COMPUTATIONAL INDENTIFICATION OF NOVEL BACE 1 MODULATORS FOR ALZHEIMER'S DISEASE: A STRUCTURE DRIVEN DRUG DISCOVERY APPROACH

A Dissertation
Submitted in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

in

BIOTECHNOLOGY

by

Sanjana Yadav

(24/MSCBIO/33)

Under the supervision of

Prof. Pravir Kumar

Department of Biotechnology

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Bawana Road, Delhi -110042, INDIA

May, 2026

ACKNOWLEDGEMENT

At the time of submitting my M.Sc. Dissertation, I would First and foremost like to express my heartfelt gratitude to the Almighty God for providing me with the strength, patience, and perseverance that was constantly required to successfully complete my Project work.

I would like to convey my sincere appreciation to my respected mentor, Prof. Pravir Kumar, Department of Biotechnology, Delhi Technological University, who gave me this chance to work on this project under Himself. I was able to successfully complete my dissertation work due to his continuous guidance, valuable input, and steadfast support throughout the course of this dissertation work. His insightful direction and encouragement have been instrumental and played a major role in completion and proper shaping of this study.

I would also like to thank the faculty and staff of my department for their continuous support and for creating an environment that encouraged effective learning and research.

I am also deeply thankful and would like to impart my Appreciation to Ms. Nishita Singh, Ms. Apurva, working under my mentor, for their persistent assistance, guidance and supervision. I truly value their willingness to share their knowledge with me. Their support, discussions, and practical insights played a crucial role in the successful completion of my thesis work.

I would like to convey my deepest appreciation to my family members for their Unrelenting support and encouragement throughout this journey. Their patience, understanding, and belief in my abilities have been a strong source of stimulus for me, especially during challenging times. Their emotional support and positive reinforcement played a central role in helping me remain focused and complete this work successfully.

I'm also highly grateful to my friend Harshita Yadav and my Sisters for their support and camaraderie throughout the completion of my thesis. Their presence made this whole journey not only just manageable but also less stressful and more enjoyable.

Finally, I offer my heartfelt appreciation to everyone who were involved, directly or indirectly, in bringing this thesis to completion.

SANJANA YADAV
(24/MSCBIO/33)

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

DECLARATION

I Sanjana Yadav 24/MSCBIO/33 hereby certify that the work which is being presented in the thesis entitled "Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach" in partial fulfillment of the requirements for the award of the Degree of Master of Science, submitted in the Department of Biotechnology, Delhi Technological University is an authentic record of my own work carried out during the period from 2024 to 2026 under the supervision of Prof. Pravir Kumar.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **Sanjana Yadav** (24/MSCBIO/33) has carried out her research work presented in this thesis entitled **"Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach"** for the award of Master of Science from Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Date:

Prof. Pravir Kumar
Supervisor
Department of Biotechnology
Delhi Technological University

Prof. Yasha Hasija
Head of Department
Department of Biotechnology
Delhi Technological University

"COMPUTATIONAL INDENTIFICATION OF NOVEL BACE 1 MODULATORS FOR ALZHEIMER'S DISEASE: A STRUCTURE DRIVEN DRUG DISCOVERY APPROACH"

SANJANA YADAV
24/MSCBIO/33

ABSTRACT

Aim: Amyloid beta deposition in the neural tissue, which leads to neural detrioration and escalating cognitive deterioration are major signatures of Alzheimer disease, a chronic neurodegenerative disorder. Since BACE1 sparks the amyloidogenic pathway fueling Amyloid beta generation and plaque augmentation, inhibiting BACE 1 is a promising approach to curb its accumulation in AD. This study used a computational approach where 5HTZ a co crystallised molecular structure which showcases a potent inhibitor bound or complexed at the catalytic pocket or binding site of the BACE1 was utilized as reference structure to find promising BACE 1 inhibitors. The primary objective was to shortlist better suited compounds for BACE 1 inhibition and thus Alzheimer treatment to that was formely found. Here, a large number of compounds were Shortlisted on the basis of structural similarity to the reference molecule which Were filtered by ADME analysis. A carefully chosen compound library was docked into the catalytic pocket of BACE1 after protein preparation and active-site mapping in order to assess interaction potency and interaction stability. Top hits having strong interactions with key residues were profiled for ADME, including drug –likeness, GI absorption, BBB permeability and CNS relevant characteristics. Numerous compounds demonstrated advantageous pharmacokinetic profiles and binding energies, making them prime candidates for pharmacological and biochemical verification. These findings supports the use of computational techniques to accelerate the discovery of novel BACE1 modulators for the disease. Further they also highlight promising lead candidates to carry out additional trial and confirmation.

Keywords— Alzheimers disease, BACE1, amyloid beta generation, virtual screening, ADME analysis, catalytic pocket, active site mapping, pharmacokinetic profiles, computational strategies.

Result: During the course of this study, an initial pool of 400 structurally similar compounds was screened, from which ADME evaluation refined the selection to 81 candidates, all exhibiting blood–brain barrier permeability. Subsequent molecular docking analysis revealed that five of these compounds demonstrated stronger binding affinities than the reference compound ,with the top-performing compound showing a binding affinity of −8.7 kcal/mol.

Conclusion: From the set of five shortlisted candidates, Compound 5 demonstrated the strongest performance, showing the highest binding affinity along with the ability to cross the blood–brain barrier. Further validation of these results through in vivo studies is recommended to confirm its potential.

List of Publications

1. Conference Paper

Title of paper - *"Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach"*

Author Names – Sanjana Yadav and Prof. Pravir Kumar

Name of the Conference - Second International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies CE2CT-2026 (IEEE Conference Record #68422)

Date of Conference - 02–04 July 2026 at Graphic Era Hill University Bhimtal, Nainital, Uttarakhand,India

Indexing - IEEE

Status of Paper – Accepted

Date of Acceptance – 24th April ,2026

TABLE OF CONTENTS

LIST OF FIGURE

LIST OF TABLES

LIST OF ABBREVIATIONS

AD	Alzheimer's Disease
Aβ	Amyloid Beta
BACE1	Beta-site APP Cleaving Enzyme 1
APP	Amyloid Precursor Protein
ADME	Absorption, Distribution, Metabolism, Excretion
Ach	Acetylcholine
CNS	Central Nervous System
AICD	APP Intracellular Domain
γ-secretase	Gamma-secretase
HTVS	High-Throughput Virtual Screening
EA Dock DSS	Evolutionary Algorithm for Docking with Dihedral Space Sampling
SP	Standard Precision
XP	Extra Precision
LY	Eli Lilly Code
JNJ	Johnson & Johnson
MK	Merck
RG	Roche / Genentech
BI	Boehringer Ingelheim
PF	PFizer
E	Eisai
CNP	Cerveau/Novartis
AutoDock	Automated Docking
PyRx	Python Prescription
GOLD	Genetic Optimisation for Ligand Docking

GLIDE	Grid-based Ligand Docking with Energetics
Surflex	Surface based flexible molecular docking
MCDock	Monte Carlo Docking
LeDock	Linux Docking
rDock	Real Dock
ICM	Internal Coordinates Mechanics
Cdcker	CHARMM Based Docker
FRED	Fast Rigid Exhaustive Docking
MOL2	Molecule format version 2
SDF	Structure Data File
PDB	Protein Data Bank
PDBQT	Protein Data Bank, Partial Charge (Q), & Atom Type (T)
GPF	Grid Parameter File
3D	Three Dimensional
OPLS3	Optimized Potentials for Liquid Simulations 3
AMBER	Assisted Model Building with Energy Refinement
UFF	Universal Force Field
MMFF94x	Merck Molecular Force Field 94x
pH	Potential of Hydrogen
ΔG_bind	Gibbs Free Energy of Binding
PyMol	Python Molecular Graphics
RMSD	Root Mean Square Deviation
BBB	Blood-Brain Barrier
TPSA	Topological Polar Surface Area

CYP2D6	Cytochrome P450 2D6
P-glycoproteins	Permeability Glycoproteins
MD	Molecular Dynamics
NPT	Number, Pressure, Temperature Ensemble
ADME	Absorption, Distribution, Metabolism, Excretion
RCSB	Research Collaboratory for Structural Bioinformatics
BIOVIA	Scientific Software Platform
ASP	Extra Precision
SMILES	Simplified Molecular Input Line Entry System
RO5	Rule of Five
DA	Docking Analysis
LOGP	Partition Coefficient
GI	Gastrointestinal
PAINS	Pan Assay Interference Compounds
Asp	Asparatic Acid
MPO	Multi-Parameter Optimization
PUBCHEM	Public Chemical Database
GUI	Graphical User Interface
PPI	Protein-Protein Interaction
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
CID	Compound Identifier
LOGKP	Skin Permeability Coefficient
BRENKS	Structural Alerts Filter
PSEN1	Presenilin 1
PSEN2	Presenilin 2
APO E	Apolipoprotein E

STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
LIPO	Lipophilcity
INSOLU	Insolubilty
INSATU	Insaturation
FLEX	Flexibilty
LogKp	Logarithm for skin Permeation
HIA	Human Intestinal Absorption
PGP	P- Glycoprotein Substrate

1. INTRODUCTION

Alzheimer's disease (AD) , an advancing brain deterioration or degradation , is attributed or associated with intellectual problems (like perception, understanding , learning) ,recollection or reminiscence impairment issues moreover changes in behavioural as well as psychological domains of life which are sufficiently enough to substantially disrupt the instrumental or functional activities of daily life .The pathological trajectory typically occurs slowly and cause significant problems in different areas of thinking. These include episodic memories of the individual's life, managerial functions that help in making decisions, problem solving as well as mental flexibilty and also it affects communication skills. The aggregation of amyloid-β (Aβ) peptides within the brain, that results in the formation of extracellular plaques, represents a core pathological hallmark of Alzheimer disease which leads to neuronal dysfunction and progressive cell death [1]. The production of Aβ peptides occurs through enzymatic cleavage of amyloid precursor protein (APP) by proteases that occurs through a process called as the amyloidogenic pathway [2]. Beta-secretase 1 (BACE1) initiates the first enzymatic event in the amyloidogenic cascade of APP, which results in production of amyloid-beta (Aβ) and formation of characteristic extracellular plaques[3]. BACE1, which is a proteolytic enzyme and belongs to a member of aspartic protease uses an aspartate residue within its catalytic mechanism. The enzyme contains domains. The extracellular domain of the enzymes contains two crucial catalytic aspartate residues (residue 93-96 and 289-292) that are essential for its proteolytic activity. The catalytic domain is optimally oriented that allows it to effectively cleave APP at the β-cleavage site, starting the amyloidogenic cascade that leads to Alzheimer's disease pathogenesis . Since Increased BACE1 activity is been linked with enhanced amyloid deposition and progression of Alzheimer's disease, it thus makes BACE1 an important therapeutic target [3].

Several small-molecule inhibitors have been engineered to inhibit BACE1 enzymatic activity one such is (3e,5s)-5-{3-Chloro-5-[5-(Prop-1-Yn-1-Yl)pyridin-3-Yl]thiophen-2-Yl}-2,5-Dimethyl-1,2,4-Thiadiazinan-3-Imine 1,1-Dioxide (6JJ). It is one of the most thoroughly researched and effective small – molecule inhibitor that have been developed to block BACE1 enzymatic activity . The protein data bank's structural datasets provide essential knowledge about protein–ligand interaction, underpinning evidence-based rational drug design methodologies [4].

Structure-based drug discovery, particularly through computational approaches, enables rapid screening of compounds and accurate prediction of protein-ligand interactions, thereby facilitating the efficient identification of lead compounds for therapeutic development while incurring substantially reduced costs and time investments relative to conventional experimental methodologies [5].

