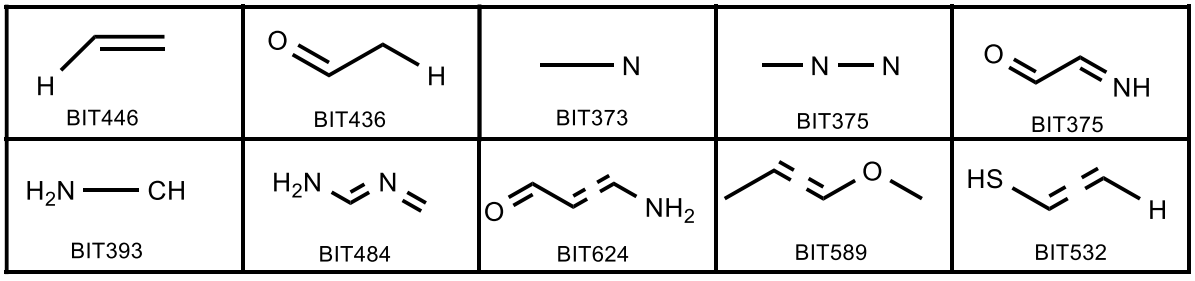


Figure 4.2 Top 10 important PubChem fingerprints

4.3.3 Screening of *in-house* library

The *in-house* library of compounds was screened virtually using XGB classifier build using PubChem fingerprints. The cut-off was kept to 0.5 which is the default value for several binary classification algorithms. The compounds having score more than the cut-off were marked

Table 4.1 Performance of classification models on the training set

Fingerprints	XGBoost classifier				LightGBM classifier				RF classifier			
	Accuracy	**Precision**	**Recall**	**F1**	**Accuracy**	**Precision**	**Recall**	**F1**	**Accuracy**	**Precision**	**Recall**	**F1**
MACCS	0.91	0.85	0.98	0.91	0.96	0.96	0.98	0.97	0.88	0.81	0.98	0.89
ECFP-4	0.91	0.90	0.98	0.91	0.99	0.99	1.00	0.99	0.88	0.87	0.97	0.91
ECFP-6	0.91	0.90	0.98	0.91	0.99	0.99	1.00	1.00	0.76	0.78	0.96	0.86
PubChem	0.95	0.94	1.00	0.97	0.98	0.98	0.99	0.99	0.86	0.85	0.99	0.91
Estate	0.83	0.81	1.00	0.89	0.86	0.97	0.96	0.91	0.78	0.78	1.00	0.87

Table 4.2 Performance of classification models on the test set

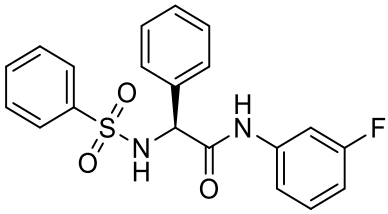
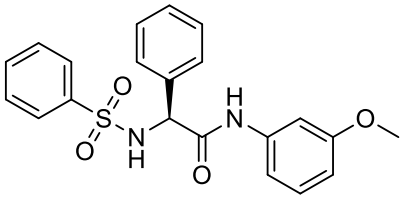
Fingerprints	RF classifier				XGBoost classifier				LightGBM classifier			
	Accuracy	**Precision**	**Recall**	**F1**	**Accuracy**	**Precision**	**Recall**	**F1**	**Accuracy**	**Precision**	**Recall**	**F1**
MACCS	0.84	0.81	0.98	0.90	0.88	0.84	0.99	0.91	0.85	0.87	0.93	0.90
ECFP-4	0.84	0.86	0.93	0.89	0.88	0.87	0.93	0.90	0.90	0.92	0.92	0.93
ECFP-6	0.79	0.87	0.93	0.88	0.88	0.87	0.93	0.90	0.88	0.92	0.92	0.92
PubChem	0.86	0.84	0.98	0.91	0.87	0.87	0.97	0.92	0.88	0.89	0.94	0.92
Estate	0.80	0.79	1.00	0.88	0.82	0.80	1.00	0.89	0.79	0.82	0.92	0.87

Table 4.3 Performance of ML models on validation set

Classifier	Descriptors	Accuracy	Precision	Recall	F1 score
RF	MACCS	0.85	0.83	0.98	0.90
	ECFP-4	0.86	0.86	0.93	0.89
	ECFP-6	0.75	0.78	0.93	0.85
	PubChem	0.83	0.82	0.98	0.89
	Estate	0.80	0.78	1.00	0.88
XGB	MACCS	0.87	0.83	0.97	0.89
	ECFP-4	0.87	0.88	0.94	0.91
	ECFP-6	0.87	0.88	0.94	0.91
	PubChem	**0.89**	**0.88**	**0.99**	**0.93**
	Estate	0.83	0.82	0.97	0.89
LightGBM	MACCS	0.82	0.86	0.90	0.88
	ECFP-4	0.88	0.85	0.89	0.87
	ECFP-6	0.89	0.85	0.89	0.87
	PubChem	0.88	0.89	0.96	0.92
	Estate	0.80	0.83	0.91	0.87

active and were selected as hit. The screening resulted in the identification of two virtual hits i.e., compound-28 (*(S)-(+)-N-(3-fluorophenyl)-2-phenyl-2-(phenylsulfonamido) acetamide*) and compound-37 *((S)-(+)-N-(3-methoxyphenyl)-2-phenyl-2-(phenylsulfonamido) acetamide*), which were previously reported for acetylcholinesterase (AChE) and butyrylcholinesterase (BChE) activity [82]. The summary of the reported properties of both the hits has been given in **Table 4.4**.

Table 4.4 Summary of reported properties for identified hits

Compound code	Structure	% Inhibition at a concentration of 50 µM	
		BChE	**AChE**
28		01.23±0.84	05.82 ± 0.78
37		41.77±0.62	14.03 ± 0.85

4.3.4 *In-vitro* BACE-1 inhibitory activity

The identified hits were evaluated for their BACE-1 inhibition using FRET-based assay kit. The molecules were initially screened to determine the percentage inhibition at 1 µM and then they were screened at different concentrations to determine their IC_{50} values. The compound 28, containing 3-fluorophenyl group, showed IC_{50} of 0.431±0.006 µM and the compound 37, containing 3-methoxyphenyl group showed IC_{50} value of 0.272±0.019 µM (**Table 4.5)**.

Table 4.5 Summary of in-vitro and docking result of ligands with BACE-1 (PDB ID-6EQM)

Compound code	**hBACE—1 IC_{50}(µM)±S.D.[a]**	**Binding energy (Kcal/mol)**	**Ligand efficiency (Kcal/mol)**	**Interactions (PDB ID-6EQM)**
Compound 28	0.431±0.006	-7.66	-0.284	Arg307 (H-bond), Thr232 (H-bond), Gly230 (H-bond), Tyr71 (Pi-Pi T-shaped), Ala396 (Pi-Sigma)
Compound 37	0.272±0.019	-7.58	-0.271	Asp32 (H-bond), Leu30 (Pi-alkyl), Phe108 (Pi-Sulfur), Asp228 (Pi-anion), Tyr198 (Pi-Pi T-shaped)

[a]Data expressed in Mean±S.D.(n=3)

4.3.5 Docking study

The grid validation was performed by redocking the co-crystallized ligand and calculating the RMSD between the docked pose and co-crystallized ligand. The RMSD value was found to be 0.389 Å. The docked pose and co-crystallized ligand have been represented in **Figure S3 of appendix** The docking study revealed that the compound 28 and compound 37 had binding energy of -7.66 and -7.58 Kcal $mol^{-1,}$ respectively. Their interaction diagram revealed that the compound 28 showed H-bond interaction with Arg307 and Thr232. The compound 37 showed interaction with the catalytic dyad i.e., Asp32 and Asp228 via. H-bond and Pi-anion interactions, respectively. The summary of docking result containing binding energy, ligand efficiency and interactions has been represented in **Table 4.5**.

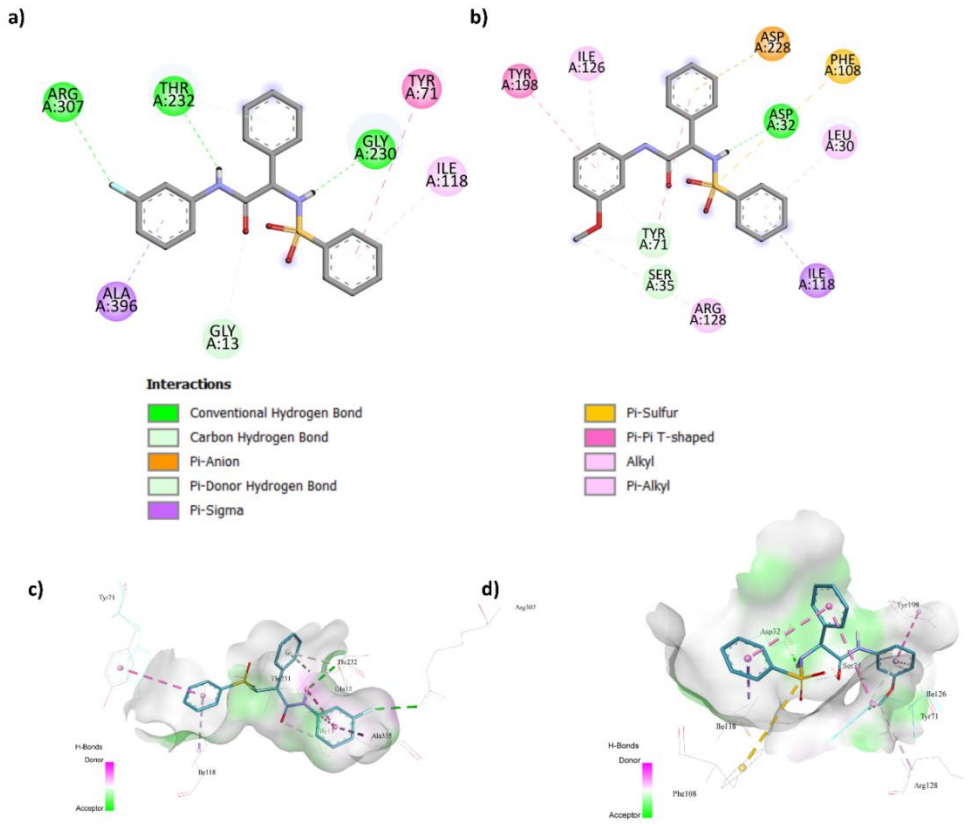


Figure 4.3 2D interaction diagram of (a) compound 28 and (b) compound 37 and 3D interaction diagram of (c) compound 28 and (d) compound 37.

4.4 Conclusion

BACE-1 is a promising target for the treatment of AD. Several sulphonamide-based BACE-1 inhibitors have shown potential for decelerating the long-term progression of AD. Drug discovery pipelines are extremely long and complicated process. In this study, a ML model was developed using to classify the BACE-1 inhibitors. The classification was based on the range of IC_{50} value. The compounds having IC_{50} value less than 500 nM were marked as active and the compound having IC_{50} value more than 500 nM were marked as inactive. The best ML model had accuracy, precision, recall and F1 score of 0.89, 0.88, 0.99 and 0.93 on the validation set. The model was built using the XGboost algorithm on PubChem fingerprints. The model was used to screen the *in-house* library of potential sulphonamides as BACE-1 inhibitors. Upon screening, we obtained two hits,

i.e., Compound 28 and compound 37, which were previously reported as weak AChE and BuChE inhibitors. Both compounds were evaluated for their *in-vitro* BACE-1 activity. The compound 28 showed an IC_{50} value of 0.431±0.006 µM and compound 37 showed an IC_{50} value of 0.272±0.019 µM. The docking study revealed compound 37 interacted with the catalytic dyad of BACE-1, i.e., Asp32 and Asp228. Thus, the developed model has shown reliable prediction and further studies can be done on the identified hits to make them potential lead molecules.

Chapter 5

Natural-Language Processing (NLP) based feature extraction technique in Deep-Learning model to predict the Blood-Brain-Barrier permeability of molecules

Summary:

Blood-Brain-Barrier (BBB) permeability is one of the critical factors in the success and failure of CNS drug development. The most accurate method of measuring BBB permeability involves clinical experiments, which are labour-intensive and time-consuming. Thus, numerous efforts were made to use artificial intelligence (AI) to predict molecules′ BBB permeability. Most of the previous models are based on calculated descriptors and molecular fingerprints. In the present work, we have developed an NLP-based feature extraction technique in Deep-Learning models to predict BBB permeability. We have used the B3DB database and generated SELFIES to extract features from the molecules. We have employed word level and N-gram tokenization to represent words into numeric vectors. The extracted features were fed into several Artificial Neural Network (ANN) and Bi-directional Long Short-Term Memory (LSTM) models. The model, ANN-10 built using ANN and 6-gram tokenization, performed best on the independent test set. The accuracy, precision, recall, F1, specificity and AUC of ROC scores were found to be 0.89, 0.91, 0.91, 0.91, 0.85 and 0.90. Thus, the developed model can be used for the early screening of CNS drugs.

5 Natural-Language Processing (NLP) based feature extraction technique in Deep-Learning model to predict the Blood-Brain-Barrier permeability of molecules

5.1 Introduction

The blood-brain-barrier (BBB) has been a bottleneck for discovering neurotherapeutics. Many drug molecules fail in the drug discovery and development phase due to the inability to cross the BBB. The two well-described barriers are the vascular BBB and the blood-cerebrospinal fluid (blood-CSF) barrier. The former consists primarily of the capillary bed, and the latter consists primarily of the choroid plexus [83]. The BBB regulates the entry of substances into the brain that are critical to CNS function. The barrier is created by the tight junctions between the cerebral endothelial cells, the epithelial cells of choroid plexus and the cells of the arachnoid epithelium [83]. The BBB only allows lipophilic molecules with a molecular weight of less than 500 Da from the bloodstream to enter the brain; other molecules require specific transport systems, which help move across membranes [84]. There are five basic mechanisms involved in the movement across the membranes. The first is passive (or simple) diffusion, in which the drug molecules travel based on the concentration gradient from a higher drug concentration to a lower concentration. Usually, molecules with high lipophilicity and low molecular weight can quickly diffuse across the BBB. Blood gases and several drugs, such as anaesthetics, crosses BBB in this manner [85]. Secondly, carrier-mediated transporters involve active and facilitated diffusion. These are ATP-binding transporters, such as P-glycoprotein (Pgp). Pgp is expressed naturally on the plasmatic membranes of endothelial cells. It protects the brain from harmful substances. Thirdly, the solute carriers (SLC) facilitate the movement of solutes bi-directionally across the cell membrane. SLC transport is carried out by concentration gradients or electrochemical gradients [86].

Some drugs, such as L-DOPA, enter the brain through such transporters [87]. Fourthly, receptor-mediated transcytosis (RMT) includes binding the macromolecules, such as proteins and peptides, to the receptors. This system utilizes the vesicular transport of endothelial cells to transport the substrates in the brain [88]. Finally, transendothelial migration or diapedesis of leukocytes, where the leukocytes penetrate the BBB directly through the cytoplasm without the disruption of tight junctions. Once the diapedesis of a leukocyte commits, it does not go back to the same cell [86] [89] **(Figure 5.1)**.

The *in-vivo* experiments to study the BBB permeability are one of the most accurate and reliable methods, but it is expensive and time-consuming. It is also challenging to perform such experiments for high throughput screening (HTS) in the early research stages.

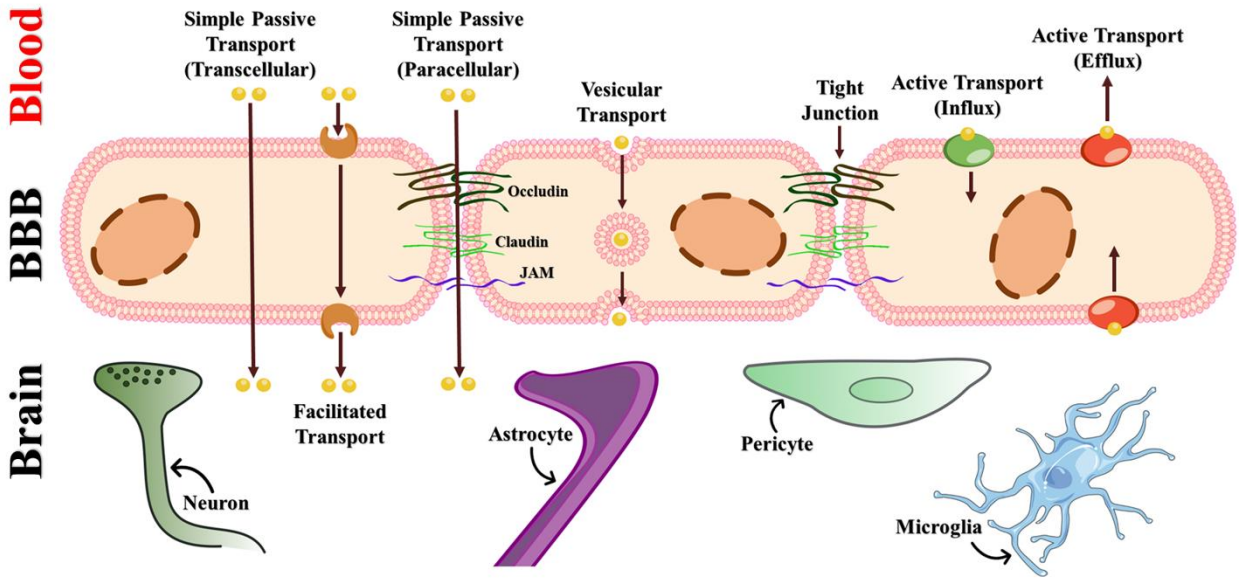


Figure 5.1 Basic mechanism involved in the transport of drugs across the BBB

Deep-Learning is a subset of machine learning which employs more than three layers of neural networks to learn from a massive amount of data. Deep-Learning eliminates most of the data preprocessing usually involved with machine learning. The neural networks can ingest and process unstructured data like texts or images and automate feature

extraction. In the present study, we have applied Artificial Neural Network (ANN) and Recurrent Neural Network (RNN) to predict the BBB permeability of a drug.

There are two types of *in-silico* models to predict BBB permeability, one is quantitative, and the other is qualitative. The quantitative models predict the logarithm of the drug's bain-plasma concentration ratio (LogBB), whereas the qualitative model predicts whether a compound is BBB permeable (BBB+) or not (BBB−). In the past several classification models have been build using machine learning and deep-learning algorithms. Kumar *et al*. developed the DeePred-BBB model using 3971 compounds with the help of physicochemical properties and substructure fingerprints. The model showed an accuracy of 98.07% using deep neural network (DNN), but the results were not validated on an external validation set [90].

Similarly, Alsenan *et al*. developed a Recurrent Neural Network (RNN) model using the dataset of 2350 compounds which was acquired from Wang et al. showed overall accuracy of 96.53% and a specificity score of 98.08 % [91]. A lightBBB model was developed by Shaker et al., which was trained using Light Gradient Boosting Machine (LGBM) on 7162 compounds and showed an overall accuracy of 93% and specificity of 77% [92]. The dataset used in the study consisted of several duplicates and molecules that could not be recognized by RDKit [93].

Another critical factor while developing predictive models is the use of molecular descriptors and fingerprints. The descriptors significantly affect the model's performance and pose a problem while deploying the model to the use case. It causes dependency on other packages and software to calculate the descriptors and then feed them into the model, thus making it difficult to use. One way to overcome this is to apply Natural Language Processing (NLP) on the chemical identifiers to extract meaningful information for the model. In a study performed by Parakkal *et al.*, they used a pre-trained model

Mol2Vec to generate the NLP-based descriptors. The AUC of the best model was found to be 0.96 [94]. In the present work, we have employed different NLP techniques to extract information from the molecules and then applied some deep-learning algorithms to train the model and get reliable predictions.

5.2 Methods

The deep-learning models were built on Python 3.10.7 using standard libraries and TensorFlow v2.11.0.

5.2.1 Dataset

The B3DB dataset, curated by Meng *et al.*, was used to develop the models. The dataset consists of 7807 ligands. This is the largest BBB dataset we know. The chiral specifications of the molecules were retained by using isomeric SMILES [93]. The dataset was checked for any missing or duplicated records. After careful curation, the dataset consisted of 4956 BBB permeable and 2851 BBB non-permeable compounds (Fig.2). The class label for BBB permeable was assigned as '1' and BBB non-permeable as '0'.

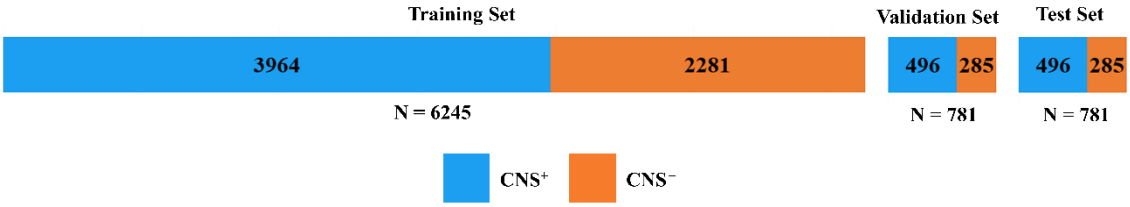


Figure 5.2 Distribution of dataset into training, validation and test set

5.2.2 Features extractions

The SMILES were converted to a more robust Self-Referencing Embedded String (SELFIES) to describe the molecular graph in a machine language. Each SELFIES corresponds to a valid molecule; every molecule can be described as SELFIES [95]. As deep-learning models work on differentiable functions which can process only the

numerical inputs in an array or tensor form. So the SELFIES need to be converted into numeric form before applying deep-learning algorithms. The SELFIES were converted into numerical tensors using vectorization. The SELFIES are split into units (called tokens), and every unit is indexed. The process of feature extraction from SELFIES has been summarized in **Figure 5.3.** For example, N-methylpropan-2-amine can be represented as 'CNC(C)C' in terms of SMILES. The same will be converted into SELFIES, which will be represented as '[C][N][C][Branch1][C][C][C]'. The SELFIES were split between '][' to generate units viz., '[C]', '[N]', '[C]', '[Branch1]', '[C]', '[C]', '[C]'. Now these individual units are tokenized.

The tokenization of SELFIES was done in two ways (i) Word level tokenization and (ii) N-gram tokenization. In the case of word-level tokenization, the sequence of the string is retained; such types of models are also called 'sequence models'. On the other hand, in N-gram tokenization, the tokens are a group of N-consecutive words; for example, '[C] [N]' is a 2-gram token. Similarly, 1 to 6-gram tokens were generated. This approach treats the input as a set, discarding the original order; such models are called 'bag-of-words models. Once the SELFIES is split into tokens then they are hashed into a fixed binary vector. An index of all the unique tokens found in the training set is made and a number is assigned to them. If any new token comes up while tokenizing training set, it will be assigned as 'UNK'. In case of N-gram tokenization every combination is taken as individual vocabulary and the same is indexed. Now considering the above example the feature vocabulary will contain following unique tokens i.e. '[C]', '[N]' and '[Branch1]'. These tokens will be given a specific index number which will be used to assign the index corresponding to token present in the molecules of training and test dataset. The tokens were then one-hot encoded, or multi-hot encoded to create numeric vectors. More information was added by using the Term frequency-inverse document frequency (TF-

IDF) normalization. TF-IDF is a metric that weights a given term by taking term frequency, how many times the string pattern appears in a current molecule divided by the measure of frequency in the entire dataset [96]. Essentially, it measures the importance of a feature by comparing its frequency within a molecule with the entire training dataset.

$$TF(w) = \frac{No.of\ times\ the\ string\ w\ appears\ in\ a\ molecular\ representation}{Total\ number\ of\ string\ in\ a\ molecular\ representation}$$

$$IDF(w) = log\frac{Total\ number\ of\ molecules\ in\ training\ set}{Number\ of\ molecules\ comtaining\ the\ string\ w}$$

So, the weight of a string 'w' in molecule 'm' is given by the following TF-IDF weighting:

$$weight\ (w,m) = TF(w,m) * IDF(w)$$

There also exists an advanced string representation method called word embedding. Contrary to the previously mentioned techniques, word embeddings are low-dimensional and dense floating point vectors. In addition to being dense, word embeddings are structured representations, and their representation is learned from the data. They encode semantic relationships between features based on their contextual usage.

5.2.3 Neural Network Architecture

In the present study, we employed ANN and RNN to predict BBB permeability. ANN is a type of deep-learning model which employs multiple perceptrons at each layer. This is also called a feed-forward neural network because the direction of input processing is only in the forward direction [97]. On the other hand, RNNs are more complex; they save the output of a processing layer and feed the result back into the model. Bi-directional Long Short-Term Memory (LSTM) is a type of RNN model that can process the sequence information in both forward and backward directions.

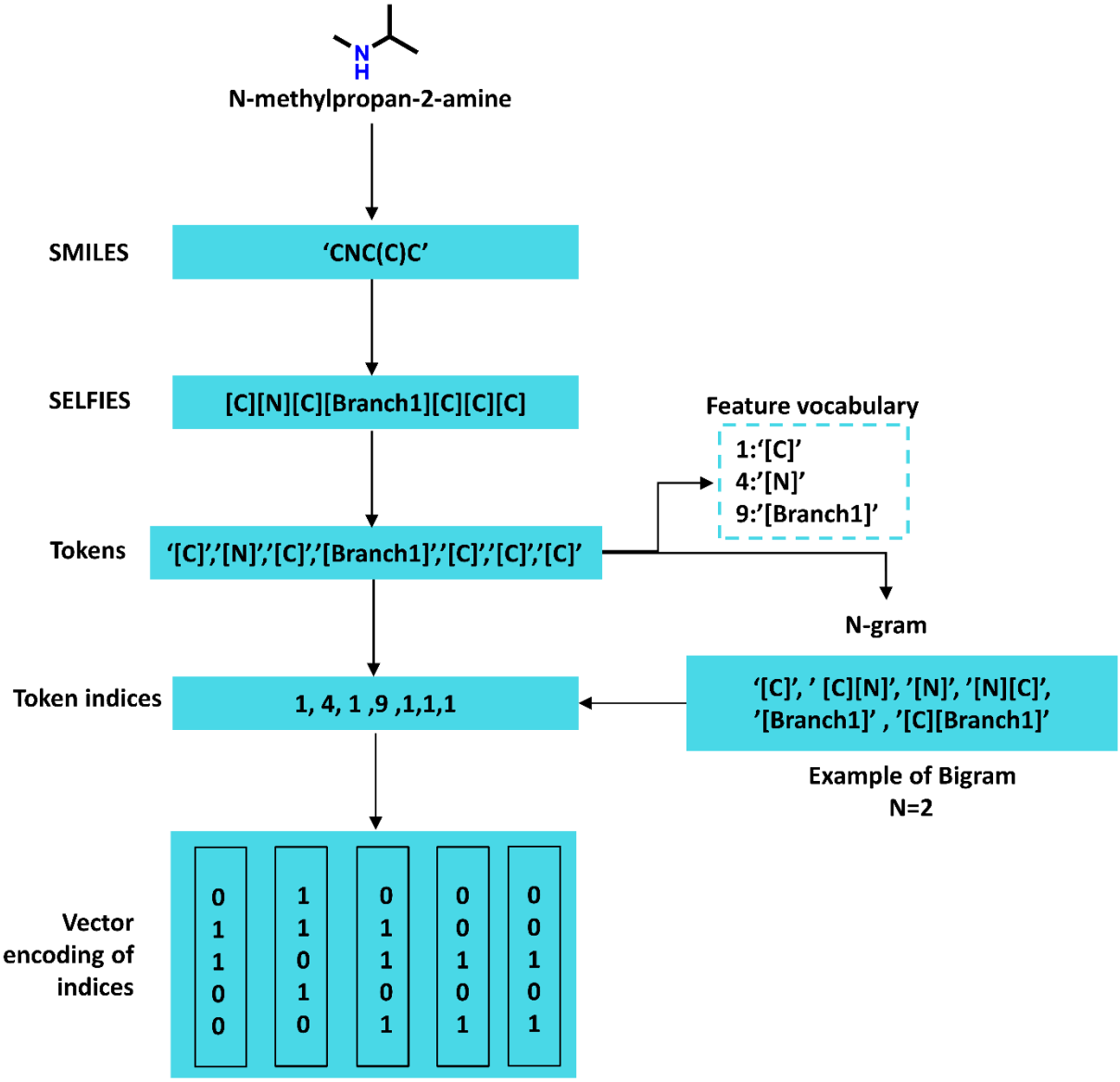


Figure 5.3 Steps of feature extraction from SELFIES

5.2.3.1 Input layer

The features extracted from the previous section were used as an input layer in the deep-learning models.

5.2.3.2 Hidden layer

The hidden layer in the case of ANN was made up of 6 layers, out of which three were dense with 512 units and three were dropout layers with a dropout rate of 0.2. The models were also tested with a dropout rate of 0.2, 0.3, 0.4 and 0.5. Dropout regularization was

used to prevent the models from overfitting. The dropout layers randomly drop the nodes as per the dropout rate. The Rectifier linear unit (ReLU) activation function was used for the dense layers **(Figure 5.4(a)).** The number of hidden layers was optimized (2-10 hidden layers); any reduction in the number of hidden layers led to a reduction in model performance, and an increment in the number of hidden layers did not show any improvements. Similarly, the number of neurons was also optimized (128, 256,512, 1024 and 2048).

In the case of the bidirectional LSTM model, the inputs were fed into an embedding layer which was then connected with a bidirectional LSTM layer of 128 units. The activation function in the forward direction was Tanh, and the recurrent activation function was selected as sigmoid. The LSTM layer was then connected with the dropout layer with a dropout rate of 0.2 **(Figure 5.4(b)).**

5.2.3.3 Output layer

The output layer comprised a dense layer with one neuron and 'sigmoid' as an activation function. The output layer gives predictions in the form of prediction probability which was then converted into classes (P>0.5 as BBB+ and P<0.5 as BBB-). The summary of parameters of both ANN and bi-directional LSTM model has been mentioned in **Table 5.1** and **Table 5.2,** respectively.

5.2.4 Model parameters

The models were trained with the RMSProp optimizer, and the learning rate was set to 0.001. Several learning rates were tried (0.01, 0.05, 0.001, 0.005 and 0.0001) but the best result was obtained with learning rate of 0.001. The optimizer minimizes the loss value from the binary cross-entropy cost function. Early stopping was used to prevent the model from overfitting. While performing early stopping, the patience was set to 25 epochs.

Patience is the number of epochs after which the training will stop if there is no improvement in the validation loss. The batch size was kept at 32.

5.2.5 Evaluating the performance of DL models

To evaluate the performance of binary classification models, we used the following metrics:

(a) Accuracy = $\frac{True\ positives + True\ negatives}{True\ positives + True\ negative + False\ positives + False\ negatives}$

(b) Precision=$\frac{True\ positives}{True\ positive + False\ positives}$

(c) Recall= $\frac{True\ positives}{True\ positive + False\ Negatives}$

(d) F1 score= $2\frac{Precision * Recall}{Precision + Recall}$

(e) Specificity= $\frac{True\ negatives}{False\ positive + True\ negatives}$

(f) The area under the ROC curve (AUC of ROC) - This is a standard binary classification evaluation metric. It demonstrates the model's ability to split the two classes through sensitivity and specificity.

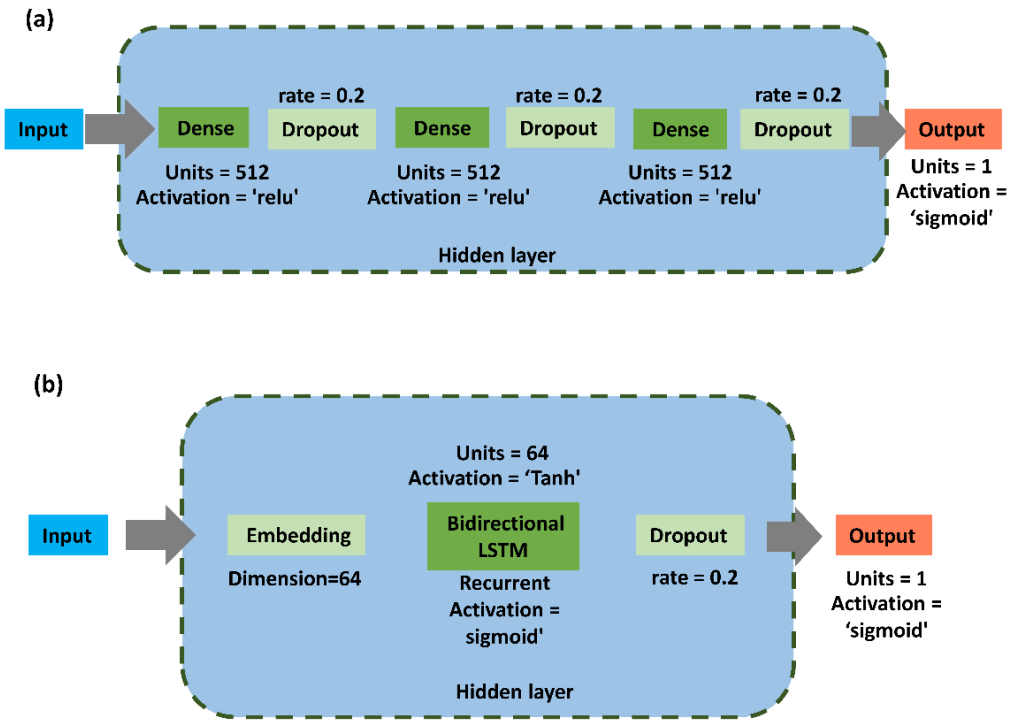


Figure 5.4 Neural network architecture of (a) ANN (b) LSTM

Table 5.1 Summary of the architecture of ANN model

Layer	Type	Output shape	No. of Parameters
Input	Input	(None, 235)	0
Hidden	Dense	(None, 512)	120832
	Dropout	(None, 512)	0
	Dense	(None, 512)	262656
	Dropout	(None, 512)	0
	Dense	(None, 512)	262656
	Dropout	(None, 512)	0
Output	Dense	(None, 1)	513

Table 5.2 Summary of the architecture of LSTM Model

Layer	Type	Output shape	No. of Parameters
Input	Input	(None, 235)	0
Hidden	Embedding	(None, 235, 64)	5952
	Bidirectional	(None, 128)	66048
	Dropout	(None, 128)	0
Output	Dense	(None, 1)	129

5.3 Result and Discussion

In the present work we have converted SMILES to SELFIES. SELFIES possess several advantages over SMILES, such as its robust molecular string representation. SELFIES never produces any invalid molecule, the combinations used in SELFIES alphabet map to a chemically relevant graph. In the case of SMILES, there is syntactic invalidity of unbalanced parentheses or ring identifiers. In the present work, if the feature were extracted based on SMILES, there could be a possibility that there may be a string with open parentheses but no corresponding closed parentheses, thus leading to an invalid molecular graph.

The text vectorization process led to the extraction of 91 unique features from the SELFIES representation. Any new feature which might appear in the test set will be assigned with 'UNK'. A list of all the extracted features has been summarized in a table (**Table S7 of appendix).** The performance of 17 deep-learning models (10-ANN and 7-bidirectional LSTM) was compared to select the best model and text representation technique.

Table 5.3 summarizes the result of loss and accuracy for the training set and test set of ANN models. It can be observed that the models with N-gram tokenization performed better than the models with word-level tokenization. The models viz., ANN-5, ANN-7, ANN-9 and ANN-10 showed an accuracy of 0.88 on the validation set. On the other hand, bidirectional LSTM models did not perform well as compared to the ANN models. The model LSTM-4, built using word-level tokenization with tf-idf encoding, showed the best accuracy among the bidirectional LSTM models. The accuracy of LSTM-4 on the validation set was found to be 0.84. Thus, the four models, ANN-5, ANN-7, ANN-9 and ANN-10, which performed better, were selected to evaluate their performance on an independent test set. The plot of loss and accuracy across epochs for the chosen models has been given in Figure 5.5**.** The plots for the remaining models have been given in the SI **(Figure S4-S10 of appendix)**. The summary of performance models on the test set has been summarized in **Table-5.** The raw confusion matrix has been shown in **Figure S10 of appendix**. The model ANN-10 showed the best accuracy, precision, and F1 score on the test set. The AUC of ROC plot has been represented in Figure 5.5**.** Thus, the model ANN-10 was selected from our study and used to make the predictions.