This study used a structure-based computational approach to identify potential inhibitors of BACE1 using the Co crystallised structure 5HTZ as a reference model. Selected compounds were evaluated through molecular docking and ADME analysis to identify molecules with promising binding characteristics and pharmacokinetic properties. Here, screening based on structure similarity, ADME analysis, protein-ligand docking Were performed.

2. REVIEW OF LITERATURE

Alzheimer's disease (AD) is responsible for roughly two-thirds of all dementia instances in people aged 65 and above, making it the leading cause of dementia. It begins slowly and keeps increasing in severity. This condition influences conduct and mental abilities. Forgetfulness typically manifests first, followed by issues with understanding, communication, focus, logic, and problem-solving. While AD itself isn't usually terminal, it impairs general health, heightening vulnerability to life-threatening issues [6].

AD is considered a multifaceted disease. Two main theories, the cholinergic and amyloid hypotheses, have been proposed as the potential triggers. Additionally, several risk elements such as elderly status, genetic predisposition, traumatic brain injuries, circulatory issues, microbial or pathogenic infections, and factors related to environment or sorroundings plays a role in the disease [2]. At now, about 50 million people all over the world are striken by this Alzheimer's disease, and this number is projected to double every five years, reaching approximately 152 million by 2050. AD's impact is far-reaching, affecting patients, caregivers, and nations, with forecasted yearly expenses of US$1 trillion. Right now, there's no cure, but treatments are available to manage symptoms [7].

2.1. Hypothesis

As discussed above two crucial disease mechanism explanations try to explain AD's pathology. These are *The Cholinergic Hypothesis* that points to dropped acetylcholine (ACh) measures in neural tissue, due to neuronal loss in the Nucleus Basalis of Meynert, as a major player in AD. ACh is vital for cognition, and its deterioration - to some extent due to beta-amyloid's negative impact on cholinergic function – contributes to role in AD. *The Amyloid Hypothesis,* presently a prominent concept, implies that amyloid beta (Aβ) peptides, mainly Aβ42, build up due to abnormal APP processing, leading to toxic amyloid aggregates that harm neurons. This concept is especially relevant for inherited AD [6].

2.2.Mechanism

2.2.1. BACE1 in Alzheimer's Disease: Mechanism-

Beta-site APP cleaving enzyme 1 is a cell membrane- associated aspartyl protease chiefly localized in CNS neurons. This enzyme serves as a pivotal driver in the development and progression of the of Alzheimer's disease by executing the primary and indispensable scission of the amyloid precursor protein (APP), triggering the amyloid-β peptide production cascade.

Disease Progression Dynamics -
At Normal Physiological State: BACE1 concentrations are maintained at a baseline level through strict cellular homeostatic mechanisms.

Alzheimer's Progression: A significant surge in protein expression and catalytic efficiency occurs, accelerating the neurodegenerative cycle.

2.2.2. Amyloidosis-

In AD, amyloidosis refers to the buildup of abnormal proteins (amyloid plaques) in the neural tissue, damaging cells and disrupting mental faculties. Tau tangles are another defining aspect, and together they drive the disease.

2.2.3. APP modification or Hydrolysis-

Amyloid precursor protein (APP) is a membrane-bound protein found in elevated levels in neurons. It can be degraded or cut by two different and contrasting mechanisms, these are:

(a) Non-amyloidogenic pathway
In this track, α-secretase cuts APP within the amyloid-β (Aβ) section. This cut ceases Aβ from generating . Because of this, the pathway is seen as protective and helpful for brain cells.

(b) Amyloidogenic pathway (BACE1-dependent)
In the second route, β-secretase (BACE1) starts the process. It cleaves APP at the β-site. This step is very important. It leads to the creation of amyloid-β peptides. Therefore, BACE1 controls whether harmful Aβ fragments are produced.

2.2.4.Phased action of BACE1-

Initiation: BACE1 identifies and cleaves APP at a specific site, releasing APPβ and generating C99.
C99 generation: C99 stays membrane-bound, functioning as the proximal source for Aβ production.
γ-secretase processing: γ-secretase further processes C99, yielding Aβ peptides (Aβ40, Aβ42) and AICD.

2.2.5. Synthesis and Buildup of Aβ-

Aβ42 is one of those peptides that are most likely to buildup and oligomerise. These peptides build up outside of cells and create:

Highly poisonous soluble oligomers

Fibrils that are insoluble

Plaques made of amyloid

These aggregates cause following pathological events, impede with synaptic signalling, and compromise or weaken neural communication.

2.2.6. Repercussions of BACE1functioning

Amplified BACE1 functioning causes more and more amyloid-β production or generation, which Facilitates to

Breakdown in neuron-to-neuron signaling

Energy crisis in the brain

Triggering of the brain's immune activity

Build-up of damaging free radicals

2.2.7. Implications in drug and therapy targeting

BACE1 is important because it carries out the very first step in the process that causes to amyloid-β formation. As this step is the one that sets up the whole pathway in flow, scientists see BACE1 as a key point where the disease process could be interrupted.

If the activity of BACE1 is lessened, the idea is to act before harmful changes build up in the brain, rather than trying to fix the damage later. In simpler terms, targeting BACE1 focuses on stopping the problem at its starting point.

By doing this, researchers aim to:

- **Lower the amount of amyloid-β formed in the brain from the starting**
- **Slow down or limit the buildup of plaques that impede with normal brain function**

Overall, this approach is designed to reduce the chain reaction of events that eventually leads to mnemonic deficit and decline in thinking ability in Alzheimer's disease [8].

2.3. Clinical Advancement: Amyloid-beta reducing agents

The notion of focusing on BACE1 suppressing compounds arose from the amyloid cascade hypothesis, which pushed pharmaceutical companies to invest significant resources into developing and fast-tracking small-molecule inhibitors as potential disease-modifying treatments for Alzheimer's disease. However, findings from initial medical trials were largely disappointing. This caused to the discontinuation of many of these candidates from development pipelines. The following part examines the progression and Ensuing decline of BACE1 inhibitors in clinical trials, with the aim of understanding the key reasons behind their current lack of success.

Here we will discuss some of BACE1 modulators.
These are:

LY2811376: First BACE1 inhibitor in clinical trials, showed Aβ reduction, but discontinued due to toxicity concerns.

LY2886721: Second-gen BACE1 inhibitor, showed Aβ reduction, but terminated due to liver toxicity.

RG7129: Discontinued due to liver toxicity.

BI 1181181: Showed Aβ reduction, but terminated due to skin reactions.

JNJ-54861911 (Atabecestat): Showed promise, but discontinued due to liver toxicity.

LY3314814 (Lanabecestat): Showed no benefit, discontinued due to futility.

MK-8931 (Verubecestat): Showed no benefit, discontinued due to safety concerns.

E2609 (Elenbecestat): Showed Aβ reduction, but discontinued due to unfavorable risk-benefit ratio.

CNP-520 (Umibecestat): Showed Aβ reduction, but discontinued due to cognitive decline and brain atrophy.

LY3202626: This modulator does not showed any significant effect thus it was discontinued shortly.

PF-06751979: Showed promise, but development stopped because of the reason of the Pfizer's shift in focus.

Many BACE1 modulators have failed in medical trials due to safety concerns or lack of efficacy [9].

2.4. Binding Mode Prediction for Structure-Based Drug Design

Molecular docking is a key part of modern drug discovery pipelines. Molecular docking serves as a cornerstone in the development of molecular based medicine approaches. computational binding studies or Molecular docking has become a key technique in pharmaceutical research. Molecular docking Appreciably contributes to the Spotlighting of lead compounds.

Definition

The use of computation of structure complexes designed up of two or more interconnected molecules is referred to as the process of molecular docking. Foreseeing an intriguing three-dimensional framework is the primary objective of molecular docking [10]

What does docking do

We can investigate the conduct of small molecules that reside in the site of binding of proteins of interest and shed light on underlying biochemical operations by utilising the conventional molecular docking protocol to represent the relationship across a bioactive compounds or ligands and protein or a macromolecular target at the atomic scale. Foreseeing the ligand's conformation as well as its position and orientation within these sites (often referred to as pose) and determining the affinity for binding are the two basic steps in the process of docking. These two steps are associated with scoring schemes and sampling techniques, respectively [11].

For the sake of in silico or computational docking methods, a number of commercial and several of digital tools and logic system are available. The medication research and scholarly fields are presently using these tools and programs that have been created. Auto Dock Vina, Discovery Studio, pyrx, Surflex, AutoDock GOLD, Glide, MCDock, MOE-Dock, FlexX, DOCK, LeDock, rDock, ICM, Cdcker, LigandFit, FRED, and UCSF Dock are some of the most popular docking programs, With the greatest scores among these programs that are AutoDock Vina, Glide, and AutoDock GOLD have been called to be the best options. Furthermore to this, contingent on the observed binding poses, a subset of of these software programs have been reliable in forecasting Root Mean Square Deviations (RMSDs) between 1.5 and 2 Å [12].

2.5. Tools used for molecular docking

2.5.1.TheAutoDock-

A very well-known open-source software package for simulations of docks is AutoDock. It reinforces numerous functions for scoring and offers adaptable binding ligand management. AutoDock is a favourite because of high level of accuracy and ability to adapt. Educational institutions as well as research establishments regularly make tremendous use of it. A variant that is quicker as well as reliable is Auto Dock Vina.

2.5.2. Pyrx

PyRx is a free docking software alongside a minimalist graphical user interface. It use AutoDock Vina for docking-related simulations. The PyRx procedure is well-known for its user-friendly design and graphics features. It is often associated in academic settings for online evaluation. PyRx backs a wide range of types of files and docking-based coding protocols.

2.5.3. GOLD

GOLD is a well-known docking technology application that the can handle an array of ligand substances. It offers a range of functions for scoring and facilitates peptide-ligand docking procedures. GOLD is known for its exactness and dependable nature. It is often associated in both industry and academia. GOLD'S genetic algorithm approach is highly valued.

2.5.4. Glide

Glide is a professional docking tool with great accuracy and rapidity. This tool was made by schrodinger. It was released in 2004 by the part of Maestro software suite. It has grown scoring skills and adaptable ligand governance. Glide is commonly used in the pharmaceutical sector for online screenings. It functions very well in conjunction with other Schrodinger-based techniques. Glide's exactness defines a benchmark for docking ability. It offers three accuracy levels, that are-

HTVS mode for fast screening of huge libraries.

SP mode used for docking jobs

XP adds penalties for bad contacts and strong anchor are rewarded [12].

2.5.5. Swiss Dock

The computational tool SwissDock was came in 2011. It came out of the Swiss Drug Design project. The original version of this swiss dock was free to use. No account was needed. It runs on the EADock DSS docking engine. The tool handles setup on its own. It prepares both the ligand and the protein automatically. All calculations happen on a remote server. The user's computer does not need strong specs. Results show up on a web page. The page is interactive. Ligand poses can be viewed in 3D on the protein surface. Since 2011, SwissDock has been widely used. More than 530,000 people have accessed it. These users come from roughly 200 countries. Over 710,000 docking jobs have been run. Usage increased a lot during the SARS-CoV-2 pandemic. By January 2024, Clarivate recorded that about 1,200 papers had cited the tool SwissDock [13].

Beside this there are other tools too which are highly beneficial.