Table 5.3 Summary of the performance of ANN models on the Training set and Validation set

Model	Tokenization Layer	Output type	Training Loss	Validation Loss	Training Accuracy	Validation Accuracy
ANN-1	Word level	int	0.39	0.62	0.87	0.83
ANN-2	Word level	Multi-hot encoding	0.33	0.58	0.85	0.82
ANN-3	Word level	Count	0.25	0.41	0.87	0.86
ANN-4	Word level	tf-idf	0.24	0.41	0.88	0.85
ANN-5	**N-gram (N=1)**	**tf-idf**	**0.28**	**0.40**	**0.88**	**0.88**
ANN-6	N-gram (N=2)	tf-idf	0.25	0.68	0.97	0.82
ANN-7	**N-gram (N=3)**	**tf-idf**	**0.20**	**0.82**	**0.97**	**0.88**
ANN-8	N-gram (N=4)	tf-idf	0.28	0.55	0.92	0.85
ANN-9	**N-gram (N=5)**	**tf-idf**	**0.22**	**1.20**	**0.97**	**0.88**
ANN-10	**N-gram (N=6)**	**tf-idf**	**0.23**	**1.10**	**0.97**	**0.88**

Table 5.4 Summary of the performance of LSTM models on the Training set and Validation set

Model No.	Tokenization Layer	Output type	Training Loss	Validation Loss	Training accuracy	Validation Accuracy
LSTM-1	Word level	int	0.24	0.38	0.89	0.83
LSTM-2	Word level	Multi-hot encoding	0.64	0.66	0.8	0.77
LSTM-3		Count	0.34	0.44	0.84	0.83
LSTM-4		tf-idf	0.38	0.46	0.85	0.84
LSTM-5	N-gram (N=1)	tf-idf	0.37	0.45	0.84	0.83
LSTM-6	N-gram (N=2)	tf-idf	0.56	0.61	0.82	0.81
LSTM-7	N-gram (N=3)	tf-idf	0.49	0.49	0.82	0.80

Table 5.5 Summary of performance on the test set

Model	Accuracy	Precision	Recall	F1 Score	Specificity	AUC of ROC
ANN-5	0.87	0.87	0.93	0.90	0.87	0.92
ANN-7	0.88	0.90	0.90	0.90	0.85	0.88
ANN-9	0.86	0.85	0.94	0.89	0.89	0.86
ANN-10	**0.89**	**0.91**	**0.91**	**0.91**	**0.85**	**0.90**

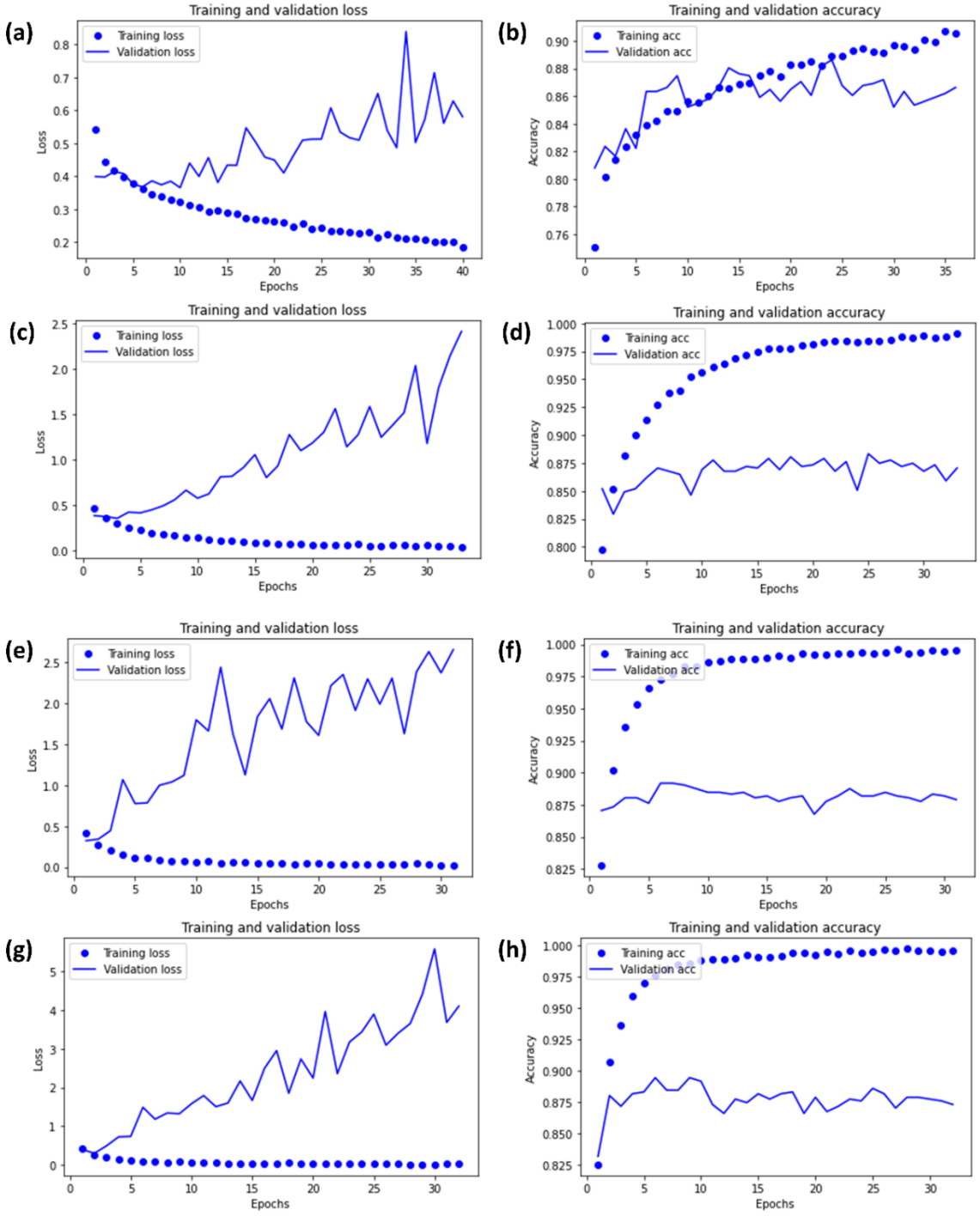


Figure 5.5 Performance of Model while training: (a) Training loss and validation loss for ANN-5 (b) Training and Validation accuracy for ANN-5 (c) Training loss and validation loss for ANN-7(d) Training and Validation accuracy for ANN-7 (e) Training loss and valid

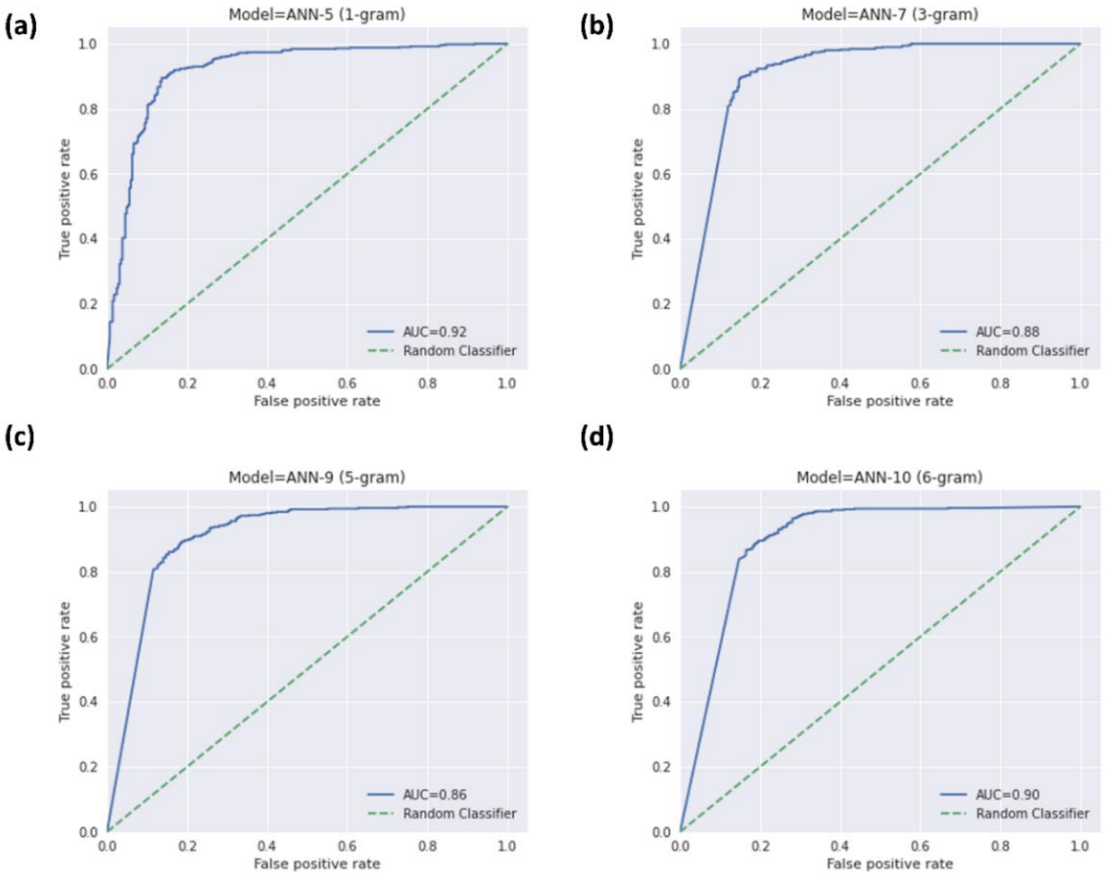


Figure 5.6 ROC curve of models on the test set. (a) ANN-5 (b) ANN-7 (c) ANN-9 (d) ANN-10

Fast model overtraining can be observed in the developed models, i.e. a difference in the model's performance on training and test set. This can be due to a large and underrepresented dataset and complex model. The implications of fast model overtraining can be significant such as poor generalization, increased sensitivity to noise and reduced model interpretability.

The performance of our best classification model was compared with the TEMPO model of Ghosh *et al.* [98]. The comparison in terms of accuracy has been summarized in Table 5.6**.** Our model was also compared with the BBB-score model developed by Gupta *et al.* [99]. Table 5.7 summarizes the accuracy score of our ANN-10 model and BBB-score

model. Shaker *et al.* developed a LightBBB model using the LightGBM algorithm to predict BBB permeability, having an accuracy of 89% [100].

Table 5.6 Accuracy of the TEMPO model and our best ANN-10 model

Model	Accuracy on CNS test set	Accuracy on Non-CNS test set
TEMPO	72.4%	71.5%
ANN-10	90.2%	87.6%

Table 5.7 Accuracy of BBB-score model and our best ANN-10 model

Model	Accuracy on CNS test set	Accuracy on Non-CNS test set
BBB-score	80%	72%
ANN-10	90%	88%

An online web server has been developed and is publicly available for predicting the BBB permeability of desired compounds. The user needs to enter the SMILES of the compound, and the server returns whether the query molecule is BBB permeable or not.

5.4 Conclusion

In the presented work, we have proposed an NLP-based deep-learning model built using ANN and LSTM to predict the BBB permeability of the drug. The B3DB dataset was used to build the model. Several NLP based feature extraction techniques were employed, and the models which were built with N-gram tokenization performed better than other models. The ANN models performed better than the bidirectional LSTM model. The accuracy, precision, recall, F1, specificity and AUC of ROC score of the best model (ANN-10) was found to be 0.89, 0.91, 0.91, 0.85 and 0.90, respectively. The data and

code can be accessed from https://github.com/ravisingh15/BBB-Deeplearning. This tool can be helpful in early drug-discovery research to identify molecules which can cross the BBB.

Chapter 6

Design, synthesis and biological evaluation of *N-benzylpiperidines* as potential multitargeted ligands for the treatment of Alzheimer's Disease

Summary:

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by the accumulation of amyloid-β (Aβ) plaques and neurofibrillary tangles, leading to progressive cognitive impairment. The multi-faceted pathology of AD involves cholinergic deficits and increased activity of enzymes such as AChE and BACE-1. This study aimed to develop novel multi-target directed ligands (MTDLs) targeting both AChE and BACE-1 through a structure-based drug design approach utilizing N-*benzylpiperidine* scaffolds. Molecular docking studies revealed promising binding affinities and ligand efficiencies for the designed compounds against the target enzymes. *in-vitro* assays demonstrated potent inhibitory activities against AChE, and BACE-1, with notable structure-activity relationships. Selected compounds exhibited mixed type of AChE inhibition, dose-dependent inhibition of propidium iodide displacement from the AChE peripheral anionic site (PAS), and favorable blood-brain barrier permeation. Additionally, compounds **72** and **77** effectively inhibited AChE-induced $A\beta_{1-42}$ aggregation. *In-vivo* studies in rats revealed cognition-enhancing effects, reduction in AChE activity, and increased acetylcholine levels for compounds **72** and **77**. Furthermore, these compounds demonstrated protective effects against $A\beta_{1-42}$-induced memory deficits in the Morris water maze test. Overall, the findings support the potential of these MTDLs as promising therapeutic candidates for AD, warranting further preclinical and clinical evaluation.

6 Design, synthesis and biological evaluation of *N-benzylpiperidines* as potential multitargeted ligands for the treatment of Alzheimer's Disease

6.1 Introduction

Dementia has emerged as a significant health and societal dilemma in contemporary times, and its impact is poised to deepen further as the global population ages. The number of individuals affected by dementia, with AD being the most prevalent form, is projected to surge dramatically from 35 million to a staggering 135 million by 2050. Currently, 60% of these cases are found in low- and middle-income countries, a percentage that is expected to increase to 71% by 2050 [101].

The absence of a cure exacerbates the challenges posed by AD. Although there are medications available, including three AChEI and memantine, they can only alleviate dementia symptoms and are incapable of halting the degenerative progression.

Recently, a novel approach known as multi-targeted inhibition has been embraced, aiming to simultaneously target multiple enzymes with a single molecule. This strategy is based on the understanding that AD is a complex disorder with involvement of multiple factors. Consequently, the design and development of multitarget-directed ligands hold great promise in effectively slowing down the progression of the disease rather than merely addressing its symptoms.

Among the various targets, AChE inhibition stands out as the most promising for improving cognitive dysfunctions exacerbated in AD. Additionally, BACE-1 inhibition is another crucial target, playing a significant role in cleaving APP, generating Aβ aggregates, and forming senile plaques. Furthermore, as Aβ is associated with oxidative damage, molecules with BACE-1 inhibitory potential may also exhibit substantial antioxidant effects.

Donepezil, the most potent AChE inhibitor, contains the N-benzylpiperidine nucleus as its core group, which binds to the catalytic active site (CAS) of AChE. The scientific literature also supports modifying the other terminal part (indanone nucleus) to enhance BACE-1 inhibitory potential.

Based on the information presented, it is proposed that ligands containing the N-benzylpiperidine nucleus could hold multi-functional enzyme inhibitory potential against both AChE and BACE-1, making them promising candidates for addressing AD.

6.2 Design aspects

The design strategy for multitarget-directed *N-benzylpiperidins* was based on hybrid pharmacophore approach. The cholinergic and Amyloid-ß ($Aß_{1-42}$) hypothesis are key approaches to treat AD. A set of MTDLs was generated as the structural template of donepezil, the most potent and widely prescribed drug for treating advanced-stage AD [102]. An exploration of the Structure-Activity Relationship (SAR) of donepezil revealed a substantial reduction in AChE inhibitory activity upon modifications or substitutions in the benzyl portion of the *N-benzylpiperidine* moiety. Additionally, the *N-benzylpiperidine* ring of donepezil was found to extend into the catalytic anionic site (CAS), while the terminal indanone nucleus was positioned in the Peripheral Anionic Site (PAS) of AChE [102]. Furthermore, the *N-benzylpiperidine* ring contains a basic nitrogen moiety that becomes protonated at physiological pH, enhancing affinity towards AChE and the aspartate dyad of BACE-1 and demonstrating significant blood-brain barrier (BBB) permeability due to acid-base equilibrium (23). These encouraging findings motivated the selection of the *N-benzylpiperidine* nucleus as a core group for designing novel hybrid compounds (**Figure 6.1**). The *N-benzylpiperidine* was chosen as an initial prompt and then was the molecule was allowed to grow and new molecules were generated using CReM structure generation tool (**Figure 6.2**) [103]. The generated molecules were then

subject to ML-based screening using previously developed models. This led to the identification of a hit, which was further optimized.

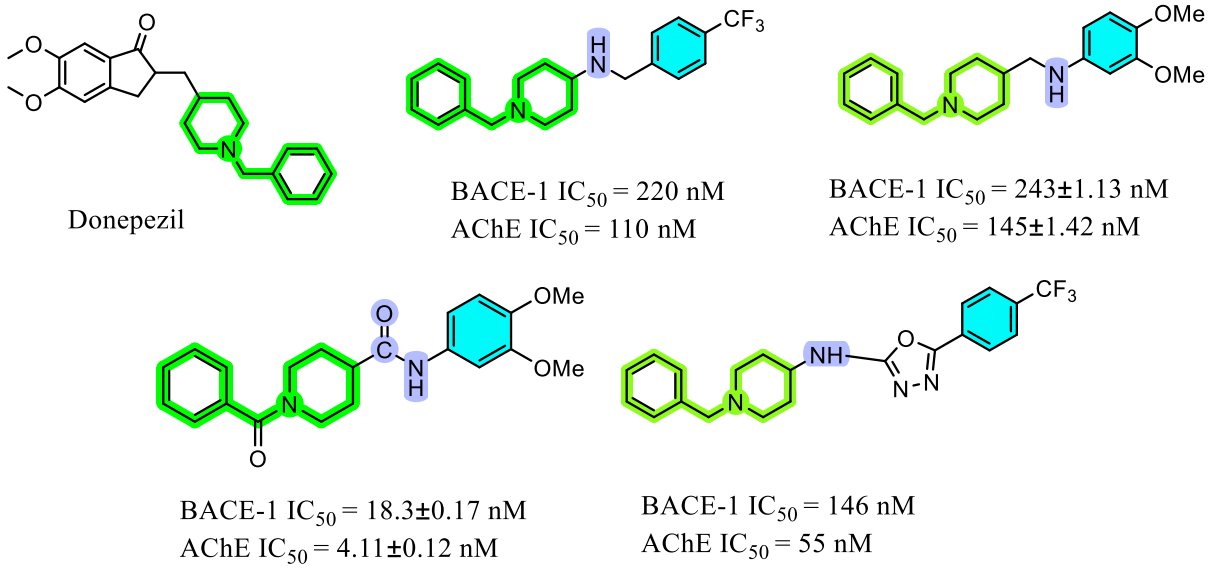


Figure 6.1 N-benzylpiperidines as AChE and BACE-1 inhibitors

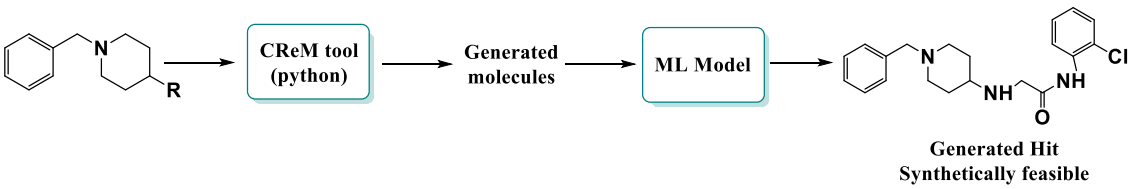


Figure 6.2 Workflow of molecular generation tool

6.3 Material and methods

6.3.1 Docking studies

The designed molecules were docked against AChE and BACE-1 using Autodock. The 3D coordinates of ligands were created using Rdkit based on the SMILES strings of the compounds. Subsequently, energy minimization was performed utilizing the MMFF94s forcefield, and the results were then transformed into Mol2 format using Open Babel. Subsequently, Autodock Tools-1.5.6 was used to convert the ligands to PDBQT format.

6.3.1.1 Docking study with AChE

PLIP was employed to pinpoint residues involved in non-covalent interactions with donepezil within the chosen PDB file (4ey7). These identified residues were used as a basis to create a grid box encompassing the active site. Autogrid 4.0 was then employed to compute grid maps detailing interaction energies concerning various atom types found in ligands (A, C, HD, NA, N, OA, S, Br, Cl, and I). The grid dimensions were configured at 74 × 70 × 82, with a grid point spacing of 0.375 Å. The Lamarckian Genetic Algorithm (LGA) was employed for conformational search. The number of GA runs was set to 50, the population size was set to 150, the maximum number of generations was set to 270000 and the maximum number of evaluations was set to 2500000.

6.3.1.2 Docking study with BACE-1

The docking protocol has been described in section **4.2.7**.

6.3.2 Molecular property and toxicity prediction

Molecular characteristics, including cLogP, cLogS, hydrogen bond acceptor, donor, and drug-likeness, were examined using DataWarrior in accordance with Lipinski's rule of five. Our previous model was utilized to perform BBB permeability *in-silico*. Hazardousness, mutagenicity, tumorigenicity, and irritancy of compounds in the investigation were also forecast using OSIRIS DataWarrior [81].

6.3.3 *In-silico* ADME property analysis

The *in-silico* ADME property analysis was performed using the SWISS ADME web-server accessible at http://www.swissadme.ch [104]. The following properties were calculated viz. Aqueous solubility, GI absorption, BBB permeant, CYP2C9 inhibitor, CYP2D6 inhibitor, CYP3A4 inhibitor and PAINS alerts.

6.3.4 Synthesis and Characterization

We procured analytical grade chemicals and solvents from various suppliers, including Avera, Alfa Acer, Merck, Sigma Aldrich, and SD. To monitor the progress of chemical reactions, we employed thin-layer chromatography (TLC) using pre-coated silica gel 60 F254 plates from Merck KGaA. Visualization was achieved using either ultraviolet light at 254 nm or iodine vapors. To purify the compounds, we utilized column chromatography with silica gel (60-120 mesh size). Melting points were determined using an automated melting point system (Barnstead Electrothermal, UK).

For spectroscopic analysis, we recorded IR(ATR) spectra on a Bruker FTIR spectrometer Alpha T (Germany). Additionally, 1H and 13C NMR spectra were acquired in DMSO-d6 on a Bruker 500 FTNMR spectrometer operating at frequencies of 500 and 125 MHz, respectively. The coupling constant (J) was measured in Hz, and the chemical shift (δ) was estimated in ppm.

Mass spectra were collected using a UPLC/XEVO G2-XS QTOF apparatus equipped with APCI and ESI multi-mode ionization sources. To assess compound purity, we employed an Agilent 1260 Infinity II Quaternary LC system with a DAD HS G7115A detector operating at a wavelength of 315 nm. The isocratic mobile phase consisted of acetonitrile (phase A) and methanol (phase B) in a 1:9 ratio, delivered by quaternary pumps at a flow rate of 1 ml/min. Samples of 20 µL were injected into the HPLC column, and detection was performed at 315 nm wavelength using an SPD10AVP detector.

The purification process utilized a CLC C18 column (5µ, 25 cm × 4.6 mm i.d) with a CLC ODS guard column (4 cm × 4.6 mm i.d).

6.3.4.1 General procedure for synthesis of derivatives of *2-chloro-N-phenylacetamide* (30-57)

Substituted aniline (1.7 mmol, 1 equivalent) was dissolved in DCM (5ml) and triethyl amine (2.5 equivalent) was added. The solution was placed over the ice bath, and drop by drop, chloroacetylchloride (1.7 mmol, 1 equivalent) was added to the solution with continuous stirring. The reaction mixture was stirred at room temperature for 1 hr. Precipitate was obtained after the addition of chloroacetylchloride in the RBF. The reaction was quenched with water (30 mL), and the mixture was extracted with DCM and washed with saturated ammonium chloride. The combined organic layer was dried over sodium sulfate. The organic layer was evaporated under reduced pressure through a rotary evaporator. The intermediate formed was checked through TLC using the ethylacetate and hexane (7:3) as the mobile phase.

***2-chloro-N-phenylacetamide* (30)**

White Solid (0.499g, 93.75% yield) M.P:163-167°C, Rf=0.86 (EA/Hex,7:3,v/v). IR(ATR): 1651 (C=0 stretching), 3345 (Secondary NH stretching), 2956 (C-H stretching), 1170 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.62 – 8.59 (NH,s, 1H), 7.55 – 7.49 (Ar-H,m, 2H), 7.47 – 7.40 (Ar-H,m, 2H), 7.17 – 7.11 (Ar-H,m, 1H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 137.76, 129.31, 123.72, 121.46, 41.74.

***2-chloro-N-(2-nitrophenyl)acetamide* (31)**

Grey Solid (0.670g, 96.76% yield) M.P: 98-100°C, Rf=0.83 (EA/Hex,7:3,v/v). IR(ATR) : 1650 (C=0 stretching), 3210(Secondary NH stretching), 2977(C-H stretching), 1240 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.68 – 9.64 (NH,s, 1H), 8.28 – 8.22 (Ar-H,m, J = 9.0, 1.4 Hz, 1H), 7.97 – 7.91 (Ar-H,m, J = 7.5, 1.5 Hz, 1H), 7.75 – 7.68 (Ar-

H,td, J = 7.8, 1.5 Hz, 1H), 7.42 – 7.35 (Ar-H,m, 1H), 4.23 – 4.20 (CH_2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.03, 136.51, 135.90, 133.67, 127.13, 124.07, 124.04, 41.74.

***2-chloro-N-(3-nitrophenyl)acetamide* (32)**

Pale Yellow Solid (0.652g, 98.54% yield) M.P: 113-117°C, Rf=0.83 (EA/Hex, 7:3,v/v). IR (ATR) : 1647 (C=0 stretching), 3310(Secondary NH stretching), 2951(C-H stretching), 1242 (C-N stretching) cm^{-1}. ^1H NMR (500 MHz, DMSO-d6) δ 9.68 – 9.65 (s, 1H), 8.52 – 8.48 (t, J = 2.2 Hz, 1H), 8.46 – 8.40 (m, J = 7.9, 1.9, 1.0 Hz, 1H), 7.76 – 7.70 (m, J = 7.9, 2.4, 0.9 Hz, 1H), 7.60 – 7.53 (t, J = 8.0 Hz, 1H), 4.23 – 4.19 (s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 148.70, 139.64, 129.27, 127.01, 120.28, 115.89, 41.74.

***2-chloro-N-(4-nitrophenyl)acetamide* (33)**

Yellow Solid (0.596g, 91% yield) M.P: 100-103°C, Rf=0.88 (EA/Hex,7:3,v/v). IR(ATR) : 1652(C=0 stretching), 3218(Secondary NH stretching), 2951(C-H stretching), 1223 (C-N stretching) cm^{-1}. ^1H NMR (500 MHz, DMSO-d6) δ 9.48 – 9.44 (NH,s, 1H), 8.22 – 8.16 (Ar-H,m, 2H), 7.83 – 7.77 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 145.42, 144.14, 125.54, 120.66, 41.74.

***N-(2-bromophenyl)-2-chloroacetamide* (34)**

Grey Solid (0.575g, 98.61% yield) M.P: 88-92°C, Rf=0.9 (EA/Hex,7:3,v/v). IR(ATR) : 1654 (C=0 stretching), 3243(Secondary NH stretching), 2950(C-H stretching), 1215 (C-N stretching) cm^{-1}. ^1H NMR (500 MHz, DMSO-d6) δ 8.94 – 8.91 (NH,s, 1H), 7.70 – 7.64 (Ar-H,m, J = 7.7, 1.4 Hz, 1H), 7.56 – 7.50 (Ar-H,m, J = 8.1, 1.4 Hz, 1H), 7.35 – 7.28 (Ar-H,td, J = 7.4, 1.4 Hz, 1H), 7.27 – 7.20 (Ar-H,m, 1H), 4.23 – 4.20 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.03, 138.33, 133.51, 128.47, 125.51, 123.39, 113.90, 41.74.

***N-(3-bromophenyl)-2-chloroacetamide* (35)**

White Solid (0.498g, 95.78%yield) M.P:110-114°C , Rf=0.75 (EA/Hex,7:3,v/v). IR(ATR) : 1661 (C=0 stretching), 3276(Secondary NH stretching), 2951(C-H stretching), 1246 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.18 – 9.14 (NH,s, 1H), 7.89 – 7.85 (Ar-H,t, J = 2.1 Hz, 1H), 7.46 – 7.40 (Ar-H,m, J = 7.9, 2.2, 1.3 Hz, 1H), 7.33 – 7.27 (Ar-H,m, J = 7.7, 2.1, 1.2 Hz, 1H), 7.26 – 7.19 (Ar-H,d, J = 15.6 Hz, 1H), 4.23 – 4.19 (CH2,s, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 167.69, 139.30, 130.42, 127.62, 122.67, 121.91, 120.63, 41.74.

***N-(4-bromophenyl)-2-chloroacetamide* (36)**

Grey Solid (0.440g, 93.34% yield) M.P: 118-121°C, Rf=0.86 (EA/Hex, 7:3,v/v). IR (ATR): 1653 (C=0 stretching), 3338(Secondary NH stretching), 2955(C-H stretching), 1157 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.16 – 9.12 (NH,s, 1H), 7.53 – 7.43 (Ar-H,m, 4H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 136.21, 131.79, 122.19, 118.25, 41.74.

***2-chloro-N-(2-fluorophenyl)acetamide* (37)**

White Solid (0.378g, 89.21% yield) M.P: 165-168°C, Rf= 0.8(EA/Hex, 7:3,v/v).IR(ATR) : 1650 (C=0 stretching), 3303(Secondary NH stretching), 2962(C-H stretching), 1173 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.34 – 9.31 (NH,s, 1H), 7.50 – 7.44 (Ar-H,m, J = 7.7, 2.2, 1.1 Hz, 1H), 7.35 – 7.25 (Ar-H,m, 2H), 6.92 – 6.84 (Ar-H,m, J = 9.9, 7.7, 2.1, 1.1 Hz, 1H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.03, 156.32, 125.85, 125.69, 125.19, 124.28, 117.10, 41.74.

***2-chloro-N-(3-fluorophenyl)acetamide* (38)**

White Solid (0.595g, 99.10 % yield) M.P: 161-163°C, Rf=0.83 (EA/Hex,7:3,v/v).IR(ATR) : 1672 (C=0 stretching), 3309(Secondary NH stretching), 2975(C-H stretching), 1161 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.34 – 9.31 (NH,s, 1H), 7.50 – 7.44 (Ar-H,m, J = 7.7, 2.2, 1.1 Hz, 1H), 7.35 – 7.25 (Ar-

H,m, 2H), 6.92 – 6.84 (Ar-H,m, J = 9.9, 7.7, 2.1, 1.1 Hz, 1H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 163.12, 140.00, 130.05, 117.04, 112.15, 109.08, 41.74.

***2-chloro-N-(4-fluorophenyl)acetamide* (39)**

Pale Pink Solid (0.403g, 90.75% yield) M.P: 179-182°C, Rf=0.77 (EA/Hex,7:3,v/v). IR(ATR) : 1667 (C=0 stretching), 3272(Secondary NH stretching), 2950(C-H stretching), 1165 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.21 – 9.17 (NH,s, 1H), 7.53 – 7.47 (Ar-H,m, 2H), 7.19 – 7.11 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 159.97, 134.73, 122.58, 116.47, 41.74.

***2-chloro-N-(o-tolyl)acetamide* (40)**

Pale Pink Solid (0.499g, 96.58% yield) M.P: 132-135°C, Rf=0.78 (EA/Hex,7:3,v/v). IR(ATR):1644 (C=0 stretching), 3268(Secondary NH stretching), 2973(C-H stretching), 1172 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.25 – 8.21 (NH,s, 1H), 7.51 – 7.41 (Ar-H,m, 2H), 7.28 – 7.19 (Ar-H,m, 2H), 4.24 – 4.20 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.03, 137.43, 131.80, 130.03, 127.54, 124.96, 122.76, 41.74, 17.35.

***2-chloro-N-(m-tolyl)acetamide* (41)**

Dark Brown Solid (0.321g, 88.46% yield) M.P: 142-144°C, Rf=0.77 (EA/Hex,7:3,v/v). IR (ATR) : 1648 (C=0 stretching), 3239(Secondary NH stretching), 2970(C-H stretching), 1176 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.68 – 8.65 (NH,s, 1H), 7.51 – 7.43 (Ar-H,m, 2H), 7.18 – 7.12 (Ar-H,d, J = 7.6 Hz, 1H), 6.97 – 6.91 (Ar-H,m, 1H), 4.23 – 4.19 (CH2,s, 2H), 2.31 – 2.27 (CH3,d, J = 0.7 Hz, 3H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 139.50, 138.32, 128.94, 126.50, 120.29, 119.22, 41.74, 21.21.

***2-chloro-N-(p-tolyl)acetamide* (42)**

Grey Solid (0.428g, 91.17% yield) M.P: 149-151°C, Rf=0.88 (EA/Hex,7:3,v/v). IR(ATR) : 1658 (C=0 stretching), 3281(Secondary NH stretching), 2951(C-H stretching), 1233 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.12 – 9.09 (NH,s, 1H), 7.45 – 7.39 (Ar-H,m, 2H), 7.16 – 7.10 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H), 2.37 – 2.33 (CH3,s, 3H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 136.58, 133.31, 129.75, 119.33, 41.74, 21.13.

***2-chloro-N-(2-chlorophenyl)acetamide* (43)**

Dark Brown Solid (0.378g, 88.47% yield) M.P: 134-136°C, Rf=0.85 (EA/Hex,7:3,v/v).IR(ATR) : 1646 (C=0 stretching), 3337(Secondary NH stretching), 2954(C-H stretching), 1237 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.97 – 8.94 (NH,s, 1H), 7.50 – 7.45 (Ar-H,m, J = 7.9, 1.4 Hz, 1H), 7.42 – 7.35 (Ar-H,m, 1H), 7.28 – 7.22 (Ar-H,m, 1H), 7.11 – 7.04 (Ar-H,m, J = 8.3, 7.1, 1.4 Hz, 1H), 4.23 – 4.20 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.03, 135.66, 130.62, 128.23, 128.05, 126.11, 124.73, 41.74.

***2-chloro-N-(3-chlorophenyl) acetamide* (44)**

Grey Solid (0.448g, 95.76% yield) M.P: 174-177°C, Rf=0.89 (EA/Hex,7:3,v/v).IR(ATR) : 1654 (C=0 stretching), 3220(Secondary NH stretching), 2958(C-H stretching), 1200 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.18 – 9.15 (NH,s, 1H), 7.56 – 7.51 (Ar-H,t, J = 2.1 Hz, 1H), 7.47 – 7.41 (Ar-H,m, J = 8.0, 2.1, 1.2 Hz, 1H), 7.35 – 7.28 (Ar-H,d, J = 15.8 Hz, 1H), 7.19 – 7.13 (Ar-H,m, J = 7.9, 2.2, 1.3 Hz, 1H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 139.54, 134.34, 129.90, 124.84, 120.45, 120.38, 41.74.

***2-chloro-N-(4-chlorophenyl)acetamide* (45)**

Pale Brown Solid (0.403g, 90.54% yield), M.P:149-151°C, Rf=0.78 (EA/Hex,7:3,v/v). IR(ATR) : 1649 (C=0 stretching), 3253(Secondary NH stretching), 2971(C-H stretching),

1156 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.27 – 9.24 (NH,s, 1H), 7.62 – 7.57 (Ar-H,m, 2H), 7.38 – 7.33 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 136.53, 129.24, 128.63, 122.21, 41.74.

***2-chloro-N-(2-methoxyphenyl) acetamide* (46)**

Grey Solid (0.475g, 95.10 % yield) M.P: 132-133°C, Rf=0.73 (EA/Hex,7:3,v/v). ^{1}H NMR (500 MHz, DMSO-d6) *δ* 8.32 (dd, *J = 7.9, 1.5 Hz*, 1H), 8.28 (s, 1H), 7.09 (td, *J = 8.0*, 1.5 Hz, 1H), 7.03 (td, *J = 8.0*, 1.4 Hz, 1H), 6.93 (dd, *J = 8.0*, 1.3 Hz, 1H), 4.22 (s, 2H), 3.93 (s, 3H). ^{13}C NMR (125 MHz, DMSO-d6) *δ* 165.29, 150.52, 128.66, 122.28, 121.76, 120.75, 110.66, 55.69, 43.01.

***2-chloro-N-(3-methoxyphenyl) acetamide* (47)**

Brown Solid (0.410 g, 98.14 % yield) M.P: 137-139°C, Rf=0.78 (EA/Hex,7:3,v/v).^{1}H NMR (500 MHz, DMSO-d6) *δ* 9.24 (s, 1H), 7.30 – 7.28 (m, 1H), 7.28 (d, *J = 1.5 Hz*, 1H), 7.26 – 7.22 (m, 1H), 7.08 – 7.06 (m, 1H), 4.21 (s, 2H), 3.82 (s, 3H). ^{13}C NMR (125 MHz, DMSO-d6) *δ* 165.30, 160.33, 139.80, 130.02, 113.46, 111.24, 106.17, 55.59, 43.38.

***2-chloro-N-(4-methoxyphenyl) acetamide* (48)**

Brown Solid (0.572g, 94.10 % yield) M.P: 134-135°C, Rf=0.83 (EA/Hex,7:3,v/v). ^{1}H NMR (500 MHz, DMSO-d6) *δ* 9.26 (s, 1H), 8.33 (d, *J = 1.3 Hz*, 1H), 8.31 (d, *J = 1.5 Hz*, 1H), 6.97 (d, *J = 1.3 Hz*, 1H), 6.96 (d, *J = 1.5 Hz*, 1H), 4.21 (s, 2H), 3.80 (s, 3H). ^{13}C NMR (125 MHz, DMSO-d6) *δ* 164.92, 156.42, 133.16, 122.30, 112.27, 55.35, 43.57.

***2-chloro-N-(2-cyanophenyl)acetamide* (49)**

Grey Solid (0.499g, 95.21% yield)M.P: 140-142°C, Rf= 0.82(EA/Hex,7:3,v/v). IR(ATR): 1642 (C=0 stretching), 3307 (Secondary NH stretching), 2963(C-H stretching), 1244 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.90 – 8.86 (NH,s, 1H), 7.75 – 7.70 (Ar-H,m, J = 6.6, 1.3 Hz, 1H), 7.69 – 7.64 (Ar-H,m, J = 7.9, 1.4 Hz, 1H), 7.49 – 7.42 (Ar-H,m, J = 8.0, 6.9, 1.4 Hz, 1H), 7.27 – 7.20 (Ar-H,m, 1H), 4.23 – 4.20 (CH2,s,

2H). 13C NMR (125 MHz, DMSO-d6) δ 168.03, 141.08, 133.62, 132.59, 122.41, 121.15, 117.68, 97.59, 41.74.

***2-chloro-N-(3-cyanophenyl)acetamide* (50)**

Brown Solid (0.378g, 89.52% yield) M.P: 167-171°C, Rf=0.79 (EA/Hex,7:3,v/v). IR(ATR): 1664 (C=0 stretching), 3264(Secondary NH stretching), 2972(C-H stretching), 1215 (C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.29 – 9.25 (NH,s, 1H), 7.93 – 7.88 (Ar-H,m, 1H), 7.83 – 7.76 (Ar-H,m, 1H), 7.47 – 7.40 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 139.27, 130.61, 129.90, 127.19, 123.07, 119.29, 111.12, 41.74. ***2-chloro-N-(4-cyanophenyl)acetamide* (51)**

Brown Solid (0.403g, 90.75% yield) M.P: 147-150°C, Rf=0.89 (EA/Hex,7:3,v/v). IR(ATR) : 1670 (C=0 stretching), 3340(Secondary NH stretching), 2964(C-H stretching), 1226 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.18 – 9.15 (NH,s, 1H), 7.75 – 7.69 (Ar-H,m, 2H), 7.64 – 7.58 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.69, 140.74, 131.79, 121.81, 119.12, 105.57, 41.74. MS (EI+): m/z calculated for C9H7ClN2O, found -194.02 (M^+).