S.No	Tools	Logo	Definition & Purpose	Key Feature	Scoring Function	Input Formats	Output Formats	License	Platform
1	AutoDock 4		Versatile open-source docking program.	LGA algorithm Flexible ligands	Semi-empirical	PDBQT	DLG, PDBQT	Open Source	Windows/Linux /macOS
2	AutoDock Vina		Successor optimized for speed.	Multi-core Fast scoring	Hybrid scoring	PDBQT	PDBQT logs	Open Source	Windows/Linux /macOS
3	Glide		HTVS and lead optimization.	SP & XP modes	GlideScore	PDB, MAE	MAE, SDF	Commercial	Linux/Windows
4	GOLD		Genetic algorithm docking.	Consensus scoring	ChemPLP	PDB, MOL2	MOL2, SDF	Commercial	Windows/Linux
5	SwissDock		Web-based docking service.	Easy setup	CHARMM-based	PDB, MOL2	PDF, CSV	Free Academic	Cloud/Web
6	PyRx		GUI for AutoDock /Vina.	No coding	AutoDock/Vina	PDB, SDF	CSV, SDF	Open Source	Windows/Linux /macOS
7	Discovery Studio		Integrated docking suite.	3D visualization	CDOCKER	PDB, MOL2	PDB, DSY	Commercial	Windows/Linux

TABLE1. Molecular Docking Tools

2.6. Formats of files for the Molecular Docking system

File formats for receptors and ligands become specific based on the docking programs used. The file format provides a standardised way to represent ligands and receptor proteins on a molecular level. This ensures that, not with standing the software utilities, the different types of molecular docking software are coordinated. Some of the file formats that are being used are-

MOL2: MOL2 files contain detailed information about a molecule's structure. This includes the 3D coordinates of each atom, the type of each atom, bond details, and partial charge distribution.

SDF: A structured data file format for storing molecular data, like chemical structures and properties.

PDB: Protein Data Bank format,a standard for storing proteins 3D structures and other biomolecules.

PDBQT: An extension of PDB, adding partial charges and atom types, often used with Auto Dock.

XYZ: A simple format for storing Cartesian coordinates of atoms in a molecule.

Key differences for docking-

Ligands - usually SDF, MOL2 or PDBQT

Proteins - PDB for structure , then converted into PDBQT for docking.

Visualisation – XYZ is common in visualization [14].

PROTEIN FORMATS

FORMAT + EXTENSION	USED FOR + DESCRIPTION
PDB (.pdb)	Protein 3D structure – Stores atomic coordinates and residues
PDBQT (.pdbqt)	Protein docking – Contains charges and torsions
MOL2 (.mol2)	Protein structure – Supports atom types and partial charges

LIGAND FORMATS

FORMAT + EXTENSION	USED FOR + DESCRIPTION
SDF (.sdf)	Multiple ligands – Stores multiple molecules with coordinates
MOL (.mol)	Single ligand – 2D/3D ligand structure
MOL2 (.mol2)	Ligand structure – Supports charges and docking preparation
PDB (.pdb)	Ligand 3D – 3D coordinates of ligand
PDBQT (.pdbqt)	Docking ligand – Charges and torsions for docking
SMILES (.smi/.txt)	Text format – Line notation of molecules

GRID FORMATS

FORMAT + EXTENSION	USED FOR + DESCRIPTION
GPF (.gpf)	Grid parameters – Defines docking grid center and size
CONF/TXT (.txt/.conf)	Docking settings – Configuration file for docking

DOCKING OUTPUT FORMATS

FORMAT + EXTENSION	USED FOR + DESCRIPTION
PDBQT (.pdbqt)	Docked poses – Protein-ligand docking output
PDB (.pdb)	Docked complex – Used for visualization and analysis
DLG (.dlg)	Docking log – Stores docking energies and logs
CSV/TXT (.csv/.txt)	Results – Scores, ranking and analysis

VISUALIZATION / OTHER FORMATS

FORMAT + EXTENSION	USED FOR + DESCRIPTION
DSV (.dsv)	Discovery Studio – Saved workspace and views
PNG/JPG (.png/.jpg)	Images – Figures and interaction diagrams

Fig1. Molecular Docking – File Formats

2.7. Molecular Docking Pipeline

An in-depth examination of a typical structure-centered medication design and biological docking pipeline procedure. Here's a brief synopsis and description of each phase for clarity:

2.7.1. Preparation phase –

Target Protein Preparation:

We obtained the protein's 3D crystal structure from data bank of protein that is PDB. Then Remove the water molecules and ligands, and add missing atoms followed by assigning the correct protonation states to side chains, and then minimizing the energy with forcefields like OPLS3.
Ligand Target Protein Preparation: Get the protein's atomic-resolution or tertiary crystal structure from PDB. Remove solvent or hydration molecules and chemical entities, then add elementary particles or missing atoms, assign the correct protonation states to side chains, and minimise energy with force fields like OPLS3 or AMBER.

Ligand Initialization:

Create 3D structures for the molecules, refine their configurations with force fields such as UFF or MMFF94x, and assign suitable ionisation states and tautomer forms based on physiologically relevant pH.

Grid setup:

To specify interaction site, create a grid map that is focused on the active pocket, which is typically centred on residues that interact with the co-crystallized ligand.

2.7.2. Docking Process-

Molecular Docking Simulation:

Place ligands in the active site, allowing flexible ligand docking (receptor rigid/semi-flexible)

Scoring and Ranking:

Rank poses using binding free energy (ΔGbind) via scoring functions (e.g., GlideScore, AutoDock Vina)

Interaction Analysis:Identify key amino acid interactions (H-bonds, hydrophobic contacts) using PyMOL or Discovery StudioValidation

Re-docking:
Re-dock native co-crystallized ligand; aim for RMSD < 2.0 Å to confirm protocol accuracy and precision.

2.7.3.ADME/Pharmacokinetics Process-

Pharmacological profile assessment or examination is an important step of event in the progression of development of drugs and its related process.

Key Parameters Evaluated:

Absorption: Lipinski Rule of 5

Distribution: Blood-brain barrier (BBB) permeabilityTopological polar surface area (TPSA)

Metabolism: Dampening of Cytochrome P450 biochemical catalysts (e.g., CYP2D6)

Excretion: Renal clearance or P-glycoprotein substrate behavior.

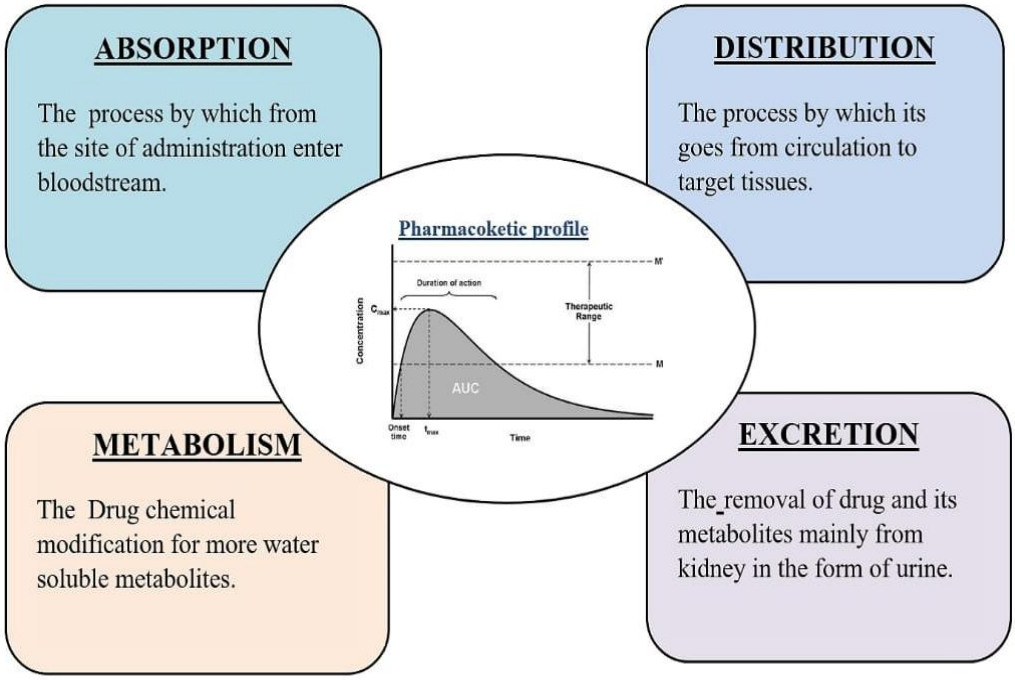


Fig 2. Pharmacokinetic Profile

2.7.4. Post-Processing & Selection

Molecular Dynamics (MD) Simulation which is Optional but Recommended.MD Simulation is a program which is utilsed for the analysis of molecular interactions. Perform MD simulations (e.g., 50 ns) using NPT ensembles to confirm the robustness of the protein–binding partner adduct over time.

Candidate Selection: Prioritize compounds that show high docking scores (strong binding) and acceptable ADME profiles (drug-likeness) for in vitro/in vivo testing [15]

3. METHODOLOGY

3.1. Retrieval and Validation of Target Protein Structure

The three-dimensional co-crystallized structure of human β-secretase 1 BACE1, that was bound to inhibitor 6JJ was fetched from the public data base that is Protein Data Bank (RCSB PDB: rcsb.org) under accession code ID(PDB ID: 5HTZ). This structure was selected based on three following criteria-

1) High resolution of 1.95 Å which was ensuring accurate side-chain and ligand placement.

2) Human origin to avoid differences between species.

3) The presence of a well-defined catalytic site with a co-crystallized small-molecule inhibitor. The structure was determined by X-ray crystallography at 100 K and deposited with an R-free value of 0.234, indicating good model-to-data agreement. The asymmetric unit contains a single BACE1 chain with residues 43–454 resolved.

For docking purpose, chain A was taken as the biological unit.

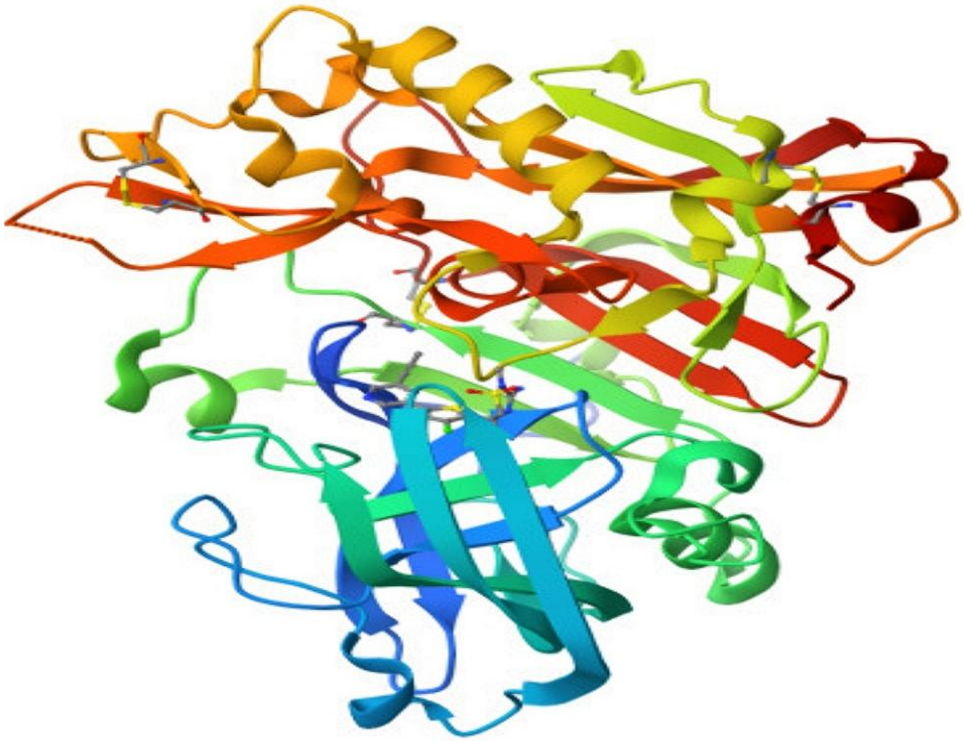


Fig 3. Three-dimensional co-crystallized structure of human β-secretase 1(BACE1)

3.2. Preparation of the BACE1 Protein

The downloaded PDB file cannot be used for docking directly. It needs cleaning. Thus the structure was unlocked in BIOVIA Discovery Studio Visualizer.

First all hydreated molecules bounded to protein was removed because the Crystal waters can block the binding site of protein. Then ions and other buffer molecules that were not required were also deleted or removed from protein configuration. It was followed by removal or deletion of the co-crystallized ligand 6JJ. The 6JJ ligand was saved separately for later use. This step made the active site empty. After cleaning, hydrogen atoms were also deleted. pH was set to 7.4 because this matches the human body. I checked the two key aspartates: Asp32 and Asp228.

One should be charged and one neutral in the active enzyme. I made sure the protonation was correct. Then I added charges to the protein. I used Kollman charges. This is the standard for AutoDock. Finally, the prepared protein was saved as a PDBQT file.

PDBQT is the format Vina needs.

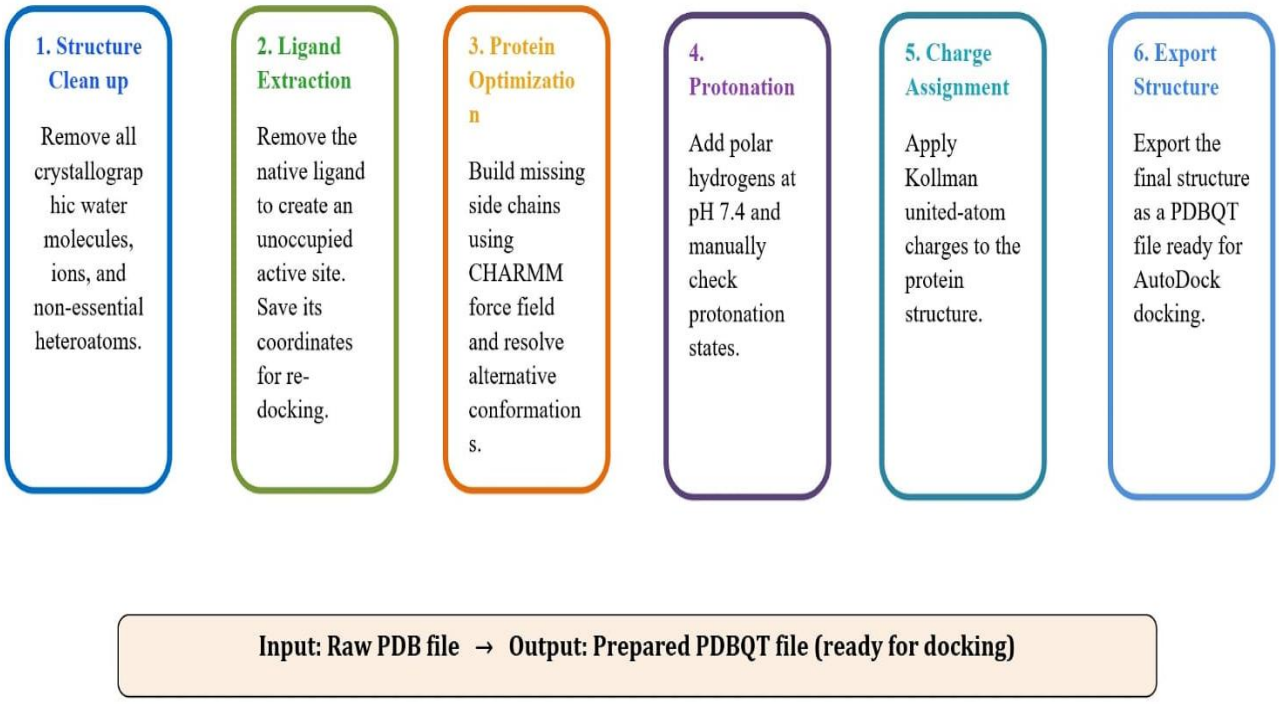


Fig 4. Protein Preparation Workflow

3.3. Selection of Ligands-

3.3.1. Reference Ligand Selection-

Several BACE1 inhibitors with experimentally validated activity has been reported in the literature. These inhibitors usually interact with the catalytic aspartate dyad (Asp32 and Asp228) at the active site. Among these, compound 6JJ from PDB ID 5HTZ has been very much studied and its binding mode is well-characterised. Therefore, 6JJ was taken or chosen as the reference ligand for validating docking protocols and as a query molecule for similarity-based screening.

3.3.2. Database Mining via Ligand-Based Virtual Screening-

To identify novel BACE1 modulators, ligand-based virtual screening was performed using the tool Swiss Similarity (swisssimilarity.ch). This web tool allows rapid screening of multiple chemical libraries by 2D and 3D molecular similarity. The SMILES notation of 6JJ was used as the query structure to screen the ZINC drug-like library. Compounds with high structural and similarity to 6JJ were retrieved.

3.3.3. ADME and Drug-Likeness Filtering:

All the hits from Swiss Similarity were subjected to pharmacokinetic and drug-likeness evaluation using Swiss ADME (swissadme.ch). The following filters were applied as essential criteria for CNS-targeted drugs-

Blood-Brain Barrier (BBB) uptake: Mandatory for Alzheimer's disease targets.

Lipinskis Rule of Five (Ro5): Including a formula weight more than 500 Da, LogP $\leq$ 5, no more than 5 hydrogen bond donors, and no more than 10 hydrogen bond acceptors.

PAINS filter: Remove Pan-Assay Interference Compounds with promiscuous activity. Brenk filter: To exclude compounds with toxic or chemically reactive fragments. CNS MPO score and GI absorption: Preferred high for oral bioavailability.

Only compounds that passed all these preferred filters and showed no violations were retained for docking. This ensured that selected ligands had or possessed favorable physicochemical and ADME properties for CNS penetration.

3.4. Preparation of Ligands

The 2D structures of all shortlisted candidate ligands, including the reference molecule 6JJ, were fetched in SDF format from the Pub Chem database. To ensure compatibility with AutoDock Vina, all ligand structures were processed using Open Babel GUI.

The procedure involved –

Minimisation of energy

Conversion of the ligands to PDBQT format.

3.5. Molecular Docking

Structure-based molecular docking was carried out using EasyDockVina 2.2. It is a user-friendly interface for AutoDock Vina that helps or facilitates docking process. The docking protocol was validated by repeat docking the bonded or bound bioactive compound that is 6JJ into the BACE1 active domain.

For virtual screening-

Prepared BACE1 PDBQT file was set as the receptor. The directory containing all prepared ligand PDBQT files was loaded. A computational grid was outlined surrounding the vicinity of the active region site that was centered at coordinates X = 25.7262, Y = 17.8873, Z = 27.6607 with dimensions 32.55 Å × 30.23 Å × 25.00 Å, ensuring complete coverage of the catalytic pocket and adjacent subpockets. Docking was run with exhaustiveness = 8 and num_modes = 9 to balance speed and accuracy. The output included individual PDBQT files for each ligand pose and a consolidated Excel spreadsheet containing binding affinity scores (kcal/mol).

More pessimistic binding energies implies more pronounced forecasted association.

Highest-ranked -scoring entities were selected for further analysis of binding interactions.

3.6. Protein-Protein Association Network Generation

To evaluate the biological relevance of BACE1 in Alzheimer's disease pathways, a protein-protein interaction (PPI) map was created using STRING knowledgebase having the version v12.0 (string-db.org). BACE1 (gene: BACE1) was used as the input query for map with *Homo sapiens* that were selected as the query organism

The following parameters were applied-

Confidence score cutoff: High confidence (0.700) was used to ensure only well-supported interactions were displayed.

Interaction category: Comprehensive string association including collectively physical and functional correlations.

Max number of interactors: 1st shell limited to 20 to focus on direct partners.

The resulting network was analyzed for enrichment in KEGG pathways and GO terms. This step confirmed BACE1's central role in the amyloidogenic pathway via interactions with APP, PSEN1, PSEN2, and APOE, thereby validating the target selection.

4. RESULTS & DISCUSSION

4.1. Protein association network analysis of BACE1 receptor and other AD responsible genes

PPI network created or generated using STRING Database shown that BACE1 is highly connected with the proteins associated within the amyoidegenic pathway which are involved in Alzheimer's disease progression making it suitable and strong target . Edges depict protein-protein associations. Total no. of edges and nodes are 109 and 21 respectively. Red coloured node represent query protein that is BACE1 and other nodes depict predicted functional partners. Known Interactions were shown in light blue and pine line while predicted in dark blue ,red and green (shown in figure 6)[16].

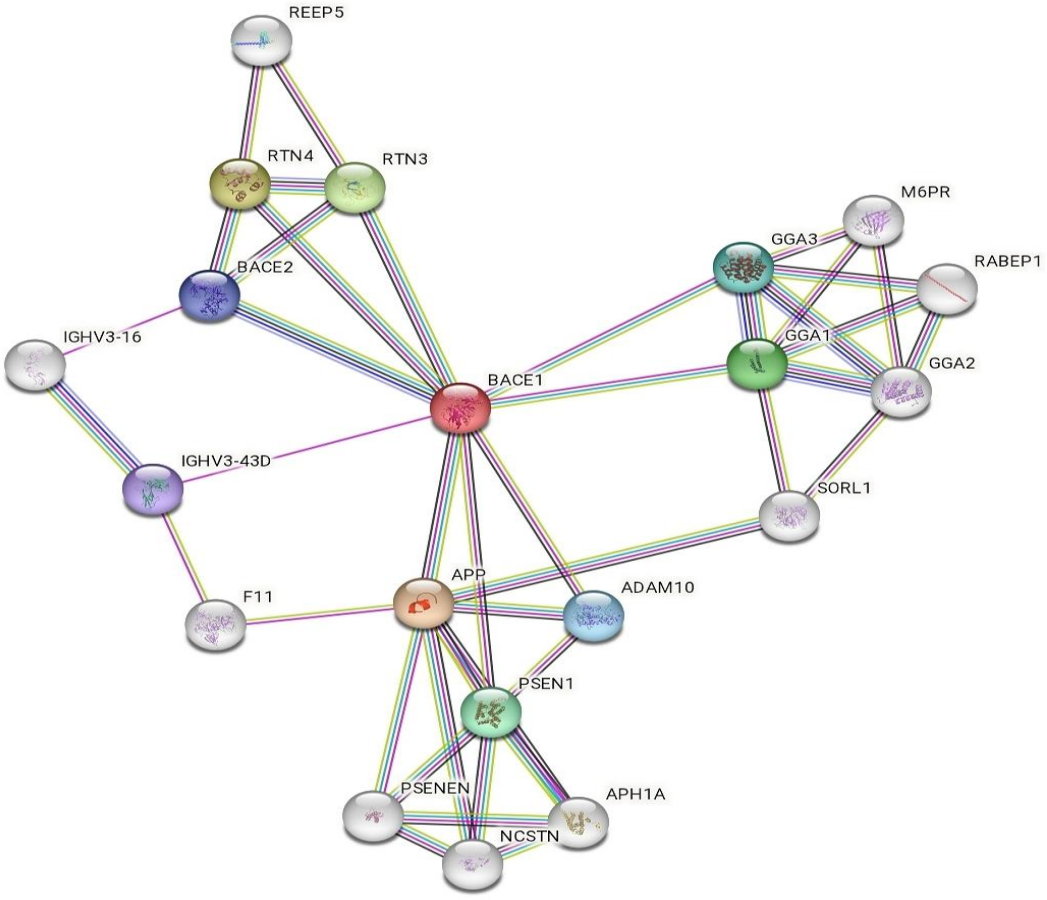


Fig 5. PPI network of BACE1 with other protein associated with AD using STRING database