***2-chloro-N-(2-(trifluoromethyl) phenyl) acetamide* (52)**

Pale Brown Solid (0.310 g, 94.54% yield), M.P:157-158°C, Rf=0.64 (EA/Hex,6:4,v/v). ^{1}H NMR (500 MHz, DMSO-d6) δ 9.66 (s, 1H), 7.95 (dd, J = 7.1, 1.5 Hz, 1H), 7.71 (dt, J = 10.6, 1.4 Hz, 1H), 7.48 – 7.44 (m, 1H), 7.37 (td, J = 7.2, 1.4 Hz, 1H), 4.22 (s, 3H). 13C NMR (125 MHz, DMSO-d6) δ 165.38, 137.85, 129.31, 127.50, 124.58, 122.34, 121.94, 121.04, 42.96.

***2-chloro-N-(4-(trifluoromethyl) phenyl) acetamide* (53)**

Pale Brown Solid (0.372 g, 97.41% yield), M.P:151-154°C, Rf=0.71 (EA/Hex,6:4,v/v) 1H NMR (500 MHz, DMSO-d6) δ 9.15 (s, 1H), 7.75 (d, J = 7.3 Hz, 2H), 7.56 (dq, J =

7.3, 1.3 Hz, 2H), 4.21 (s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 164.92, 138.69, 126.84, 126.80, 125.52, 122.64, 120.48, 120.45, 43.57.

***2-chloro-N-(3-(trifluoromethyl) phenyl) acetamide* (54)**

Pale Brown Solid (0.410 g, 94.46 % yield), M.P:154-156°C, Rf=0.60 (EA/Hex,6:4,v/v) ^{1}H NMR (500 MHz, DMSO-d6) δ 9.32 (s, 1H), 7.86 (t, J = 2.0 Hz, 1H), 7.58 (dt, J = 7.1, 2.0 Hz, 1H), 7.45 – 7.35 (m, 2H), 4.21 (s, 2H). 13C NMR (125 MHz, DMSO-d6) δ 165.29, 138.08, 131.68, 129.51, 124.88, 121.34, 120.50, 116.90, 43.44.

***N-(2-acetylphenyl)-2-chloroacetamide* (55)**

Pale Yellow Solid (0.375g, 89.54% yield) M.P: 154-157°C, Rf=0.9 (EA/Hex,7:3,v/v). IR(ATR): 1658 (C=0 stretching), 3264(Secondary NH stretching), 2962(C-H stretching), 1222 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.51 – 9.47 (NH,s, 1H), 8.41 – 8.35 (Ar-H,m, J = 8.1, 1.4 Hz, 1H), 7.82 – 7.77 (Ar-H,m, J = 8.1, 1.5 Hz, 1H), 7.56 – 7.49 (Ar-H,td, J = 7.9, 1.5 Hz, 1H), 7.24 – 7.18 (Ar-H,m, 1H), 4.23 – 4.20 (CH2,s, 2H), 2.66 – 2.63 (CH3,s, 3H). 13C NMR (125 MHz, DMSO-d6) δ 201.96, 168.03, 138.23, 133.18, 129.86, 129.71, 123.95, 122.96, 41.74, 28.28.

***N-(3-acetylphenyl)-2-chloroacetamide* (56)**

Grey Solid (0.321g, 87.33% yield) M.P: 178-181°C, Rf=0.79 (EA/Hex,7:3,v/v). IR(ATR) : 1646 (C=0 stretching), 3226(Secondary NH stretching), 2956(C-H stretching), 1182 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.38 – 9.34 (NH,s, 1H), 8.00 – 7.96 (Ar-H,t, J = 2.2 Hz, 1H), 7.75 – 7.70 (Ar-H,m, J = 7.9, 2.2, 1.3 Hz, 1H), 7.67 – 7.61 (Ar-H,m, J = 8.0, 2.2, 1.3 Hz, 1H), 7.45 – 7.38 (Ar-H,d, J = 16.0 Hz, 1H), 4.23 – 4.19 (CH,s, 2H), 2.57 – 2.53 (CH3,s, 3H).. 13C NMR (125 MHz, DMSO-d6) δ 197.37, 167.69, 138.63, 138.05, 128.79, 126.64, 124.50, 121.11, 41.74, 27.79.

***N-(4-acetylphenyl)-2-chloroacetamide* (57)**

Grey Solid (0.403g, 90.87% yield) M.P: 167-172°C, Rf=0.88 (EA/Hex,7:3,v/v). IR(ATR) : 1651 (C=0 stretching), 3218(Secondary NH stretching), 2967(C-H stretching), 1188 (C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.18 – 9.14 (NH,s, 1H), 7.89 – 7.83 (Ar-H,m, 2H), 7.54 – 7.48 (Ar-H,m, 2H), 4.23 – 4.19 (CH2,s, 2H), 2.60 – 2.56 (CH3,s, 3H). 13C NMR (125 MHz, DMSO-d6) δ 197.14, 167.69, 143.26, 131.62, 130.19, 119.59, 41.74, 27.79.

6.3.4.2 General procedure for the synthesis of *N-benzyl piperidine* analogues

Aniline intermediate (1 equivalent) was dissolved in Acetonitrile (10 ml) containing Triethylamine (2 equivalent) and 4-Amino benzyl piperidine (1 equivalent). The reaction mixture was heated up to 70℃ for 6 h, followed by a workup with water-ethylacetate system and the organic layer was evaporated under reduced pressure through a rotary evaporator. The product formed was checked through TLC using the ethylacetate and hexane (7:3) as the mobile phase. The product was purified by column chromatography using ethylacetate and hexane as a mobile phase.

***2-((1-benzylpiperidin-4-yl) amino)-N-phenylacetamide* (59)**

White solid (0.15g, 65.36% yield). M.P: 152-155°C,R_f =0.42 (EA/Hex,7:3,v/v). IR (ATR): 1649(C=0 stretching), 3271(Secondary NH stretching), 2962(C-H stretching), 1214(C-N stretching) cm^{-1}. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.99 – 8.96 (CONH,s, 1H), 7.55 – 7.49 (Ar-H,m, 2H), 7.47 – 7.40 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 5H), 7.17 – 7.11 (Ar-H,m, J = 6.9, 1.1 Hz, 1H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (CH2,d, J = 5.8 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 168.41, 138.23, 137.76, 129.42, 128.92, 128.14,

126.55, 123.72, 121.57, 63.66, 58.62, 51.44, 50.81, 30.97. MS (EI+): m/z calculated for $C_{20}H_{25}N_3O$ 324.2031, found - 324.2028 $(M+H)^+$.

2-((1-benzylpiperidin-4-yl)amino)-N-(2-nitrophenyl)acetamide **(60)**

Yellow Solid (0.230 g, 61.66% yield) M.P: 160-164°C, Rf=0.39 (EA/Hex,7:3,v/v). IR(ATR): 1664(C=0 stretching), 3148(Secondary NH stretching), 2921(C-H stretching), 1175(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.75 – 9.71 (CONH,s, 1H), 8.51 – 8.46 (Ar-H,m, J = 7.7, 1.4 Hz, 1H), 8.28 – 8.22 (Ar-H,m, J = 9.0, 1.4 Hz, 1H), 7.75 – 7.68 (Ar-H,td, J = 7.6, 1.3 Hz, 1H), 7.42 – 7.35 (Ar-H,m, J = 9.2, 7.9, 1.4 Hz, 1H), 7.33 – 7.23 (Ar-H,m, 5H), 3.56 – 3.51 (CH2,m, 4H), 3.40 – 3.32 (NH,m, J = 7.3, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.99, 138.23, 136.51, 135.90, 133.67, 128.81, 127.95, 127.13, 126.55, 124.07, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}N_4O_3$ 369.1882, found 369.1892 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl)amino)-N-(3-nitrophenyl)acetamide **(61)**

Pale Yellow Solid (0.262 g, 63.53% yield)M.P: 179-185°C, Rf=0.49 (EA/Hex,7:3,v/v). IR(ATR) : 1670(C=0 stretching),3247(Secondary NH stretching),2954(C-H stretching),1197(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.94 – 9.91 (CONH,s, 1H), 8.52 – 8.48 (Ar-H,t, J = 2.2 Hz, 1H), 8.46 – 8.40 (Ar-H,m, J = 7.9, 1.9, 1.0 Hz, 1H), 7.76 – 7.70 (Ar-H,m, J = 7.9, 2.4, 0.9 Hz, 1H), 7.60 – 7.53 (Ar-H,m, 1H), 7.33 – 7.23 (Ar-H,m, 5H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (d, J = 5.8 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR

(125 MHz, DMSO-d6) δ 168.41, 148.70, 139.64, 138.23, 129.27, 128.81, 127.95, 127.01, 126.55, 120.28, 115.89, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}N_4O_3$ 369.1882, found 369.1869 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(4-nitrophenyl)acetamide* (62)**

Yellow Solid (0.256g, 62.71% yield) M.P: 170-173°C, Rf=0.4 (EA/Hex,7:3,v/v). IR(ATR): 1644(C=0 stretching),3267(Secondary NH stretching),2924(C-H stretching),1173(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.64 – 9.60 (CONH,s, 1H), 8.22 – 8.16 (Ar-H,m, 2H), 7.83 – 7.77 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 5H), 3.56 – 3.52 (CH2,d, J = 0.7 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 145.42, 144.14, 138.23, 128.81, 127.95, 126.55, 125.54, 120.66, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}N_4O_3$ 369.1882, found- 369.1909 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(2-bromophenyl)acetamide* (63)**

White Solid (0.173g, 68.41% yield) M.P: 150-154°C, Rf=0.4 (EA/Hex,7:3,v/v). IR(ATR) : 1642(C=0 stretching),3233(Secondary NH stretching),2954(C-H stretching),1178(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.52 – 9.48 (CONH,s, 1H), 7.69 – 7.64 (Ar-H,m, J = 7.7, 1.4 Hz, 1H), 7.56 – 7.50 (Ar-H,m, J = 8.1, 1.4 Hz, 1H), 7.35 – 7.20 (Ar-H,m, 7H), 3.56 – 3.50 (CH2,m, 4H), 3.40 – 3.32 (NH,m, J = 7.2, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.99, 138.33, 138.23, 133.51, 128.81, 128.47, 127.95, 126.55, 125.51, 123.39, 113.90, 63.66, 58.62,

51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}BrN_3O$ 402.1181, found - 402.1169 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl)amino)-N-(3-bromophenyl)acetamide **(64)**

White Solid (0.166, 67.60% yield) M.P: 180-183°C, Rf=0.39 (EA/Hex,7:3,v/v). IR(ATR): 1654(C=0 stretching), 3367 (Secondary NH stretching), 2977(C-H stretching),1184 (C-N stretching) cm^{-1}. ^1H NMR (500 MHz, DMSO-d6) δ 9.52 – 9.48 (CONH,s, 1H), 7.69 – 7.64 (Ar-H,m, J = 7.7, 1.4 Hz, 1H), 7.56 – 7.50 (Ar-H,m, J = 8.1, 1.4 Hz, 1H), 7.35 – 7.20 (Ar-H,m, 7H), 3.56 – 3.50 (CH2,m, 4H), 3.40 – 3.32 (NH,m, J = 7.2, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 139.30, 138.23, 130.42, 128.81, 127.95, 127.62, 122.67, 121.91, 120.63, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}BrN_3O$ 402.1181, found 401.1174 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl)amino)-N-(4-bromophenyl)acetamide **(65)**

White Solid (0.192, 63.72% yield) M.P:176-178°C, Rf=0.36 (EA/Hex, 7:3,v/v). IR(ATR): 1658(C=0 stretching),3255(Secondary NH stretching),2977(C-H stretching),1187(C-N stretching) cm^{-1}.^1H NMR (500 MHz, DMSO-d6) δ 9.58 – 9.54 (CONH,s, 1H), 7.53 – 7.43 (Ar-H,m, 4H), 7.33 – 7.23 (Ar-H,m, 5H), 3.56 – 3.52 (CH2,d, J = 0.7 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 138.23, 136.21, 131.79, 128.81, 127.95, 126.55, 122.19, 118.25, 63.66, 58.62,

51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}BrN_3O$ 402.1181, found – 402.1156 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl)amino)-N-(2-fluorophenyl)acetamide **(66)**

White Solid (0.232g,61.70% yield)M.P: 162 – 165°C, Rf=0.38 (EA/Hex,7:3,v/v). IR(ATR) : 1647(C=0 stretching),3354(Secondary NH stretching),2978(C-H stretching),1192(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.84 (s, 1H), 8.29 – 8.19 (m, 2H), 7.89 (dd, J = 5.2, 1.1 Hz, 1H), 7.77 (d, J = 5.5 Hz, 2H), 7.71 – 7.61 (m, 2H), 7.41 (t, J = 4.6 Hz, 1H), 7.18 (t, J = 5.1 Hz, 1H), 5.13 (s, 1H), 4.68 (s, 2H), 4.37 (s, 2H), 3.14 – 3.00 (m, 4H), 2.90 (dd, J = 6.8, 3.2 Hz, 1H), 2.46 – 2.09 (m, 4H). 13C NMR (125 MHz, DMSO-d6) δ 171.63, 163.58, 140.77, 139.13, 130.89, 129.20, 127.26, 115.34, 110.09, 106.44, 62.68, 55.00, 52.22, 50.57, 32.51. MS (EI+) m/z: calculated for $C_{20}H_{24}FN_3O$ 342.1937, found - 342.1969 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl)amino)-N-(3-fluorophenyl)acetamide **(67)**

White Solid (0.212g, 66.77% yield) M.P: 110-114°C, Rf=0.43 (EA/Hex,7:3,v/v). IR(ATR) : 1652(C=0 stretching),3362(Secondary NH stretching),2978(C-H stretching),1192(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.70 – 9.67 (CONH,s, 1H), 7.50 – 7.44 (Ar-H,m, J = 7.7, 2.2, 1.1 Hz, 1H), 7.35 – 7.23 (Ar-H,m, 7H), 6.92 – 6.84 (Ar-H,m, J = 9.9, 7.7, 2.1, 1.1 Hz, 1H), 3.56 – 3.52 (CH2,d, J = 0.7 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 163.12, 140.00, 138.23, 130.05, 128.81, 127.95, 126.55, 117.04, 112.15, 109.08, 63.66, 58.62, 51.2550.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}FN_3O$ 342.1937, found- 341.19 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(4-fluorophenyl)acetamide* (68)**

Brown Solid (0.248g,64.27% yield) M.P: 156-162°C, Rf=0.39 (EA/Hex,7:3,v/v). IR(ATR): 1661 (C=0 stretching), 3257 (Secondary NH stretching), 2975 (C-H stretching), 1200 (C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.62 – 9.59 (CONH,s, 1H), 7.53 – 7.47 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 5H), 7.19 – 7.11 (Ar-H,m, 2H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (CH2,d, J = 5.8 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 159.97, 138.23, 134.73, 128.81, 127.95, 126.55, 122.58, 122.58, 116.47, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}FN_3O$ 342.1937, found- 342.1948 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(o-tolyl)acetamide* (69)**

White Solid (0.15g,44.93% yield) M.P: 138-142°C, Rf=0.36 (EA/Hex,7:3,v/v). IR(ATR) : 1645(C=0 stretching),3207(Secondary NH stretching),2920(C-H stretching),1217(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 8.73 – 8.70 (CONH,s, 1H), 7.51 – 7.41 (Ar-H,m, 2H), 7.33 – 7.19 (Ar-H,m, 7H), 3.56 – 3.50 (CH2,m, 4H), 3.40 – 3.32 (NH,m, J = 7.2, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH3,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.99, 138.23, 137.43, 131.80, 130.03, 128.81, 127.95, 127.54, 126.55, 124.96, 122.76, 63.66, 58.62, 51.25, 50.81, 30.87, 17.35. MS (EI+): m/z calculated for $C_{21}H_{27}N_3O$ 338.2188, found -338.2201 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(m-tolyl)acetamide* (70)**

Brown Solid (0.175 g, 52.41% yield) M.P: 165-170°C, Rf=0.38 (EA/Hex,7:3,v/v). IR(ATR) : 1650(C=0 stretching),3269(Secondary NH stretching),2921(C-H stretching),1180(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.31 – 9.28 (CONH,s, 1H), 7.51 – 7.43 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 5H), 7.19 – 7.12 (Ar-H,t, J = 7.6 Hz, 1H), 6.97 – 6.91 (Ar-H,m, 1H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (CH2,d, J = 5.8 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH3,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 139.50, 138.32, 138.23, 128.94, 128.81, 127.95, 126.55, 126.50, 120.29, 119.22, 63.66, 58.62, 51.25, 50.81, 30.87, 21.21. MS (EI+): m/z calculated for $C_{21}H_{27}N_3O$ 338.2188, found- 338.2196 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(p-tolyl)acetamide* (71)**

Brown Solid (0.160 g, 47.35% yield) M.P: 155-158°C, Rf=0.43 (EA/Hex,7:3,v/v). IR(ATR) : 1648(C=0 stretching),3284(Secondary NH stretching),2978(C-H stretching),1182(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.56 – 9.52 (CONH,s, 1H), 7.45 – 7.39 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 5H), 7.16 – 7.10 (Ar-H,m, 2H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (CH2,d, J = 5.8 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH3,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 138.23, 136.58, 133.31, 129.75, 128.81, 127.95, 126.55, 119.33, 63.66, 58.62, 51.25, 50.81, 30.87, 21.13. MS (EI+): m/z calculated for $C_{21}H_{27}N_3O$ 338.2188, found- 338.2194 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(2-chlorophenyl)acetamide* (72)**

Dark Brown Solid (0.192g,71.55% yield) M.P: 167-172°C, Rf=0.47 (EA/Hex,7:3,v/v). IR(ATR) : 1651(C=0 stretching),3212(Secondary NH stretching),2921(C-H stretching),1172(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.00 – 8.97 (CONH,s, 1H), 7.50 – 7.44 (m, J = 7.9, 1.4 Hz, 1H), 7.42 – 7.35 (m, 1H), 7.33 – 7.22 (m, 6H), 7.11 – 7.04 (m, J = 8.2, 7.1, 1.4 Hz, 1H), 3.56 – 3.50 (m, 4H), 3.40 – 3.32 (NH,m, J = 7.2, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.99, 138.23, 135.66, 130.62, 128.81, 128.23, 128.05, 127.95, 126.55, 126.11, 124.73, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}ClN_3O$ 358.1686, found - 358.1671 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(3-chlorophenyl)acetamide* (73)**

Dark Brown Solid (0.24g, 72.16% yield) M.P: 168-170°C, Rf=0.35 (EA/Hex,7:3,v/v). IR(ATR) : 1646(C=0 stretching),3259(Secondary NH stretching),2975(C-H stretching),1216(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.39 – 9.35 (CONH,s, 1H), 7.56 – 7.52 (Ar-H,t, J = 2.2 Hz, 1H), 7.47 – 7.41 (Ar-H,m, J = 8.1, 2.1, 1.2 Hz, 1H), 7.35 – 7.23 (Ar-H,m, 6H), 7.19 – 7.13 (Ar-H,m, J = 7.9, 2.2, 1.3 Hz, 1H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (CH2,d, J = 5.8 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 139.54, 138.23, 134.34, 129.90, 128.81, 127.95, 126.55, 124.84, 120.45, 120.38, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{20}H_{24}ClN_3O$ 358.1686, found - 358.1664 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl) amino)-N-(4-chlorophenyl)acetamide* (74)**

Pale Brown Solid (0.378 g, 60.19% yield) M.P: 176-178°C, R_f=0.45 (EA/Hex,7:3,v/v). IR(ATR): 1654 (C=0stretching), 3258 (Secondary NH stretching), 2977 (C-H stretching), 1171 (C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.58 – 9.54 (CONH,s, 1H), 7.62 – 7.56 (Ar-H,m, 2H), 7.38 – 7.23 (Ar-H,m, 5H), 3.56 – 3.52 (CH2,d, J = 0.7 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 168.41, 138.23, 136.53, 129.24, 128.81, 128.63, 127.95, 126.55, 122.21, 63.66, 58.62, 51.25, 50.81,30.87. MS (EI+): m/z calculated for $C_{20}H_{24}ClN_3O$ 358.1686, found -358.1662 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl) amino)-N-(2-methoxyphenyl) acetamide **(75)**

Pale White Solid (0.318 g, 64.10% yield) M.P: 196-198°C, R_f=0.36 (EA/Hex,7:3,v/v). ^{1}H NMR (500 MHz, DMSO-d6) *δ* 8.75 (s, 1H), 8.32 (dd, *J = 7.9, 1.5 Hz*, 1H), 7.33 – 7.23 (m, 5H), 7.09 (td, *J = 8.0, 1.5 Hz*, 1H), 7.03 (td, *J = 8.0, 1.4 Hz*, 1H), 6.93 (dd, *J = 8.1, 1.5 Hz*, 1H), 3.93 (s, 3H), 3.54 (s, 2H), 3.52 (d, *J = 5.7 Hz*, 2H), 3.40 – 3.32 (m, 1H), 3.02 – 2.94 (m, 1H), 2.80 – 2.71 (m, 2H), 2.53 – 2.44 (m, 2H), 1.96 – 1.86 (m, 2H), 1.69 – 1.59 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) *δ* 169.70, 150.12, 137.93, 129.34, 128.89, 128.82, 127.68, 122.28, 121.79, 120.79, 110.67, 62.76, 55.69, 54.72, 52.81, 50.59, 31.73. MS (EI+): m/z calculated for $C_{21}H_{27}N_3O_2$ 354.2137, found -354.2108 $[M+H]^+$.

2-((1-benzylpiperidin-4-yl) amino)-N-(3-methoxyphenyl) acetamide **(76)**

Pale White Solid (0.401 g, 69.48% yield) M.P: 197-199°C, R_f=0.37 (EA/Hex,7:3,v/v) ^{1}H NMR (500 MHz, DMSO-d6) *δ* 9.48 (s, 1H), 7.31 – 7.26 (m, 7H), 7.24 (d, *J = 7.9 Hz*, 1H), 7.07 (dt, *J = 7.5, 1.8 Hz*, 1H), 3.82 (s, 3H), 3.54 (s, 2H), 3.47 (d, *J = 5.9 Hz*, 2H),

3.40 – 3.32 (m, 1H), 3.02 – 2.94 (m, 1H), 2.80 – 2.71 (m, 2H), 2.53 – 2.44 (m, 2H), 1.96 – 1.86 (m, 2H), 1.69 – 1.59 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 169.73, 160.33, 140.19, 137.93, 130.02, 128.89, 128.82, 127.68, 113.55, 111.24, 106.09, 62.76, 55.59, 54.72, 52.81, 50.71, 31.73. MS (EI+): m/z calculated for $C_{21}H_{27}N_3O_2$ 354.2137, found - 354.2109 (M+).

***2-((1-benzylpiperidin-4-yl) amino)-N-(4-methoxyphenyl) acetamide* (77)**

Pale White Solid (0.271 g, 60.50% yield) M.P: 199-202°C, R_f=0.42 (EA/Hex,7:3,v/v) ^{1}H NMR (500 MHz, DMSO-d6) δ 9.59 (s, 1H), 7.98 (d, *J = 8.6 Hz*, 2H), 7.33 – 7.23 (m, 5H), 6.97 (d, *J = 8.6 Hz*, 2H), 3.80 (s, 3H), 3.54 (s, 2H), 3.47 (d, *J = 5.9 Hz*, 2H), 3.40 – 3.32 (m, 1H), 3.02 – 2.94 (m, 1H), 2.80 – 2.71 (m, 2H), 2.53 – 2.44 (m, 2H), 1.96 – 1.86 (m, 2H), 1.69 – 1.59 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 169.79, 156.42, 137.93, 133.42, 128.89, 128.82, 127.68, 122.22, 112.26, 62.76, 55.35, 54.72, 52.81, 50.65, 31.73. MS (EI+): m/z calculated for $C_{21}H_{27}N_3O_2$ 354.2137, found 354.2168 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(2-cyanophenyl)acetamide* (78)**

White Solid (0.189 g, 58.45% yield) M.P: 177-182°C, Rf=0.37 (EA/Hex,7:3,v/v). IR(ATR) : 1653(C=0 stretching),3278(Secondary NH stretching),2964(C-H stretching),1175(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.46 – 9.42 (CONH,s, 1H), 7.75 – 7.70 (Ar-H,m, J = 6.6, 1.3 Hz, 1H), 7.67 – 7.62 (Ar-H,m, J = 7.9, 1.4 Hz, 1H), 7.49 – 7.42 (Ar-H,m, J = 8.0, 6.9, 1.4 Hz, 1H), 7.33 – 7.20 (Ar-H,m, 6H), 3.56 – 3.50 (CH2,m, 4H), 3.40 – 3.32 (NH,m, J = 7.2, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 167.99, 141.08, 138.23, 133.62, 132.59, 128.81, 127.95, 126.55, 122.41, 121.15, 117.68, 97.59, 63.66, 58.62, 51.25,

50.81, 30.87. MS (EI+): m/z calculated for $C_{21}H_{24}N_4O$ 349.1984, found 349.1972 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(3-cyanophenyl)acetamide* (79)**

White Solid (0.185g, 57.48 % yield) M.P: 175-182°C, Rf=0.44 (EA/Hex,7:3,v/v). IR(ATR): 1661(C=0 stretching), 3273(Secondary NH stretching), 2964(C-H stretching), 1184(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.42 – 9.39 (CONH,s, 1H), 7.91 – 7.87 (Ar-H,m, 1H), 7.83 – 7.76 (Ar-H,m, 1H), 7.47 – 7.40 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 3H), 3.56 – 3.53 (CH2,s, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 139.27, 138.23, 130.61, 129.90, 128.81, 127.95, 127.19, 126.55, 123.07, 119.29, 111.12, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{21}H_{24}N_4O$ 349.1984, found- 349.1977 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl)amino)-N-(4-cyanophenyl)acetamide* (80)**

White Solid (0.249g, 66.55% yield) M.P: 175-178°C, Rf=0.39 (EA/Hex,7:3,v/v). IR(ATR): 1656 (C=0 stretching),3 228 (Secondary NH stretching), 2962 (C-H stretching), 1199 (C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.61 – 9.57 (CONH,s, 1H), 7.75 – 7.69 (Ar-H,m, 2H), 7.64 – 7.58 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 5H), 3.56 – 3.52 (CH2,d, J = 0.7 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 168.41, 140.74, 138.23, 131.79, 128.81, 127.95, 126.55, 121.81,

119.12, 105.57, 63.66, 58.62, 51.25, 50.81, 30.87. MS (EI+): m/z calculated for $C_{21}H_{24}N_4O$ 349.1984, found -349.1969 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl) amino)-N-(2-(trifluoromethyl) phenyl) acetamide* (81)**

White Solid (0.184g, 55.27% yield) M.P: 185-187°C, Rf=0.40 (EA/Hex,7:3,v/v).^{1}H NMR (500 MHz, DMSO-d6) δ 9.07 (s, 1H), 7.95 (dd, *J = 7.1, 1.5 Hz*, 1H), 7.71 (dt, *J = 10.6, 1.4 Hz*, 1H), 7.48 – 7.44 (m, 1H), 7.37 (td, *J = 7.2, 1.4 Hz*, 1H), 7.33 – 7.23 (m, 5H), 3.56 – 3.50 (m, 4H), 3.40 – 3.32 (m, 1H), 3.02 – 2.94 (m, 1H), 2.80 – 2.71 (m, 2H), 2.53 – 2.44 (m, 2H), 1.96 – 1.86 (m, 2H), 1.69 – 1.60 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 169.76, 137.93, 137.57, 129.31, 128.89, 128.82, 127.68, 127.50, 124.58, 122.13, 121.94, 120.79, 62.76, 54.72, 52.81, 50.60, 31.73. MS (EI+): m/z calculated for $C_{21}H_{24}F_3N_3O$, 392.1905, found- 392.1901 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl) amino)-N-(3-(trifluoromethyl) phenyl) acetamide* (82)**

White Solid (0.177g, 52.50% yield) M.P: 181-184°C, Rf=0.36 (EA/Hex,7:3,v/v).

^{1}H NMR (500 MHz, DMSO-d6) δ 9.48 (s, 1H), 8.08 (t, *J = 2.1 Hz*, 1H), 7.83 (dt, *J = 7.0, 1.9 Hz*, 1H), 7.45 – 7.35 (m, 2H), 7.33 – 7.23 (m, 5H), 3.54 (s, 2H), 3.47 (d, *J = 5.9 Hz*, 2H), 3.40 – 3.32 (m, 1H), 3.02 – 2.94 (m, 1H), 2.80 – 2.71 (m, 2H), 2.53 – 2.44 (m, 2H), 1.96 – 1.86 (m, 2H), 1.69 – 1.59 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 169.70, 138.91, 137.93, 131.69, 129.46, 128.89, 128.82, 127.68, 124.88, 121.34, 120.47, 116.91, 62.76, 54.72, 52.81, 50.65, 31.73. MS (EI+): m/z calculated for $C_{21}H_{24}F_3N_3O$, 392.1905, found- 392.1897 $[M+H]^+$.

***2-((1-benzylpiperidin-4-yl) amino)-N-(4-(trifluoromethyl) phenyl) acetamide* (83)**

White Solid (0.198g, 61.48% yield) M.P: 185-187°C, Rf=0.36 (EA/Hex,7:3,v/v).

^{1}H NMR (500 MHz, DMSO-d6) δ 9.57 (s, 1H), 7.75 (d, *J = 7.3 Hz*, 2H), 7.56 (dd, *J = 7.3, 1.5 Hz*, 2H), 7.33 – 7.23 (m, 5H), 3.54 (s, 2H), 3.47 (d, *J = 5.9 Hz*, 2H), 3.40 – 3.32 (m, 1H), 3.02 – 2.94 (m, 1H), 2.80 – 2.71 (m, 2H), 2.53 – 2.44 (m, 2H), 1.96 – 1.86 (m,

2H), 1.69 – 1.59 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6) δ 169.81, 139.06, 137.93, 128.89, 128.82, 127.68, 126.85, 126.81, 125.52, 122.64, 120.39, 120.33, 62.76, 54.72, 52.81, 50.65, 31.73. MS (EI+): m/z calculated for $C_{21}H_{24}F_3N_3O$, 392.1905, found-392.1890 $[M+H]^+$.

N-(2-acetylphenyl)-2-((1-benzylpiperidin-4-yl)amino)acetamide **(84)**

Pale Pink (0.203g, 58.29% yield) M.P: 151-154°C, Rf=0.45 (EA/Hex,7:3,v/v).IR(ATR) : 1667(C=0 stretching),3367(Secondary NH stretching),2920(C-H stretching),1217(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 8.41 – 8.35 (Ar-H,m, J = 8.0, 1.3 Hz, 1H), 7.82 – 7.77 (Ar-H,m, J = 8.1, 1.5 Hz, 1H), 7.56 – 7.49 (Ar-H,td, J = 7.9, 1.6 Hz, 1H), 7.33 – 7.24 (CONH,Ar-H,m, 3H), 7.24 – 7.18 (Ar-H,td, J = 7.9, 1.4 Hz, 1H), 3.56 – 3.50 (CH2,m, 4H), 3.40 – 3.32 (NH,m, J = 7.1, 5.8 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.66 – 2.63 (CH3,s, 3H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 201.96, 167.99, 138.23, 133.18, 129.86, 129.71, 128.81, 127.95, 126.55, 123.95, 122.96, 63.66, 58.62, 51.25, 50.81, 30.87, 28.28. MS (EI+): m/z calculated for $C_{22}H_{27}N_3O_2$, 366.2137, found-366.2121 $[M+H]^+$.

N-(3-acetylphenyl)-2-((1-benzylpiperidin-4-yl)amino)acetamide **(85)**

Grey Solid (0.245g, 60.17% yield) M.P: 167-172°C, Rf=0.40 (EA/Hex,7:3,v/v). IR(ATR) : 1672(C=0 stretching),3354(Secondary NH stretching),2978(C-H stretching),1177(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.61 – 9.57 (CONH,s, 1H), 7.97 – 7.92 (Ar-H,t, J = 2.1 Hz, 1H), 7.75 – 7.70 (Ar-H,m, J = 7.9, 2.2, 1.3 Hz, 1H), 7.67 – 7.61 (Ar-H,m, J = 8.0, 2.2, 1.3 Hz, 1H), 7.45 – 7.38 (Ar-H,t, J = 8.0 Hz, 1H), 7.33 – 7.23 (Ar-H,m, 3H), 3.56 – 3.52 (CH2,d, J = 0.8 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 –

2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.57 – 2.53 (CH3,s, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.6, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 197.37, 168.41, 138.63, 138.23, 138.05, 128.81, 128.79, 127.95, 126.64, 126.55, 124.50, 121.11, 63.66, 58.62, 51.25, 50.81, 30.87, 27.79. MS (EI+): m/z calculated for $C_{22}H_{27}N_3O_2$ 366.2137, found- 366.2144 $[M+H]^+$.

N-(4-acetylphenyl)-2-((1-benzylpiperidin-4-yl)amino)acetamide **(86)**

Grey Solid (0.164g, 54.35% yield) M.P: 169-174°C, Rf=0.44 (EA/Hex,7:3,v/v). IR(ATR) : 1650(C=0 stretching), 3285(Secondary NH stretching), 2961(C-H stretching),1197(C-N stretching) cm^{-1}.^{1}H NMR (500 MHz, DMSO-d6) δ 9.61 – 9.58 (CONH,s, 1H), 7.89 – 7.83 (Ar-H,m, 2H), 7.54 – 7.48 (Ar-H,m, 2H), 7.33 – 7.23 (Ar-H,m, 3H), 3.56 – 3.52 (CH2,d, J = 0.8 Hz, 2H), 3.49 – 3.44 (CH2,d, J = 5.7 Hz, 2H), 3.40 – 3.32 (NH,m, J = 7.1, 6.0, 5.3 Hz, 1H), 3.02 – 2.94 (CH,dp, J = 7.1, 4.1 Hz, 1H), 2.80 – 2.71 (CH2,m, J = 12.6, 8.1, 5.4 Hz, 2H), 2.60 – 2.56 (CH3,s, 2H), 2.53 – 2.44 (CH2,m, J = 12.8, 8.2, 5.5 Hz, 2H), 1.96 – 1.86 (CH2,m, J = 12.5, 8.2, 5.5, 4.2 Hz, 2H), 1.69 – 1.59 (CH2,m, 2H). 13C NMR (125 MHz, DMSO-d6) δ 197.14, 168.41, 143.26, 138.23, 131.62, 130.19, 128.81, 127.95, 126.55, 119.59, 63.66, 58.62, 51.25, 50.81, 30.87, 27.79. MS (EI+): m/z calculated for $C_{22}H_{27}N_3O_2$ 366.2137, found -366.2149 (M+).

6.3.5 *In-vitro* studies

6.3.5.1 *In-vitro* cholinesterase inhibitory activity

Ellman's technique was utilized to perform the *in-vitro* studies on BChE and AChE inhibition for synthesized compounds [80, 82, 105]. *Equine* BChE and *Electrophorus electricus* AChE (CAS No. 9000-81-5 and 9001-08-1) were obtained from Sigma Aldrich, while acetylthiocholine iodide (ATCI, CAS No. 1866-15-5), butyrylthiocholine

iodide (BTCI, CAS No. 1866-16-6), and 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB, CAS No. 69-78-3) were purchased from Himedia.

The compounds were first tested for AChE and BChE inhibition at a concentration of 50 μM. Subsequently, the compounds showing promising % inhibition were further evaluated to determine their IC_{50}. Six concentrations of inhibitors (50, 500, 1000, 5000, 10000, and 50000 nM) were prepared from a stock solution of 1 mg/ml in methanol, with appropriate dilutions in the assay solution. The enzyme assays were carried out in phosphate buffer saline (PBS, pH 7.4).

In the assay, 10 μL of the working solutions of inhibitors and 100 μL of DTNB solution (0.005 M) prepared in PBS were mixed in a 96-well plate and incubated for 10 minutes. After incubation, 50 μL of AChE (0.5 U mL−1) or 50 μL of BChE (0.5 U mL−1) was added, followed by a further incubation of 15 minutes. Finally, 30 μL of the substrate, either ATCI (0.00235 M) or BTCI (0.00235 M) prepared in PBS, was added to the reaction mixture. The formation of the yellow-colored 5-thio-2-nitrobenzoate anion was measured at 1-minute intervals for 20 minutes at an absorbance of 415 nm using the Synergy HTX multimode reader (BioTek, USA). A blank containing 10 μL of methanol instead of the test compound was used for comparison, and the entire procedure was performed in triplicate. The IC_{50} values were determined using GraphPad Prism 5.

The enzyme kinetics analysis was carried out for compounds 34 and 37 to determine their inhibition mechanisms, using their respective IC_{50} values as a basis. To conduct the enzyme kinetic study, five different concentrations of each inhibitor were employed. For compound 34, the concentrations were 15, 30, 60, 120, and 240 nM, while for compound 37, the concentrations were 10.5, 21, 42, 84, and 168 nM. These concentrations were selected to cover two concentrations higher and two concentrations lower than the IC_{50} of each compound.

In the enzyme kinetic study, each inhibitor concentration was incubated with six different concentrations of ATCI (0.25, 1, 2, 3, 4, and 5 μM). The reaction mixture's absorbance was monitored for 25 minutes at 415 nm using the Synergy HTX multimode reader (BioTek, USA). A blank containing 10 μL of methanol instead of the test compound was used for comparison.