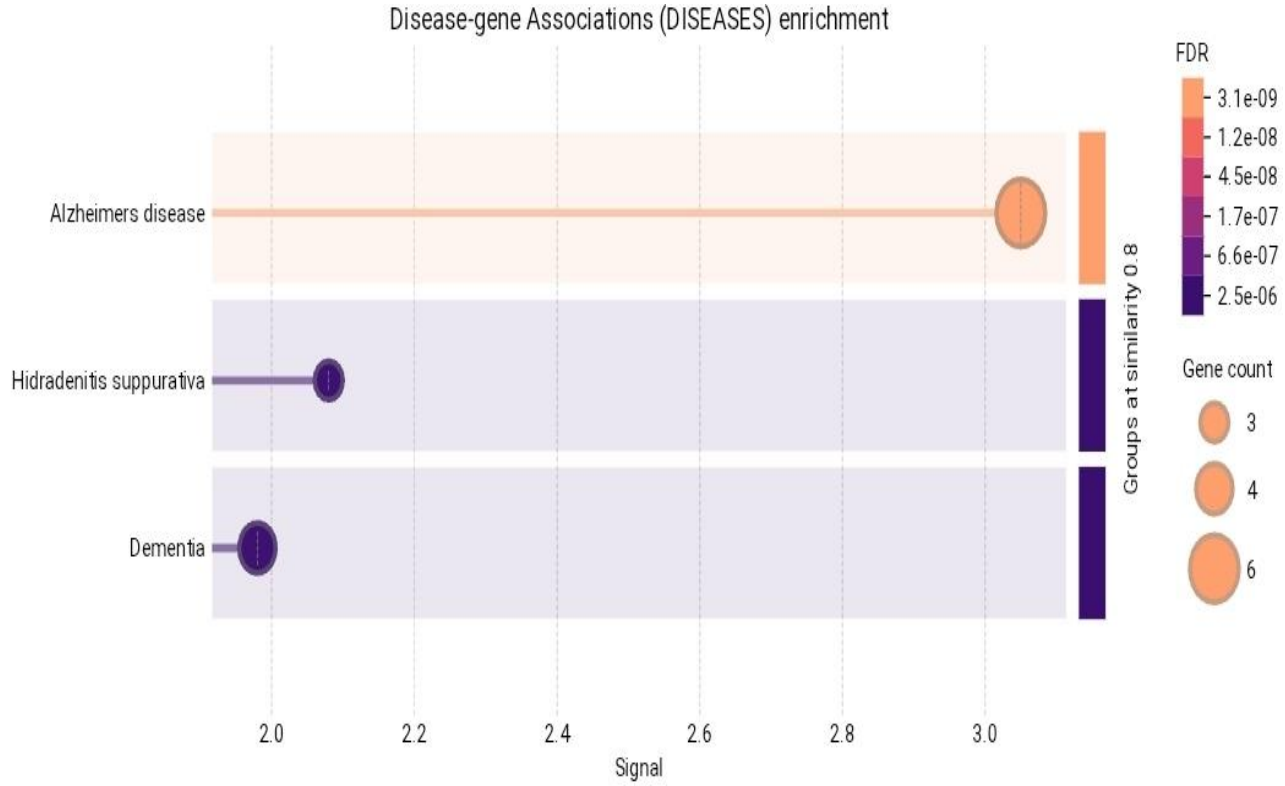


Fig 6. BACE 1 Associated Diseases

4.2. Initial screening and ADME analysis

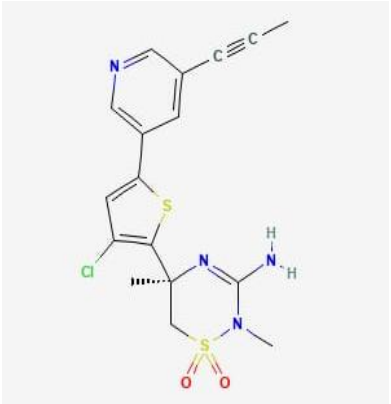
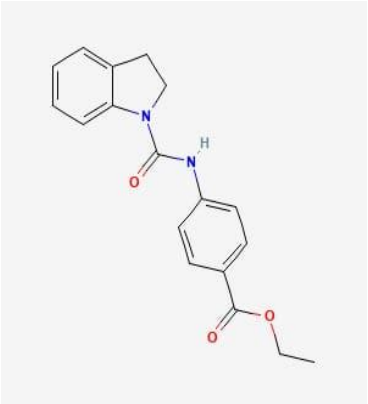
Compounds were shortlisted by Swiss Similarity on the basis of how similar they are structurally to reference molecule. As a result of this search 323 compounds were obtained with a similarity score above 0.85. These compounds were then further filtered using ADME analysis for checking BBB permeability, Lipinski's rule of five, PAINS, and Brenk alerts. This was performed using Swiss ADME. Following ADME filtering, only compounds that passed all checks were moved on to molecular docking studies. All the 81 compounds shortlisted Were BBB permeable, had zero Lipinski violations, no PAINS Alerts and no Brenk alert.

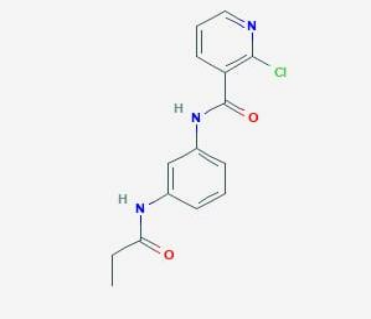
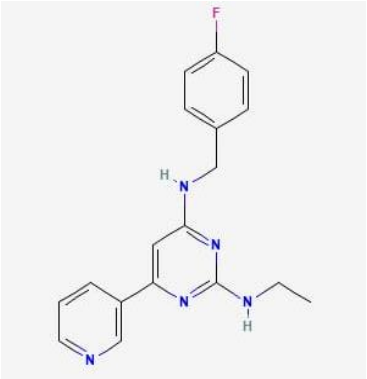
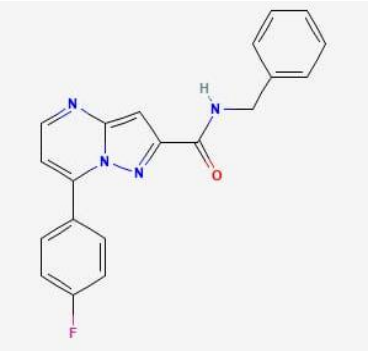
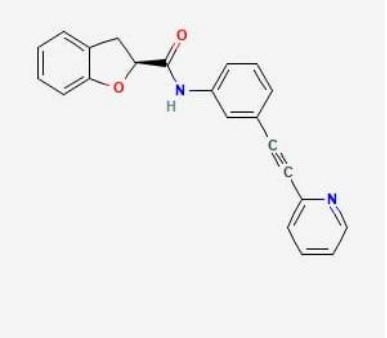
4.3. Docking studies

Molecular modeling was carried out to appraise the BACE1's binding affinity with the docked candidates. Docking analysis of all 81 compounds was carried out .This docking analysis or examination thus led to the identification of the five promising candidates with association strength exceeding that of the reference ligand that is (-8.1 kcal/mol). The association strength of these selected compounds spanned from -8.5 kcal/mol to -8.7 kcal/mol.

Notably, Compound [5] (PubChem CID 52472027) exhibited the maximum affinity of binding of about -8.7 kcal/mol, demonstrating the most favourable interaction with the BACE1 catalytic site (Table 2).

TABLE 2. BINDING AFFINITIES OF LIGANDS

Compound Name	**PubChem CID**	**2D Chemical Structure**	**Binding Affinity (kcal/mol)**
Reference	68111516		-8.1
Compound 1 ethyl 4-[(2,3-dihydro-1H-indol-1-ylcarbonyl)amino]benzoate	971721		-8.2

Compound 2 2-chloro-N-[3-(propanoylamino)phenyl]pyridine-3-carboxamide	17102411		-8.4
Compound3 2-N-ethyl-4-N-[(4-fluorophenyl)methyl]-6-pyridin-3-ylpyrimidine-2,4-diamine	124278181		-8.5
Compound 4 N-benzyl-7-(4-fluorophenyl)pyrazolo[1,5-a]pyrimidine-2-carboxamide	4866947		-8.6
Compound 5 **(**2S)-N-[3-(2-pyridin-2-ylethynyl)phenyl]-2,3-dihydro-1-benzofuran-2-carboxamide	52472027		-8.7

4.4. Detailed ADME analysis

The results of Swiss ADME analysis for these five Compounds depicted favourable observations for multiple Drug-likeliness factors. All these compounds were BBB Permeable with zero Lipinski violations rendering them useful For drug development Procedure. There were zero PAINS alerts for all five of these compounds which meant that chances of false Positives were considerably less or even zero. Zero Brenk alerts were seen, which further also Solidifieds this observation which was consistent with all the five Compounds. Further ADME evaluation utilized filters like-

TPSA value

GI absorption.
Consensus log P value
and logKp Value
which were also considered to aid in determining the Viability of the result.

TABLE 3. ADME ANALYSIS OF LIGANDS

Compound No.	BBB Permeable	Lipinski Violation	Gastrointestinal Absorption	Consensus Log Kp (cm/s)
1	Yes	0	High	-6.12
2	Yes	0	High	-6.52
3	Yes	0	High	-5.87
4	Yes	0	High	-6.15
5	Yes	0	High	-5.64

BBB Permeable = Blood-Brain Barrier permeability **Lipinski Violation** = Number of rule violations

GI Absorption = Predicted gastrointestinal absorption **Log Kp** = Skin permeation coefficient

4.5. ADME analysis of Compound 3

The ADME analysis for the molecule yields an overall favourable profile for to be a potential or promising drug candidate. The molecule satisfy Lipinski,s rule of five and have 0 violations and other drug likeness filters (Ghose,Veber,etc). It also shows good oral bioavailabilty of about 0.55 score with lipophilicity of about 3.39, which balances solubility as well as membrane permeability also. It shows high gastrointestinal absorbtion and BBB permeability thus it can reach to CNS effectively. There were no PAINS or Brenk alerts detected and it also showed positive lead likeness, thus supporting it as a potential lead candidate.

The Boiled egg representation of the molecule is shown Fig 7.