To determine the enzyme inhibition mechanism, the velocity of the enzyme reaction, obtained from the product formed, was utilized to calculate Vmax and Km using Michaelis-Menten nonlinear regression graph. Additionally, the Lineweaver-Burk reciprocal plots were employed to analyze the enzyme inhibition mechanism.

The Ki value, indicative of the inhibitor's binding affinity to the enzyme, was obtained from the Dixon plot of the slope of the double reciprocal Lineweaver-Burk plot against the concentrations of the test compounds. The entire enzyme kinetic assay was performed in triplicate to ensure accuracy and consistency of the results.

6.3.5.2 *In-vitro* BACE-1 inhibition assay

The identified hits were evaluated for their BACE-1 inhibition potential using a fluorescence resonance energy transfer (FRET) based BACE-1 fluorescence assay kit (Catalog No. CS0010, Sigma-Aldrich). The kit consists of fluorescent assay buffer, stop solution, substrate (7-Methoxycumarin-4-acetyl [Asn670, Lue671]-Amyloid βA4 Precursor Protein 770 Fragment 667-676-(2,4 dinitrophenyl) Lys-Arg-Arg amide trifluoroacetate salt) and BACE-1 enzyme. Different concentrations of test compounds were prepared. The fluorescence intensity was measured immediately after the addition of BACE-1 enzyme with the wavelength of excitation and emission was set at 320 nm and 405 nm, respectively. All the measurements were performed in triplicate. The percentage inhibition was calculated using the following formulae: $[(Io–Ii)/Io] \times 100$, where Io and Ii are the fluorescence intensities obtained in the absence and presence of

an inhibitor, respectively and the IC_{50} values were calculated using linear regression graph (GraphPad Prism 5.1, GraphPad Software Inc.) [106].

6.3.5.3 *In-vitro* blood-brain barrier permeation assay

Porcine brain lipid (PBL, CAS No. 475995-51-8) and dodecane were obtained from Avanti polar lipids, Alabaster and Avra Synthesis, Hyderabad, respectively. The donor and acceptor with PVDF membrane (pore size 0.45 mm) microplates (Cat No. MAIPNTR10 and MATRNPS50) were purchased from Merck Millipore. The BBB permeability was determined by the parallel artificial membrane permeation assay (PAMPA) of compounds **64, 70, 75, 77, 80** and **86** [107]. The acceptor plate was covered with 4 µL of PBL dissolved in dodecane at a concentration of 20 mg/ml. Then, it was hydrated with 200 µL of PBS at pH 7.4. To prepare the working solution of each compound, a 200-fold dilution of a 5 mg/mL stock solution in DMSO with PBS (pH 7.4) was made, resulting in a final concentration of 25 µg/ml for each compound. Afterward, 200 µL of the working solution was added to each well of the donor plate. The donor plate was placed over the acceptor plates, and the entire setup was incubated for 18 hours. To determine the compound's concentration in the acceptor, donor, and reference wells, UV spectroscopy (HTX multimode reader, BioTek, USA) was employed. Each sample was scanned at least five different wavelengths and in three independent runs.

In a previous study, the validity of this protocol was confirmed by testing nine commercial drugs (Verapamil, Diazepam, Progesterone, Atenolol, Dopamine, Lomefloxacin, Alprazolam, Chlorpromazine, and Oxazepam) with known blood-brain barrier (BBB) permeability [80].

Following formula was used to calculate the permeability-

$$P_e = C \times -ln\left(1 - \frac{[OD]_{acceptor}}{[OD]_{equilibrium}}\right)$$

$$C = \left(\frac{(V_D - V_A)}{(V_D + V_A)a \times t}\right)$$

Where $OD_{acceptor}$ is the absorbance of Acceptor Solution minus Blank, $OD_{equilibrium}$ is the absorbance of the Equilibrium Standard minus Blank, and, using an 18 hour incubation, C = 7.72×10^{-6} cm/s.

6.3.5.4 Propidium iodine displacement assay

The Propidium iodide displacement assay serves as a valuable technique for evaluating the binding of a compound to the peripheral site of AChE. This is achieved by competitively displacing propidium iodide. In the experiment, AChE (5U) was mixed with test compounds (final concentrations 10 µM and 50 µM, 150 µl) and incubated for 6 hours at 25 °C. Another set of AChE samples was incubated without the test compounds to serve as controls. Following the incubation period, propidium iodide (final concentration 1 µM, 50 µl) was added to reach a final assay volume of 200 µl. After a 10-minute incubation, the fluorescence intensity was measured using a fluorescence plate reader (BioTek Synergy HTX) with excitation and emission wavelengths set at λ_{ex} = 535 nm and λe_m = 595 nm, respectively.

The percentage inhibition was calculated by following expression: $100 - (IF_i/IF_0 \times 100)$, where IF_i and IF_0 are the fluorescence intensities with and without inhibitor, respectively. Each assay was performed in triplicates, as three separate experiments [108].

6.3.5.5 *In-vitro* $A\beta_{1-42}$ aggregation inhibition activity

The $A\beta_{1-42}$ peptide for assessing $A\beta_{1-42}$ aggregation inhibition activity through the thioflavin T assay was obtained from Sigma-Aldrich, USA. A stock solution of $A\beta_{1-42}$ (1 mg/mL) was prepared by dissolving 1 mg of the peptide in 80 µL of 1% NH_4OH and making up the volume to 1 mL with PBS 7.4. Working concentrations of $A\beta_{1-42}$ (10 µM) were then prepared using PBS 7.4.

For the self-induced $A\beta_{1-42}$ aggregation inhibition assay, $A\beta_{1-42}$ (10 μM: 2 μl) was incubated with or without test compounds (5 μM, 10 μM, and 20 μM: 2 μl) at 37 °C for 48 hours. After the incubation period, 178 μl of 20 μM thioflavin T (ThT) was added to the 96-well plates containing the samples. Blank readings were taken using PBS buffer 7.4 instead of inhibitors. Fluorescence intensity was measured at 450 nm (excitation) and 485 nm (emission) using a microplate reader. The resulting data were plotted as normalized fluorescence intensity (NFI) with or without test compounds at different concentrations [109-111].

6.3.6 *In-vivo* evaluation of compounds

6.3.6.1 Animals, housing, and materials

Wistar rats weighing between 200–250 g of both sexes were obtained from the Institute of Medical Sciences, Banaras Hindu University, Varanasi. The rats were housed at the Animal House, Department of Pharmaceutical Engineering and Technology, Indian Institute of Technology (BHU), Varanasi. The animals were kept under controlled environmental conditions with a temperature of 25 ± 2 °C, humidity of 65 ± 5% RH, and 12 h light/dark cycles. They were provided with unrestricted access to commercial feed and water. The research protocol for behavioral studies was approved by the institutional animal ethics committee under the reference number 'IIT (BHU)/IAEC/2022/008'.

6.3.6.2 Acute oral toxicity evaluation

To assess toxicity, the evaluation followed the OECD 423 guidelines [112]. Acute toxicity studies were conducted on female rats (n = 6) by administering the drugs orally. These rats were closely observed for up to 24 hours to detect any behavioral changes, seizures, diarrhea, or mortality. Subsequently, they were kept under normal observation for an additional 14 days [113]. After the 14-day period, the animals were humanely euthanized, and their organs (brain, liver, kidney, and heart) were isolated. Thin

transverse sections (10 μm thick) were obtained using a cryostat (SLEE MEV, Germany), mounted on glass slides, and stained with hematoxylin and eosin. The slides were then examined under a bright-field microscope (Magnus MLX plus) at 10X resolution to identify any potential damage to the organs [109].

6.3.6.3 Scopolamine-induced amnesia model for testing cognition enhancement in rats

The animals were segregated into nine groups, with six animals in each group. The following treatments were used in the study: (I) control, (II) SCO (5 mg/Kg), (III) SCO + compound 72 (5 mg/Kg), (IV) SCO + compound 72 (10 mg/Kg), (V) SCO + compound 72 (20 mg/Kg), (VI) SCO + compound 77 (5 mg/Kg), (VII) SCO + compound 77 (10 mg/Kg), (VIII) SCO + compound 77 (20 mg/Kg), and (IX) SCO + DNP (5 mg/Kg). DNP and SCO were dissolved in distilled water, while the investigational compounds were suspended in 0.5% SCMC prior to the administration. SCO was administered through intraperitoneal injection (i.p.) and other compounds were administered through the oral route (p.o.) using oral gavage. The compounds, except SCO, were administered for seven days, while SCO was administered on the seventh day to induce amnesia. The behavioural experiments were performed half-hour after the SCO administration [114, 115].

6.3.6.3.1 Y-maze test

The Y-maze is a device consisting of three identical wooden arms (A, B, and C) separated by a 120˚ angle. Its purpose is to assess intermediate working and spatial memory. On the seventh day of the treatment, the test compounds 72, 77, and DNP were evaluated. Initially, a training session was conducted with one arm closed off using a wooden partition. The animal was allowed to enter the maze with its head facing towards the centre. This training occurred after dosing and four hours before the test session, during which the animal explored the maze for 15 minutes.

The test session took place half an hour after the administration of SCO (the substance being tested). In this session, the previously closed arm was opened, and the rodent was free to move around the maze for five minutes. The researchers recorded the sequence of arm entries. Repeated entries in the same arm indicated memory impairment, whereas novel arm entries and spontaneous alterations in three-consecutive components (ABC, BCA, CAB, but not ABA) were considered indicators of memory improvement [116, 117].

After each session, the maze was cleaned with 70% ethanol to eliminate any potential olfactory clues that might influence subsequent results. The % spontaneous alteration was calculated as:

$$\%\ Spontaneous\ alteration = \frac{Number\ of\ alteration}{(total\ arm\ entries - 2)} \times 100$$

6.3.6.3.2 *Ex vivo* and biochemical analysis

After the experiments, the animals were euthanized, and their brains were removed. The neurochemical analysis was focused on the hippocampus and prefrontal cortex (PFC). Tissue homogenates were prepared by homogenizing the brain tissues in 10 mM phosphate-buffered saline (PBS) with a pH of 7.4. Subsequently, centrifugation was carried out at 15000 rpm for 15 minutes at 4°C to obtain the supernatant, which was used for further analysis.

ChE (Cholinesterase) activity was determined using the Ellman method, employing acetylthiocholine iodide (ATCI) and butyrylthicholine iodide (BTCI) as substrates. Initially, 10 µL of the supernatant was diluted with 100 µL of PBS, followed by the addition of freshly prepared substrate solutions (5 mM) and incubated for 5 minutes. Afterward, 1.5 mM of DTNB solution was added, and the absorbance was measured at 415 nm using a Synergy HTX multimode reader (BioTek, USA) against a blank.

The activity of the catalase (CAT) enzyme, responsible for converting harmful H_2O_2 into water and oxygen, was measured in tissue homogenates. This was achieved by mixing 10 µL of the supernatant with 150 µL of PBS, followed by the addition of 250 µL of H_2O_2 (160 mM) and incubating for 1 minute at 37°C. Then, 1.5 mL of a stopping solution containing dichromate/acetic acid (5% $K_2Cr_2O_7$/glacial acetic acid; 1:3 v/v) was added, and the reaction mixture was boiled for 15 minutes. The green color, a result of dichromate oxidation to chromic (III) sulfate, was compared with a control mixture lacking the enzyme. Absorbance was measured at 570 nm using the Synergy HTX multimode reader (BioTek, USA) against a blank.

Superoxide dismutase (SOD) activity was determined using Markland's method, based on the autoxidation of pyrogallol. In 10 µL of tissue homogenate, 200 µL of 0.1 M Tris-HCl with 1 mM EDTA at pH 8.2 were added, followed by the addition of 50 µL of 4.5 mM pyrogallol solution prepared in 1 µM HCl. Absorbance was measured after 1 minute at 325 nm wavelength using the Synergy HTX multimode reader (BioTek, USA) against a blank. A control sample without tissue supernatant was used to determine the enzyme activity. The experiments were performed in triplicates, and enzyme activities were normalized with respect to the control group.

6.3.6.4 Aβ-induced AD phenotypic model: Morris water maze test

To evaluate memory and learning impairment induced by $A\beta_{1-42}$, a neurotoxic model was employed using the Morris water maze test [11,39]. Rats (n=6) were randomly assigned to five groups: normal control (sham), $A\beta_{1-42}$ control (negative control), 72 (20 mg/kg p.o.), 77 (20 mg/kg p.o.), and DPZ (5 mg/kg p.o.). Anesthesia was induced using an i.p. injection of a ketamine and xylene mixture (90 and 9 mg/kg, respectively). Subsequently, rats were secured on a stereotaxic apparatus, and the head was fixed with ear bars after incising and retracting the scalp. Stereotaxic coordinates were set at bregma (-0.5 mm

anteroposterior, +1.2 mm mediolateral, -3.2 dorsoventral, and incision bar set at -3.3 mm) [11]. $A\beta_{1-42}$ sterile saline solution (4 μM, 5 μl) was then infused into all groups, except the sham group, using a Hamilton microsyringe at 2 μl/min. The syringe was left in place for an additional 5 minutes to prevent solution efflux. In the sham group, saline was infused instead of $A\beta_{1-42.}$

After seven days, treatment with the test drug VA10 and DPZ commenced, administered orally once a day for the next seven days. Memory and learning capacity were assessed through behavioral studies using a test apparatus divided into four quadrants filled with opaque water (25 ± 2 °C). The water maze, with dimensions of 62 cm height, 32 cm depth, and 121 cm diameter, was used. In training trials conducted over five days, the animals were placed in one quadrant, facing the wall, and given 120 seconds to locate the hidden platform (four trials per day with a 10-minute intra-trial interval). On the sixth day, before the probe trial, the platform was removed, and animals were allowed to swim for 120 seconds to assess spatial memory. Data from observed values (escape latency in training trials, time spent on the platform, and number of entries to the platform zone) were analyzed using GraphPad Prism 9.

6.3.6.5 Statistical analyses

All values were expressed as the mean ± standard error of the mean (SEM). One-way ANOVA followed by multiple comparison posthoc test was performed.

6.4 Results and discussion

6.4.1 Rationale of drug design & *in-silico* optimization

The aim of this research was to design and novel Multi-Target Directed Ligands (MTDLs) for AD utilizing *N-benzylpiperidine* scaffolds [118]. The active site of AChE is situated within a narrow gorge, reaching a depth of 20 Å, and is characterized by two crucial residues: Trp86 at the catalytic anionic site (CAS) and Trp286 at the Peripheral Anionic

Site (PAS) [119, 120]. Considering the distinctive architecture of the active site, our molecular design incorporates two aromatic systems linked by a piperidine and a flexible chain enhancer acetamide linker. This design facilitates π–π stacking interactions with the two vital residues of AChE, namely Trp86 and Trp286 . These interactions contribute to a snug fit of the ligand within the active site, promoting a more favorable free energy of complexation. The compounds for the study were designed by a hybrid pharmacophore approach and molecule generation technique (**Figure 6.3**).

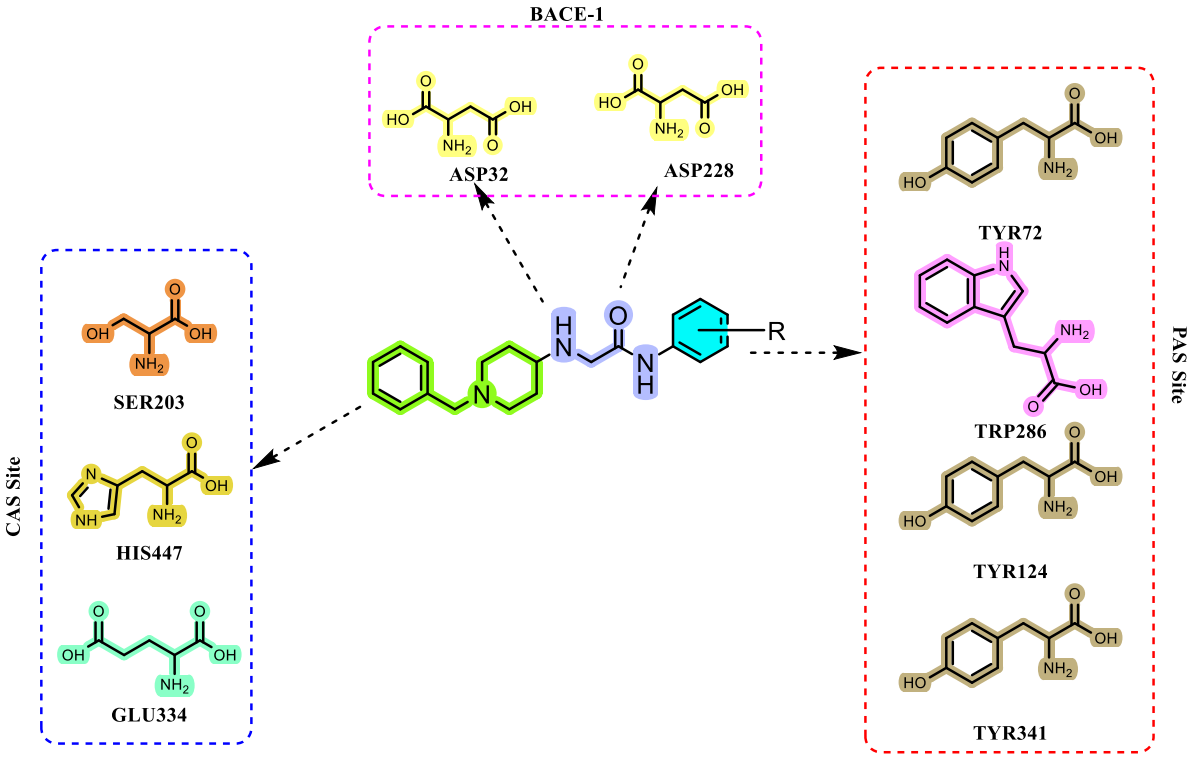
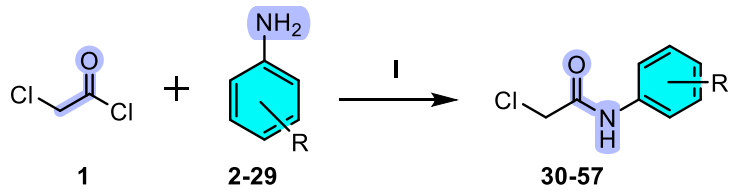


Figure 6.3 Interaction of N-benzylpiperidines with; PAS and CAS site of AChE and BACE-1.

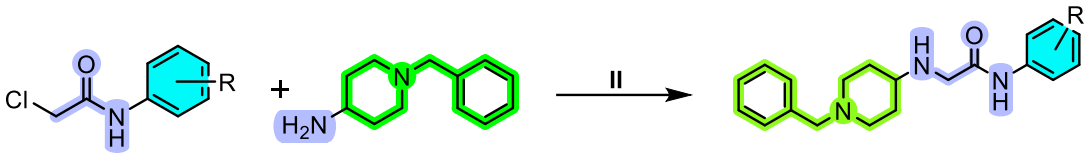
6.4.2 Chemistry

Step-I Synthesis of 2-chloro-N-phenylacetamide derivatives



1 2-29 30-57

Step-II Synthesis of N-benzyl piperidine derivatives



30-57 58 59-86

2, 30, 59 = -H	9, 37, 66 = -2F	16, 44, 73 = -3Cl	23, 51, 80 = -4CN
3, 31, 60 = -2NO_2	10, 38, 67 = -3F	17, 45, 74 = -4Cl	24, 52, 81 = -2CF_3
4, 32, 61 = -3NO_2	11, 39, 68 = -4F	18, 46, 75 = -2OCH_3	25, 53, 82 = -3CF_3
5, 33, 62 = -4NO_2	12, 40, 69 = -2CH_3	19, 47, 76 = -3OCH_3	26, 54, 83 = -4CF_3
6, 34, 63 = -2Br	13, 41, 70 = -3CH_3	20, 48, 77 = -4OCH_3	27, 55, 84 = -2$COCH_3$
7, 35, 64 = -3Br	14, 42, 71 = -4CH_3	21, 49, 78 = -2CN	28, 56, 85 = -3$COCH_3$
8, 36, 65 = -4Br	15, 43, 72 = -2Cl	22, 50, 79 = -3CN	29, 57, 86 = -4$COCH_3$

Figure 6.4 Reaction scheme (I). TEA/DCM, 1- 5 ℃, 1 hr (II). DIPEA/THF, reflux, 24hr

6.4.3 Docking studies

The docking studies with compounds 59 to 87, including DNP and CNP520, against BACE-1 (PDB ID- 6EQM) and AChE (PDB ID- 4EY7) revealed notable binding energies and ligand efficiencies. For BACE-1, the compounds exhibited binding energies ranging from -7.26 to -10.18 Kcal mol^{-1}, with corresponding ligand efficiencies between -0.29 and -0.377 Kcal mol^{-1}. Compound 72 displayed the strongest binding affinity with a remarkably low binding energy of -10.18 Kcal mol^{-1} and a ligand efficiency of -0.377 Kcal mol^{-1}. On the other hand, the interactions with AChE demonstrated binding energies from -7.27 to -9.67 Kcal mol^{-1} and ligand efficiencies between -0.269 and -0.371 Kcal mol^{-1}. Compound 77 exhibited the highest binding affinity for AChE with a binding

energy of -9.67 Kcal mol^{-1} and a ligand efficiency of -0.358 Kcal mol^{-1}. Overall, the docking results suggest that the investigated compounds have the potential to interact with both BACE-1 and AChE, indicating their possible role as dual inhibitors (**Table 6.1**).

Table 6.1 Summary of docking studies of designed compounds with AChE and BACE-1

Compound No.	**Docking result with AChE (PDB ID- 4EY7)**		**Docking result with BACE-1 (PDB ID- 6EQM)**	
	Binding energy (Kcal mol^{-1})	Ligand efficiency (Kcal mol^{-1})	Binding energy (Kcal mol^{-1})	Ligand efficiency (Kcal mol^{-1})
59	-8.38	-0.349	-7.29	-0.304
60	-8.83	-0.353	-7.86	-0.314
61	-8.73	-0.349	-7.64	-0.306
62	-8.81	-0.352	-7.55	-0.302
63	-8.65	-0.346	-8.25	-0.330
64	-9.20	-0.368	-8.08	-0.323
65	-9.04	-0.362	-7.73	-0.309
66	-8.68	-0.347	-7.40	-0.296
67	-8.27	-0.331	-7.26	-0.290
68	-8.57	-0.343	-7.21	-0.288
69	-8.95	-0.358	-7.38	-0.295
70	-9.27	-0.371	-7.51	-0.300
71	-8.84	-0.354	-8.37	-0.335
72	-7.27	-0.269	-10.18	-0.377
73	-8.02	-0.297	-8.84	-0.327
74	-8.57	-0.317	-7.83	-0.29
75	-9.16	-0.339	-8.78	-0.325
76	-8.79	-0.326	-8.65	-0.32
77	-9.67	-0.358	-8.53	-0.316
78	-9.49	-0.365	-7.44	-0.286
79	-9.29	-0.357	-8.65	-0.333
80	-9.48	-0.365	-8.28	-0.318
81	-8.06	-0.288	-7.49	-0.268
82	-8.43	-0.301	-7.35	-0.263
83	-8.43	-0.301	-7.44	-0.266
84	-9.09	-0.350	-7.28	-0.280
85	-8.29	-0.319	-7.39	-0.284
86	-8.65	-0.333	-8.15	-0.313
DNP	-8.47	-0.247	--	--
CNP520	--	--	-8.32	-0.323

6.4.4 Molecular property and toxicity prediction

The **Table 6.2** encapsulates information on diverse chemical compounds, encompassing compound ID, molecular weight, cLogP, hydrogen bond acceptors, hydrogen bond donors, and polar surface area. Noteworthy trends include a consistent count of 4 hydrogen bond acceptors and 2 hydrogen bond donors, alongside molecular weight variations. Compounds exhibit cLogP values ranging from 1.3879 to 3.1578. Notably, compounds with IDs 65 and 86 are flagged for high tumorigenicity, with compound 86 additionally marked as highly mutagenic (**Table 6.2**).

6.4.5 *In-silico* ADME prediction analysis

The **Table 6.3** provides a detailed overview of several compounds, each identified by a unique Compound ID. The compounds share commonalities in terms of their high aqueous solubility and gastrointestinal absorption, suggesting favorable characteristics for absorption in the digestive system. Moreover, all the compounds demonstrate the ability to permeate the blood-brain barrier, indicating potential central nervous system activity. In terms of enzyme inhibition, none of the compounds act as inhibitors of CYP2C9, while all of them inhibit CYP2D6 and CYP3A4. This information is crucial in understanding potential drug interactions, as these enzymes play vital roles in drug metabolism. The absence of Pan-Assay Interference Compounds (PAINS) alerts suggests that the compounds do not possess structural elements known to interfere with common screening assays, enhancing their reliability in drug development. Overall, the detailed characteristics provided in the table offer valuable insights into the pharmacological and physicochemical properties of these compounds, which can be vital for further drug development and assessment of potential therapeutic applications.

Table 6.2 Physicochemical properties and predicted toxicities of synthesized compounds

Compound ID	Mol. weight	cLogP	HA	HD	Polar Surface Area	Mutagenic	Tumorigenic
59	323.439	2.3095	4	2	44.37	None	None
60	368.436	1.3879	7	2	90.19	None	None
61	368.436	1.3879	7	2	90.19	None	None
62	368.436	1.3879	7	2	90.19	None	None
63	402.335	3.0347	4	2	44.37	None	None
64	402.335	3.0347	4	2	44.37	None	None
65	402.335	3.0347	4	2	44.37	None	High
66	341.429	2.4103	4	2	44.37	None	None
67	341.429	2.4103	4	2	44.37	None	None
68	341.429	2.4103	4	2	44.37	None	None
69	337.465	2.6534	4	2	44.37	None	None
70	337.465	2.6534	4	2	44.37	None	None
71	337.465	2.6534	4	2	44.37	None	None
72	357.884	2.9155	4	2	44.37	None	None
73	357.884	2.9155	4	2	44.37	None	None
74	357.884	2.9155	4	2	44.37	None	None
75	353.464	2.2395	5	2	53.60	None	None
76	353.464	2.2395	5	2	53.60	None	None
77	353.464	2.2395	5	2	53.60	None	None
78	348.449	2.1451	5	2	68.16	None	None
79	348.449	2.1451	5	2	68.16	None	None
80	348.449	2.1451	5	2	68.16	None	None
81	391.436	3.1578	4	2	44.37	None	None
82	391.436	3.1578	4	2	44.37	None	None
83	391.436	3.1578	4	2	44.37	None	None
84	365.475	2.1807	5	2	61.44	None	None
85	365.475	2.1807	5	2	61.44	None	None
86	365.475	2.1807	5	2	61.44	High	High

(HA- H-bond acceptor, HD- H-bond donor)

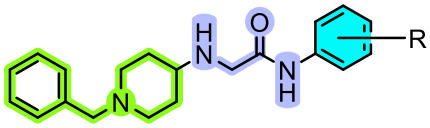
Table 6.3 *In-silico* ADME properties

Compound ID.	Aqueous solubility	GI absorption	BBB permeant	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	PAINS alerts
59	Soluble	High	Yes	No	Yes	No	0
60	Soluble	High	No	No	Yes	Yes	0
61	Soluble	High	No	No	Yes	Yes	0
62	Soluble	High	No	No	Yes	Yes	0
63	Moderately soluble	High	Yes	No	Yes	Yes	0
64	Moderately soluble	High	Yes	No	Yes	Yes	0
65	Moderately soluble	High	Yes	No	Yes	Yes	0
66	Soluble	High	Yes	No	Yes	No	0
67	Soluble	High	Yes	No	Yes	No	0
68	Soluble	High	Yes	No	Yes	No	0
69	Soluble	High	Yes	No	Yes	No	0
70	Soluble	High	Yes	No	Yes	No	0
71	Soluble	High	Yes	No	Yes	No	0
72	Moderately soluble	High	Yes	No	Yes	Yes	0
73	Moderately soluble	High	Yes	No	Yes	Yes	0
74	Moderately soluble	High	Yes	No	Yes	Yes	0
75	Soluble	High	Yes	No	Yes	No	0
76	Soluble	High	Yes	No	Yes	Yes	0
77	Soluble	High	Yes	No	Yes	Yes	0
78	Soluble	High	Yes	No	Yes	Yes	0
79	Soluble	High	Yes	No	Yes	Yes	0
80	Soluble	High	Yes	No	Yes	Yes	0
81	Moderately soluble	High	Yes	No	Yes	Yes	0
82	Moderately soluble	High	Yes	No	Yes	Yes	0
83	Moderately soluble	High	Yes	No	Yes	Yes	0
84	Soluble	High	Yes	No	Yes	Yes	0
85	Soluble	High	Yes	No	Yes	Yes	0
86	Soluble	High	Yes	No	Yes	Yes	0

6.4.6 *In-vitro* evaluation

The summary of results against eeAChE, eqBuChe and hBACE-1 has been reported in **Table 6.4.**

Table 6.4 Inhibitory potencies of compounds against eeAChE, eqBuChE and hBCAE-1



S. No	Comp.	R	eeAChE IC_{50} ± SEM (µM)	eqBuChE (% inhibition ± SEM)	hBACE-1 IC_{50} ± SEM (µM)
1	59	-H	2.144±0.083	22.46±1.079	
2	60	-2NO_2	0.165±0.04	14.63±1.053	1.384±0.064
3	61	-3NO_2	4.85±0.098	26.03±1.081	
4	62	-4NO_2	1.756±0.084	24.96±2.073	
5	63	-2Br	2.197±0.073	23.13±3.037	
6	64	-3Br	0.421±0.077	20.39±1.032	0.298±0.032
7	65	-4Br	0.804±0.075	5.620±4.074	
8	66	-2F	0.530±0.096	11.55±0.087	
9	67	-3F	4.168±0.068	17.36±2.072	
10	68	-4F	3.218±0.018	25.15±3.098	
11	69	-2CH_3	1.823±0.020	16.90±0.095	
12	70	-3CH_3	2.274±0.068	21.72±4.095	
13	71	-4CH_3	20.889±0.080	8.455±0.863	
14	72	-2Cl	0.117±0.030	19.29±3.048	0.322±0.028
15	73	-3Cl	16.206±0.075	11.90±1.025	
16	74	-4Cl	1.913±0.088	31.69±3.081	
17	75	-2OCH_3	0.413±0.010	5.325±0.560	1.936±0.144
18	76	-3OCH_3	1.168±0.079	22.98±2.054	
19	77	-4OCH_3	0.102±0.066	21.62±2.064	0.241±0.077
20	78	-2CN	0.837±0.022	12.86±1.081	
21	79	-3CN	2.819±0.084	20.70±1.065	
22	80	-4CN	0.148±0.081	27.30±0.094	0.844±0.072
23	81	-2CF_3	1.455±0.083	8.030±0.091	
24	82	-3CF_3	7.011±0.066	15.72±2.052	
25	83	-4CF_3	0.689±0.039	24.07±0.059	
26	84	-2$COCH_3$	1.609±0.079	21.42±0.062	
27	85	-3$COCH_3$	1.894±0.066	21.00±0.082	
28	86	-4$COCH_3$	0.348±0.009	7.463±4.034	2.002±0.427
29	DNP	--	0.037±0.01	26.96±2.051	

Data expressed as Mean ± SEM

6.4.6.1 Cholinesterase inhibition assay

The analysis of AChE inhibition reveals diverse trends among the tested compounds, shedding light on the structure-activity relationship (SAR). Compound 77 with a methoxy group at *para* position, stands out as a potent AChE inhibitor with a remarkably low IC_{50} of 0.102 ± 0.066 µM. Compound 60 with a nitro group at *ortho* position also showed significantly low IC_{50} of 0.165 ± 0.04 µM, suggesting that the electron-withdrawing nitro group enhances inhibitory potency. Similarly, Compound 72 containing a chloro snbstitution at *ortho* position, exhibits strong AChE inhibition with an IC_{50} of 0.117 ± 0.030 µM, emphasizing the favourable impact of certain electron-withdrawing substituents. The compounds with methyl substitution at either *ortho*, *meta*, and *para* position did not show good AChE inhibition potential as their IC_{50} values were over 1 µM.

The halogen substituents in Compounds 63, 64, 65, and 72 show diverse effects on AChE inhibition. While the Compound 63 with bromo substitution on *ortho* position and Compound 65 with bromo substitution at *para* position exhibit moderate inhibition whereas, Compound 64 having bromo substitution at *meta* position and Compound 72 with chloro substitution *ortho* position at demonstrate stronger inhibitory effects, suggesting a structural dependency on halogen substituents.

Each compound's unique structure contributes to its specific inhibitory activity against AChE, illustrating the complex interplay between R groups and inhibitory potency. These observations underscore the importance of considering both electron-withdrawing and electron-donating substituents in designing AChE inhibitors and offer valuable insights for further drug development efforts in the context of neurodegenerative disorders.

The 2D interaction diagram of Compound 60, 64, 72, 77 and 80 with crustal structure of AChE (PDB ID- 4ey7) has been shown in **Figure 6.6.** The interaction diagram shows the interaction of the selected ligands with important residues of AChE.

The enzyme kinetics study was performed to understand the type of inhibition. The enzyme kinetic study involved varying substrate concentrations (ranging from 0.25 to 5 mM) while maintaining a constant enzyme concentration to elucidate the mechanism of enzyme inhibition (**Figure 6.5**). The analysis utilized a double reciprocal Lineweaver-Burk plot, revealing that both compounds **72** and **77** exhibit a non-competitive mode of AChE inhibition. The converging lines in the third quadrant of the plot indicate that these inhibitors display a higher affinity for the enzyme-substrate complex ($\alpha < 1$) rather than the free enzyme (**Figure 6.5**).

To further quantify this interaction, Dixon plots were employed to determine the inhibition constants (Ki). The calculated inhibition constants for compounds **72** and **77** were found to be 171.99 and 203.02 nM, respectively, as illustrated in **Figure 6.5**. This data provides valuable insights into the inhibitory mechanism and potency of these compounds against AChE, emphasizing their ability to interact preferentially with the enzyme-substrate complex rather than the free enzyme during the inhibition process.

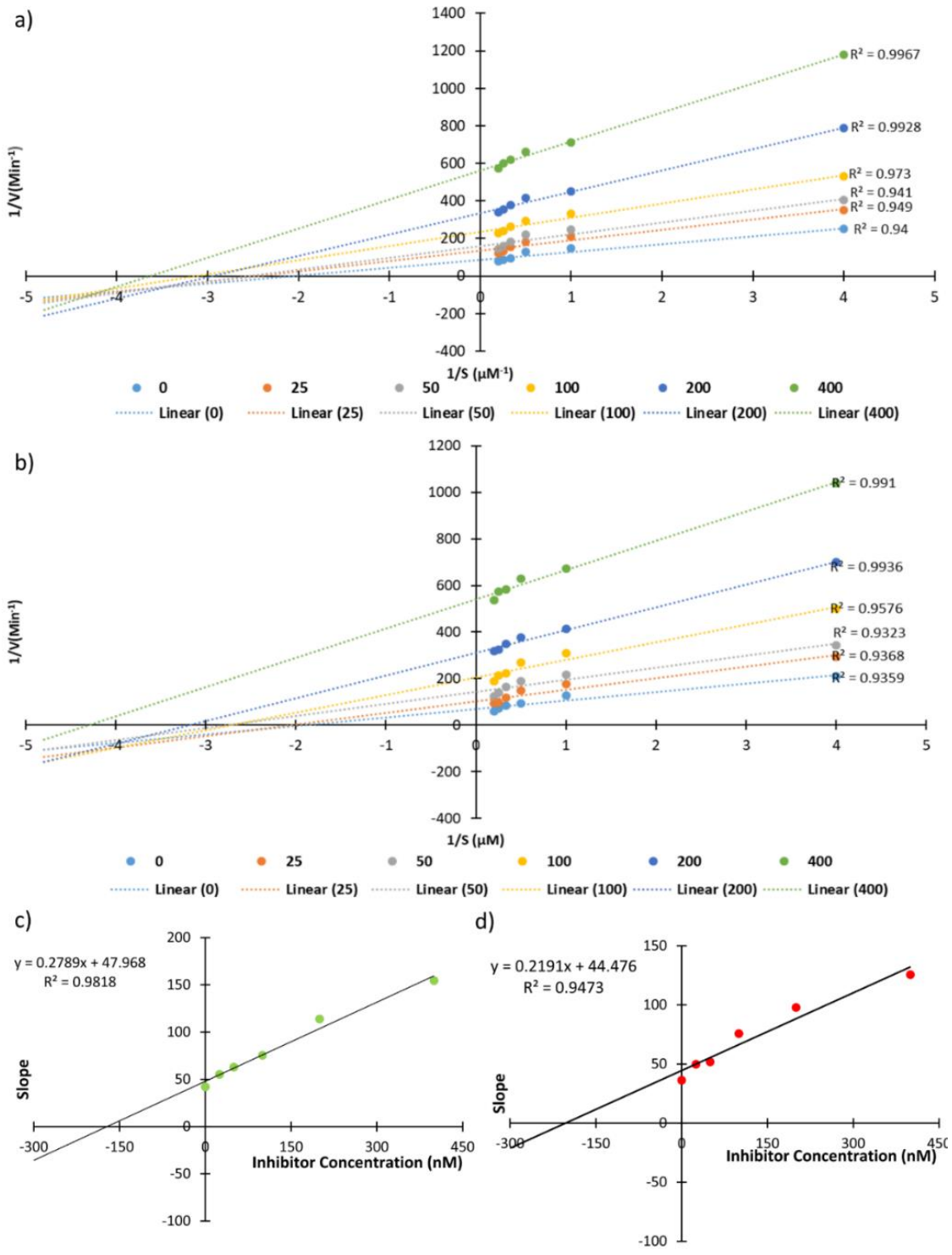


Figure 6.5 Lineweaver Burk double reciprocal plot of compounds (a) 72 and (b) 77. Dixon plot for compound (c) 72 and (d) 33 for Ki calculation

6.4.6.2 *In-vitro h*BCAE-1 inhibition assay

On the basis of AChE inhibition activity, seven compounds i.e., Compound 60, 64, 72, 75, 77, 80 and 86 were selected hBACE-1 inhibition assay. The result showed that

compound 72 and compound 77 showed good hBACE-1 inhibitory potential as compared to other five compounds (**Table 6.4**). The 2D interaction diagram of compound 72 and 77 has been shown in **Figure 6.7**. The diagrams shows interaction of ligands with important residues of BACE-1 enzyme.

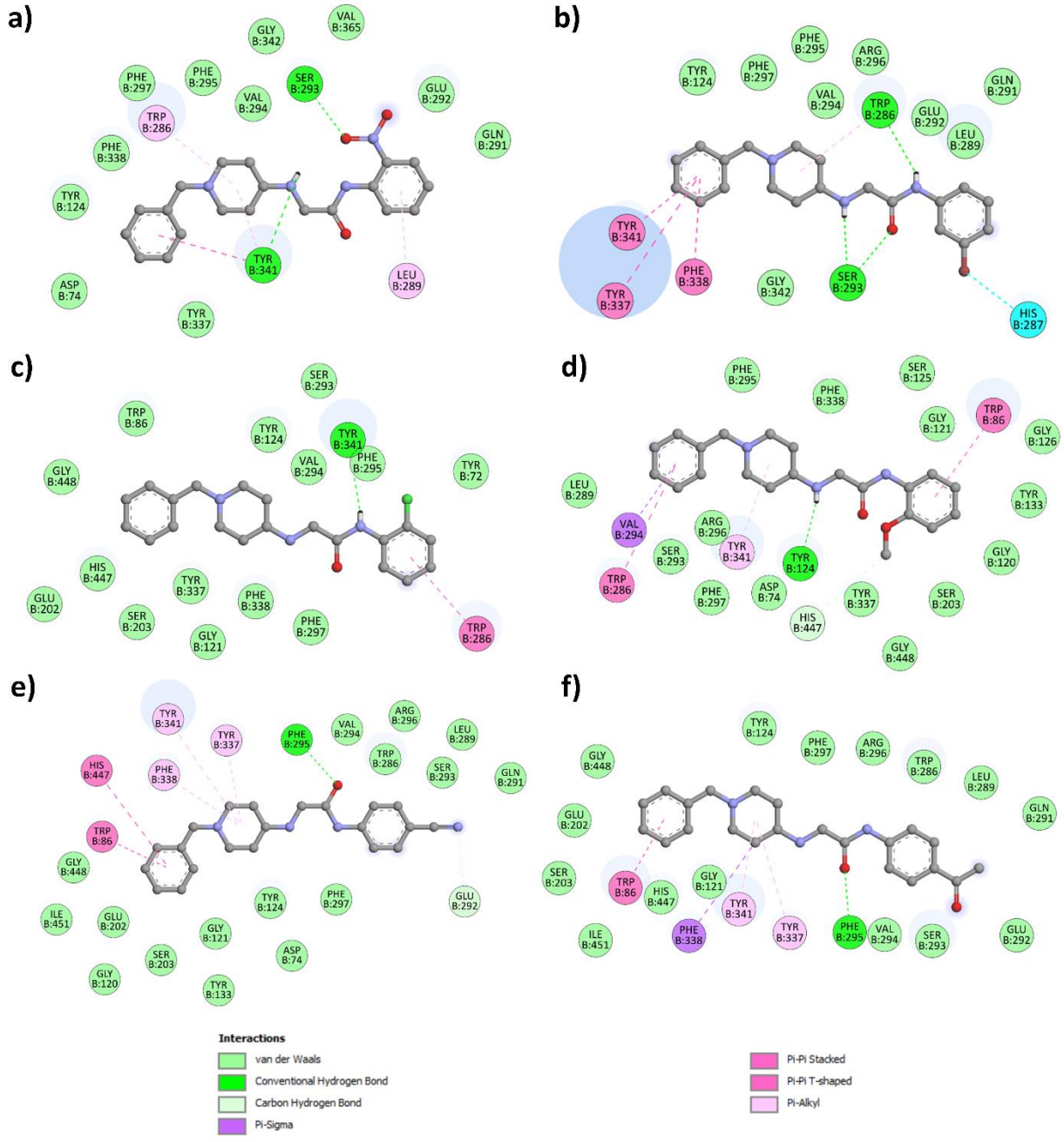


Figure 6.6 2D interaction diagram of ligands with AChE (PDB id- 4ey7). a) Compound 60, b) Compound 64, c) Compound 72, d) Compound 77, e) Compound 80 f) Compound 86

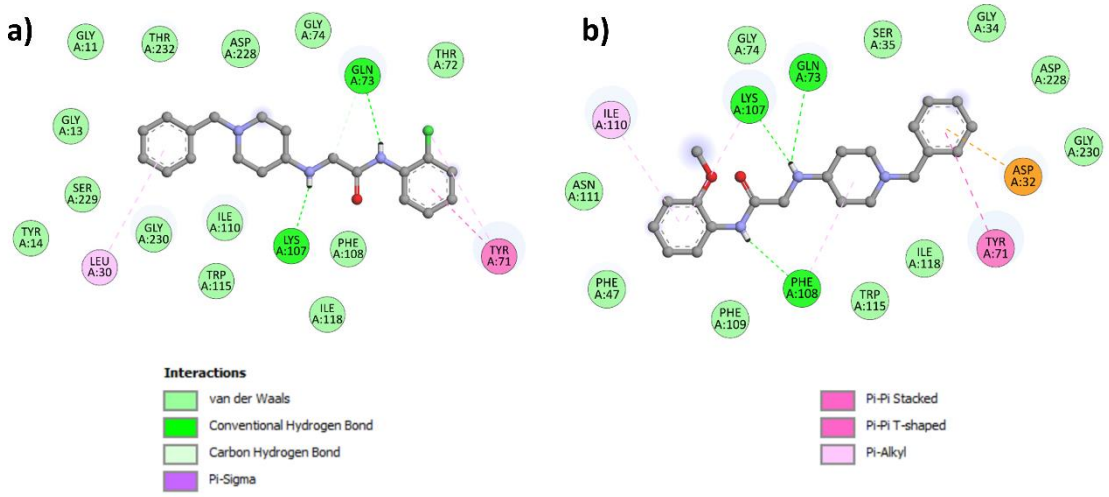


Figure 6.7 2D interaction diagram of ligands with BACE-1 (PDB id- 6eqm)

6.4.6.3 Propidium iodide (PI) displacement assay

Propidium iodide displacement assay was performed on AChE to test the ability of compounds to displace propidium with reference to the donepezil at 10 and 50 µM. Molecular docking studies revealed substantial interactions between compounds **72** and **77** and the PAS residues. Additionally, both enzyme inhibition assays and PAMPA experiments demonstrated potent inhibition of AChE and significant BBB permeation. Consequently, the binding affinity of compounds **72** and **77** to the PAS was assessed using a propidium iodide displacement assay at concentrations of 10 and 50 µM. The interaction of compounds **72** and **77** with PAS-AChE led to a reduction in fluorescence intensity. The results indicated that at 10 µM, compound **72** exhibited a 12.35% inhibition of propidium iodide displacement, while compound **77** showed a slightly higher inhibition of 15.84%. At the higher concentration of 50 µM, compound **72** demonstrated a 16.90% inhibition, and compound **77** exhibited a further increase to 21.28%. These findings suggest that both compounds exhibit a dose-dependent inhibition of propidium

iodide displacement from the AChE PAS. As a reference, the standard drug donepezil displayed a higher inhibition, with values of 19.75% at 10 µM and 32.80% at 50 µM. Overall, the results imply that compounds **72** and **77** have a notable impact on the interaction between propidium iodide and the AChE PAS, indicative of their potential as modulators of AChE function. The results obtained from the propidium iodide displacement assay align with the findings from the molecular docking studies of compounds **72** and **77 (Table 6.5)**.

Table 6.5 Propidium iodide displacement assay.

Compound	**Propidium iodide displacement from AChE PAS (% inhibition)[a]**	
	At 10 µM	At 50 µM
72	12.35 ± 1.81	16.90 ± 0.78
77	15.84 ± 2.14	21.28 ± 1.74
Donepezil	19.75 ± 2.92	32.80 ± 1.77

[a] The results are reported in Mean ± SEM (n=3)

6.4.6.4 *In-vitro* blood-brain barrier permeation assay

The PAMPA assay results provide insights into the blood-brain barrier (BBB) permeation potential of various compounds. Compound 60 exhibited a permeability value of 2.907 x 10^{-6} cm/s, suggesting a moderate likelihood of crossing the BBB (CNS±). In contrast, compounds **64, 72,** and **77** demonstrated higher permeability values of 7.435 x 10^{-6}cm/s, 5.713 x 10^{-6} cm/s, and 4.453 x 10^{-6} cm/s, respectively, indicating a positive prediction for BBB permeation (CNS+). Compound 80, with a permeability value of 1.989 x 10^{-6} cm/s, falls within the moderate range (CNS±). These findings imply that compounds 64, 72, and 77 exhibit enhanced BBB permeation potential, making them promising candidates for further investigation in the context of central nervous system drug development (**Table 6.6**).

Table 6.6 Permeability Pe ($10^{-}6$ cm s^{-1}) data for selected potent compounds from the PAMPA-BBB assay along with their BBB Penetration prediction.

Compound code	Pe (10^{-6} cm s^{-1})	Prediction
60	2.907±0.741	CNS±
64	7.435±0.966	CNS+
72	5.713±0.620	CNS+
77	4.453±0.806	CNS+
80	1.989±0.730	CNS±

All data were expressed as Mean ± SD for experiment performed in triplicates.
Compounds with Pe > 4.324×10^{-6} cm s^{-1} could cross the BBB (CNS+), Pe < 1.846×10^{-6} cm s^{-1} could not cross the BBB (CNS−) and 1.846×10^{-6} cm s^{-1} < Pe < 4.324×10^{-6} cm s^{-1} showed uncertain BBB permeation (CNS ±).

6.4.6.5 *In-vitro* AChE induced $A\beta_{1-42}$ aggregation inhibition assay

The accumulation and aggregation of $A\beta_{1-42}$ are recognized as significant detrimental factors in AD. Findings from the PI displacement assay underscore the notable binding affinity of compounds **72** and **77** with PAS-AChE. Importantly, it is well-established that the PAS binding of the inhibitor not only leads to AChE inhibition but also plays a crucial role in preventing the aggregation of $A\beta_{1-42}$ [50]. The thioflavin T assay was employed to assess the anti-$A\beta_{1-42}$ aggregation activity of the compounds in self-induced experiments. Tests were conducted at three distinct concentration ratios of $A\beta_{1-42}$ and the inhibitor (10:5 μM, 10:10 μM, and 10:20 μM, respectively), and the results were quantified as normalized fluorescence intensity (NFI). The outcomes revealed a concentration-dependent inhibition of $A\beta_{1-42}$ aggregation, with maximum effectiveness observed at a 20 μM inhibitor concentration (**Figure 6.8**). Overall, the graph indicates that Compounds 72 and 77 exhibited concentration-dependent inhibition of Amyloid beta aggregation, with maximum inhibition at 20 μM concentration, comparable to the standard drug Donepezil. This suggests the potential of these compounds in mitigating $A\beta_{1-42}$ aggregation, a key pathological hallmark of AD.

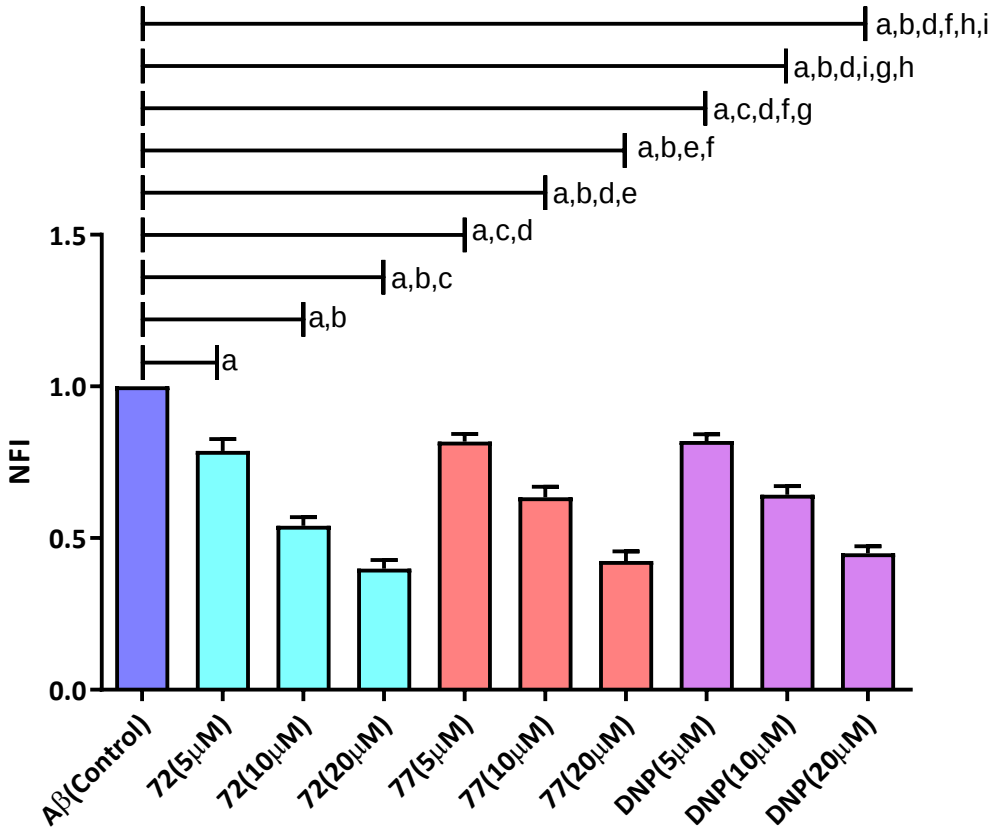


Figure 6.8 Effect of compounds 72 and 77 on $A\beta_{1-42}$ aggregation. [a] $p<0.05$ vs. Control, [b]$p<0.05$ vs. 72(5μM), [c]$p<0.05$ vs. 72(10μM), [d]$p<0.05$ vs. 72(20μM), [e]$p<0.05$ vs. 77(5μM), [f]$p<0.05$ vs. 77 (10μM), [g]$p<0.05$ vs. 77(20μM), [h]$p<0.05$ vs. DNP(5μM), [i]$p<0.05$ vs. DNP(10μM).

6.4.7 *In-Vivo* evaluation

6.4.7.1 Acute Oral Toxicity Study

Healthy female Wistar rats were used in acute oral toxicity studies to evaluate the safety of compounds **72** and **77**. The rats treated with compounds **72** and **77** did not display any signs of toxicity, such as mortality, changes in body color, weight loss, or abnormal behavior (**Table 6.7**). No significant differences were observed in biochemical parameters between the normal control group and the drug-treated rats, including creatinine, creatine kinase myocardial band (Ck-MD), alkaline phosphatase (ALP), aspartate transaminase (AST), and alanine transaminase (ALP) (**Table 6.8**). Additionally,

microscopic examination of major organs such as the brain, liver, kidney, and heart tissues revealed no adverse reactions in response to the drug therapy at doses up to 300 mg/kg when observed at 10X magnification (**Figure 6.9**).

Table 6.7 Effect of Single-Dose Oral Administration of Compound 72 and Compound 77

No. of days	Normal control (Weight in grams	Compound 72 administered rats (weight in grams)	Compound 77 administered rats (weight in grams)
0th day	214.18 ± 17.27	211.09 ± 14.25	216.49± 11.74
7th day	220.42 ± 16.36	216.38 ± 15.34	218.62± 09.16
14th day	224.76 ± 14.82	219.41 ± 14.72	222.71± 10.61

All values are expressed in Mean ± SD (n=6)

Table 6.8 Effect of Oral Administration of Compound 72 and 77

Biochemical Parameters	Normal Control	Compound 72 (300 mg/kg, po)	Compound 77 (300 mg/kg, po)
ALP (U/L)	63.64 ± 4.67	62.42 ± 2.90	64.96 ± 3.53
AST (U/L)	128.52 ± 5.31	125.25± 8.71	130.69± 4.48
ALT (U/L)	45.62 ± 3.73	47.52± 3.32	51.97 ± 4.09
Creatinine (mg/dl)	0.64 ± 0.07	0.66 ± 0.03	0.67 ± 0.14
Ck-MD (U/L)	664.47 ± 21.25	668.12 ± 19.47	661.91 ± 18.44

All values are expressed in Mean ± SD (n=6)

6.4.7.2 Scopolamine-induced amnesia model for testing cognition enhancement in rats

6.4.7.2.1 Y-maze test

The study aimed to assess the cognitive-enhancing effects of compounds **72** and **77** in comparison to the established cognitive enhancer, donepezil, using a scopolamine-induced Y-maze test in healthy male Wistar rats. Scopolamine is commonly used to induce cognitive impairment by affecting cholinergic function. The standard dose of donepezil (5 mg/kg, p.o.) and varying doses (5, 10, and 20 mg/kg, p.o.) of compounds **72** and **77** were administered to separate groups of rats for seven consecutive days. On the seventh day, scopolamine hydrobromide (0.5 mg/kg, i.p.) was administered to induce

cognitive impairment, and the Y-maze test was conducted to assess spontaneous alternations. Improved cognitive function was indicated by an increased percentage of spontaneous alternation.

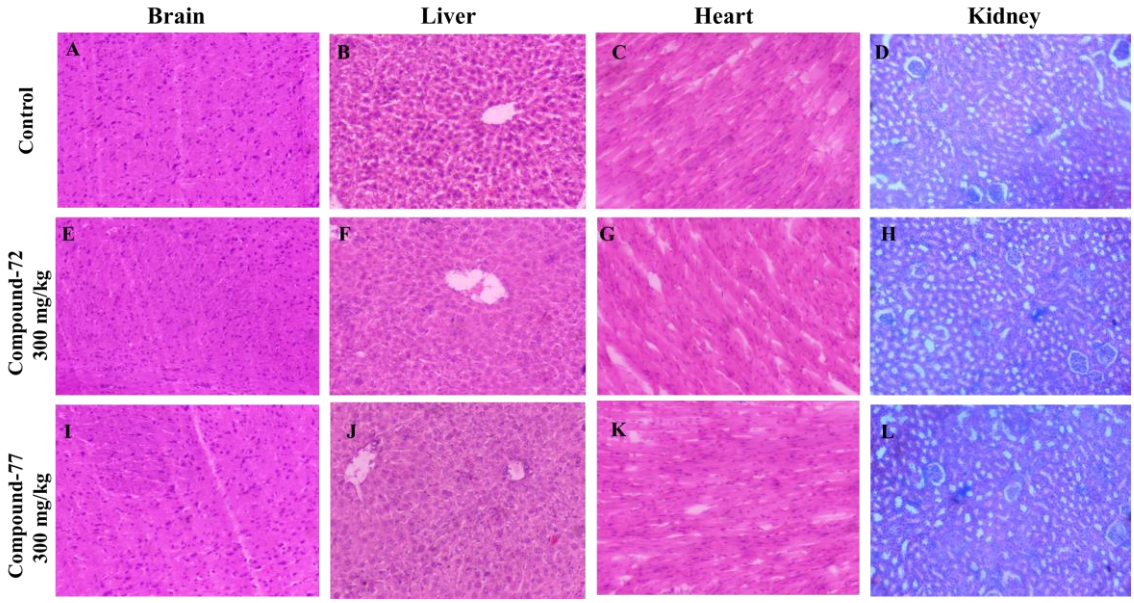


Figure 6.9 Acute oral toxicity study. Effects of the normal control group (A-D), Compound-72 (E-H) and Compound-77 (I-L) treatment on the Brain, Liver, Heart, and Kidney, respectively

Scopolamine administration led to a significant reduction in % spontaneous alternation compared to the control group, confirming the induction of cognitive deficit (Figure 5A, $p < 0.05$). Donepezil treatment exhibited significantly higher % spontaneous alternations compared to scopolamine (**Figure 6.10 (a)**, $p < 0.05$). Moreover, compounds **72** and **77** demonstrated a dose-dependent improvement in scopolamine-induced cognitive deficit, reaching maximum efficacy at 20 mg/kg. Treatment with both compounds resulted in significantly higher % spontaneous alternations compared to scopolamine, except for compound **77** at the 5 mg/kg dose (**Figure 6.10 (a)**, $p < 0.05$). Notably, at the highest dose (20 mg/kg), compounds **72** and **77** showed no statistically significant difference in % spontaneous alternation compared to donepezil (**Figure 6.10 (a),** ns).

6.4.7.2.2 *Ex-vivo* and biochemical analysis

The *ex-vivo* investigation of AChE activity was conducted using the Ellman assay, and the findings indicated a significant increase in AChE activity induced by scopolamine (**Figure 6.10 (a)**, $p < 0.05$). This elevation was attributed to extensive substrate hydrolysis compared to the control group. However, treatment with compounds **72, 77,** and donepezil exhibited notable reductions in AChE activity (**Figure 6.10 (b)**, $p < 0.05$) compared to the scopolamine-induced condition.

Furthermore, a dose-dependent elevation in acetylcholine (ACh) levels was noted in the groups treated with the drug (**Figure 6.10 (c)).** Notably, at a dose of 20 mg/kg, compound **72**, compound **77**, and donepezil exhibited no statistically significant difference in ACh levels (**Figure 6.10 (c),** ns). The behavioral and neurochemical assessment data implied that compounds **72** and **77** effectively inhibited AChE activity, leading to an increased availability of ACh in the brain.

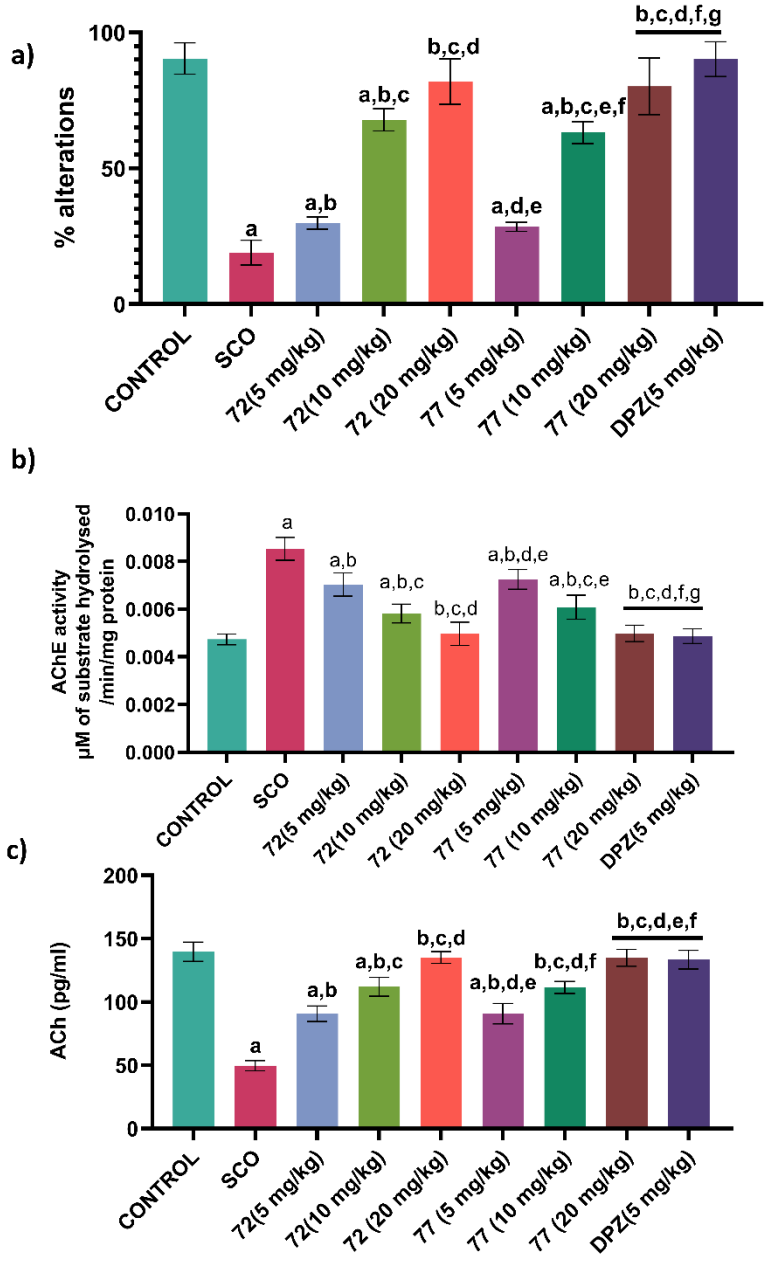


Figure 6.10 Effect of compounds 72 and 77 on scopolamine-induced cognition and memory impairment. (a) Effect of compounds 72 and 77 on % spontaneous alterations. (b) Effect of compounds 72 and 77 on AChE activity (c) Effect of compounds 72 and 77 on ACh levels. [a]p<0.05 vs. control; [b]p<0.05 vs. SCO; [c]p<0.05 vs. Compound 72 (5 mg/kg); [d]p<0.05 vs. Compound 72 (10 mg/kg); [e]p<0.05 vs. compound 72 (20 mg/kg); [f]p<0.05 vs. Compound 77 (5 mg/kg); [g]p<0.05 vs. Compound 77 (10 mg/kg) One way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine]

6.4.7.3 Aβ-induced AD phenotypic model: Morris water maze test

The primary pathological condition in AD is the cerebral aggregation of $A\beta_{1-42}$. $A\beta_{1-42}$ plaques initiate toxicity and neurodegeneration by activating various biochemical cascades, such as inflammation, oxidative stress, and apoptosis, leading to the deterioration of cognitive and memory functions [121]. Cholinergic activity may decrease due to increased AChE activity around amyloid plaques. Previous studies have indicated that the accumulation of Aβ reduces ACh levels in the AD brain by upregulating the expression of AChE [122]. The intracerebroventricular injection (ICV) of $A\beta_{1-42}$ into the rat brain mimics AD-like behavior, resembling human AD [123]. $A\beta_{1-42}$ deposits cause neuronal inflammation and microglial activation, potentially resulting in learning and memory deficits in rats [124].

To assess learning and memory, behavioral studies were conducted using the Morris water maze test, i.e., escape latency time (ELT) and the time spent in the platform zone. During training trials, the escape latency time for rats decreased gradually, and the $A\beta_{1-42}$ group (negative control) took more time to find the platform (**Figure 6.11 (a)**). After completion of last training trial, memory retention was predicted by a special probe trial test, with the removal of the platform. The cognitive impairment induced by Aβ administration affirmed a significant extension of ELT (**Figure 6.11 (a**)) in the model group of animals compared to the control over 5 days. The treatment with compound **72** and **77** drastically reduced ELT compared to the model group of animals. The time spent in the platform zone was greater in the drug treated group than the $A\beta_{1-42}$ group (**Figure 6.11(b)** $P<0.05$). The compound **72** (20 mg/kg) and **77** (20 mg/kg) showed a non-significant difference (**Figure 6.11 (b),** ns) with Donepezil treated group in time spent in the platform quadrant in the probe trial.

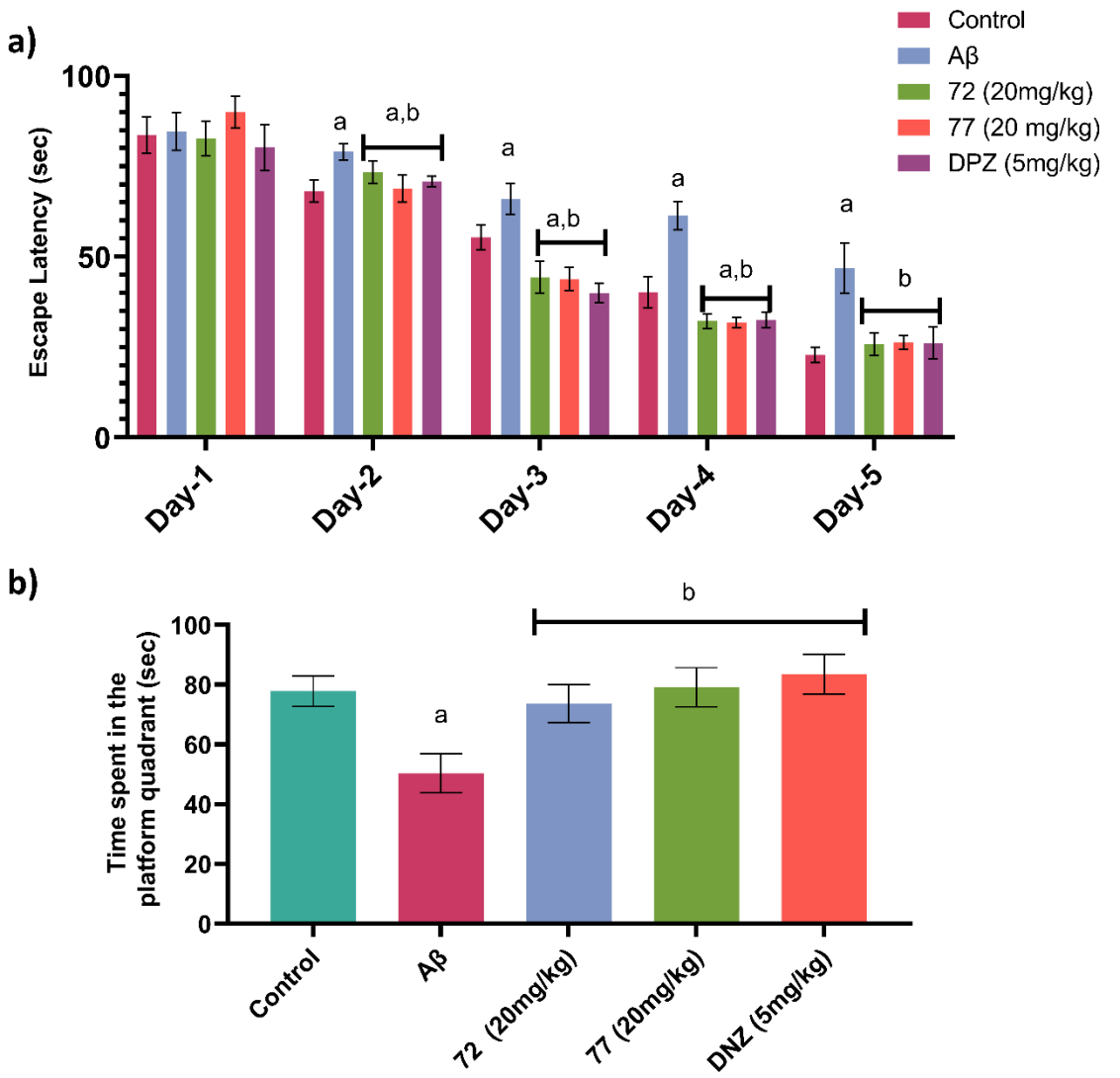


Figure 6.11 Protective effect of compounds 72, 77, and donepezil on $A\beta_{1-42}$-induced memory deficits analyzed by the Morris water maze test. (a) Escape latency during the training trials in the MWM tests; (b) time spent in the platform quadrant in the probe trial $^{a}p< 0.05$ vs. control and $^{b}p < 0.05$ vs. $A\beta_{1-42}$. One-way ANOVA followed by Newman - Keuls posthoc test.

6.5 Conclusions

This study employed a hybrid pharmacophore-based drug design approach to develop novel MTDLs targeting both AChE and BACE-1 for the treatment of AD. The design strategy focused on utilizing *N-benzylpiperidine* scaffolds to facilitate interactions with key residues in the active sites of the target enzymes. Through a combination of molecular

modeling, docking studies, molecular property predictions, and *in-vitro* assays, the potential of synthesized compounds to serve as dual inhibitors of AChE and BACE-1 was evaluated.

The docking studies revealed promising binding affinities and ligand efficiencies for the designed compounds against AChE and BACE-1. Additionally, molecular property predictions indicated favorable physicochemical properties and predicted toxicities for most compounds, with certain exceptions flagged for tumorigenic and mutagenic traits. *In- silico* ADME analysis suggested good aqueous solubility, gastrointestinal absorption, and BBB permeation for most compounds, along with minimal potential for drug interactions.

In-vitro evaluation demonstrated significant inhibitory activity of the synthesized compounds against AChE, and hBACE-1, with notable structure-activity relationships observed. Compounds featuring electron-withdrawing groups exhibited enhanced inhibitory potency, while compounds with electron-donating groups showed varied inhibitory activities. Further investigation revealed mixed inhibition mechanisms for selected compounds against AChE and dose-dependent inhibition of propidium iodide displacement from the AChE PAS. Moreover, the PAMPA assay indicated several compounds' promising blood-brain barrier permeation potential.

Finally, *in-vitro* assays assessing $A\beta_{1-42}$ aggregation inhibition highlighted the potential of compounds **72** and **77** to prevent AChE-induced aggregation, suggesting their therapeutic relevance in combating AD pathology.

Overall, the findings from this study support the potential of the synthesized MTDLs as promising candidates for further development for the treatment of AD. Further preclinical and clinical studies are warranted to validate these compounds' efficacy, safety, and therapeutic potential *in-vivo*.

Chapter 7

Summary and Conclusion

7 Summary and Conclusion

Alzheimer's disease (AD) is a multifactorial neurodegeneration that causes dementia and other neuropsychiatric symptoms such as apathy, depression, and shifts in sleep patterns. A significant degeneration of the hippocampus and cerebral cortex results in the loss of short and long-term memory. There are several hypotheses for the neurodegeneration process featuring β-amyloids, tau protein, cholinergic system etc. The exact cause of AD is still unknown, but these abnormalities are believed to play a central role in its progression. Several potential targets are being explored for therapeutic intervention:

- **Acetylcholinesterase (AChE) inhibitors:** These drugs work by increasing the levels of acetylcholine, a neurotransmitter important for memory and learning, in the brain. While they don't prevent disease progression, they may help manage symptoms like memory loss and improve cognitive function in the early stages.

- **BACE-1 inhibitors**: β-site amyloid precursor protein cleaving enzyme 1 (BACE-1) is an enzyme that plays a crucial role in the production of amyloid beta. By inhibiting BACE-1, these drugs aim to reduce the formation of amyloid beta plaques and slow down the disease process.

- **Amyloid beta aggregation inhibitors:** These drugs target the clumping of amyloid beta proteins, preventing them from forming plaques. This approach could potentially halt or even reverse the accumulation of these harmful deposits in the brain.

Progress is being made in the treatment of AD, the US Food and Drug Administration (FDA) has approved five cognitive-enhancing drugs, one agent for reduction of agitation in AD, and two disease-modifying therapies.

Conventional wet laboratory testing, validations, and synthetic procedures are costly and time-consuming for drug discovery. Advancements in artificial intelligence (AI) techniques have revolutionized their applications to drug discovery. Combined with accessible data resources, AI techniques are changing the landscape of drug discovery. In the past decades, a series of AI-based models have been developed for various steps of drug discovery. These models have been used as complements of conventional experiments and have accelerated the drug discovery process.

In the current work, the potential of AI is explored in various stages of drug discovery for AD.

Firstly, the research aimed to develop accurate machine learning models for classifying BACE-1 inhibitors. The BACE-1 is a transmembrane aspartyl-protease that cleaves amyloid precursor protein (APP) at the β-site. The sequential proteolytic cleavage of APP, first by β-secretase and then by γ-secretase complex, leads to the production and release of amyloid-β peptide, a pathological hallmark of AD. BACE-1 inhibitors are reported to possess considerable potential in decreasing the level of amyloid-β in brain and preventing the progression of AD. A classification study has been conducted on 3536 diverse BACE-1 inhibitors, obtained from Binding DB database, by extracting two types of descriptors, that is molecular property (Mordred) and fingerprints (Pubchem, MACCS and KRFP). Furthermore, based on the descriptors, various machine learning algorithms such as Naïve Bayesian (NB), nearest known neighbours (kNN), support vector machine (SVM), random forest (RF) and gradient-boosted algorithms (XGB) were applied to develop classification models. The performance of models was evaluated by using accuracy, precision, recall and F1 score of test set. The best NB, kNN, SVM, RF and XGB classifiers had F1 score of 0.74, 0.85, 0.86, 0.87 and 0.87, respectively. The diverse 3536 BACE-1 inhibitors were clustered into 11 subsets, and the structural features of each subset were evaluated. The

important fragments present in active and inactive compounds were also identified. The model developed in the study would serve as a valuable tool for the designing of BACE-1 inhibitors, and also in virtual screening of molecules to identify these.

Secondly, the research leveraged machine learning to screen an *in-house* database of previously synthesized sulphonamide containing compounds for potential BACE-1 inhibitors. In the present work, we have developed an ML model based on the sulfonamides dataset. The best ML model was built using the XGBoost algorithm on PubChem fingerprints. The model had an accuracy, precision, recall and F1 score of 0.89, 0.88, 0.99 and 0.93, respectively, on the validation set. The same model was used to screen the database of previously synthesized and reported in-house compounds. The screening resulted in the identification of two hits, i.e., compound **28** and compound **37**. Both the compounds were screened for their BACE-1 inhibitor activity. The IC_{50} value of compound 28 was found to be 0.431 ± 0.06 μM, and compound 37 showed an IC_{50} value of 0.272 ± 0.019 μM. The docking study revealed that compound 37 also showed interactions with the catalytic dyad of BACE-1, i.e., Asp32 and Asp228.