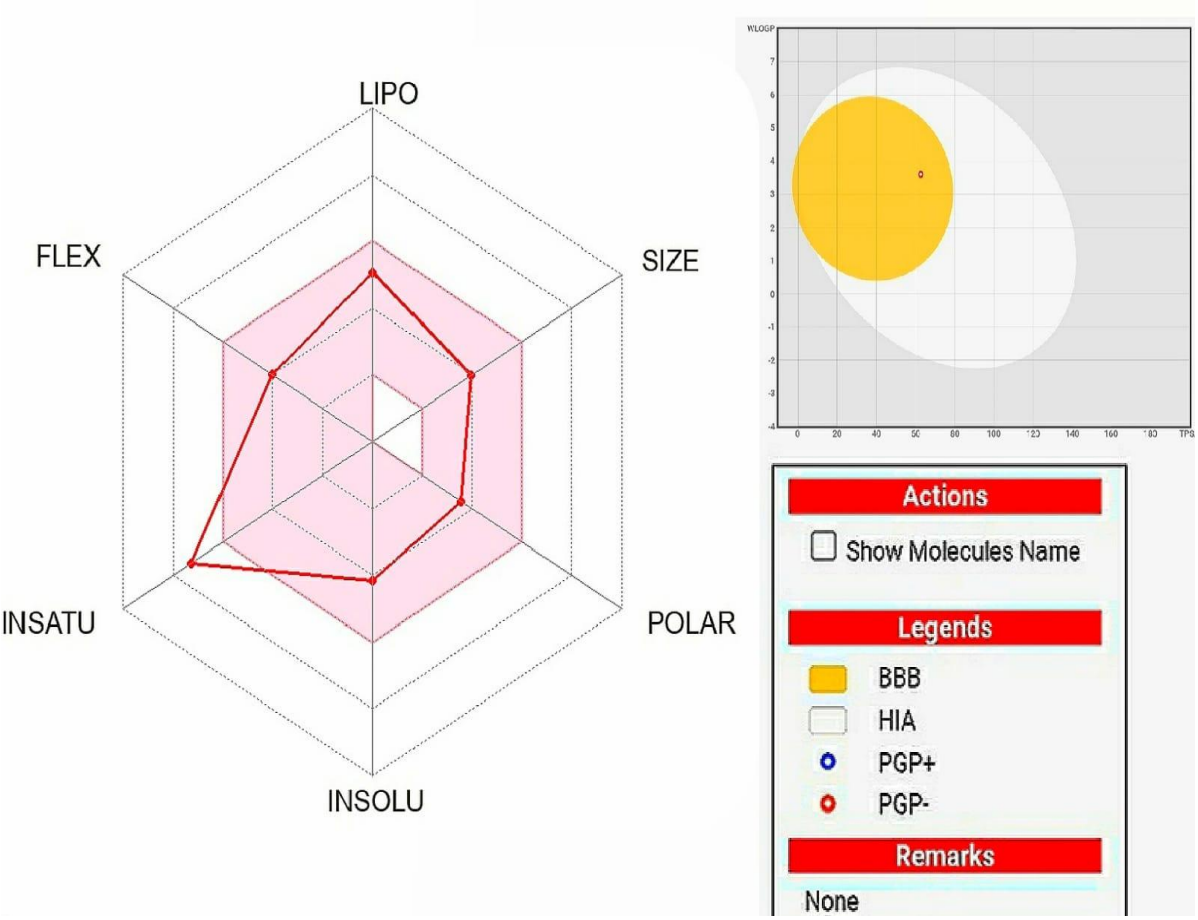


Fig7.. Boiled egg visualization for ADME profiling 2-N-ethyl-4-N-[(4-fluorophenyl) methyl]-6-pyridin-3-ylpyrimidine-2, 4-diamine

5. CONCLUSION

This study focused on finding new inhibitors for BACE1, a key protein in Alzheimer's disease. BACE1 helps create amyloid plaques in the brain. Blocking BACE1 is a promising way to slow down the disease.

We started by downloading the tertiary framework or configuration of human BACE1 from the Data Bank of proteins which had The PDB ID 5HTZ. We cleaned and prepared this protein for docking using Discovery Studio.

Next, we searched for compounds similar to the known BACE1 inhibitor 6JJ. Swiss Similarity was used for this. Over 2,500 compounds were found. Swiss ADME filtered these further. Only compounds that could cross the blood-brain barrier were kept. After applying Lipinski's rules and removing toxic compounds, 85 final ligands remained. We docked these 85 compounds to BACE1 using the docking tool that is the Easy Dock Vina.

The top hits had binding scores of -8.2 kcal/mol or better. This means they fit the BACE1 pocket well. We also checked BACE1's protein network using STRING Database. BACE1 connects to APP, PSEN1, and PSEN2 – all important proteins in the disease of Alzheimer's.

Overall, this work identified new BACE1 inhibitor candidates. These compounds are drug-like and brain-penetrant that is they can pass in brain. They are good starting points for further testing. Lab experiments or testing are needed next to check their actual effects on BACE1.

This computational approach is useful for finding Alzheimer's drug leads or drug candidates. As more BACE1 structures are solved, even better inhibitors can be designed. With further optimization, these hits could become lead compounds that can be used for therapeutics approach.

REFERENCES

[1] Zhang, Y., Chen, H., Li, R., Sterling, K., and Song, W., "Amyloid β-based therapy for Alzheimer's disease: challenges, successes and future," *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, Art. no. 225, 2020, doi: 10.1038/s41392-020-00372-7.

[2] Z. Breijyeh and R. Karaman, "Comprehensive review on Alzheimer's disease: Causes and treatment," *Molecules*, vol. 25, no. 24, p. 5789. doi: 10.3390/molecules25245789, **2020**.

[3] J. A. H. R. Claassen, "β-secretase inhibitor; a promising novel therapeutic drug in Alzheimer's disease," Front. Aging Neuroscience , vol. 6, Jul. 2014, Art. No. 165, doi: 10.3389/fnagi.2014.00165

[4] J. D. Scott et al., "Discovery of the 3-Imino-1,2,4-thiadiazinan 1,1-Dioxide Derivative Verubecestat (MK-8931)—A β-Site Amyloid Precursor Protein Cleaving Enzyme 1 Inhibitor for the Treatment of Alzheimer's Disease," J. Med. Chem., vol. 59, no. 23, pp. 10435–10450, Nov. 2016, doi: 10.1021/acs.jmedchem.6b00307.

[5] J. Carlsson and A. Luttens, "Structure-based virtual screening of vast chemical space as a starting point for drug discovery," Curr. Opin. Struct. Biol., vol. 87, Aug. 2024, Art. No. 102829, doi: 10.1016/j.sbi.2024.102829 .

[6] Lui F, Tsao JW. Alzheimer Disease. [Updated 2024 Feb 12]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from https://www.ncbi.nlm.nih.gov/books/NBK499922/

[7] S. J. Miller, R. Logan, C. Zhang, and B. P. Hafler, "The amyloid-beta wave hypothesis of Alzheimer's disease," *Frontiers in Cellular and Infection Microbiology*, vol. 15, p. 1723095. doi: 10.3389/fcimb.2025.1723095, 2025.

[8] S. L. Cole and R. Vassar, "The Alzheimer's disease β-secretase enzyme, BACE1," *Molecular Neurodegeneration*, vol. 2, p. 22. doi: 10.1186/1750-1326-2-22, 2007.

[9] F. H. Bazzari and A. H. Bazzari, "BACE1 inhibitors for Alzheimer's disease: The past, present and any future?" *Molecules*, vol. 27, no. 24, p. 8823. doi: 10.3390/molecules27248823, 2022.

[10] Y. Chang, B. A. Hawkins, J. J. Du, P. W. Groundwater, D. E. Hibbs, and F. Lai, "A guide to in silico drug design," *Pharmaceutics*, vol. 15, no. 1, p. 49. doi:10.3390/pharmaceutics15010049, 2022.

[11] Meng, X. Y., Zhang, H. X., Mezei, M., and Cui, M., "Molecular docking: a powerful approach for structure-based drug discovery," *Current Computer-Aided Drug Design*, vol. 7, no. 2, pp. 146–157, Jun. 2011, doi: 10.2174/157340911795677602.

[12] P. C. Agu, C. A. Afiukwa, U. O. Orji, E. M. Ezeh, I. H. Ofoke, C. O. Ogbu, E. I. Ugwuja, and P. M. Aja, "Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management," *Scientific Reports*, vol. 13, no. 1, p. 13398. doi: 10.1038/s41598-023-40160-2, 2023.

[13] Khan, S. U., Ahemad, N., Chuah, L. H., and Naidu, R., “Illustrated step by step protocol to perform molecular docking: Human estrogen receptor complex with 4-hydroxytamoxifen as a case study,” *Progress in Drug Discovery & Biomedical Science*, vol. 3, no. 1, 2020, doi: 10.36877/pddbs.a0000054.

[14] Krishnankutty, G., “Molecular Docking: Principle, Steps, Types, Tools, Models, Uses,” Dec. 22, 2025. [Online]. Available: https://microbenotes.com/molecular-docking-principle-steps-types-tools-models-uses/

[15] Mohan, A., Rendine, N., Mohammed, M. K. S., *et al*., “Structure-based virtual screening, *in silico* docking, ADME properties prediction and molecular dynamics studies for the identification of potential inhibitors against SARS-CoV-2 Mpro,” *Molecular Diversity*, vol. 26, pp. 1645–1661, Jun. 2022, doi: 10.1007/s11030-021-10298-0.

[16] D. Szklarczyk et al., "STRING v12.0: protein–protein association networks, with increased coverage and integration," Nucleic Acids Res., vol. 51, no. D1, pp. D408–D414, Jan. 2023, doi: 10.1093/nar/gkaa1074.

List of Publications

1. Conference Paper

Title of paper - *"Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach"*

Author Names – Sanjana Yadav and Prof. Pravir Kumar

Name of the Conference - Second International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies CE2CT-2026 (IEEE Conference Record #68422)

Date of Conference - 02–04 July 2026 at Graphic Era Hill University Bhimtal, Nainital, Uttarakhand,India

Indexing - IEEE

Status of Paper – Accepted

Date of Acceptance – 24th April ,2026

Acceptance Notification of Paper ID "1812" for International Conference “CE2CT-2026" (IEEE Record #68422) Inbox

Microsoft CMT 24 Apr

to me, ce2ct2026

Dear Authors,

Congratulations !!!

On behalf of the Technical Program Committee (TPC) of CE2CT-2026, we are pleased to inform you that your paper, Paper ID: 1812, entitled “Computational Identification of Novel BACE1 modulators for Alzheimer’s Disease: A Structure-driven Drug Discovery approach”, has been ACCEPTED as a Regular Paper for PHYSICAL PRESENTATION at the Second International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (IEEE Conference Record #68422), scheduled to be held from 02-04 July 2026 at Bhimtal, Nainital, Uttarakhand, India. All accepted and presented papers will be submitted for inclusion into IEEE Xplore digital library subject to meeting IEEE Xplore’s scope and quality requirements.

Computational Identification of Novel BACE1 modulators for Alzheimer's Disease: A Structure-driven Drug Discovery approach

Sanjana Yadav
Dept. of Biotechnology
Molecular Neuroscience and Functional Genomics Laboratory
Delhi Technological University
Delhi 110042- India
sanjanayadavbtl@gmail.com

Pravir Kumar
Dept. name of Biotechnology
Molecular Neuroscience andFunctional Genomics Laboratory
Delhi Technological University
Delhi 110042- India
pravirkumar@dtu.ac.in

Abstract— **Amyloid beta deposition in the neural tissue , which leads to neural detrioration and escalating cognitive deterioration are major signatures of Alzheimer disease, a chronic neurodegenerative disorder. Since BACE1 sparks the amyloidogenic pathway fueling Amyloid beta generation and plaque augmentation, inhibiting BACE 1 is a promising approach to curb its accumulation in AD. This study used a computational approach where 5HTZ a co crystallised molecular structure which showcases a potent inhibitor bound or complexed at the catalytic pocket or binding site of the BACE1 was utilized as reference structure to find promising BACE 1 inhibitors. The primary objective was to shortlist better suited compounds for BACE 1 inhibition and thus Alzheimer treatment to that was formely found. Here, a large number of compounds were Shortlisted on the basis of structural similarity to the reference molecule which Were filtered by ADME analysis. A carefully chosen compound library was docked into the catalytic pocket of BACE1 after protein preparation and active-site mapping in order to assess interaction potency and interaction stability. Top hits having strong interactions with key residues were profiled for ADME, including drug –likeness , GI absorption ,BBB permeability and CNS relevant characteristics. Numerous compounds demonstrated advantageous pharmacokinetic profiles and binding energies, making them prime candidates for pharmacological and biochemical verification. These findings supports the use of computational techniques to accelerate the discovery of novel BACE1 modulators for the disease. Further they also highlight promising lead candidates to carry out additional trial and confirmation.**

Keywords—**Alzheimers disease, BACE1, amyloid beta generation, virtual screening, ADME analysis, catalytic pocket, active site mapping, pharmacokinetic profiles , computational strategies.**

I. INTRODUCTION

Alzheimer's disease (AD) , an advancing brain deterioration or degradation , is attributed or associated with intellectual problems (like perception, understanding , learning) ,recollection or reminiscence impairment issues moreover changes in behavioural as well as psychological domains of life which are sufficiently enough to substantially disrupt the instrumental or functional activities of daily life .The pathological trajectory typically occurs slowly and cause significant problems in different areas of thinking. These include episodic memories of the individual's life, managerial functions that help in making decisions, problem solving as well as mental flexibilty and also it affects communication skills. The aggregation of amyloid-β (Aβ) peptides within the brain, that results in the formation of extracellular plaques, represents a core pathological hallmark of Alzheimer disease which leads to neuronal dysfunction and progressive cell death [1]. The production of Aβ peptides occurs through proteolytic processing of amyloid precursor protein (APP) that occurs through a process called as the amyloidogenic pathway [2]. Beta-secretase 1 (BACE1) initiates the first enzymatic event in the amyloidogenic cascade of APP, which results in production of amyloid-beta (Aβ) and formation of characteristic extracellular plaques[3]. BACE1, which is a proteolytic enzyme and belongs to a member of aspartic protease uses an aspartate residue within its catalytic mechanism. The enzyme contains domains. The extracellular domain of the enzymes contains two crucial catalytic aspartate residues (residue 93-96 and 289-292) that are essential for its proteolytic activity. The catalytic domain is optimally oriented that allows it to effectively cleave APP at the β-cleavage site, starting the amyloidogenic cascade that leads to Alzheimer's disease pathogenesis . Since Increased BACE1 activity is been linked with enhanced amyloid deposition and progression of Alzheimer's disease, it thus makes BACE1 an important therapeutic target [3].

Several small-molecule inhibitors have been engineered to inhibit BACE1 enzymatic activity one such is (3e,5s)-5-{3-Chloro-5-[5-(Prop-1-Yn-1-Yl)pyridin-3-Yl]thiophen-2-Yl}-2,5-Dimethyl-1,2,4-Thiadiazinan-3-Imine 1,1-Dioxide (6JJ). It is one of the most thoroughly researched and effective small – molecule inhibitor that have been developed to block BACE1 enzymatic activity . The protein data bank's structural datasets provide essential knowledge about protein – ligand interaction, underpinning evidence-based rational drug design methodologies [4].

Structure-based drug discovery, particularly through computational approaches, enables rapid screening of compounds and accurate prediction of protein-ligand interactions, thereby facilitating the efficient identification of lead compounds for therapeutic development while incurring substantially reduced costs and time investments relative to conventional experimental methodologies [5].

This study used a structure-based computational approach to identify potential inhibitors of BACE1 using the Co crystallised structure 5HTZ as a reference model. Selected compounds were evaluated through molecular docking and ADME analysis to identify molecules with promising binding characteristics and pharmacokinetic properties. Here, screening based on structure similarity, ADME analysis, protein-ligand docking Were performed.

II. METHODOLOGY

A. *Retrieval of Target protein*

The three-dimensional co-crystallised structure was obtained. We got this structure from the global archive for 3D biological macromolecular structures (https://www.rcsb.org/) using (PDB ID: 5HTZ). This particular compound structure on PDB was resolved via X-ray diffraction at 1.95Å .

B. *Preparation of target protein*

The protein structure was prepared prior to docking using discovery studio by removing crystallographic water molecules and other non-essential heteroatoms and appending Polar hydrogen and charges along with removing the native ligand from the structure . The prepared protein file was then stored in PDBQT format [6] .

C. *Ligands selection*

•*Reference Selection:* Several BACE1 inhibitors with documented efficacy have been identified in literature, targeting the catalytic aspartate residues within the active site pocket. Among these, the one discussed above that is 6JJ has been extensively studied and co-crystallized with BACE1, making it a suitable reference molecule for computational screening. Consequently, it was selected as the reference ligand for this investigation.

•*Database Mining:* Swiss Similarity (http://www.swisssimilarity.ch/) is a web-based platform designed for virtual screening of multiple small-molecule libraries through ligand-based similarity searches. This tool helps in high-throughput screening (HTS) of chemical databases. In this study, Swiss Similarity was used to search the ZINC drug library in which structure-based similarity analysis was done with reference molecule which served as the query molecule in SMILES format [7].

•*Ligand Selection Based on ADME Analysis*: The abbreviation ADME which encompasses Absorption, Distribution, Metabolism, and Excretion are pharmacokinetic parameters that are essential in pharmaceutical research and drug development. When finding potential lead compounds many numbers of essential criteria must be evaluated. These includes CNS penetration, Lipinski rule (Ro5) , PAINS (Pan-Assay Interference Compounds), and Brenk filters. These parameters serve as standard threshold requirements or basic checklist that candidate ligands must satisfy to proceed to molecular docking studies [8]. Swiss ADME (http://www.swissadme.ch/) is a computational tool used for predicting pharmacokinetic properties of small molecules. In this study swiss adme was used to check if compounds could cross blood - brain barrier , followed lipinski rule of five and other essential filters [9] . Compounds that passed all these criteria were selected and moved on to molecular docking simulations.

D. *Preparation of ligands*

The 3D molecular structures of all shortlisted candidate ligands along with reference molecule were saved in sdf file type after sourcing from pubchem. All obtained molecular models, along with reference ligand and candidate compounds, were then converted into PDBQT format by OpenBabel GUI to ensure their compatibility with molecular docking software or docking analysis.

E. *Molecular docking*

Molecular docking simulations were performed using EasyDockVina 2.2 It is computational tool that provides a simplified interface to dock multiple ligands against a single protein target. For docking simulations, the directory containing the BACE1 protein file was specified, followed by the directory containing all ligand files. A search space along with centeroid (X = 25.7262, Y = 17.8873, Z = 27.6607) was created with dimensions of 32.5476 Å × 30.2322 Å × 25.0000 Å to ensure adequate coverage. The docking procedure generated individual PDBQT output files for each ligand, along with a comprehensive Excel spreadsheet with binding affinities scores This binding affinity scores helps in comparison of binding potential of compounds to target protein indicating their effectiveness.

F. *Protein-Protein association network generation*

Computational assessment of protein network was also carried out to support BACE1's biological role and associations with Alzheimer's disease. This network was constructed using STRING Database by putting BACE1 as target input and *Homo sapiens* as target organism and then further analysed.

II. RESULT AND DISCUSSION

A. *Protein association network analysis of BACE1 receptor and other AD responsible genes*

PPI network created or generated using STRING Database shown that BACE1 is highly connected with the proteins associated within the amyoidegenic pathway which are involved in Alzheimer's disease progression making it suitable and strong target . Edges depict protein-protein associations. Total no. of edges and nodes are 109 and 21 respectively. Red coloured node represent query protein that is BACE1 and other nodes depict predicted functional partners. Known Interactions were shown in light blue and pine line while predicted in dark blue ,red and green (shown in figure 1)[10].

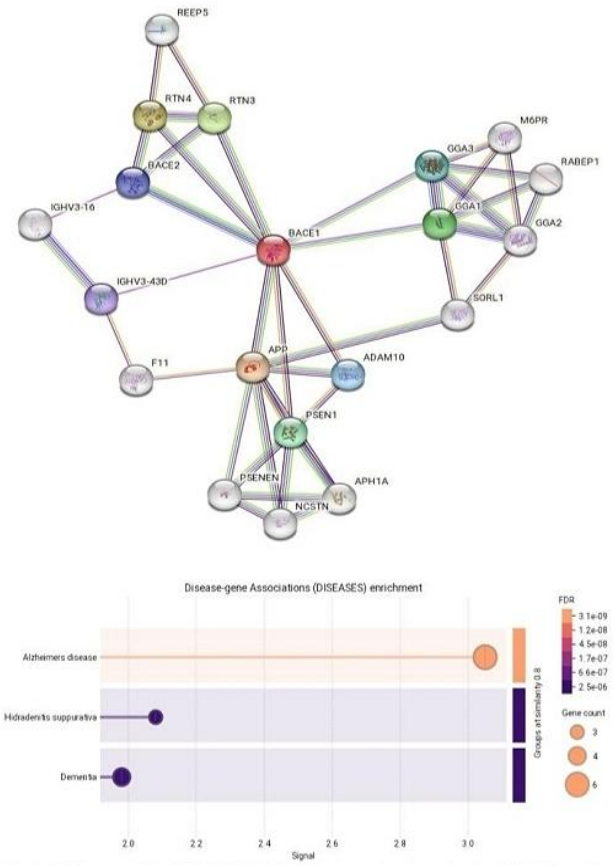


Fig1. PPI network of BACE1 with other protein associated with AD using STRING database

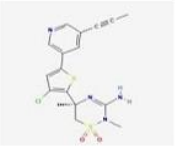
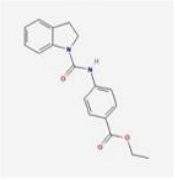
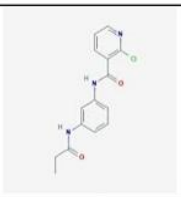
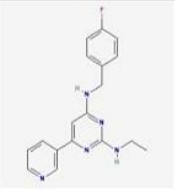
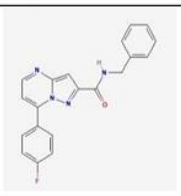
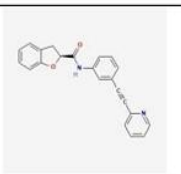
B. *Initial screening and ADME analysis*

Compounds were shortlisted by Swiss Similarity on the basis of how similar they are structurally to reference molecule. As a result of this search 323 compounds were obtained with a similarity score above 0.85. These compounds were then further filtered using ADME analysis for checking BBB permeability, Lipinski's rule of five, PAINS, and Brenk alerts. This was performed using Swiss ADME. Following ADME filtering, only compounds that passed all checks were moved on to molecular docking studies. All the 81 compounds shortlisted Were BBB permeable, had zero Lipinski violations, no PAINS Alerts and no Brenk alert.

C. *Docking studies*

Molecular modeling was carried out to appraise the BACE1's binding affinity with the docked candidates. Docking analysis of all 81 compounds was carried out .This

TABLE I . BINDING AFFINITIES OF LIGANDS

Compound Name	**PubChem CID**	**2D Chemical Structure**	**Binding Affinity (kcal/mol)**
Reference	68111516		-8.1
Compound 1 ethyl 4-[(2,3-dihydro-1H-indol-1-ylcarbonyl)amino]benzoate	971721		-8.2
Compound 2 2-chloro-N-[3-(propanoylamino)phenyl]pyridine-3-carboxamide	17102411		-8.4
Compound3 2-N-ethyl-4-N-[(4-fluorophenyl)methyl]-6-pyridin-3-ylpyrimidine-2,4-diamine	124278181		-8.5
Compound 4 N-benzyl-7-(4-fluorophenyl)pyrazolo[1,5-a]pyrimidine-2-carboxamide	4866947		-8.6
Compound 5 (2S)-N-[3-(2-pyridin-2-ylethynyl)phenyl]-2,3-dihydro-1-benzofuran-2-carboxamide	52472027		-8.7

docking analysis or examination thus led to the identification of the five promising candidates with binding affinities exceeding that of the reference ligand that is (-8.1 kcal/mol). The binding affinities of these selected compounds spanned from -8.5 kcal/mol to -8.7 kcal/mol. Notably, Compound [5] (PubChem CID 52472027) exhibited the maximum affinity of binding of about -8.7 kcal/mol, demonstrating the most favourable interaction with the BACE1 catalytic site (Table I).