Next, the research investigated a novel approach for predicting Blood-Brain Barrier (BBB) permeability of molecules using Natural Language Processing (NLP) and Deep Learning. Blood-Brain-Barrier (BBB) permeability is one of the critical factors in the success and failure of CNS drug development. The most accurate method of measuring BBB permeability involves clinical experiments, which are labour-intensive and time-consuming. Thus, numerous efforts were made to use artificial intelligence (AI) to predict molecules′ BBB permeability. Most of the previous models are based on calculated descriptors and molecular fingerprints. In the present work, we have developed an NLP-based feature extraction technique in Deep-Learning models to predict BBB permeability. We have used the B3DB database and generated SELFIES to extract features from the

molecules. We have employed word level and N-gram tokenization to represent words into numeric vectors. The extracted features were fed into several Artificial Neural Network (ANN) and Bi-directional Long Short-Term Memory (LSTM) models. The model, ANN-10 built using ANN and 6-gram tokenization, performed best on the independent test set. The accuracy, precision, recall, F1, specificity and AUC of ROC scores were found to be 0.89, 0.91, 0.91, 0.91, 0.85 and 0.90. Thus, the developed model can be used for the early screening of CNS drugs.

Finally, a series of *N-benzylpiperidines* were designed and evaluated as multi-target directed ligand for the treatment of AD. The compounds were synthesized using suitable reactions and characterised using sophisticated analytical techniques. The compounds were subjected to molecular docking studies and Absorption, distribution, metabolism, elimination and toxicity (ADMET) prediction using AI models. The compounds showed good AChE and BACE-1 inhibitory potential. Based on *in-vitro* evaluations two compounds i.e. Compound **72** *(2-((1-benzylpiperidin-4-yl)amino)-N-(2-chlorophenyl)acetamide)* and Compound **77** *(2-((1-benzylpiperidin-4-yl)amino)-N-(4-methoxyphenyl)acetamide)* were selected for Aβ aggregation inhibition and *in-vivo* study. The compounds did not show any signs of toxicity upon oral administration. Both the compounds showed dose-dependent increase in % spontaneous alternations in scopolamine-induced amnesic rodents and demonstrated a substantial improvement of the cognitive and spatial memory impairment in $A\beta_{1-42}$ treated rats. Both *in-vitro* and *in-vivo* investigations evidenced that compound **72** and **77** were potential leads for the discovery of safe and effective agents for the treatment of AD.

Chapter 8
References

8 References

1. DeTure, M.A. and D.W. Dickson, *The neuropathological diagnosis of Alzheimer's disease.* Molecular Neurodegeneration, 2019. **14**(1): p. 32.
2. Dugger, B.N. and D.W. Dickson, *Pathology of Neurodegenerative Diseases.* 2017. **9**(7).
3. Slanzi, A., et al., *In vitro Models of Neurodegenerative Diseases.* 2020. **8**.
4. Manera, V., et al., *A Feasibility Study with Image-Based Rendered Virtual Reality in Patients with Mild Cognitive Impairment and Dementia.* PLOS ONE, 2016. **11**(3): p. e0151487.
5. Adnet, T., et al., *Pharmacotechnical Development of a Nasal Drug Delivery Composite Nanosystem Intended for Alzheimer's Disease Treatment.* 2020. **12**(3): p. 251.
6. Chintamaneni, M. and M. Bhaskar, *Biomarkers in Alzheimer's Disease: A Review.* ISRN Pharmacology, 2012. **2012**: p. 984786.
7. Bagheri, S., et al., *Probable Reasons for Neuron Copper Deficiency in the Brain of Patients with Alzheimer’s Disease: The Complex Role of Amyloid.* 2022. **10**(1): p. 6.
8. WuLi, W., et al., *Chinese herbs and acupuncture to improve cognitive function in Alzheimer's disease.* Tzu Chi Medical Journal, 2021. **33**(2).
9. van Gils, A.M., et al., *Assessing the Views of Professionals, Patients, and Care Partners Concerning the Use of Computer Tools in Memory Clinics: International Survey Study.* JMIR Form Res, 2021. **5**(12): p. e31053.
10. Burns, A., *Diagnosis and management of Alzheimer's disease.* Dialogues in Clinical Neuroscience, 2000. **2**(2): p. 129-138.
11. Association, A.s. Diagnosis of Alzheimer's Disease [cited 2023 05-08]; Available from: https://www.alz.org/alzheimers-dementia/diagnosis/medical_tests.
12. Saint-Aubert, L., et al., *Cortical florbetapir-PET amyloid load in prodromal Alzheimer's disease patients.* EJNMMI Res, 2013. **3**(1): p. 43.
13. Yeo, J.M., et al., *A systematic review and meta-analysis of 18F-labeled amyloid imaging in Alzheimer's disease.* 2015. **1**(1): p. 5-13.
14. Librizzi, D., et al. *Clinical Relevance of [18F]Florbetaben and [18F]FDG PET/CT Imaging on the Management of Patients with Dementia*. Molecules, 2021. **26**, DOI: 10.3390/molecules26051282.
15. Hampel, H., et al., *Core candidate neurochemical and imaging biomarkers of Alzheimer's disease∗∗This paper was presented in part by the 1st author at the 10th International Conference of Alzheimer's Disease and Related Disorders (ICAD), Madrid, Spain, July 2006, as an invited plenary lecture.* Alzheimer's & Dementia, 2008. **4**(1): p. 38-48.
16. Hippius, H. and G. Neundörfer, *The discovery of Alzheimer's disease.* Dialogues in Clinical Neuroscience, 2003. **5**(1): p. 101-108.
17. Cáceres, C., et al. *The Major Hypotheses of Alzheimer's Disease: Related Nanotechnology-Based Approaches for Its Diagnosis and Treatment*. Cells, 2023. **12**, DOI: 10.3390/cells12232669.
18. MacLeod, R., et al., *The role and therapeutic targeting of α-, β- and γ-secretase in Alzheimer's disease.* Future Sci OA, 2015. **1**(3): p. Fso11.
19. Roy, R.G., P.K. Mandal, and J.C. Maroon, *Oxidative Stress Occurs Prior to Amyloid Aβ Plaque Formation and Tau Phosphorylation in Alzheimer's Disease: Role of Glutathione and Metal Ions.* ACS Chemical Neuroscience, 2023. **14**(17): p. 2944-2954.

20. Sohrabi, M., et al., *IGF-1R Inhibitor Ameliorates Neuroinflammation in an Alzheimer's Disease Transgenic Mouse Model.* 2020. **14**.
21. Corraliza-Gomez, M., et al., *Insulin-degrading enzyme (IDE) as a modulator of microglial phenotypes in the context of Alzheimer's disease and brain aging.* Journal of Neuroinflammation, 2023. **20**(1): p. 233.
22. Ullah, R., et al., *Abnormal amyloid beta metabolism in systemic abnormalities and Alzheimer's pathology: Insights and therapeutic approaches from periphery.* Ageing Research Reviews, 2021. **71**: p. 101451.
23. Chistiakov, D.A. and A.A. Chistiakov, *α-Synuclein-carrying extracellular vesicles in Parkinson's disease: deadly transmitters.* Acta Neurologica Belgica, 2017. **117**(1): p. 43-51.
24. Satarker, S., et al. *Astrocytic Glutamatergic Transmission and Its Implications in Neurodegenerative Disorders*. Cells, 2022. **11**, DOI: 10.3390/cells11071139.
25. Armato, U., et al., *Calcium-sensing receptor antagonist (calcilytic) NPS 2143 specifically blocks the increased secretion of endogenous Aβ42 prompted by exogenous fibrillary or soluble Aβ25–35 in human cortical astrocytes and neurons—Therapeutic relevance to Alzheimer's disease.* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2013. **1832**(10): p. 1634-1652.
26. Haake, A., et al., *An update on the utility and safety of cholinesterase inhibitors for the treatment of Alzheimer's disease.* Expert Opinion on Drug Safety, 2020. **19**(2): p. 147-157.
27. Yiannopoulou, K.G. and S.G. Papageorgiou, *Current and Future Treatments in Alzheimer Disease: An Update.* Journal of Central Nervous System Disease, 2020. **12**: p. 1179573520907397.
28. van Dyck, C.H., et al., *Lecanemab in Early Alzheimer's Disease.* New England Journal of Medicine, 2022. **388**(1): p. 9-21.
29. Vaz, M., et al., *Role of Aducanumab in the Treatment of Alzheimer's Disease: Challenges and Opportunities.* Clin Interv Aging, 2022. **17**: p. 797-810.
30. Vora, L.K., et al., *Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design.* Pharmaceutics, 2023. **15**(7).
31. Arabi, A.A., *Artificial intelligence in drug design: algorithms, applications, challenges and ethics.* Future Drug Discovery, 2021. **3**(2): p. FDD59.
32. Volkamer, A., et al., *Machine learning for small molecule drug discovery in academia and industry.* Artificial Intelligence in the Life Sciences, 2023. **3**: p. 100056.
33. van Tilborg, D., A. Alenicheva, and F. Grisoni, *Exposing the Limitations of Molecular Machine Learning with Activity Cliffs.* Journal of Chemical Information and Modeling, 2022. **62**(23): p. 5938-5951.
34. Paul, D., et al., *Artificial intelligence in drug discovery and development.* Drug Discov Today, 2021. **26**(1): p. 80-93.
35. Vora, L.K., et al., *Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design.* 2023. **15**(7): p. 1916.
36. Askin, S., et al., *Artificial Intelligence Applied to clinical trials: opportunities and challenges.* Health and Technology, 2023. **13**(2): p. 203-213.
37. Kumar, D., et al., *Secretase inhibitors for the treatment of Alzheimer's disease: Long road ahead.* European Journal of Medicinal Chemistry, 2018. **148**: p. 436-452.
38. Coimbra, J.R.M., et al., *Highlights in BACE1 Inhibitors for Alzheimer's Disease Treatment.* 2018. **6**(178).

39. Munro, K.M., et al., *Functions of the Alzheimer's Disease Protease BACE1 at the Synapse in the Central Nervous System.* Journal of molecular neuroscience : MN, 2016. **60**(3): p. 305-315.
40. Yan, R. and R. Vassar, *Targeting the β secretase BACE1 for Alzheimer's disease therapy.* The Lancet Neurology, 2014. **13**(3): p. 319-329.
41. Cole, S.L. and R. Vassar, *The Alzheimer's disease β-secretase enzyme, BACE1.* Molecular Neurodegeneration, 2007. **2**(1): p. 22.
42. Luo, Y., et al., *Mice deficient in BACE1, the Alzheimer's beta-secretase, have normal phenotype and abolished beta-amyloid generation.* Nature neuroscience, 2001. **4**(3): p. 231-232.
43. Butini, S., et al., *A stereoselective approach to peptidomimetic BACE1 inhibitors.* European Journal of Medicinal Chemistry, 2013. **70**: p. 233-247.
44. Vassar, R., *The β-secretase, BACE.* Journal of Molecular Neuroscience, 2001. **17**(2): p. 157-170.
45. Mullard, A., *BACE failures lower AD expectations, again.* Nature Reviews Drug Discovery, 2018. **17**(6): p. 385-385.
46. Moussa-Pacha, N.M., et al., *BACE1 inhibitors: Current status and future directions in treating Alzheimer's disease.* Medicinal Research Reviews, 2020. **40**(1): p. 339-384.
47. Yuan, J., et al., *Structure-based design of β-site APP cleaving enzyme 1 (BACE1) inhibitors for the treatment of Alzheimer's disease.* 2013. **56**(11): p. 4156-4180.
48. Ponzoni, I., et al., *QSAR Classification Models for Predicting the Activity of Inhibitors of Beta-Secretase (BACE1) Associated with Alzheimer's Disease.* Scientific Reports, 2019. **9**(1): p. 9102.
49. Huang, D., et al., *Comprehensive 3D-QSAR and binding mode of BACE-1 inhibitors using R-group search and molecular docking.* Journal of Molecular Graphics and Modelling, 2013. **45**: p. 65-83.
50. Subramanian, G., et al., *Computational Modeling of β-Secretase 1 (BACE-1) Inhibitors Using Ligand Based Approaches.* Journal of Chemical Information and Modeling, 2016. **56**(10): p. 1936-1949.
51. Kumar, V., et al., *Exploring 2D-QSAR for prediction of beta-secretase 1 (BACE1) inhibitory activity against Alzheimer's disease.* SAR and QSAR in Environmental Research, 2020. **31**(2): p. 87-133.
52. Gilson, M.K., et al., *BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology.* Nucleic Acids Research, 2015. **44**(D1): p. D1045-D1053.
53. Kotu, V. and B. Deshpande, *Chapter 14 - Feature Selection*, in *Data Science (Second Edition)*, V. Kotu and B. Deshpande, Editors. 2019, Morgan Kaufmann. p. 467-490.
54. Cai, J., et al., *Predicting DPP-IV inhibitors with machine learning approaches.* Journal of Computer-Aided Molecular Design, 2017. **31**(4): p. 393-402.
55. Khanfar, M.A. and M.O. Taha, *Elaborate ligand-based modeling coupled with multiple linear regression and k nearest neighbor QSAR analyses unveiled new nanomolar mTOR inhibitors.* J Chem Inf Model, 2013. **53**(10): p. 2587-612.
56. Noble, W.S., *What is a support vector machine?* Nature Biotechnology, 2006. **24**(12): p. 1565-1567.
57. Gedeck, P., C. Kramer, and P. Ertl, *4 - Computational Analysis of Structure–Activity Relationships*, in *Progress in Medicinal Chemistry*, G. Lawton and D.R. Witty, Editors. 2010, Elsevier. p. 113-160.
58. Natekin, A. and A. Knoll, *Gradient boosting machines, a tutorial.* 2013. **7**(21).

59. Melagraki, G., et al., *In Silico Exploration for Identifying Structure–Activity Relationship of MEK Inhibition and Oral Bioavailability for Isothiazole Derivatives.* Chemical Biology & Drug Design, 2010. **76**(5): p. 397-406.
60. Hampel, H., et al., *The β-secretase BACE1 in Alzheimer's disease.* Biological psychiatry, 2021. **89**(8): p. 745-756.
61. Vassar, R., *BACE1 inhibitor drugs in clinical trials for Alzheimer's disease.* Alzheimer's research & therapy, 2014. **6**(9): p. 1-14.
62. Hung, Y.H., A.I. Bush, and R.A. Cherny, *Copper in the brain and Alzheimer's disease.* JBIC Journal of Biological Inorganic Chemistry, 2010. **15**(1): p. 61-76.
63. Turner, R.T., et al., *Subsite specificity of memapsin 2 (β-secretase): implications for inhibitor design.* Biochemistry, 2001. **40**(34): p. 10001-10006.
64. Ghosh, A.K. and H.L. Osswald, *BACE1 (β-secretase) inhibitors for the treatment of Alzheimer's disease.* Chemical Society Reviews, 2014. **43**(19): p. 6765-6813.
65. Kennedy, M.E., et al., *The BACE1 inhibitor verubecestat (MK-8931) reduces CNS β-amyloid in animal models and in Alzheimer's disease patients.* Science translational medicine, 2016. **8**(363): p. 363ra150-363ra150.
66. Sastre, A.A., et al., *Effect of the treatment of type 2 diabetes mellitus on the development of cognitive impairment and dementia.* Cochrane Database of Systematic Reviews, 2017(6).
67. Bertini, S., et al., *Sulfonamido-derivatives of unsubstituted carbazoles as BACE1 inhibitors.* Bioorganic & Medicinal Chemistry Letters, 2017. **27**(21): p. 4812-4816.
68. Voytyuk, I., B. De Strooper, and L. Chávez-Gutiérrez, *Modulation of γ- and β-Secretases as Early Prevention Against Alzheimer's Disease.* Biological Psychiatry, 2018. **83**(4): p. 320-327.
69. Carracedo-Reboredo, P., et al., *A review on machine learning approaches and trends in drug discovery.* Computational and structural biotechnology journal, 2021. **19**: p. 4538-4558.
70. Gupta, R., et al., *Artificial intelligence to deep learning: machine intelligence approach for drug discovery.* Molecular Diversity, 2021. **25**(3): p. 1315-1360.
71. Breiman, L., *Random forests.* Machine learning, 2001. **45**(1): p. 5-32.
72. Sagi, O. and L. Rokach, *Approximating XGBoost with an interpretable decision tree.* Information Sciences, 2021. **572**: p. 522-542.
73. Babajide Mustapha, I. and F. Saeed, *Bioactive molecule prediction using extreme gradient boosting.* Molecules, 2016. **21**(8): p. 983.
74. Du, Z., et al., *Inference of gene regulatory networks based on the Light Gradient Boosting Machine.* Computational Biology and Chemistry, 2022. **101**: p. 107769.
75. Zhao, Q., et al., *Predicting complexation performance between cyclodextrins and guest molecules by integrated machine learning and molecular modeling techniques.* Acta Pharmaceutica Sinica B, 2019. **9**(6): p. 1241-1252.
76. Gilson, M.K., et al., *BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology.* Nucleic Acids Research, 2016. **44**(D1): p. D1045-D1053.
77. Berthold, M.R., et al., *KNIME-the Konstanz information miner: version 2.0 and beyond.* 2009. **11**(1): p. 26-31.
78. Swetha, R., et al., *Multifunctional hybrid sulfonamides as novel therapeutic agents for Alzheimer's disease.* Future Medicinal Chemistry, 2019. **11**(24): p. 3161-3178.
79. Ganeshpurkar, A., D. Kumar, and S.K. Singh, *Design, synthesis and collagenase inhibitory activity of some novel phenylglycine derivatives as metalloproteinase*

inhibitors. International journal of biological macromolecules, 2018. **107**: p. 1491-1500.
80. Kumar, D., et al., *Development of Piperazinediones as dual inhibitor for treatment of Alzheimer's disease.* European Journal of Medicinal Chemistry, 2018. **150**: p. 87-101.
81. Sander, T., et al., *DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis.* Journal of Chemical Information and Modeling, 2015. **55**(2): p. 460-473.
82. Ganeshpurkar, A., et al., *Identification of sulfonamide based butyrylcholinesterase inhibitors through scaffold hopping approach.* International Journal of Biological Macromolecules, 2022. **203**: p. 195-211.
83. Pardridge, W.M., *The blood-brain barrier: Bottleneck in brain drug development.* NeuroRX, 2005. **2**(1): p. 3-14.
84. Chen, Y., G. Dalwadi, and H. Benson, *Drug delivery across the blood-brain barrier.* Current drug delivery, 2004. **1**(4): p. 361-376.
85. Tosi, G., et al., *Potential Use of Polymeric Nanoparticles for Drug Delivery Across the Blood-Brain Barrier.* Current Medicinal Chemistry, 2013. **20**(17): p. 2212-2225.
86. Bellettato, C.M. and M. Scarpa, *Possible strategies to cross the blood–brain barrier.* Italian journal of pediatrics, 2018. **44**(2): p. 127-133.
87. Sanchez-Covarrubias, L., et al., *Transporters at CNS barrier sites: obstacles or opportunities for drug delivery?* Current pharmaceutical design, 2014. **20**(10): p. 1422-1449.
88. Preston, J.E., N.J. Abbott, and D.J. Begley, *Transcytosis of macromolecules at the blood–brain barrier.* Advances in pharmacology, 2014. **71**: p. 147-163.
89. Muller, W.A., *PECAM: Regulating the start of diapedesis.* Adhesion molecules: Function and inhibition, 2007: p. 201-220.
90. Kumar, R., et al., *DeePred-BBB: A Blood Brain Barrier Permeability Prediction Model With Improved Accuracy.* 2022. **16**.
91. Alsenan, S., I. Al-Turaiki, and A. Hafez, *A Recurrent Neural Network model to predict blood–brain barrier permeability.* Computational Biology and Chemistry, 2020. **89**: p. 107377.
92. Shaker, B., et al., *LightBBB: computational prediction model of blood–brain-barrier penetration based on LightGBM.* Bioinformatics, 2021. **37**(8): p. 1135-1139.
93. Meng, F., et al., *A curated diverse molecular database of blood-brain barrier permeability with chemical descriptors.* Scientific Data, 2021. **8**(1): p. 289.
94. Cherian Parakkal, S., R. Datta, and D. Das, *DeepBBBP: High Accuracy Blood-brain-barrier Permeability Prediction with a Mixed Deep Learning Model.* Molecular Informatics, 2022. **41**(10): p. 2100315.
95. Krenn, M., et al., *Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation.* 2020. **1**(4): p. 045024.
96. Erra, U., et al., *Approximate TF–IDF based on topic extraction from massive message stream using the GPU.* Information Sciences, 2015. **292**: p. 143-161.
97. Malekian, A. and N. Chitsaz, *Chapter 4 - Concepts, procedures, and applications of artificial neural network models in streamflow forecasting*, in *Advances in Streamflow Forecasting*, P. Sharma and D. Machiwal, Editors. 2021, Elsevier. p. 115-147.
98. Ghose, A.K., G.R. Ott, and R.L. Hudkins, *Technically Extended MultiParameter Optimization (TEMPO): An Advanced Robust Scoring Scheme To Calculate*

Central Nervous System Druggability and Monitor Lead Optimization. ACS Chemical Neuroscience, 2017. **8**(1): p. 147-154.
99. Gupta, M., et al., *The Blood–Brain Barrier (BBB) Score.* Journal of Medicinal Chemistry, 2019. **62**(21): p. 9824-9836.
100. Shaker, B., et al., *LightBBB: computational prediction model of blood–brain-barrier penetration based on LightGBM.* Bioinformatics, 2020. **37**(8): p. 1135-1139.
101. Livingston, G., et al., *Dementia prevention, intervention, and care: 2020 report of the Lancet Commission.* The Lancet, 2020. **396**(10248): p. 413-446.
102. Caliandro, R., et al., *Kinetic and structural studies on the interactions of Torpedo californica acetylcholinesterase with two donepezil-like rigid analogues.* 2018. **33**(1): p. 794-803.
103. Polishchuk, P., *CReM: chemically reasonable mutations framework for structure generation.* Journal of Cheminformatics, 2020. **12**(1): p. 28.
104. Daina, A., O. Michielin, and V. Zoete, *SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.* Scientific Reports, 2017. **7**(1): p. 42717.
105. Ellman, G.L., et al., *A new and rapid colorimetric determination of acetylcholinesterase activity.* Biochemical Pharmacology, 1961. **7**(2): p. 88-95.
106. Singh, R., et al., *Machine learning-based screening of in-house database to identify BACE-1 inhibitors.* Chemical Papers, 2023.
107. Di, L., et al., *High throughput artificial membrane permeability assay for blood–brain barrier.* European Journal of Medicinal Chemistry, 2003. **38**(3): p. 223-232.
108. Peauger, L., et al., *Donepezil-Based Central Acetylcholinesterase Inhibitors by Means of a "Bio-Oxidizable" Prodrug Strategy: Design, Synthesis, and in Vitro Biological Evaluation.* Journal of Medicinal Chemistry, 2017. **60**(13): p. 5909-5926.
109. Bhanukiran, K., et al., *Discovery of multi-target directed 3-OH pyrrolidine derivatives through a semisynthetic approach from alkaloid vasicine for the treatment of Alzheimer's disease.* European Journal of Medicinal Chemistry, 2023. **249**: p. 115145.
110. Sharma, P., et al., *Design and development of multitarget-directed N-Benzylpiperidine analogs as potential candidates for the treatment of Alzheimer's disease.* European Journal of Medicinal Chemistry, 2019. **167**: p. 510-524.
111. Zha, X., et al., *Novel Tacrine–Benzofuran Hybrids as Potent Multitarget-Directed Ligands for the Treatment of Alzheimer's Disease: Design, Synthesis, Biological Evaluation, and X-ray Crystallography.* Journal of Medicinal Chemistry, 2016. **59**(1): p. 114-131.
112. OECD Test %J OECD guidelines for the testing of chemicals, s., *423: Acute oral toxicity-acute toxic class method.* 2002. **4**: p. 14.
113. Majumdar, S. and S. Krishnamurthy, *In vivo toxicological evaluation of barium-doped bioactive glass in rats.* Ceramics International, 2022. **48**(22): p. 33288-33305.
114. Ngoupaye, G.T., et al., *Gladiolus dalenii lyophilisate reverses scopolamine-induced amnesia and reduces oxidative stress in rat brain.* Biomedicine & Pharmacotherapy, 2017. **91**: p. 350-357.
115. Ganeshpurkar, A., et al., *Effect of sulfonamide derivatives of phenylglycine on scopolamine-induced amnesia in rats.* Ibrain, 2023. **9**(1): p. 13-31.

116. Hritcu, L., et al., *Nicotine versus 6-hydroxy-l-nicotine against chlorisondamine induced memory impairment and oxidative stress in the rat hippocampus.* Biomedicine & Pharmacotherapy, 2017. **86**: p. 102-108.
117. Gabriel Olaiya, O., et al., *Ameliorative effects of Moringa on cuprizone-induced memory decline in rat model ofmultiple sclerosis.* Anat Cell Biol, 2018. **51**(2): p. 119-127.
118. Manzoor, S., et al., *Discovery of Quinolinone Hybrids as Dual Inhibitors of Acetylcholinesterase and Aβ Aggregation for Alzheimer's Disease Therapy.* ACS Chemical Neuroscience, 2023.
119. Sharma, P., et al., *Novel Molecular Hybrids of N-Benzylpiperidine and 1,3,4-Oxadiazole as Multitargeted Therapeutics to Treat Alzheimer's Disease.* ACS Chemical Neuroscience, 2019. **10**(10): p. 4361-4384.
120. Chen, Y., et al., *Discovery of a novel acetylcholinesterase inhibitor by structure-based virtual screening techniques.* Bioorganic & Medicinal Chemistry Letters, 2012. **22**(9): p. 3181-3187.
121. Hampel, H., et al., *The Amyloid-β Pathway in Alzheimer's Disease.* Molecular Psychiatry, 2021. **26**(10): p. 5481-5503.
122. Carvajal, F. and N. Inestrosa, *Interactions of AChE with Aβ Aggregates in Alzheimer's Brain: Therapeutic Relevance of IDN 5706.* 2011. **4**.
123. Kasza, Á., et al. *Studies for Improving a Rat Model of Alzheimer's Disease: Icv Administration of Well-Characterized β-Amyloid 1-42 Oligomers Induce Dysfunction in Spatial Memory*. Molecules, 2017. **22**, DOI: 10.3390/molecules22112007.
124. Rather, M.A., et al., *Inflammation and Alzheimer's Disease: Mechanisms and Therapeutic Implications by Natural Products.* Mediators of Inflammation, 2021. **2021**: p. 9982954.

Appendices

Tables

Table-S1 Number of features calculated and used for model building

S.No.	Name of features	No. of features calculated	No. of features used for building model
1	Modred	1613	22
2	MACCS	166	151
3	PUBCHEM	881	610
4	KRFP	4861	1753

Table-S2 List of MACCS, PUBCHEM and KRFP fingerprints used in building models

Fingerprints	Bits
MACCS	MACCSFP8, MACCSFP11, MACCSFP13, MACCSFP14, MACCSFP16, MACCSFP17, MACCSFP18, MACCSFP19, MACCSFP20, MACCSFP21, MACCSFP22, MACCSFP23, MACCSFP24, MACCSFP25, MACCSFP26, MACCSFP27, MACCSFP28, MACCSFP29, MACCSFP30, MACCSFP32, MACCSFP33, MACCSFP34, MACCSFP36, MACCSFP37, MACCSFP38, MACCSFP39, MACCSFP40, MACCSFP41, MACCSFP42, MACCSFP43, MACCSFP45, MACCSFP46, MACCSFP47, MACCSFP48, MACCSFP49, MACCSFP50, MACCSFP51, MACCSFP52, MACCSFP53, MACCSFP54, MACCSFP55, MACCSFP56, MACCSFP57, MACCSFP58, MACCSFP59, MACCSFP60, MACCSFP61, MACCSFP62, MACCSFP63, MACCSFP64, MACCSFP65, MACCSFP66, MACCSFP67, MACCSFP68, MACCSFP69, MACCSFP70, MACCSFP71, MACCSFP72, MACCSFP73, MACCSFP74, MACCSFP75, MACCSFP76, MACCSFP77, MACCSFP78, MACCSFP79, MACCSFP80, MACCSFP81, MACCSFP82, MACCSFP83, MACCSFP84, MACCSFP85, MACCSFP86, MACCSFP87, MACCSFP88, MACCSFP89, MACCSFP90, MACCSFP91, MACCSFP92, MACCSFP93, MACCSFP94, MACCSFP95, MACCSFP96, MACCSFP97, MACCSFP98, MACCSFP99, MACCSFP100, MACCSFP101, MACCSFP102, MACCSFP103, MACCSFP104, MACCSFP105, MACCSFP106, MACCSFP107, MACCSFP108, MACCSFP109, MACCSFP110, MACCSFP111, MACCSFP112, MACCSFP113, MACCSFP114, MACCSFP115, MACCSFP116, MACCSFP117, MACCSFP118, MACCSFP119, MACCSFP120, MACCSFP121, MACCSFP122, MACCSFP123, MACCSFP124, MACCSFP125, MACCSFP126, MACCSFP127, MACCSFP128, MACCSFP129, MACCSFP130, MACCSFP131,

	MACCSFP132, MACCSFP133, MACCSFP134, MACCSFP135, MACCSFP136, MACCSFP137, MACCSFP138, MACCSFP139, MACCSFP140, MACCSFP141, MACCSFP142, MACCSFP143, MACCSFP144, MACCSFP145, MACCSFP146, MACCSFP147, MACCSFP148, MACCSFP149, MACCSFP150, MACCSFP151, MACCSFP152, MACCSFP153, MACCSFP154, MACCSFP155, MACCSFP156, MACCSFP157, MACCSFP158, MACCSFP159, MACCSFP160, MACCSFP161, MACCSFP162, MACCSFP163, MACCSFP164, MACCSFP165,
PUBCHEM	PubchemFP1, PubchemFP2, PubchemFP3, PubchemFP6, PubchemFP10, PubchemFP11, PubchemFP12, PubchemFP13, PubchemFP14, PubchemFP15, PubchemFP16, PubchemFP17, PubchemFP18, PubchemFP19, PubchemFP20, PubchemFP21, PubchemFP22, PubchemFP23, PubchemFP24, PubchemFP25, PubchemFP28, PubchemFP30, PubchemFP33, PubchemFP34, PubchemFP37, PubchemFP38, PubchemFP43, PubchemFP44, PubchemFP46, PubchemFP47, PubchemFP115, PubchemFP116, PubchemFP118, PubchemFP122, PubchemFP123, PubchemFP129, PubchemFP130, PubchemFP131, PubchemFP132, PubchemFP143, PubchemFP144, PubchemFP145, PubchemFP146, PubchemFP147, PubchemFP148, PubchemFP149, PubchemFP150, PubchemFP152, PubchemFP153, PubchemFP155, PubchemFP156, PubchemFP157, PubchemFP162, PubchemFP163, PubchemFP164, PubchemFP178, PubchemFP179, PubchemFP180, PubchemFP181, PubchemFP182, PubchemFP183, PubchemFP184, PubchemFP185, PubchemFP186, PubchemFP187, PubchemFP188, PubchemFP189, PubchemFP190, PubchemFP191, PubchemFP192, PubchemFP193, PubchemFP195, PubchemFP196, PubchemFP197, PubchemFP198, PubchemFP199, PubchemFP200, PubchemFP203, PubchemFP204, PubchemFP205, PubchemFP206, PubchemFP210, PubchemFP212, PubchemFP213, PubchemFP214, PubchemFP215, PubchemFP216, PubchemFP217, PubchemFP218, PubchemFP219, PubchemFP227, PubchemFP229, PubchemFP230, PubchemFP231, PubchemFP232, PubchemFP233, PubchemFP234, PubchemFP238, PubchemFP274, PubchemFP276, PubchemFP285, PubchemFP286, PubchemFP287, PubchemFP291, PubchemFP292, PubchemFP293, PubchemFP294, PubchemFP297, PubchemFP298, PubchemFP299, PubchemFP300, PubchemFP301, PubchemFP305, PubchemFP308, PubchemFP313, PubchemFP314, PubchemFP327, PubchemFP328, PubchemFP330, PubchemFP331, PubchemFP332, PubchemFP333, PubchemFP334, PubchemFP335, PubchemFP336, PubchemFP337, PubchemFP338, PubchemFP339, PubchemFP340, PubchemFP341, PubchemFP342, PubchemFP343, PubchemFP345, PubchemFP346, PubchemFP347, PubchemFP348, PubchemFP349, PubchemFP350, PubchemFP351, PubchemFP352, PubchemFP353,

	PubchemFP354,	PubchemFP355,	PubchemFP356,	PubchemFP357,
	PubchemFP358,	PubchemFP359,	PubchemFP361,	PubchemFP362,
	PubchemFP363,	PubchemFP364,	PubchemFP365,	PubchemFP366,
	PubchemFP367,	PubchemFP368,	PubchemFP369,	PubchemFP370,
	PubchemFP371,	PubchemFP372,	PubchemFP373,	PubchemFP374,
	PubchemFP375,	PubchemFP376,	PubchemFP377,	PubchemFP378,
	PubchemFP379,	PubchemFP380,	PubchemFP381,	PubchemFP382,
	PubchemFP383,	PubchemFP384,	PubchemFP385,	PubchemFP386,
	PubchemFP387,	PubchemFP388,	PubchemFP389,	PubchemFP390,
	PubchemFP391,	PubchemFP392,	PubchemFP393,	PubchemFP394,
	PubchemFP395,	PubchemFP396,	PubchemFP397,	PubchemFP398,
	PubchemFP399,	PubchemFP400,	PubchemFP401,	PubchemFP403,
	PubchemFP404,	PubchemFP405,	PubchemFP406,	PubchemFP408,
	PubchemFP409,	PubchemFP410,	PubchemFP411,	PubchemFP412,
	PubchemFP414,	PubchemFP415,	PubchemFP416,	PubchemFP417,
	PubchemFP418,	PubchemFP419,	PubchemFP420,	PubchemFP421,
	PubchemFP422,	PubchemFP423,	PubchemFP425,	PubchemFP427,
	PubchemFP428,	PubchemFP429,	PubchemFP430,	PubchemFP431,
	PubchemFP432,	PubchemFP434,	PubchemFP435,	PubchemFP436,
	PubchemFP437,	PubchemFP438,	PubchemFP439,	PubchemFP440,
	PubchemFP441,	PubchemFP442,	PubchemFP443,	PubchemFP445,
	PubchemFP446,	PubchemFP447,	PubchemFP448,	PubchemFP449,
	PubchemFP450,	PubchemFP451,	PubchemFP452,	PubchemFP453,
	PubchemFP454,	PubchemFP455,	PubchemFP456,	PubchemFP457,
	PubchemFP458,	PubchemFP459,	PubchemFP460,	PubchemFP461,
	PubchemFP462,	PubchemFP464,	PubchemFP465,	PubchemFP466,
	PubchemFP467,	PubchemFP468,	PubchemFP470,	PubchemFP471,
	PubchemFP472,	PubchemFP473,	PubchemFP474,	PubchemFP475,
	PubchemFP476,	PubchemFP477,	PubchemFP479,	PubchemFP480,
	PubchemFP482,	PubchemFP483,	PubchemFP484,	PubchemFP485,
	PubchemFP486,	PubchemFP487,	PubchemFP488,	PubchemFP489,
	PubchemFP490,	PubchemFP491,	PubchemFP492,	PubchemFP493,
	PubchemFP494,	PubchemFP495,	PubchemFP496,	PubchemFP497,
	PubchemFP498,	PubchemFP499,	PubchemFP500,	PubchemFP501,
	PubchemFP502,	PubchemFP503,	PubchemFP504,	PubchemFP505,
	PubchemFP506,	PubchemFP507,	PubchemFP508,	PubchemFP509,
	PubchemFP510,	PubchemFP511,	PubchemFP513,	PubchemFP514,
	PubchemFP515,	PubchemFP516,	PubchemFP517,	PubchemFP518,
	PubchemFP519,	PubchemFP520,	PubchemFP521,	PubchemFP522,
	PubchemFP523,	PubchemFP524,	PubchemFP526,	PubchemFP527,
	PubchemFP528,	PubchemFP529,	PubchemFP530,	PubchemFP531,
	PubchemFP532,	PubchemFP533,	PubchemFP534,	PubchemFP535,
	PubchemFP536,	PubchemFP537,	PubchemFP538,	PubchemFP539,
	PubchemFP540,	PubchemFP541,	PubchemFP542,	PubchemFP543,

	PubchemFP544,	PubchemFP545,	PubchemFP546,	PubchemFP547,
	PubchemFP548,	PubchemFP549,	PubchemFP550,	PubchemFP551,
	PubchemFP552,	PubchemFP553,	PubchemFP554,	PubchemFP555,
	PubchemFP556,	PubchemFP557,	PubchemFP558,	PubchemFP559,
	PubchemFP560,	PubchemFP561,	PubchemFP562,	PubchemFP563,
	PubchemFP564,	PubchemFP565,	PubchemFP566,	PubchemFP567,
	PubchemFP568,	PubchemFP569,	PubchemFP570,	PubchemFP571,
	PubchemFP572,	PubchemFP573,	PubchemFP574,	PubchemFP575,
	PubchemFP576,	PubchemFP577,	PubchemFP578,	PubchemFP579,
	PubchemFP580,	PubchemFP581,	PubchemFP582,	PubchemFP583,
	PubchemFP584,	PubchemFP585,	PubchemFP586,	PubchemFP587,
	PubchemFP588,	PubchemFP589,	PubchemFP590,	PubchemFP591,
	PubchemFP592,	PubchemFP593,	PubchemFP594,	PubchemFP595,
	PubchemFP596,	PubchemFP597,	PubchemFP598,	PubchemFP599,
	PubchemFP600,	PubchemFP601,	PubchemFP602,	PubchemFP603,
	PubchemFP604,	PubchemFP605,	PubchemFP606,	PubchemFP607,
	PubchemFP608,	PubchemFP609,	PubchemFP610,	PubchemFP611,
	PubchemFP612,	PubchemFP613,	PubchemFP614,	PubchemFP615,
	PubchemFP616,	PubchemFP617,	PubchemFP618,	PubchemFP619,
	PubchemFP620,	PubchemFP621,	PubchemFP622,	PubchemFP623,
	PubchemFP624,	PubchemFP625,	PubchemFP626,	PubchemFP628,
	PubchemFP629,	PubchemFP630,	PubchemFP631,	PubchemFP632,
	PubchemFP633,	PubchemFP634,	PubchemFP635,	PubchemFP636,
	PubchemFP637,	PubchemFP638,	PubchemFP639,	PubchemFP640,
	PubchemFP641,	PubchemFP642,	PubchemFP643,	PubchemFP644,
	PubchemFP645,	PubchemFP646,	PubchemFP647,	PubchemFP648,
	PubchemFP650,	PubchemFP651,	PubchemFP652,	PubchemFP653,
	PubchemFP654,	PubchemFP655,	PubchemFP656,	PubchemFP657,
	PubchemFP658,	PubchemFP659,	PubchemFP660,	PubchemFP661,
	PubchemFP662,	PubchemFP663,	PubchemFP664,	PubchemFP665,
	PubchemFP666,	PubchemFP667,	PubchemFP668,	PubchemFP669,
	PubchemFP670,	PubchemFP671,	PubchemFP672,	PubchemFP673,
	PubchemFP674,	PubchemFP675,	PubchemFP676,	PubchemFP677,
	PubchemFP678,	PubchemFP679,	PubchemFP680,	PubchemFP681,
	PubchemFP682,	PubchemFP683,	PubchemFP684,	PubchemFP685,
	PubchemFP686,	PubchemFP687,	PubchemFP688,	PubchemFP689,
	PubchemFP690,	PubchemFP691,	PubchemFP692,	PubchemFP693,
	PubchemFP694,	PubchemFP695,	PubchemFP696,	PubchemFP697,
	PubchemFP698,	PubchemFP699,	PubchemFP700,	PubchemFP701,
	PubchemFP702,	PubchemFP703,	PubchemFP704,	PubchemFP705,
	PubchemFP706,	PubchemFP707,	PubchemFP708,	PubchemFP709,
	PubchemFP710,	PubchemFP711,	PubchemFP712,	PubchemFP713,
	PubchemFP714,	PubchemFP715,	PubchemFP716,	PubchemFP717,
	PubchemFP718,	PubchemFP719,	PubchemFP720,	PubchemFP721,

	PubchemFP722, PubchemFP723, PubchemFP725, PubchemFP726, PubchemFP727, PubchemFP728, PubchemFP729, PubchemFP730, PubchemFP731, PubchemFP732, PubchemFP733, PubchemFP734, PubchemFP735, PubchemFP736, PubchemFP737, PubchemFP738, PubchemFP739, PubchemFP740, PubchemFP741, PubchemFP742, PubchemFP743, PubchemFP744, PubchemFP746, PubchemFP747, PubchemFP748, PubchemFP749, PubchemFP750, PubchemFP751, PubchemFP752, PubchemFP753, PubchemFP754, PubchemFP755, PubchemFP756, PubchemFP757, PubchemFP758, PubchemFP759, PubchemFP760, PubchemFP761, PubchemFP762, PubchemFP763, PubchemFP764, PubchemFP765, PubchemFP766, PubchemFP767, PubchemFP768, PubchemFP769, PubchemFP770, PubchemFP771, PubchemFP772, PubchemFP773, PubchemFP776, PubchemFP777, PubchemFP778, PubchemFP779, PubchemFP780, PubchemFP781, PubchemFP782, PubchemFP783, PubchemFP784, PubchemFP785, PubchemFP786, PubchemFP788, PubchemFP789, PubchemFP790, PubchemFP791, PubchemFP792, PubchemFP793, PubchemFP794, PubchemFP795, PubchemFP796, PubchemFP797, PubchemFP798, PubchemFP799, PubchemFP800, PubchemFP801, PubchemFP802, PubchemFP803, PubchemFP804, PubchemFP805, PubchemFP806, PubchemFP807, PubchemFP809, PubchemFP810, PubchemFP811, PubchemFP812, PubchemFP813, PubchemFP814, PubchemFP815, PubchemFP816, PubchemFP817, PubchemFP818, PubchemFP819, PubchemFP820, PubchemFP821, PubchemFP822, PubchemFP823, PubchemFP824, PubchemFP825, PubchemFP826, PubchemFP827, PubchemFP828, PubchemFP829, PubchemFP830, PubchemFP831, PubchemFP832, PubchemFP833, PubchemFP834, PubchemFP835, PubchemFP836, PubchemFP839, PubchemFP840, PubchemFP842, PubchemFP860, PubchemFP861, PubchemFP863, PubchemFP868, PubchemFP875,
KRFP	KRFP1, KRFP2, KRFP7, KRFP8, KRFP10, KRFP13, KRFP14, KRFP16, KRFP17, KRFP18, KRFP19, KRFP20, KRFP21, KRFP25, KRFP33, KRFP34, KRFP36, KRFP41, KRFP45, KRFP47, KRFP52, KRFP54, KRFP56, KRFP57, KRFP67, KRFP69, KRFP72, KRFP73, KRFP74, KRFP76, KRFP91, KRFP92, KRFP93, KRFP98, KRFP99, KRFP101, KRFP102, KRFP104, KRFP113, KRFP114, KRFP115, KRFP118, KRFP119, KRFP126, KRFP133, KRFP136, KRFP138, KRFP161, KRFP168, KRFP189, KRFP195, KRFP201, KRFP202, KRFP207, KRFP208, KRFP212, KRFP214, KRFP215, KRFP216, KRFP218, KRFP223, KRFP224, KRFP225, KRFP228, KRFP232, KRFP287, KRFP296, KRFP297, KRFP298, KRFP299, KRFP300, KRFP302, KRFP303, KRFP304, KRFP305, KRFP306, KRFP307, KRFP311, KRFP312, KRFP314, KRFP315, KRFP317, KRFP326, KRFP327, KRFP328, KRFP331, KRFP336, KRFP337, KRFP338, KRFP339,

	KRFP341, KRFP342, KRFP343, KRFP344, KRFP346, KRFP347, KRFP348, KRFP349, KRFP350, KRFP351, KRFP352, KRFP354, KRFP355, KRFP357, KRFP358, KRFP359, KRFP360, KRFP361, KRFP362, KRFP363, KRFP364, KRFP366, KRFP367, KRFP368, KRFP369, KRFP370, KRFP371, KRFP375, KRFP376, KRFP381, KRFP382, KRFP383, KRFP387, KRFP388, KRFP390, KRFP391, KRFP392, KRFP394, KRFP396, KRFP397, KRFP398, KRFP399, KRFP401, KRFP403, KRFP408, KRFP409, KRFP410, KRFP412, KRFP413, KRFP414, KRFP418, KRFP425, KRFP433, KRFP434, KRFP435, KRFP437, KRFP438, KRFP442, KRFP444, KRFP447, KRFP449, KRFP453, KRFP456, KRFP457, KRFP458, KRFP459, KRFP466, KRFP467, KRFP468, KRFP469, KRFP470, KRFP476, KRFP477, KRFP480, KRFP481, KRFP483, KRFP486, KRFP487, KRFP488, KRFP492, KRFP493, KRFP494, KRFP496, KRFP497, KRFP504, KRFP506, KRFP507, KRFP508, KRFP517, KRFP520, KRFP547, KRFP548, KRFP550, KRFP551, KRFP554, KRFP555, KRFP557, KRFP558, KRFP560, KRFP561, KRFP562, KRFP565, KRFP566, KRFP568, KRFP569, KRFP570, KRFP571, KRFP573, KRFP574, KRFP576, KRFP577, KRFP578, KRFP579, KRFP580, KRFP581, KRFP582, KRFP583, KRFP584, KRFP585, KRFP586, KRFP588, KRFP589, KRFP590, KRFP591, KRFP592, KRFP593, KRFP595, KRFP596, KRFP598, KRFP599, KRFP604, KRFP605, KRFP606, KRFP607, KRFP608, KRFP610, KRFP612, KRFP613, KRFP620, KRFP621, KRFP622, KRFP623, KRFP629, KRFP630, KRFP631, KRFP638, KRFP639, KRFP640, KRFP644, KRFP645, KRFP647, KRFP648, KRFP651, KRFP652, KRFP653, KRFP654, KRFP658, KRFP659, KRFP663, KRFP664, KRFP665, KRFP669, KRFP672, KRFP677, KRFP678, KRFP679, KRFP680, KRFP681, KRFP682, KRFP683, KRFP689, KRFP690, KRFP691, KRFP692, KRFP696, KRFP697, KRFP698, KRFP699, KRFP700, KRFP703, KRFP706, KRFP707, KRFP708, KRFP748, KRFP752, KRFP753, KRFP754, KRFP758, KRFP759, KRFP762, KRFP763, KRFP768, KRFP770, KRFP771, KRFP772, KRFP773, KRFP774, KRFP775, KRFP776, KRFP778, KRFP779, KRFP780, KRFP782, KRFP786, KRFP788, KRFP790, KRFP791, KRFP797, KRFP802, KRFP804, KRFP805, KRFP806, KRFP810, KRFP811, KRFP813, KRFP820, KRFP821, KRFP822, KRFP823, KRFP826, KRFP829, KRFP833, KRFP837, KRFP839, KRFP840, KRFP841, KRFP842, KRFP848, KRFP852, KRFP862, KRFP865, KRFP866, KRFP874, KRFP885, KRFP886, KRFP887, KRFP890, KRFP891, KRFP892, KRFP900, KRFP908, KRFP910, KRFP911, KRFP912, KRFP924, KRFP936, KRFP966, KRFP972, KRFP973, KRFP983, KRFP986, KRFP987, KRFP989, KRFP992, KRFP993, KRFP994, KRFP995, KRFP997, KRFP998, KRFP1000, KRFP1001, KRFP1002, KRFP1003, KRFP1007,

	KRFP1008, KRFP1009, KRFP1013, KRFP1014, KRFP1021, KRFP1024, KRFP1030, KRFP1031, KRFP1035, KRFP1041, KRFP1042, KRFP1044, KRFP1046, KRFP1049, KRFP1050, KRFP1051, KRFP1055, KRFP1073, KRFP1074, KRFP1080, KRFP1084, KRFP1097, KRFP1098, KRFP1104, KRFP1105, KRFP1106, KRFP1107, KRFP1109, KRFP1110, KRFP1111, KRFP1115, KRFP1116, KRFP1120, KRFP1124, KRFP1125, KRFP1126, KRFP1139, KRFP1141, KRFP1144, KRFP1145, KRFP1146, KRFP1147, KRFP1148, KRFP1150, KRFP1151, KRFP1152, KRFP1153, KRFP1154, KRFP1155, KRFP1156, KRFP1157, KRFP1158, KRFP1159, KRFP1160, KRFP1161, KRFP1173, KRFP1180, KRFP1184, KRFP1193, KRFP1198, KRFP1204, KRFP1212, KRFP1218, KRFP1219, KRFP1222, KRFP1223, KRFP1225, KRFP1236, KRFP1240, KRFP1245, KRFP1250, KRFP1251, KRFP1253, KRFP1255, KRFP1256, KRFP1263, KRFP1264, KRFP1265, KRFP1266, KRFP1272, KRFP1273, KRFP1276, KRFP1277, KRFP1283, KRFP1285, KRFP1287, KRFP1288, KRFP1290, KRFP1294, KRFP1295, KRFP1299, KRFP1300, KRFP1309, KRFP1310, KRFP1311, KRFP1315, KRFP1318, KRFP1319, KRFP1404, KRFP1405, KRFP1406, KRFP1407, KRFP1408, KRFP1410, KRFP1412, KRFP1418, KRFP1426, KRFP1427, KRFP1432, KRFP1437, KRFP1438, KRFP1439, KRFP1440, KRFP1441, KRFP1442, KRFP1443, KRFP1444, KRFP1447, KRFP1448, KRFP1449, KRFP1450, KRFP1452, KRFP1454, KRFP1457, KRFP1458, KRFP1461, KRFP1463, KRFP1464, KRFP1467, KRFP1468, KRFP1470, KRFP1471, KRFP1474, KRFP1484, KRFP1500, KRFP1503, KRFP1504, KRFP1506, KRFP1508, KRFP1513, KRFP1520, KRFP1524, KRFP1525, KRFP1530, KRFP1533, KRFP1535, KRFP1536, KRFP1538, KRFP1539, KRFP1542, KRFP1546, KRFP1564, KRFP1565, KRFP1566, KRFP1567, KRFP1568, KRFP1569, KRFP1571, KRFP1572, KRFP1573, KRFP1574, KRFP1577, KRFP1578, KRFP1583, KRFP1584, KRFP1588, KRFP1589, KRFP1590, KRFP1591, KRFP1592, KRFP1594, KRFP1597, KRFP1599, KRFP1600, KRFP1604, KRFP1606, KRFP1607, KRFP1609, KRFP1610, KRFP1611, KRFP1612, KRFP1613, KRFP1614, KRFP1615, KRFP1616, KRFP1619, KRFP1620, KRFP1621, KRFP1622, KRFP1624, KRFP1633, KRFP1634, KRFP1637, KRFP1638, KRFP1642, KRFP1644, KRFP1645, KRFP1646, KRFP1647, KRFP1648, KRFP1649, KRFP1652, KRFP1653, KRFP1654, KRFP1655, KRFP1659, KRFP1661, KRFP1662, KRFP1663, KRFP1664, KRFP1670, KRFP1676, KRFP1677, KRFP1684, KRFP1686, KRFP1695, KRFP1700, KRFP1707, KRFP1724, KRFP1725, KRFP1731, KRFP1733, KRFP1734, KRFP1735, KRFP1736, KRFP1737, KRFP1748, KRFP1750, KRFP1751, KRFP1755, KRFP1756, KRFP1761, KRFP1763, KRFP1769, KRFP1770, KRFP1771, KRFP1772, KRFP1773, KRFP1775, KRFP1776, KRFP1778, KRFP1779, KRFP1786, KRFP1787, KRFP1791, KRFP1792, KRFP1793, KRFP1796, KRFP1798, KRFP1799, KRFP1800, KRFP1801, KRFP1802, KRFP1811, KRFP1814, KRFP1815, KRFP1816, KRFP1817, KRFP1819, KRFP1822, KRFP1824, KRFP1826, KRFP1828, KRFP1829,

	KRFP1834, KRFP1861, KRFP1894, KRFP1895, KRFP1905, KRFP1907, KRFP1909, KRFP1910, KRFP1913, KRFP1915, KRFP1916, KRFP1920, KRFP1922, KRFP1926, KRFP1928, KRFP1930, KRFP1931, KRFP1932, KRFP1933, KRFP1934, KRFP1937, KRFP1945, KRFP1947, KRFP1948, KRFP1951, KRFP1957, KRFP1960, KRFP1963, KRFP1965, KRFP1974, KRFP1975, KRFP1979, KRFP1980, KRFP1982, KRFP1986, KRFP1987, KRFP1989, KRFP1997, KRFP1998, KRFP1999, KRFP2000, KRFP2002, KRFP2005, KRFP2007, KRFP2013, KRFP2015, KRFP2018, KRFP2019, KRFP2020, KRFP2021, KRFP2024, KRFP2025, KRFP2026, KRFP2027, KRFP2029, KRFP2030, KRFP2034, KRFP2035, KRFP2037, KRFP2038, KRFP2039, KRFP2043, KRFP2048, KRFP2054, KRFP2066, KRFP2068, KRFP2095, KRFP2103, KRFP2104, KRFP2105, KRFP2113, KRFP2123, KRFP2124, KRFP2125, KRFP2127, KRFP2128, KRFP2130, KRFP2131, KRFP2135, KRFP2137, KRFP2149, KRFP2150, KRFP2151, KRFP2152, KRFP2154, KRFP2155, KRFP2158, KRFP2159, KRFP2161, KRFP2163, KRFP2164, KRFP2166, KRFP2167, KRFP2170, KRFP2171, KRFP2182, KRFP2183, KRFP2184, KRFP2186, KRFP2191, KRFP2192, KRFP2194, KRFP2195, KRFP2196, KRFP2197, KRFP2198, KRFP2200, KRFP2201, KRFP2203, KRFP2212, KRFP2214, KRFP2218, KRFP2220, KRFP2221, KRFP2222, KRFP2227, KRFP2235, KRFP2241, KRFP2242, KRFP2252, KRFP2259, KRFP2260, KRFP2261, KRFP2262, KRFP2263, KRFP2264, KRFP2265, KRFP2267, KRFP2277, KRFP2283, KRFP2288, KRFP2289, KRFP2305, KRFP2306, KRFP2308, KRFP2312, KRFP2314, KRFP2323, KRFP2336, KRFP2337, KRFP2349, KRFP2359, KRFP2379, KRFP2380, KRFP2386, KRFP2388, KRFP2392, KRFP2393, KRFP2398, KRFP2408, KRFP2409, KRFP2410, KRFP2413, KRFP2414, KRFP2415, KRFP2419, KRFP2420, KRFP2428, KRFP2433, KRFP2434, KRFP2443, KRFP2448, KRFP2457, KRFP2486, KRFP2507, KRFP2512, KRFP2516, KRFP2517, KRFP2524, KRFP2547, KRFP2548, KRFP2550, KRFP2560, KRFP2561, KRFP2564, KRFP2565, KRFP2576, KRFP2580, KRFP2581, KRFP2582, KRFP2583, KRFP2584, KRFP2585, KRFP2587, KRFP2589, KRFP2590, KRFP2592, KRFP2595, KRFP2596, KRFP2598, KRFP2607, KRFP2651, KRFP2665, KRFP2666, KRFP2667, KRFP2668, KRFP2672, KRFP2673, KRFP2675, KRFP2680, KRFP2682, KRFP2694, KRFP2695, KRFP2698, KRFP2699, KRFP2700, KRFP2703, KRFP2706, KRFP2710, KRFP2711, KRFP2712, KRFP2713, KRFP2714, KRFP2716, KRFP2721, KRFP2733, KRFP2736, KRFP2737, KRFP2739, KRFP2741, KRFP2750, KRFP2753, KRFP2760, KRFP2773, KRFP2777, KRFP2781, KRFP2796, KRFP2815, KRFP2816, KRFP2819, KRFP2820, KRFP2833, KRFP2835, KRFP2836, KRFP2838, KRFP2845, KRFP2848, KRFP2849, KRFP2850, KRFP2855, KRFP2856, KRFP2857, KRFP2859, KRFP2868, KRFP2870, KRFP2875, KRFP2876, KRFP2877, KRFP2882, KRFP2883, KRFP2884, KRFP2885, KRFP2886, KRFP2887, KRFP2905, KRFP2911, KRFP2913, KRFP2923, KRFP2930, KRFP2934, KRFP2949, KRFP2950, KRFP2960, KRFP2963,

	KRFP2964, KRFP2966, KRFP2967, KRFP2969, KRFP2972, KRFP2975, KRFP2976, KRFP2977, KRFP2978, KRFP2979, KRFP2981, KRFP2983, KRFP2986, KRFP2988, KRFP3008, KRFP3010, KRFP3012, KRFP3013, KRFP3017, KRFP3022, KRFP3025, KRFP3032, KRFP3034, KRFP3040, KRFP3045, KRFP3049, KRFP3050, KRFP3051, KRFP3054, KRFP3056, KRFP3057, KRFP3058, KRFP3062, KRFP3066, KRFP3078, KRFP3079, KRFP3081, KRFP3082, KRFP3084, KRFP3085, KRFP3086, KRFP3096, KRFP3099, KRFP3104, KRFP3105, KRFP3107, KRFP3111, KRFP3113, KRFP3120, KRFP3121, KRFP3131, KRFP3139, KRFP3143, KRFP3144, KRFP3145, KRFP3148, KRFP3149, KRFP3150, KRFP3151, KRFP3152, KRFP3153, KRFP3155, KRFP3156, KRFP3157, KRFP3158, KRFP3159, KRFP3160, KRFP3161, KRFP3163, KRFP3165, KRFP3166, KRFP3167, KRFP3169, KRFP3171, KRFP3173, KRFP3175, KRFP3177, KRFP3179, KRFP3180, KRFP3181, KRFP3182, KRFP3184, KRFP3188, KRFP3190, KRFP3191, KRFP3192, KRFP3195, KRFP3196, KRFP3197, KRFP3198, KRFP3200, KRFP3203, KRFP3206, KRFP3207, KRFP3208, KRFP3214, KRFP3215, KRFP3216, KRFP3218, KRFP3219, KRFP3220, KRFP3222, KRFP3223, KRFP3224, KRFP3225, KRFP3228, KRFP3229, KRFP3231, KRFP3232, KRFP3233, KRFP3255, KRFP3262, KRFP3264, KRFP3268, KRFP3276, KRFP3281, KRFP3283, KRFP3284, KRFP3285, KRFP3288, KRFP3289, KRFP3291, KRFP3293, KRFP3295, KRFP3297, KRFP3302, KRFP3303, KRFP3304, KRFP3321, KRFP3322, KRFP3325, KRFP3326, KRFP3327, KRFP3328, KRFP3331, KRFP3336, KRFP3337, KRFP3339, KRFP3346, KRFP3349, KRFP3351, KRFP3353, KRFP3355, KRFP3358, KRFP3359, KRFP3360, KRFP3361, KRFP3362, KRFP3363, KRFP3366, KRFP3367, KRFP3368, KRFP3369, KRFP3371, KRFP3373, KRFP3375, KRFP3379, KRFP3381, KRFP3382, KRFP3385, KRFP3387, KRFP3388, KRFP3389, KRFP3390, KRFP3392, KRFP3393, KRFP3394, KRFP3395, KRFP3396, KRFP3398, KRFP3399, KRFP3400, KRFP3402, KRFP3403, KRFP3404, KRFP3405, KRFP3407, KRFP3408, KRFP3409, KRFP3410, KRFP3412, KRFP3414, KRFP3415, KRFP3416, KRFP3418, KRFP3419, KRFP3420, KRFP3425, KRFP3426, KRFP3427, KRFP3428, KRFP3430, KRFP3434, KRFP3435, KRFP3436, KRFP3440, KRFP3441, KRFP3442, KRFP3443, KRFP3444, KRFP3448, KRFP3450, KRFP3452, KRFP3455, KRFP3466, KRFP3467, KRFP3473, KRFP3474, KRFP3478, KRFP3489, KRFP3490, KRFP3497, KRFP3501, KRFP3507, KRFP3509, KRFP3510, KRFP3512, KRFP3514, KRFP3515, KRFP3517, KRFP3520, KRFP3521, KRFP3523, KRFP3524, KRFP3525, KRFP3528, KRFP3529, KRFP3531, KRFP3534, KRFP3536, KRFP3537, KRFP3538, KRFP3540, KRFP3541, KRFP3542, KRFP3543, KRFP3546, KRFP3547, KRFP3548, KRFP3549, KRFP3550, KRFP3553, KRFP3554, KRFP3559, KRFP3560, KRFP3561, KRFP3562, KRFP3564, KRFP3568, KRFP3569, KRFP3572, KRFP3573, KRFP3574, KRFP3577, KRFP3578, KRFP3579, KRFP3580, KRFP3581, KRFP3582, KRFP3583, KRFP3584, KRFP3585, KRFP3586, KRFP3587,

	KRFP3590, KRFP3591, KRFP3592, KRFP3593, KRFP3594, KRFP3595, KRFP3596, KRFP3597, KRFP3598, KRFP3599, KRFP3600, KRFP3602, KRFP3605, KRFP3606, KRFP3608, KRFP3609, KRFP3610, KRFP3611, KRFP3612, KRFP3613, KRFP3614, KRFP3616, KRFP3617, KRFP3618, KRFP3619, KRFP3620, KRFP3624, KRFP3625, KRFP3634, KRFP3635, KRFP3636, KRFP3639, KRFP3640, KRFP3641, KRFP3643, KRFP3645, KRFP3646, KRFP3647, KRFP3648, KRFP3649, KRFP3651, KRFP3653, KRFP3654, KRFP3655, KRFP3656, KRFP3657, KRFP3658, KRFP3659, KRFP3660, KRFP3661, KRFP3662, KRFP3665, KRFP3667, KRFP3668, KRFP3669, KRFP3670, KRFP3671, KRFP3672, KRFP3675, KRFP3676, KRFP3678, KRFP3680, KRFP3681, KRFP3682, KRFP3683, KRFP3685, KRFP3688, KRFP3689, KRFP3690, KRFP3691, KRFP3692, KRFP3693, KRFP3696, KRFP3697, KRFP3698, KRFP3699, KRFP3701, KRFP3702, KRFP3703, KRFP3704, KRFP3705, KRFP3706, KRFP3707, KRFP3708, KRFP3709, KRFP3710, KRFP3712, KRFP3713, KRFP3716, KRFP3717, KRFP3718, KRFP3719, KRFP3721, KRFP3722, KRFP3725, KRFP3726, KRFP3727, KRFP3728, KRFP3729, KRFP3730, KRFP3731, KRFP3732, KRFP3733, KRFP3735, KRFP3736, KRFP3737, KRFP3738, KRFP3739, KRFP3740, KRFP3741, KRFP3742, KRFP3743, KRFP3744, KRFP3745, KRFP3746, KRFP3749, KRFP3750, KRFP3751, KRFP3752, KRFP3753, KRFP3756, KRFP3757, KRFP3758, KRFP3759, KRFP3760, KRFP3761, KRFP3762, KRFP3763, KRFP3764, KRFP3765, KRFP3766, KRFP3767, KRFP3768, KRFP3770, KRFP3772, KRFP3773, KRFP3775, KRFP3776, KRFP3777, KRFP3778, KRFP3779, KRFP3781, KRFP3782, KRFP3783, KRFP3784, KRFP3785, KRFP3786, KRFP3787, KRFP3788, KRFP3789, KRFP3790, KRFP3795, KRFP3800, KRFP3802, KRFP3803, KRFP3806, KRFP3809, KRFP3811, KRFP3812, KRFP3813, KRFP3814, KRFP3815, KRFP3816, KRFP3817, KRFP3818, KRFP3820, KRFP3821, KRFP3840, KRFP3843, KRFP3844, KRFP3845, KRFP3849, KRFP3851, KRFP3855, KRFP3857, KRFP3859, KRFP3861, KRFP3862, KRFP3863, KRFP3864, KRFP3866, KRFP3867, KRFP3869, KRFP3872, KRFP3873, KRFP3874, KRFP3876, KRFP3879, KRFP3880, KRFP3882, KRFP3884, KRFP3885, KRFP3887, KRFP3888, KRFP3889, KRFP3892, KRFP3893, KRFP3894, KRFP3895, KRFP3897, KRFP3898, KRFP3899, KRFP3901, KRFP3902, KRFP3905, KRFP3907, KRFP3910, KRFP3914, KRFP3915, KRFP3916, KRFP3919, KRFP3921, KRFP3922, KRFP3925, KRFP3926, KRFP3928, KRFP3929, KRFP3930, KRFP3931, KRFP3932, KRFP3933, KRFP3934, KRFP3936, KRFP3937, KRFP3938, KRFP3940, KRFP3942, KRFP3943, KRFP3944, KRFP3945, KRFP3946, KRFP3947, KRFP3948, KRFP3949, KRFP3950, KRFP3953, KRFP3955, KRFP3956, KRFP3957, KRFP3959, KRFP3961, KRFP3963, KRFP3966, KRFP3967, KRFP3968, KRFP3971, KRFP3973, KRFP3983, KRFP3985, KRFP3989, KRFP3991, KRFP3994, KRFP3995, KRFP3997, KRFP3998, KRFP3999, KRFP4000, KRFP4003, KRFP4005, KRFP4006, KRFP4007, KRFP4009, KRFP4011, KRFP4013,

	KRFP4014, KRFP4015, KRFP4018, KRFP4019, KRFP4020, KRFP4021, KRFP4022, KRFP4023, KRFP4024, KRFP4026, KRFP4027, KRFP4028, KRFP4032, KRFP4035, KRFP4045, KRFP4046, KRFP4047, KRFP4050, KRFP4051, KRFP4052, KRFP4053, KRFP4054, KRFP4055, KRFP4059, KRFP4062, KRFP4065, KRFP4066, KRFP4067, KRFP4068, KRFP4070, KRFP4071, KRFP4073, KRFP4074, KRFP4076, KRFP4077, KRFP4080, KRFP4081, KRFP4085, KRFP4087, KRFP4088, KRFP4094, KRFP4105, KRFP4113, KRFP4115, KRFP4116, KRFP4117, KRFP4121, KRFP4122, KRFP4124, KRFP4126, KRFP4127, KRFP4128, KRFP4129, KRFP4130, KRFP4139, KRFP4142, KRFP4153, KRFP4155, KRFP4159, KRFP4160, KRFP4162, KRFP4163, KRFP4166, KRFP4167, KRFP4168, KRFP4169, KRFP4173, KRFP4174, KRFP4175, KRFP4176, KRFP4177, KRFP4178, KRFP4179, KRFP4181, KRFP4182, KRFP4183, KRFP4185, KRFP4188, KRFP4189, KRFP4191, KRFP4192, KRFP4193, KRFP4197, KRFP4198, KRFP4205, KRFP4209, KRFP4218, KRFP4225, KRFP4231, KRFP4237, KRFP4238, KRFP4239, KRFP4241, KRFP4243, KRFP4245, KRFP4246, KRFP4247, KRFP4250, KRFP4251, KRFP4252, KRFP4253, KRFP4254, KRFP4255, KRFP4256, KRFP4257, KRFP4258, KRFP4259, KRFP4260, KRFP4261, KRFP4262, KRFP4264, KRFP4267, KRFP4268, KRFP4269, KRFP4270, KRFP4273, KRFP4274, KRFP4275, KRFP4277, KRFP4278, KRFP4279, KRFP4280, KRFP4281, KRFP4282, KRFP4283, KRFP4284, KRFP4285, KRFP4286, KRFP4287, KRFP4288, KRFP4289, KRFP4290, KRFP4291, KRFP4292, KRFP4293, KRFP4294, KRFP4295, KRFP4296, KRFP4301, KRFP4302, KRFP4305, KRFP4313, KRFP4314, KRFP4315, KRFP4317, KRFP4318, KRFP4320, KRFP4321, KRFP4324, KRFP4326, KRFP4328, KRFP4329, KRFP4331, KRFP4335, KRFP4336, KRFP4338, KRFP4358, KRFP4359, KRFP4360, KRFP4368, KRFP4373, KRFP4387, KRFP4391, KRFP4398, KRFP4399, KRFP4401, KRFP4409, KRFP4411, KRFP4412, KRFP4413, KRFP4425, KRFP4427, KRFP4440, KRFP4452, KRFP4457, KRFP4470, KRFP4472, KRFP4479, KRFP4484, KRFP4485, KRFP4486, KRFP4487, KRFP4489, KRFP4490, KRFP4491, KRFP4494, KRFP4495, KRFP4496, KRFP4498, KRFP4499, KRFP4500, KRFP4501, KRFP4503, KRFP4504, KRFP4508, KRFP4509, KRFP4511, KRFP4512, KRFP4514, KRFP4515, KRFP4516, KRFP4517, KRFP4519, KRFP4520, KRFP4521, KRFP4522, KRFP4523, KRFP4524, KRFP4531, KRFP4533, KRFP4538, KRFP4539, KRFP4540, KRFP4541, KRFP4542, KRFP4545, KRFP4546, KRFP4547, KRFP4550, KRFP4552, KRFP4556, KRFP4557, KRFP4558, KRFP4561, KRFP4571, KRFP4579, KRFP4581, KRFP4583, KRFP4591, KRFP4593, KRFP4594, KRFP4595, KRFP4599, KRFP4600, KRFP4601, KRFP4602, KRFP4610, KRFP4612, KRFP4626, KRFP4632, KRFP4633, KRFP4640, KRFP4642, KRFP4643, KRFP4645, KRFP4650, KRFP4651, KRFP4659, KRFP4663, KRFP4665, KRFP4666, KRFP4667, KRFP4668, KRFP4669, KRFP4671, KRFP4674, KRFP4678, KRFP4679, KRFP4680, KRFP4681, KRFP4685, KRFP4692, KRFP4695, KRFP4698,

	KRFP4701, KRFP4708, KRFP4709, KRFP4710, KRFP4711, KRFP4713, KRFP4714, KRFP4716, KRFP4717, KRFP4718, KRFP4719, KRFP4725, KRFP4728, KRFP4729, KRFP4730, KRFP4733, KRFP4734, KRFP4735, KRFP4736, KRFP4740, KRFP4746, KRFP4747, KRFP4749, KRFP4751, KRFP4752, KRFP4753, KRFP4754, KRFP4755, KRFP4757, KRFP4759, KRFP4760, KRFP4762, KRFP4763, KRFP4764, KRFP4765, KRFP4766, KRFP4767, KRFP4769, KRFP4770, KRFP4772, KRFP4774, KRFP4778, KRFP4780, KRFP4793, KRFP4796, KRFP4798, KRFP4803, KRFP4804, KRFP4805, KRFP4808, KRFP4810, KRFP4811, KRFP4812, KRFP4813, KRFP4817, KRFP4820, KRFP4821, KRFP4822, KRFP4823, KRFP4824, KRFP4826, KRFP4827, KRFP4828, KRFP4829, KRFP4830, KRFP4831, KRFP4832, KRFP4833, KRFP4834, KRFP4835, KRFP4842, KRFP4843, KRFP4852, KRFP4853, KRFP4854, KRFP4856, KRFP4857, KRFP4858, KRFP4859,

Table S3- AUC of ROC plot of models built on 3D descriptors

Model No.	Parameters	AUC of ROC score
NB-5	Type- Gaussian Var smoothing- 10^{-9}	0.80
KNN-5	N neighbours -7	0.90
SVM-5	C=0.01 Gamma=0.1 Kernel=Rbf	0.88
RF-5	N estimator=500 Max depth=15 Min samples split= 4	0.93
GB-5	N estimator=500 Max depth=4 Min samples split= 7	0.92

Table S4- Summary of parameters for Random Forest classifier

Descriptors	**Parameters**
MACCS	max_depth=20, min_samples_leaf=5, n_estimators=200, n_jobs=-1,
ECFP-4	max_depth=0, min_samples_leaf=5, n_estimators=50, n_jobs=-1,

ECFP-6	max_depth=5, min_samples_leaf=5, n_estimators=10, n_jobs=-1,
PubChem	max_depth=20, min_samples_leaf=5, n_estimators=50, n_jobs=-1,
Estate	max_depth=5, min_samples_leaf=5, n_estimators=300, n_jobs=-1,

Table S5- Summary of parameters for XGBoost classifier

Descriptors	**Parameters**
MACCS	colsample_bytree: 0.5, gamma: 0, learning_rate: 0.1, max_depth: 11, reg_lambda: 1, scale_pos_weight: 5, subsample: 0.8, tree_method: 'gpu_hist'
ECFP-4	colsample_bytree: 0.5, gamma: 0, learning_rate: 0.05, max_depth: 7, reg_lambda: 10, scale_pos_weight: 1, subsample: 0.8, tree_method: 'gpu_hist'
ECFP-6	colsample_bytree: 0.5, gamma: 0, learning_rate: 0.1, max_depth: 11, reg_lambda: 3, scale_pos_weight: 3, subsample: 0.8, tree_method: 'gpu_hist'
PubChem	colsample_bytree: 0.5, gamma: 0, learning_rate: 0.1, max_depth: 9, reg_lambda: 0, scale_pos_weight: 9,

	subsample: 0.8, tree_method: 'gpu_hist'
Estate	colsample_bytree: 0.5, gamma: 0, learning_rate: 0.1, max_depth: 7, reg_lambda: 0, scale_pos_weight: 5, subsample: 0.8, tree_method: 'gpu_hist'

Table S6- Summary of parameters for LightGBM classifier

Descriptors	**Parameters**
MACCS	num_leaves=15, reg_alpha=0.1, min_data_in_leaf=50, learning_rate=0.01, lambda_l1=0, lambda_l2=1
ECFP-4	num_leaves=30, reg_alpha=0.1, min_data_in_leaf=30, learning_rate=0.5, lambda_l1=1, lambda_l2=0
ECFP-6	num_leaves=20, reg_alpha=0.1, min_data_in_leaf=30, learning_rate=0.5, lambda_l1=1, lambda_l2=0
PubChem	num_leaves=10, reg_alpha=0.1, min_data_in_leaf=50, learning_rate=0.7, lambda_l1=1, lambda_l2=1
Estate	num_leaves=20, reg_alpha=0.1, min_data_in_leaf=30, learning_rate=0.3, lambda_l1=0, lambda_l2=1

Table S7. List of all the unique features extracted from the training dataset:

Features	['','[UNK]','C','=C','Branch1','Ring1', 'O','=Branch1', 'N','=O', 'Ring2', 'C@H1', 'C@@H1', 'Branch2', '#Branch1', 'S', '=N', '=Branch2', '#Branch2', 'F', 'Cl', 'C@@', '#C', 'C@', 'P', '/C', 'NH1', '=Ring1', 'N+1', '\\C', 'O-1', '/N', 'Br', '#N', 'I', '=S', '=N+1', '/O', '=Ring2', '\\O', 'S+1', '/C@@H1', '/C@H1', '\\C@H1', '\\N', '\\C@@H1', '=P', '=N-1', '/S', '/N+1', 'S@@+1', 'N@+1', 'N.N', 'N-1', 'C+1', '/Cl', '\\S', 'N@@+1', '/F', '-\\Ring2', 'S@+1', 'O.C', 'NH0', 'B', '=O+1', '#N+1', '\\F', '\\Cl', 'Xe', 'SiH3', 'SH0', 'S@@', 'Ring1.N', 'P+1', 'O+1', 'Ne', 'NH1-1', 'Li', 'Kr', 'CH1', 'C.N', 'C.C', 'C-1', 'Branch2.C', 'Ar', '=SH1','=P@@','/I','-/Ring2','#S@','#C-1']

Figures:

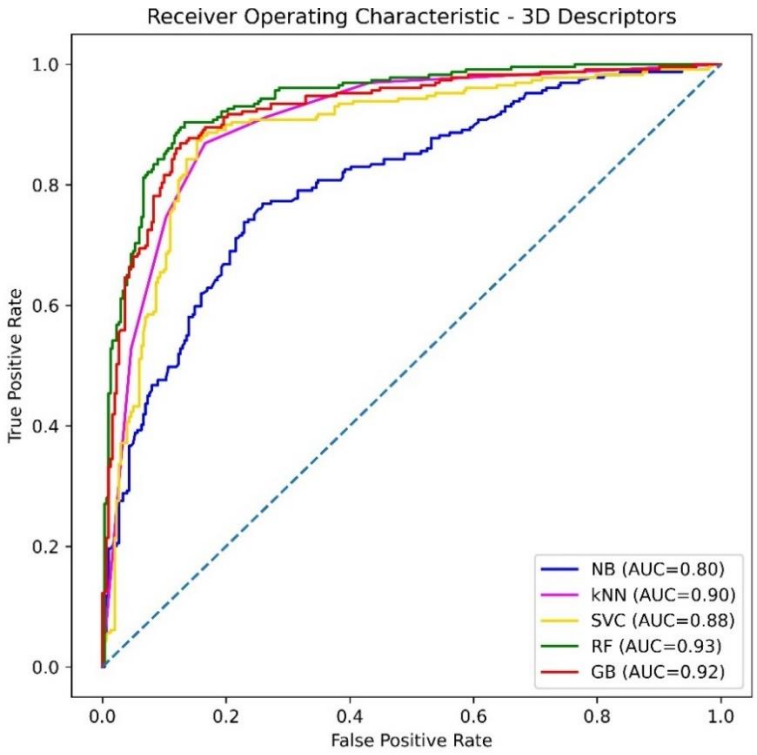


Figure S1- Receiver operating characteristics (ROC) curves of models build on 3D descriptors.

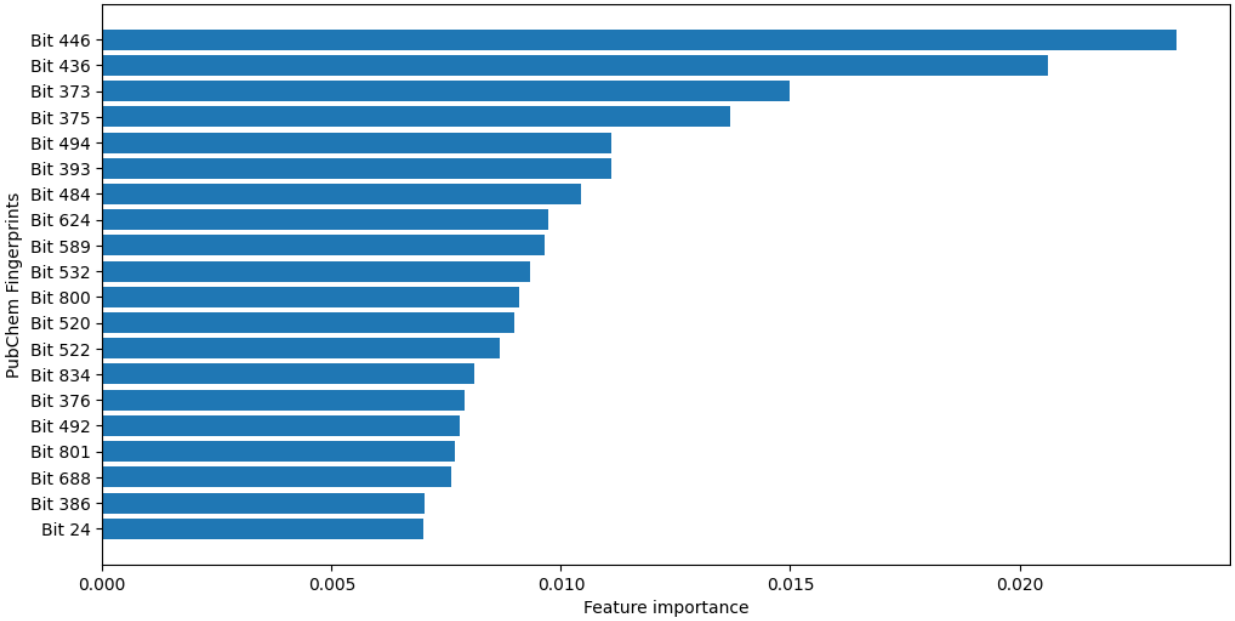


Figure S2- Feature importance of top 20 PubChem fingerprints obtained from the XGBoost model

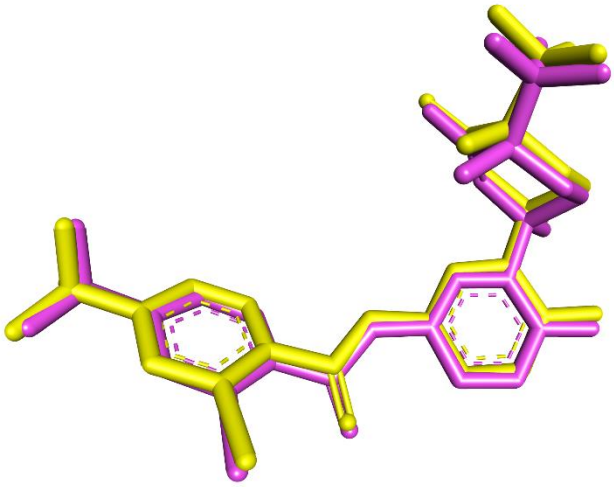


Figure S3- Superimposed structure of co-crystallized ligand (Yellow) and docked pose (Pink).

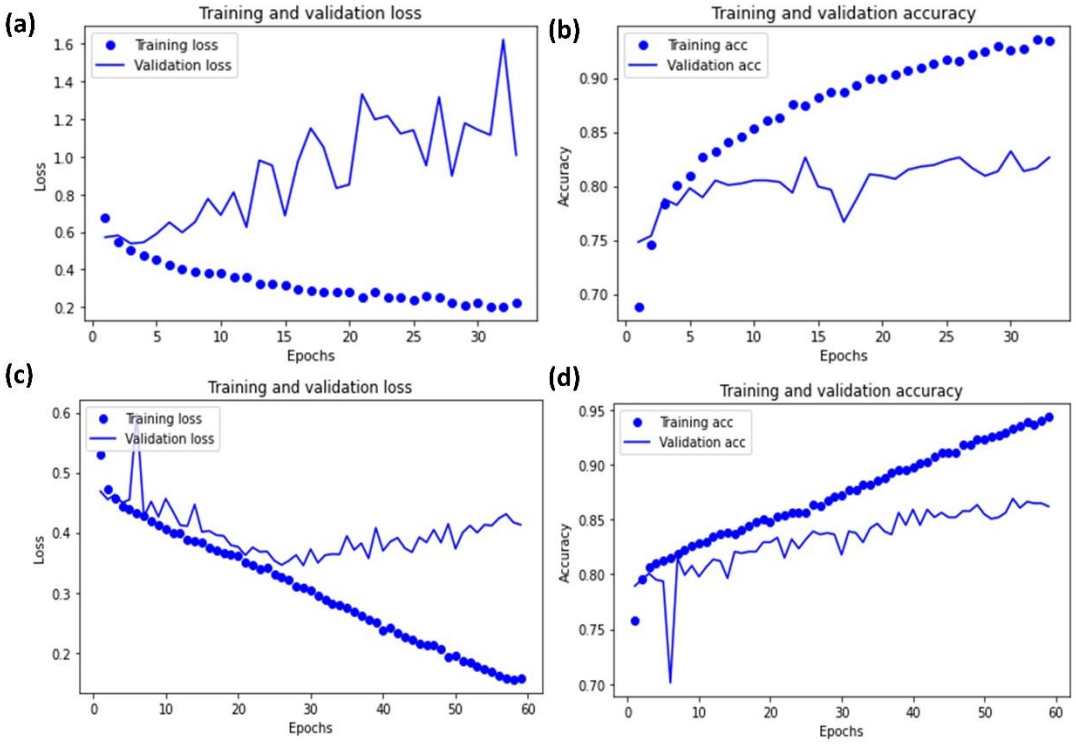


Figure S4- Performance of Model while training: (a) Training loss and validation loss for ANN-1 Model (b) Training and Validation accuracy for ANN-1 model (c) Training loss and validation loss for LSTM-1 Model (d) Training and Validation accuracy for LSTM-1 model

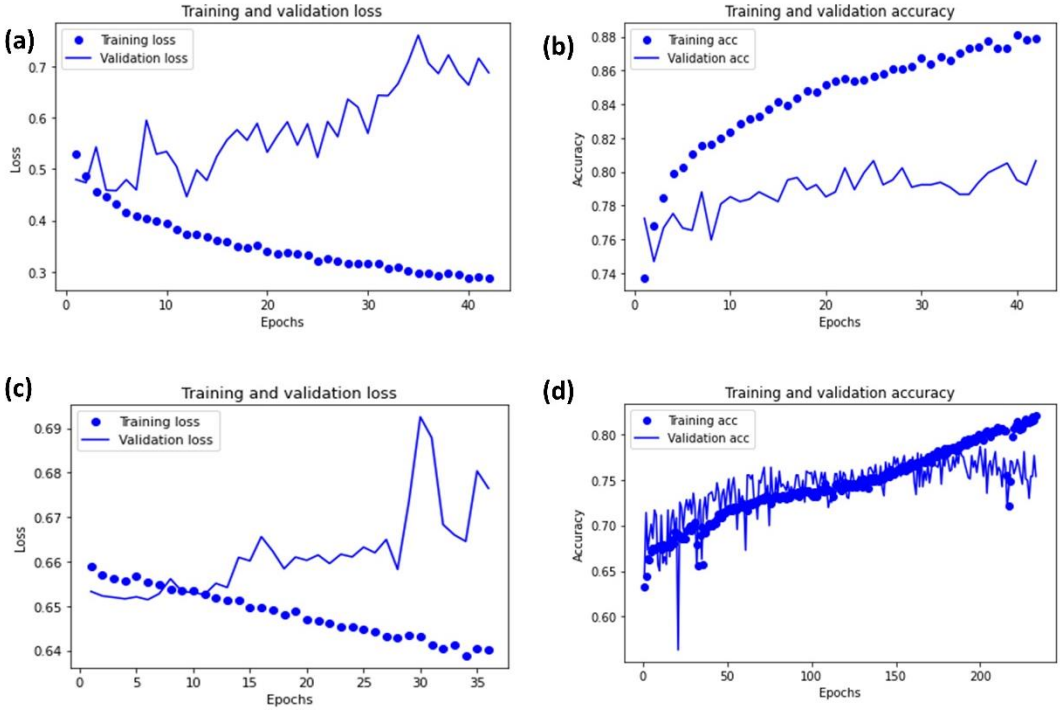


Figure S5- Performance of Model while training: (a) Training loss and validation loss for ANN-2 Model (b) Training and Validation accuracy for ANN-2 model (c) Training loss and validation loss for LSTM-2 Model (d) Training and Validation accuracy for LSTM-2 model

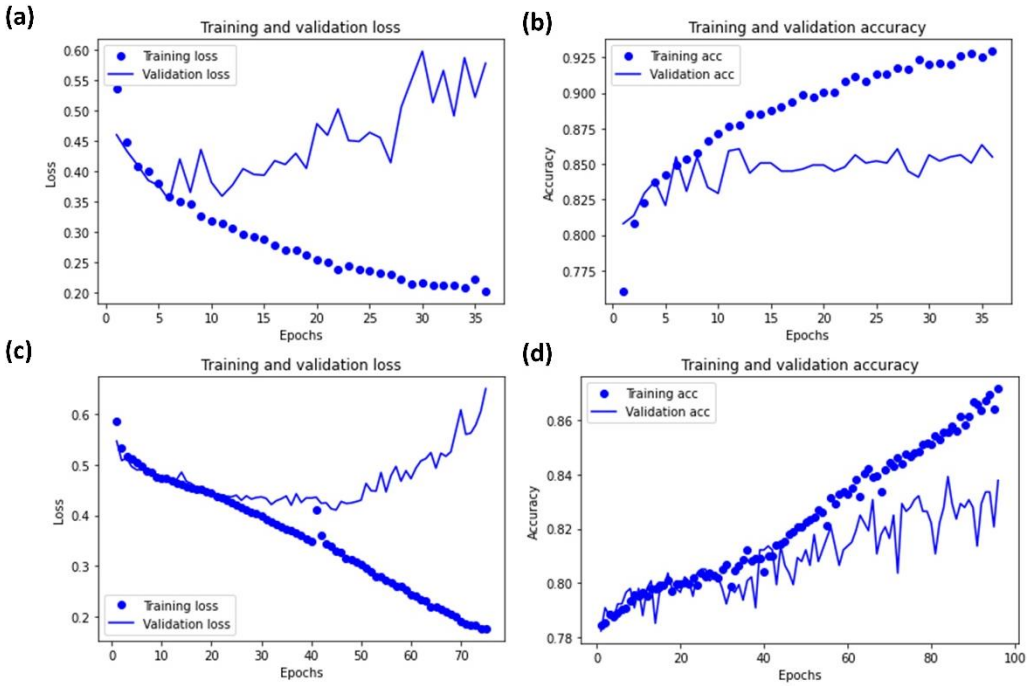


Figure S6- Performance of Model while training: (a) Training loss and validation loss for ANN-3 Model (b) Training and Validation accuracy for ANN-3 model (c) Training loss and validation loss for LSTM-3 Model (d) Training and Validation accuracy for LSTM-3 model

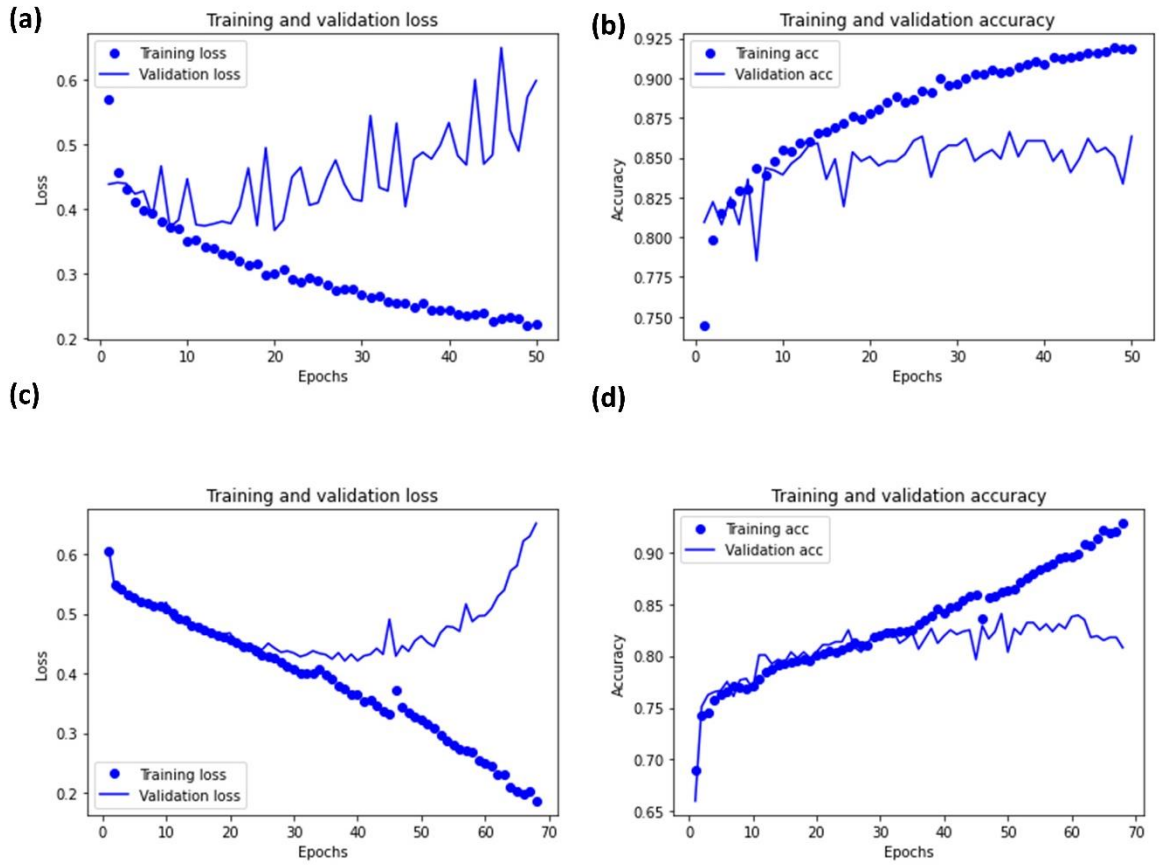


Figure S7- Performance of Model while training: (a) Training loss and validation loss for ANN-4 Model (b) Training and Validation accuracy for ANN-4 model (c) Training loss and validation loss for LSTM-4 Model (d) Training and Validation accuracy for LSTM-4 model

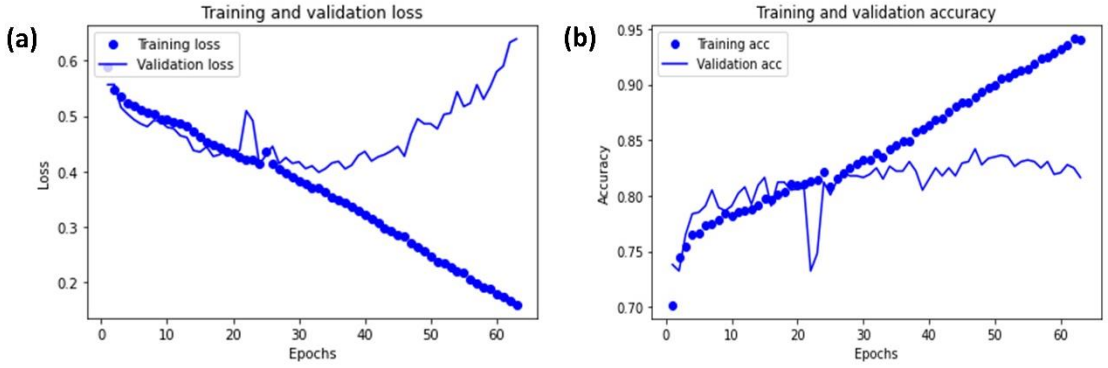


Figure S8- Performance of Model while training: (a) Training loss and validation loss for LSTM-5 Model (b) Training and Validation accuracy for LSTM-5 model

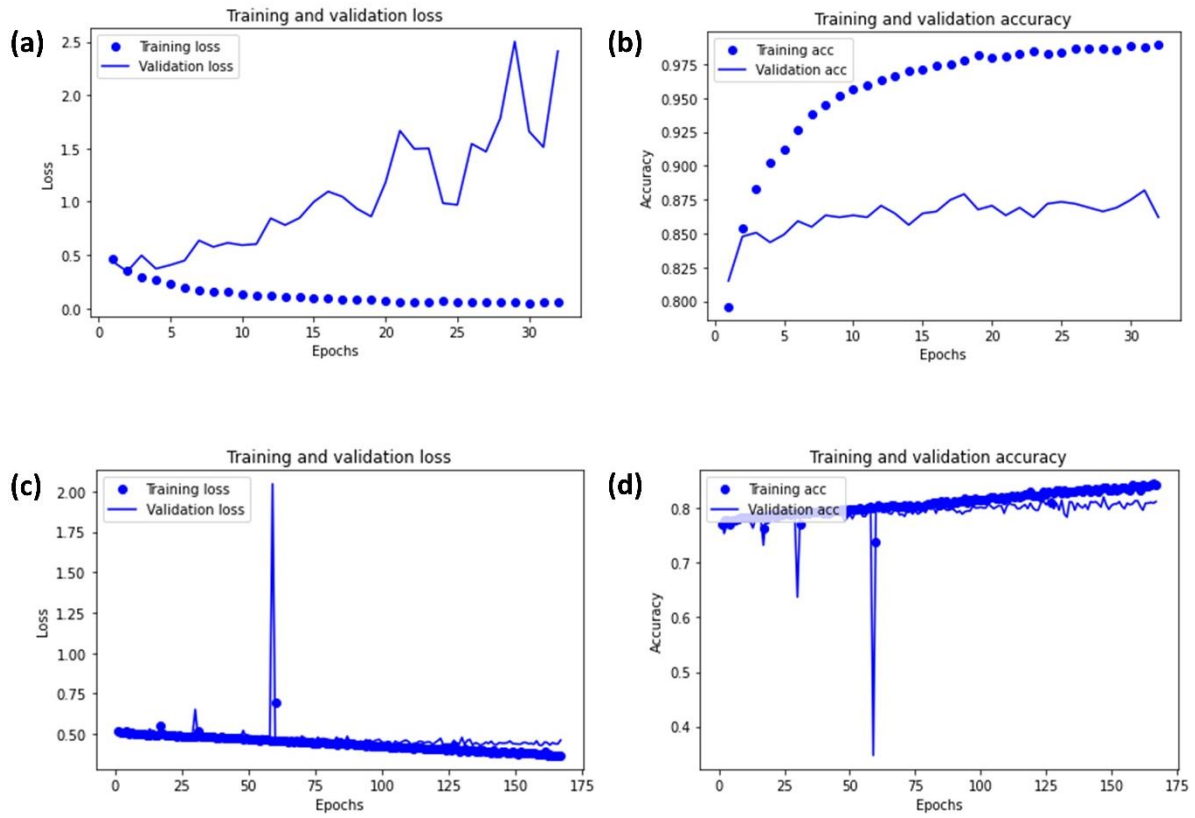


Figure S9- Performance of Model while training: (a) Training loss and validation loss for ANN-6 Model (b) Training and Validation accuracy for ANN-6 model (c) Training loss and validation loss for LSTM-6 Model (d) Training and Validation accuracy for LSTM-6 model

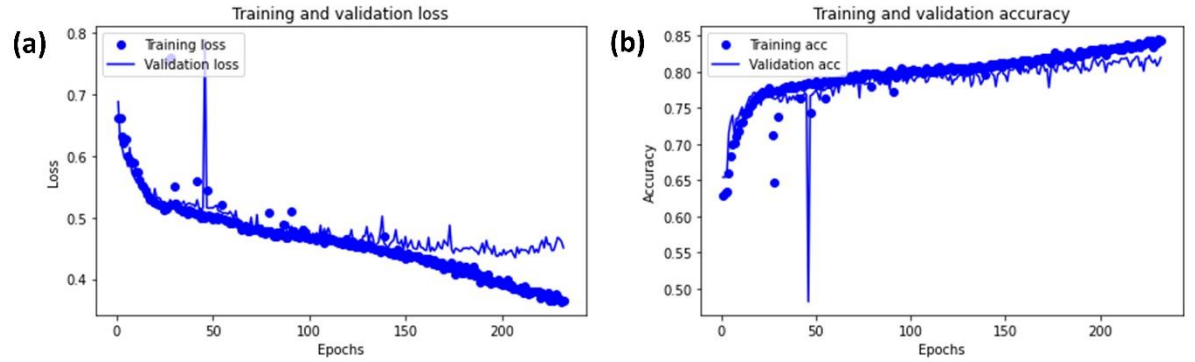


Figure S10- Performance of Model while training: (a) Training loss and validation loss for LSTM-7 Model (b) Training and Validation accuracy for LSTM-7 model

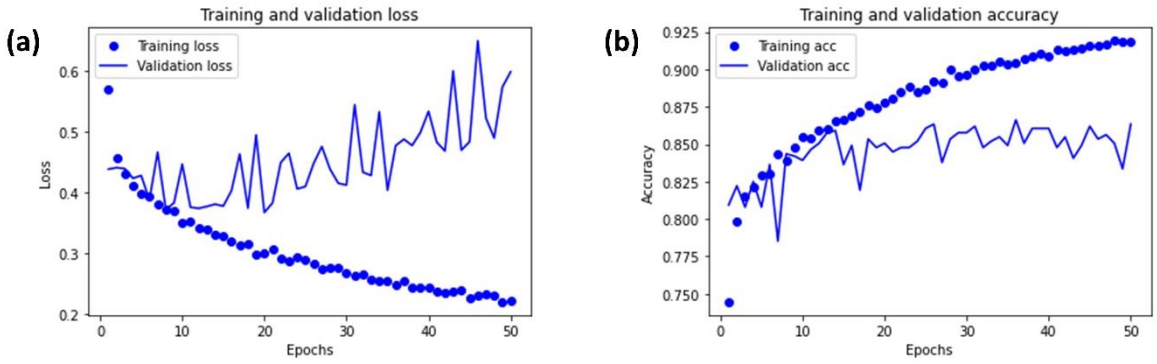


Figure S11- Performance of Model while training: (a) Training loss and validation loss for ANN-8 Model (b) Training and Validation accuracy for ANN-8 model

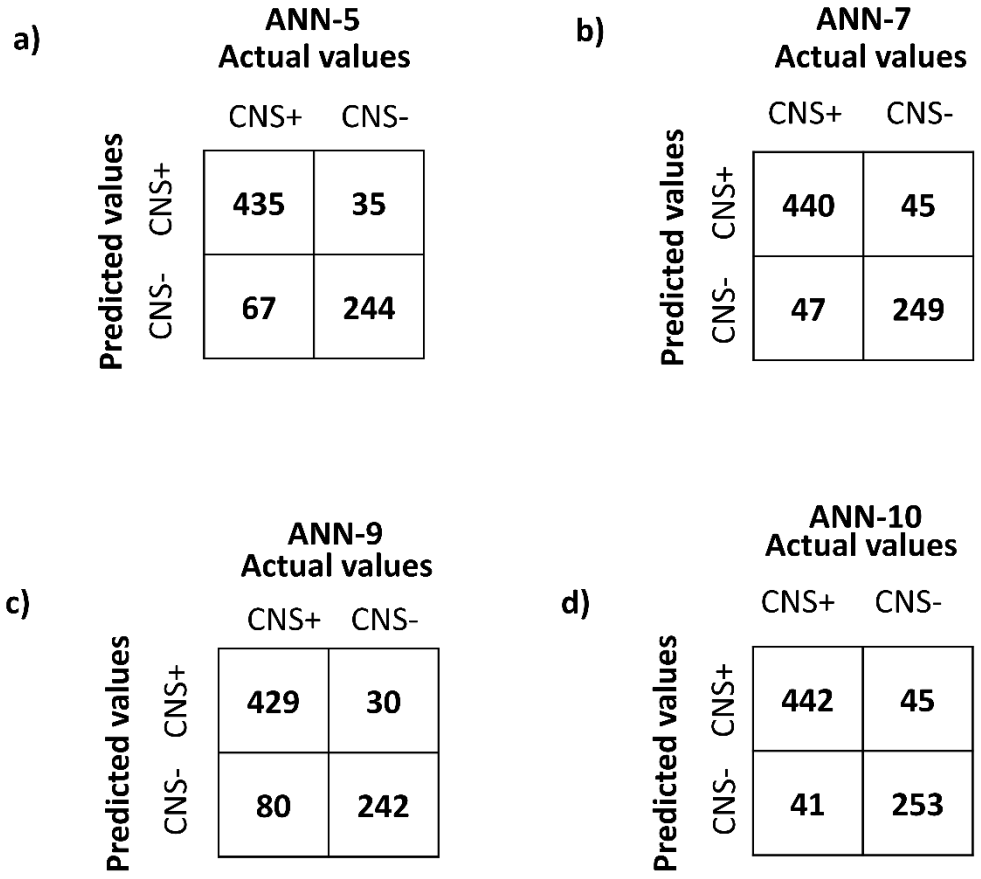
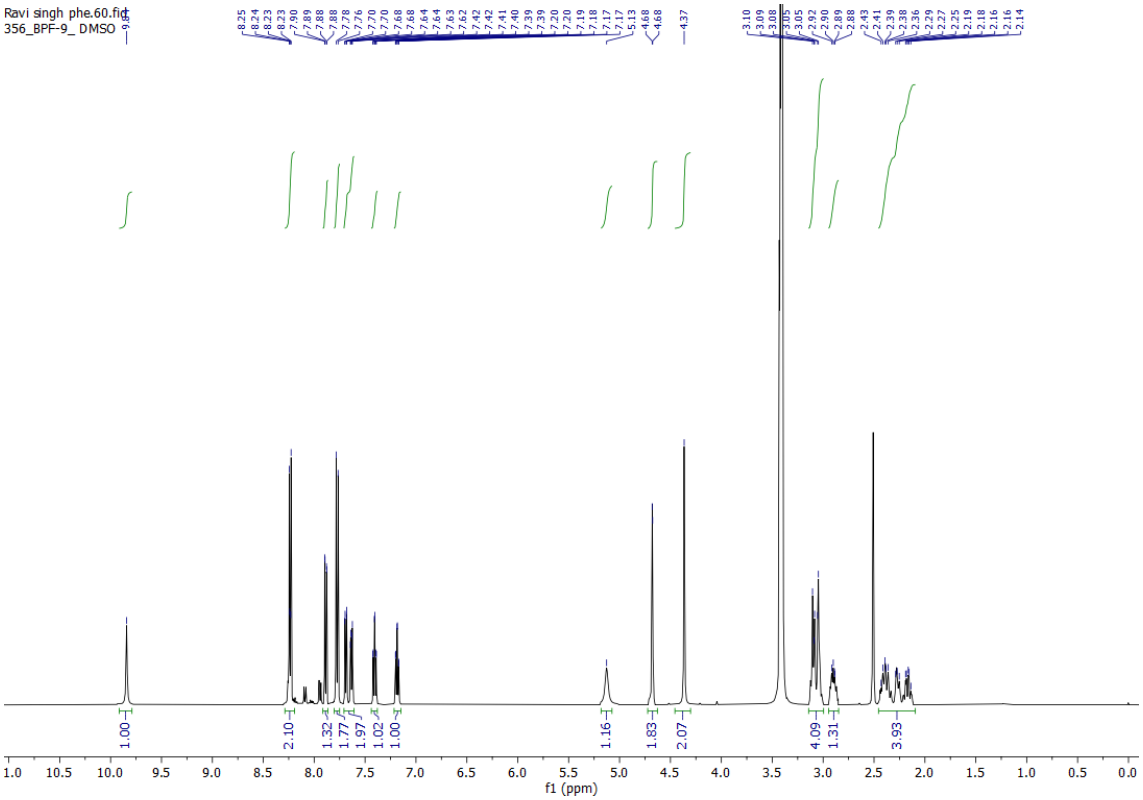


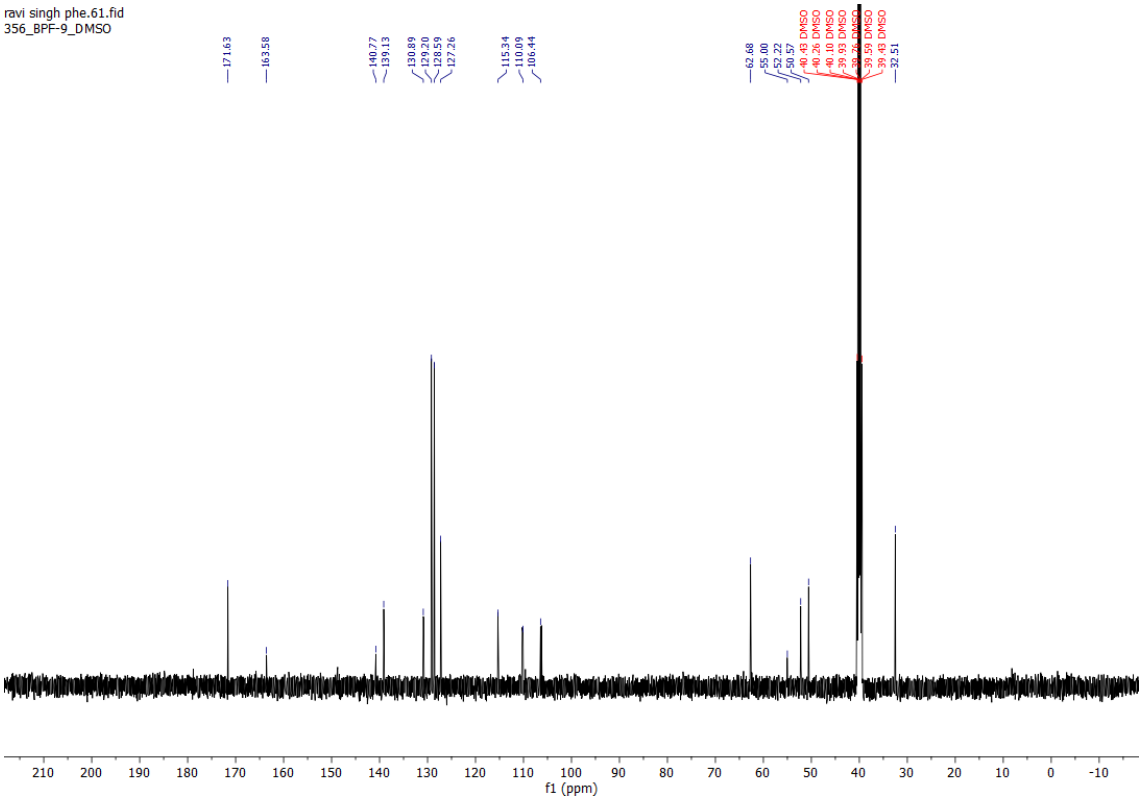
Figure S12- Confusion matrix of models on Test set: (a) ANN-5 model (b) ANN-7 (c) ANN-9 (d) ANN-10

Spectral data:

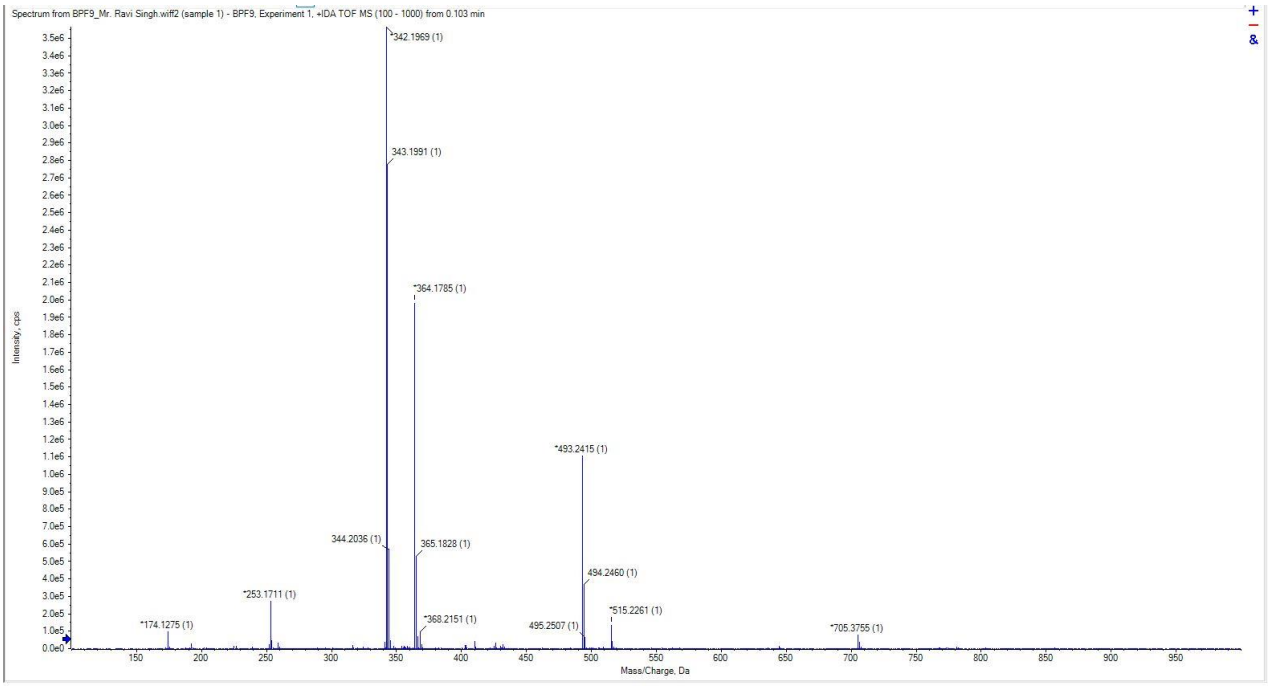
2-((1-benzylpiperidin-4-yl)amino)-N-(2-fluorophenyl)acetamide



Spectra 1. ^{1}H-NMR spectra of compound 66



Spectra 2. ^{13}C-NMR spectra of compound 66



Spectra 3. HRMS of compound 66

List of Publications

1. **Singh R**, Ganeshpurkar A, Ghosh P, Pokle AV, Kumar D, Singh Rb, Singh SK, Kumar A. 2021. Classification of beta-site amyloid precursor protein cleaving enzyme 1 inhibitors by using machine learning methods. *Chemical Biology & Drug Design*. 98(6):1079-1097.
2. **Singh R**, Anand A, Ganeshpurkar A, Ghosh P, Chaurasia T, Singh RB, Kumar D, Singh SK, Kumar A. 2023. Machine learning-based screening of in-house database to identify BACE-1 inhibitors. *Chemical Papers*. 77(11):6849-6858.
3. **Singh R**, Ghosh P, Ganeshpurkar A, Anand A, Swetha R, Singh RB, Kumar D, Singh SK, Kumar A. 2023. Natural-Language Processing (NLP) based feature extraction technique in Deep-Learning model to predict the Blood-Brain-Barrier permeability of molecules. *Molecular Informatics*. 42(10):2200271.
4. **Singh R**, Ganeshpurkar A, Singh SK, Kumar A, Design, synthesis and biological evaluation of N-benzylpiperidines as multi-target directed ligands for the treatment of Alzheimer's disease *(Manuscript under review*)