D. Detailed ADME analysis

The results of Swiss ADME analysis for these five Compounds depicted favourable observations for multiple Drug-likeliness factors. All these compounds were BBB Permeable with zero Lipinski violations rendering them useful For drug development. There were zero PAINS alerts for all Seven of these compounds which meant that chances of false Positives were considerably less. Zero Brenk alerts also Solidifieds this observation which was consistent with all five Compounds. Further ADME evaluation utilized filters like TPSA value, GI absorption. Consensus log P value and logKp Value which were also considered to aid in determining the Viability of the result.

TABLE II. ADME ANALYSIS OF LIGANDS

Compound No.	BBB Permeable	Lipinski Violation	Gastrointestinal Absorption	Consensus Log Kp
1	Yes	0	High	-6.12
2	Yes	0	High	-6.52
3	Yes	0	High	-5.87
4	Yes	0	High	-6.15
5	Yes	0	High	-5.64

E. ADME analysis of Compound 3

The ADME analysis for the molecule yields an overall favourable profile for to be a potential or promising drug candidate. The molecule satisfy Lipinski,s rule of five and have 0 violations and other drug likeness filters (Ghose,Veber,etc). It also shows good oral bioavailabilty of about 0.55 score with lipophilicity of about 3.39, which balances solubility as well as membrane permeability also. It shows high gastrointestinal absorbtion and BBB permeability thus it can reach to CNS effectively. There were no PAINS or Brenk alerts detected and it also showed positive lead likeness, thus supporting it as a potential lead candidate. The Boiled egg representation of the molecule is shown Fig 2.

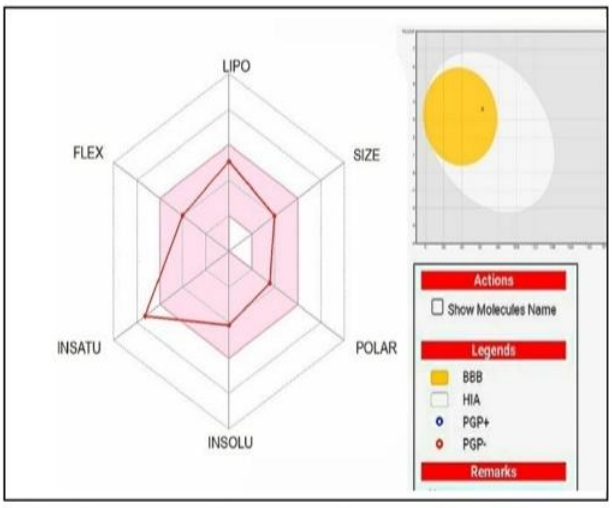


Fig2. Boiled egg visualization for ADME profiling 2-N-ethyl-4-N-[(4-fluorophenyl) methyl]-6-pyridin-3-ylpyrimidine-2, 4-diamine

ACKNOWLEDGMENT

We sincerely extend our gratitude to the senior management Of Delhi Technological University for their constant support.

REFERENCES

[1] R. Tenchov, J. M. Sasso, and Q. A. Zhou, "Alzheimer's Disease: Exploring the Landscape of Cognitive Decline," ACS Chem. Neurosci., vol. 15, no. 21, pp. 3800–3827, Oct. 2024, doi: 10.1021/acschemneuro.4c00339.

[2] E. Portelius et al., "A novel pathway for amyloid precursor protein processing," Neurobiol. Aging, vol. 32, no. 6, pp. 1090–1098, Junasli. 2011, doi: 10.1016/j.neurobiolaging.2009.06.002.

[3] J. A. H. R. Claassen, "β-secretase inhibitor; a promising novel therapeutic drug in Alzheimer's disease," Front. Aging Neurosci., vol. 6, Jul. 2014, Art. No. 165, doi: 10.3389/fnagi.2014.00165

[4] J. D. Scott et al., "Discovery of the 3-Imino-1,2,4-thiadiazinan 1,1-Dioxide Derivative Verubecestat (MK-8931)—A β-Site Amyloid Precursor Protein Cleaving Enzyme 1 Inhibitor for the Treatment of Alzheimer's Disease," J. Med. Chem., vol. 59, no. 23, pp. 10435–10450, Nov. 2016, doi: 10.1021/acs.jmedchem.6b00307.

[5] J. Carlsson and A. Luttens, "Structure-based virtual screening of vast chemical space as a starting point for drug discovery," Curr. Opin. Struct. Biol., vol. 87, Aug. 2024, Art. No. 102829, doi: 10.1016/j.sbi.2024.102829 .

[6] H. Al-Helo, S. Desai, T. Varma, P. Garg, and A. Shahiwala, —Molecular docking and dynamics of natural compounds targeting iron and erythropoietic pathways in β-thalassemia,‖ In Silico Research in Biomedicine, vol. 2, 2026, Art. no. 100220, doi: 10.1016/j.insi.2026.100220.

[7] Bragina, ME., Daina, A., Perez, MAS., Michielin, O. & Zoete, V. SwissSimilarity 2021 Web Tool: Novel Chemical Libraries and Additional Methods for an Enhanced Ligand-Based Virtual Screening Experience., Int. J. Mol. Sci, 2022, 23(2), 811.

[8] M. Ghatesse, H. Charoute, A. Raoufi, Y. Tabouktirt, I. N. Irahal, F. Hmimid, M. Kabine, N. Bourhim, and Y. Zouheir, —In silico discovery of GES-5 inhibitors via molecular docking, molecular dynamics simulation studies, and ADMET prediction,‖ Computational Biology and Chemistry, vol. 123, p. 108975, 2026, doi: 10.1016/j.compbiolchem.2026.108975.

[9] A. Daina, O. Michielin, and V. Zoete, "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry

friendliness of small molecules," Scientific Reports, vol. 7, 2017, Art. no. 42717.

[10] D. Szklarczyk et al., "STRING v12.0: protein–protein association networks, with increased coverage and integration," Nucleic Acids Res., vol. 51, no. D1, pp. D408–D414, Jan. 2023, doi: 10.1093/nar/gkaa1074.

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis "**Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach"** Total Pages **51** Name of the Scholar **Sanjana Yadav** (24/MSCBIO/33).

Supervisor:

Prof. Pravir Kumar

Department of Biotechnology

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: **Turnitin** Similarity Index: **4% ,** Total Word Count: 7968

Date:

Candidate's Signature **Signature of Supervisor**

Pravir Kumar

THESIS

Document Details

Submission ID
trn:oid:::27535:137209733

Submission Date
Apr 30, 2026, 11:17 AM GMT+0

Download Date
Apr 30, 2026, 11:29 AM GMT+0

File Name
THESIS.pdf

File Size
800.6 KB

23 Pages

4,922 Words

27,763 Characters

4% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Small Matches (less than 8 words)

Match Groups

17 Not Cited or Quoted 4%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

3% Internet sources
2% Publications
2% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

SANJANA YADAV

Contact sanjanayadavbtl@gmail.com / 7983674816

EDUCATION

Master of Science in Biotechnology
Delhi Technological University, Delhi
2024 – 2026

Bachelor of Science in Life Science
Ramjas College (University of Delhi)
2021 – 2024

Examinations Qualified

GATE BT
ICAR PG
CUET PG

CAREER OBJECTIVE

Motivated Biotechnology Master's candidate driven by scientific discovery. Qualified through national-level competitive exams with proven proficiency in laboratory techniques like cloning, cell culturing , chromatography, electrophoresis, microscopy, mass spectrometry, and data analysis with knowledge in silico tools like discovery studio , autodock, pyrx, swiss ADME, pymol, openbabel,etc. Eager to contribute to cutting-edge research or biotech innovation while expanding technical expertise. Comitted to continous learning and applying scientific knowledge to translate research findings into impactful therapeutic solutions.

EXPERIENCE

Role Name : Intern

Under the Project by ***MINISTRY OF FISHERIES, ANIMAL HUSBANDRY &DAIRYING***

Gained hands-on experience with a wide range of laboratory equipment

Successfully conducted and completed research work as part of the project under national initiative.

PROJECT /PAPER NAME:

"EFFECT OF SOLVENT ON EXTRACTION OF POLAR &NON POLAR COMPONENET ON COW DUNG SAMPLES FROM NAINITAL TO IMPROVE BIOGAS &BIOPOLYMER PRODUCTION"

Role name : Dissertation

Title: "**Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach"**

PUBLICATION

---

1. *Conference Paper:*

Sanjana Yadav and P. Kumar, *"Computational Identification of Novel BACE1 Modulators for Alzheimer's Disease: A Structure Driven Drug Discovery approach",* 2026 Second International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (IEEE Conference Record #68422), scheduled to be held from 02–04 July 2026 at Bhimtal, Nainital, Uttarakhand, India . Keywords: Alzheimers disease, BACE1, amyloid beta generation, virtual screening, ADME analysis, catalytic pocket, active site mapping, pharmacokinetic profiles, computational strategies.

CERTIFICATION & WEBINARS

Attended webinar: "Feeding the 10 Billion: A Genomics Perspective"

Certified in Artificial Intelligence in Biotechnology

Certificate Course in English Communication

Certified in Calligraphy

SKILLS & EXTRA CURRICULAR ACTIVITIES

- Proficient in Microsoft Word and Microsoft Excel

- Skilled in handling laboratory equipment and managing scientific procedures

- Proficient in PCR, Gel Electrophoresis, Spectrophotometry, ELISA

- Proficient with protein purification techniques like brafoed assay ,cell culturing techniques and its related assays.

- Familiar with DNA/RNA Extraction, SDS-PAGE, Western Blotting

- Laboratory Equipment Handling*: Centrifuge, pH Meter, Laminar Air Flow, Autoclave

- Molecular Docking & Computational Tools: AutoDockVina, PyRx, SwissADME, Discovery Studio, PyMOL

- In-silico Analysis : ADMET prediction, Drug-likeness screening (Lipinski, Ghose, Veber), Protein-ligand interaction analysis

- Bioinformatics Tools: BLAST, Chimera, ChemDraw

- Knowledge about advanced biological computing techniques such as multiple sequence alignment its exhaustive (Progressive , Iterative and Block based approaches) and heuristic approach , FASTA AND BLAST techniques its algorithm, double sequence alignment approach including PAM and BLOSUM, soft computation tachniques like fuzzy logic ,SVM, Machine learning and its types, Neural networks, genetic algorithm , all about pharmacogenomics, etc

- Strong communication and teamwork skills

- Adaptable, organized, and detail-oriented

LEADERSHIP

Have strong leadership skills .

Led a team during internship

Recipient of “Student of the Year” award

LANGUAGES

•English – Fluent

•Hindi – Fluent

•French – Basic

Undertaking

I am taking full responsibilities for this project work and the thesis written for my dissertation. This thesis is neither copied from any place or written using AI/ChatGPT assisted manner. In case of any discrepancies or falsification of information, I understand that my degree may be cancelled.

I will take responsibilities of the dissertation. If something found wrong ,take an appropriate action.

Name of Student:

Roll No.:

Signature: