

Computational, Machine Learning and Multi Target Directed Ligand (MTDL) Approaches to Identify Novel Leads for Alzheimer's Disease



Thesis submitted in partial fulfilment for the
Award of Degree

Doctor of Philosophy

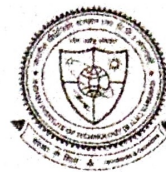
By

Powsali Ghosh, M. Pharm

DEPARTMENT OF PHARMACEUTICAL ENGINEERING & TECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY
(BANARAS HINDU UNIVERSITY)
VARANASI – 221 005

Roll No. 21161503

Year: 2025



CERTIFICATE

It is certified that the work contained in the thesis titled “**Computational, Machine Learning and Multi Target Directed Ligand (MTDL) Approaches to Identify Novel Leads for Alzheimer’s Disease**” by **Ms. Powsali Ghosh** 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.

Date: 26.6.25

Place: IIT (BHU), Varanasi

Prof. Sushil K. Singh
(Supervisor)

सुशाल कुमार सिंह / Sushil K. Singh
आचार्य / Professor
मेडिकली अभियंत्रिकी 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



DECLARATION BY THE CANDIDATE

I, **Powsali Ghosh**, 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 **Prof. Sushil K. Singh** from **December, 2021 to June, 2025** 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: 26.06.2025
Place: IIT (BHU), Varanasi

Powsali Ghosh
Powsali Ghosh

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.

Sushil K. Singh
Prof. Sushil K. Singh
(Supervisor)
सुशील कुमार सिंह / Sushil K. Singh
आचार्य / Professor
भेषजकीय 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

Sairam Krishnamurthy
Prof. Sairam Krishnamurthy
(Head of the Department)
विभागाध्यक्ष / Head
भेषजकीय 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



COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: Computational, Machine Learning and Multi Target Directed Ligand (MTDL) Approaches to Identify Novel Leads for Alzheimer's Disease.

Candidate's Name: Ms. Powsali Ghosh

Copyright Transfer

The undersigned hereby assigns to the Indian Institute of Technology (Banaras Hindu University), Varanasi, all rights under copyright that may exist in and for the above thesis submitted for the award of the Ph.D. degree.

Date: 26.06.2025

Place: IIT (BHU), Varanasi


Powsali Ghosh

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and University's copyright notice are indicated.

Acknowledgements

*At the very outset, I express my utmost respect and heartfelt reverence to **Bharat Ratna Pandit Madan Mohan Malaviya Ji** and the esteemed **Late Prof. M.L. Schroff** for their visionary and humanitarian efforts in establishing this distinguished temple of learning, which has been instrumental in nurturing generations of scholars and fostering academic excellence.*

*I wish to convey my deepest sense of gratitude to my supervisor, **Prof. Sushil Kumar Singh**, for his constant support, insightful guidance, and unwavering belief in my capabilities throughout my doctoral journey. His mentorship, constructive criticism, and invaluable advice have been a source of inspiration and have played a pivotal role in shaping both my academic and personal growth.*

*I sincerely thank **Dr. Ashok Kumar** for his valuable guidance and constructive suggestions throughout my Ph.D.*

*I am profoundly thankful to **Prof. Sairam Krishnamurthy**, Head, Department of Pharmaceutical Engineering & Technology, IIT (BHU) Varanasi, for providing the essential resources, research facilities, and a stimulating environment conducive to the successful completion of my work. I am especially grateful to him and his research group for their generous support and cooperation in facilitating the in vivo studies that formed an integral part of my doctoral research. His encouragement and support at every stage of my research have been deeply appreciated.*

*I also express my sincere gratitude to the former Heads of the Department, **Prof. S.K. Shrivastava** and **Prof. (Mrs.) S. Hemalatha**, for their constant support, motivation, and guidance, which greatly contributed to the smooth execution of my research activities.*

*I am sincerely grateful to all the esteemed faculty members of the department, **Prof. B. Mishra, Prof. Senthil Raja A, Prof. M.S. Muthu, Dr. Alakh N. Sahu, Dr. Ashish K. Agrawal, Dr. G.P. Modi, Dr. Ruchi Chawla, Dr. S.K. Jain, Dr. Vinod Tiwari, Dr. P.K. Nayak, Dr. S.K. Mishra, Dr. Rajnish, Dr. Deepak Kumar, Dr. Dinesh Kumar, Dr. Jairam Meena, and Dr. Arun Khatri**, for their constant encouragement, insightful discussions, and kind support, which have been invaluable throughout my academic journey.*

*I gratefully acknowledge the constructive feedback and valuable suggestions provided by my Research Progress Evaluation Committee members, **Dr. Shreyans Kumar Jain and Prof. Abha Mishra**, whose inputs have greatly helped in refining and strengthening my research work.*

*It is my pleasure to express my heartfelt appreciation to my friends and colleagues, **Sagorika Nag, Sanheeta, Pallavi, Pooja, Soki, Rasha, and Nilesh**, for their friendship, enriching discussions, and the constant moral support that made this journey less challenging and more fulfilling.*

*I also extend my sincere gratitude to my seniors and lab members, **Dr. Ravi Singh, Dr. Swetha Rayala, Dr. Ankit Ganeshpurkar, Dr. Gopichand Gutti, and Dr. Devendra Kumar**, for their invaluable guidance, technical support, and insightful suggestions, which contributed significantly to my research. Additionally, I wish to acknowledge the constant encouragement and support from my seniors, **Dr. Rangan Mitra, Mr. T.A. Gajendra, Dr. Anurag T.K. Baidya, Dr. Shreyasi Majumdar, Dr. Digambar Waikar, Dr. Aiswarya Chaudhuri, and Dr. Samarpita Das**, whose academic discussions, constructive feedback, and moral support have been instrumental in my academic journey.*

I gratefully acknowledge the PARAM Shivay Supercomputing Facility, the Centre for Computing and Information Services (CCIS), and the Central Instrument Facility (CIF) of IIT (BHU) for providing the necessary infrastructure, advanced instrumentation, and technical support, which were critical to the successful execution of my research.

*I am grateful to **IIT (BHU)** for providing me with a Teaching Assistantship, which not only supported me financially but also enriched my academic experience.*

*I am thankful to all the non-teaching staff of the department for their administrative assistance and technical support. In particular, I extend my gratitude to **Mr. Yashwant Singh and Mr. Anand**, whose constant help and willingness to assist in every possible way greatly contributed to my research activities.*

*I also wish to acknowledge the invaluable contribution of **laboratory animals** used during my research. Their role in advancing scientific knowledge cannot be overstated, and I express my deepest gratitude and respect for their sacrifice in the interest of pharmaceutical research.*

Above all, I am eternally indebted to my beloved parents and sister for their unconditional love, patience, and unwavering belief in me. Their constant encouragement, sacrifices, and support have been the foundation of my perseverance and success. I would also like to express my heartfelt gratitude to my husband, Lokesh, whose presence, understanding, and unwavering support have been a source of immense strength and motivation throughout this journey.

To all those who, in one way or another, contributed to the successful completion of my doctoral work, your support is deeply appreciated, and I remain ever grateful.

Date:

Place: IIT (BHU), Varanasi



(Powsali Ghosh)

Table of Contents

List of Tables.....	xvii
List of Figures.....	xix
List of Abbreviations.....	xxvii
List of Symbols.....	xxxix
Preface.....	xxxiii

Chapter 1

1. Introduction

1.1. Alzheimer's disease (AD).....	1
1.1.1. Pathophysiology of AD.....	2
1.1.2. Pathology of AD.....	6
1.1.3. Stages of AD.....	7
1.1.4. Biomarkers of AD.....	9
1.1.5. Diagnosis of AD.....	10
1.1.6. Management of AD.....	12
1.2. Computational chemistry.....	14
1.2.1. Computer-aided drug design (CADD).....	16
1.2.2. Integration of big data, AI, and ML in advanced CADD.....	20
1.3. Machine learning: approaches, algorithms, and validation.....	21
1.3.1. Approaches to ML.....	21
1.3.2. Machine learning algorithms.....	23
1.3.3. Validation of machine learning models.....	23
1.3.4. Performance metrics.....	25
1.4. Multi-target-directed ligands (MTDLs).....	27
1.4.1. Molecular hybridization in MTDLs.....	27
1.4.2. Lead optimization in MTDLs.....	28

Chapter 2

2. Literature review

2.1. CADD in drug discovery.....	29
---	-----------

2.1.1. SBDD.....	29
2.1.2. LBDD.....	30
2.1.3. Integrated SBDD-LBDD approaches.....	31
2.2. Machine learning in drug discovery.....	32
2.2.1. Applications of supervised learning in drug discovery and development.....	34
2.2.2. Applications of unsupervised learning in drug discovery.....	39
2.3. MTDLs for AD.....	41
2.3.1. Molecular hybridization-based MTDLs	41
2.3.2. Lead optimization-based MTDLs.....	43
2.4. Sulphonamides.....	44
2.4.1. Sulfonamides as anti-Alzheimer agents.....	45

Chapter 3

3. Objectives and plan of work

3.1. Objectives	49
3.2. Plan of Work.....	50
3.2.1. Computational Identification of Potential DAPK1 Inhibitors for Targeting Neuroinflammation in Alzheimer's Disease.....	50
3.2.2. Computational Screening of Coumarin Derivatives as NLRP3 Inflammasome Inhibitors for Alzheimer's Disease.....	50
3.2.3. Development of Machine Learning-based KNIME Workflow to Predict COX-2 Inhibitors.....	51
3.2.4. Development of Machine Learning-based KNIME Workflow to Predict Selective COX-2 Inhibitors.....	52
3.2.5. Development of multi-targeted Biphenyl Sulfonamide derivatives for the treatment of Alzheimer's Disease.....	52
3.2.6. Development of multi-targeted N-phenylbenzenesulfonamide derivatives for the treatment of Alzheimer's Disease.....	53

Chapter 4

4. Computational Identification of Potential DAPK1 Inhibitors for Targeting Neuroinflammation in Alzheimer's Disease

4.1. Introduction.....	55
------------------------	----

4.2. Materials & methods.....	57
4.2.1. Pharmacophoric feature mapping and pharmacophore generation and validation.....	57
4.2.2. Validation of ligand-based pharmacophore.....	58
4.2.3. Pharmacophore-based virtual screening of a database.....	58
4.2.4. Drug-likeness and pan-assay interference compounds (PAINS) filtering.....	59
4.2.5. Duplicate removal based on InChI.....	61
4.2.6. Protein preparation.....	61
4.2.7. Ligand preparation for high-throughput virtual screening and docking studies.....	62
4.2.8. Grid generation and validation.....	62
4.2.9. HTVS and docking studies.....	63
4.2.10. <i>In silico</i> ADMET prediction.....	63
4.2.11. Molecular dynamics simulation study.....	64
4.3. Result and discussion.....	65
4.3.1. Pharmacophore modelling and validation	65
4.3.2. Drug-likeness filter.....	66
4.3.3. Duplicate removal based on InChI.....	67
4.3.4. Homology modelling and validation.....	67
4.3.5. HTVS and docking studies.....	68
4.3.6. <i>In silico</i> ADME predictions.....	71
4.3.7. <i>In silico</i> toxicity predictions.....	73
4.3.8. Molecular dynamics simulation study.....	73
4.4. Conclusion.....	82

Chapter 5

5. Computational Screening of Coumarin Derivatives as NLRP3 Inflammasome Inhibitors for Alzheimer's Disease	
5.1. Introduction.....	85
5.2. Materials and methods.....	88

5.2.1. Dataset preparation (library construction & filtering)	88
5.2.2. Drug-likeness and PAINS filtering.....	89
5.2.3. Protein structure selection and preparation.....	90
5.2.4. Docking-based high-throughput virtual screening (HTVS)	91
5.2.5. Molecular docking studies.....	92
5.2.6. <i>In silico</i> BBB permeability and ADMET prediction.....	93
5.2.7. MD simulations.....	93
5.2.8. Binding free energy calculations.....	94
5.3. Results and discussion.....	95
5.3.1. HTVS and molecular docking studies.....	95
5.3.2. <i>In silico</i> BBB permeability and ADMET prediction.....	99
5.3.3. Docking-based protein-ligand interaction studies.....	101
5.3.4. MD simulation study.....	104
5.4. Conclusion.....	115

Chapter 6A

6.A. Development of Machine Learning-based KNIME Workflow to Predict COX-2 Inhibitors

6A.1. Introduction.....	117
6A.2. Materials and Methods.....	121
6A.2.1. Computational details.....	121
6A.2.2. Database of COX-2 inhibitors.....	121
6A.2.3. Data curation.....	121
6A.2.4. Compound designation as actives and inactives	122
6A.2.5. Dataset splitting and distribution analysis.....	123
6A.2.6. Molecular descriptors and fingerprints calculation.....	123
6A.2.7. Feature engineering and feature selection.....	124
6A.2.8. Individual classification model generation.....	124
6A.2.9. Ensemble model generation.....	130
6A.2.10. Validation of models' performance.....	132
6A.2.11. Applicability domain.....	134
6A.2.12. Clustering of COX-2 inhibitor dataset.....	135

6A.3. Results and discussion.....	136
6A.3.1. Data curation.....	136
6A.3.2. Active and inactive designation.....	136
6A.3.3. Chemical space exploration and dataset distribution.....	137
6A.3.4. Feature selection.....	138
6A.3.5. Data partitioning and distribution.....	141
6A.3.6. Data classification and model performance evaluation.....	142
6A.3.7. Clustering of COX-2 inhibitor dataset.....	150
6A.3.8. Automated prediction of COX-2 inhibitors using KNIME workflow.....	154
6A.3.9. Comparison with previous studies.....	156
6A.3.10. Key messages and insights.....	158
6A.3.11. Data and software availability.....	159
6A.4. Conclusion.....	159

Chapter 6B

6B. Development of Machine Learning-based KNIME Workflow to Predict Selective COX-2 Inhibitors

6B.1 Introduction.....	161
6B.2. Materials and methods.....	164
6B.2.1. Computational details.....	164
6B.2.2. Database of COX-1 inhibitors.....	164
6B.2.3. Data curation.....	165
6B.2.4. Compound designation as actives and inactive	165
6B.2.5. Dataset splitting and distribution analysis.....	165
6B.2.6. Molecular descriptors and fingerprints calculation.....	165
6B.2.7. Feature engineering and feature selection.....	165
6B.2.8. Individual classification model generation.....	165
6B.2.9. Ensemble model generation.....	166
6B.2.10. Validation of models' performance.....	166
6B.2.11. Applicability domain.....	166
6B.2.12. Clustering of COX-1 inhibitor dataset.....	166
6B.3. Results and discussion.....	166

6B.3.1. Data curation.....	166
6B.3.2. Active and inactive designation.....	167
6B.3.3. Chemical space exploration and dataset distribution.....	167
6B.3.4. Feature selection.....	168
6B.3.5. Data partitioning and distribution.....	171
6B.3.6. Data classification and model performance evaluation.....	172
6B.3.7. Clustering of COX-1 inhibitor dataset.....	177
6B.3.8. Automated prediction of selective COX-2 inhibitors using KNIME workflow.....	181
6B.4. Conclusion.....	183

Chapter 7

7. Development of Multi-targeted Biphenyl Sulfonamide Derivatives for the treatment of AD

7.1. Introduction.....	185
7.2. Rationale of drug design.....	187
7.3. Materials and methods.....	191
7.3.1. Chemistry.....	191
7.3.2. <i>In silico</i> studies.....	192
7.3.3. <i>In vitro</i> evaluation	194
7.3.4. <i>In vivo</i> evaluation.....	201
7.4. Results and discussion.....	208
7.4.1. Chemistry.....	208
7.4.2. <i>In silico</i> studies.....	221
7.4.3. <i>In vitro</i> evaluation.....	226
7.4.4. <i>In vivo</i> evaluation.....	236
7.5. Conclusion.....	242

Chapter 8

8. Development of Multi-targeted N-phenylbenzenesulfonamide Derivatives for the treatment of Alzheimer's Disease

8.1. Introduction.....	245
8.2. Rationale of drug design.....	246

8.3. Materials and methods.....	248
8.3.1. Chemistry.....	248
8.3.2. <i>In silico</i> studies.....	249
8.3.3. <i>In vitro</i> evaluation	250
8.3.4. <i>In vivo</i> evaluation.....	252
8.4. Results and discussion.....	254
8.4.1. Chemistry.....	254
8.4.2. <i>In silico</i> studies.....	267
8.4.3. <i>In vitro</i> evaluation.....	271
8.4.4. <i>In vivo</i> evaluation.....	282
8.5. Conclusion.....	287
 Chapter 9	
9. Summary and conclusions.....	291
References.....	295
Appendix.....	321
List of Publications.....	391

List of Tables

Table 1.1	FDA approved drug treatments for AD.....	12
Table 4.1	The parameters of the Lipinski's rule of five and REOS filter.....	59
Table 4.2	Summary of docking parameters used in AutoDock 4.2.....	63
Table 4.3	The features as well as EF and GH scores of the pharmacophore models of DAPK1 inhibitor.....	66
Table 4.4	Structure validation results from MolProbity before and after Homology modelling.....	68
Table 4.5	Number of compounds subjected to HTVS and different stages of docking.....	69
Table 4.6	Third stage docking results displaying mean binding energy, cluster size percentage, and binding efficiency of ligands.....	70
Table 4.7	The results of the ADME test with SwissADME.....	72
Table 4.8	Predicted <i>in silico</i> toxicities of selected compounds.....	73
Table 4.9	Energy contributions of protein-ligand complexes in MM-PBSA assay.....	82
Table 4.10	ZINC ID, 2D structure, IUPAC name, and SMILES of identified hits.....	83
Table 5.1	Molecules subjected to HTVS and molecular docking.....	97
Table 5.2	Results of the third stage of docking displaying mean binding energy, binding efficiency, and cluster size percentage.....	98
Table 5.3	SwissADME results of the selected compounds.....	100
Table 5.4	PreADMET results of the selected compounds.....	100
Table 5.5	Hydrophobic, hydrogen, cation- π , and π - π stacking interactions and amino acid residues for the compounds with the target protein.....	102
Table 5.6	Energy contributions of protein-ligand complexes in MM-GBSA assay.....	114
Table 5.7	Energy contributions of protein-ligand complexes in MM-PBSA assay.....	114
Table 5.8	MolPort ID, 2D structure, IUPAC name, and SMILES of identified hits.....	116
Table 6.1	Description of the final molecular features selected for ML model building.....	139

Table 6.2	Number of fingerprints calculated and used for machine learning model building.....	140
Table 6.3	Performance results for the final ensemble model.....	146
Table 6.4	Key outcomes and contributions of the study.....	158
Table 6.5	Description of the final molecular features selected for ML model building.....	170
Table 6.6	Number of features calculated and used for ML model building.....	170
Table 6.7	Performance results for the final ensemble model.....	176
Table 7.1	Docking scores of designed compounds with AChE and BACE1.....	223
Table 7.2	Selected molecular and ADMET property prediction results of synthesised compounds.....	225
Table 7.3	Inhibitory potencies of compounds against <i>ee</i> AChE, <i>eq</i> BuChE and <i>h</i> BCAE1.....	227
Table 7.4	PAMPA permeability P_e (10^{-6} cm s $^{-1}$) data for selected potent compounds and their corresponding BBB Penetration prediction.....	229
Table 7.5	Propidium iodide displacement assay results.....	229
Table 7.6	Effect of Single-Dose Oral Administration of Compound 34	236
Table 8.1	Docking scores of designed compounds with AChE and BACE1.....	268
Table 8.2	Selected molecular and ADMET property prediction results of synthesized compounds.....	270
Table 8.3	Inhibitory potencies of compounds against <i>ee</i> AChE, <i>eq</i> BuChE and <i>h</i> BCAE1.....	273
Table 8.4	PAMPA permeability P_e (10^{-6} cm s $^{-1}$) data for selected potent compounds and their corresponding BBB Penetration prediction.....	274
Table 8.5	Propidium iodide displacement assay results.....	275
Table 8.6	Effect of Single-Dose Oral Administration of Compound 78	282

List of Figures

Figure 1.1	Macroscopic and microscopic features of AD.....	7
Figure 1.2	Clinical stages of AD based on the degree of cognitive impairment.....	8
Figure 1.3	FDA approved drugs for the treatment of AD.....	12
Figure 4.1	(a) DAPK1 kinase domain in complex with a small molecule inhibitor (PDB ID-4TXC), (b) Chemical structure of compound 38G (PDB ID-4TXC).....	57
Figure 4.2	Knime workflow for drug-likeness filtering.....	60
Figure 4.3	Knime workflow for PAINS Filtering.....	60
Figure 4.4	Selected pharmacophore models with pharmacophoric features (HA: H-bond acceptor, HY: hydrophobic, A: Aromatic, HD: H-bond donor)	65
Figure 4.5	Superimposition of co-crystallised (green) and docked 38G (lavender)	69
Figure 4.6	(a) RMSD of the protein backbone atoms of the protein-ligand complexes (ZINC000020650468 (blue), ZINC000006755051 (orange), ZINC000020648330 (grey) and the native protein (yellow). (b) RMSD of heavy atoms of the ligands (ZINC000020650468 (blue), ZINC000006755051 (orange), and ZINC000020648330 (grey)	78
Figure 4.7	Residue-wise RMSF deviation of the protein-ligand complexes ZINC000020650468 (blue), ZINC000006755051 (orange), ZINC000020648330 (grey) and the native protein (yellow).....	79
Figure 4.8	Radius of gyration of protein backbone of protein-ZINC000020650468 complex (blue), protein-ZINC000006755051 complex (orange), protein-ZINC000020648330 complex (grey), and the native protein (yellow)	77
Figure 4.9	Number of intermolecular H-Bonds between ligand and protein during the entire simulation run (a) ZINC000020650468 (b) ZINC000006755051 (c) ZINC000020648330.....	80
Figure 4.10	Radial Distribution Function graphs of the ligands and important residues (a) LYS42, (b) VAL96, (c) ASP161.....	81
Figure 5.1	(a) Schematic representation of the structural domains of the NLRP3 inflammasome.....	85
	(b) 3D cartoon depiction of the NLRP3 inflammasome with bound ADP, omitting the pyrin domain. The domains are colour-coded as their labels and the protein surface outline is shown in light grey.....	86
Figure 5.2	Distribution of the initial dataset of coumarin derivatives based on the source of the molecules.....	88

Figure 5.3	Knime workflow for drug-likeness and PAINS filtering.....	89
Figure 5.4	NLRP3 NACHT domain in complex with a small molecule inhibitor (PDB ID 7ALV) with (left) and without (right) visible protein surface.....	90
Figure 5.5	Chemical structure of compound RM5 (PDB ID 7ALV) (left) and compound MCC950 (right)	90
Figure 5.6	2D protein–ligand interaction diagram visualized by BIOVIA Discovery Studio Visualizer (left) and LigPlot+ (right)	95
Figure 5.7	Superimposition of co-crystallised (green) and redocked RM5 (lavender)	96
Figure 5.8	Alignment of co-crystal protein 7ALV (green) to homology modelled protein.....	97
Figure 5.9	BOILED-Egg plots of compounds analyzed by SwissADME.....	101
Figure 5.10	2D and 3D protein–ligand interaction diagrams of (a) MolPort-050-872-358 , (b) MolPort-050-884-068 , (c) MolPort-051-135-630 , and (d) protein– MCC950 complexes.....	103
Figure 5.11	(a) RMSD of the C α of the protein– MolPort-050-872-358 (blue), protein– MolPort-050-884-068 (red), protein– MolPort-051-135-630 (green), and protein– MCC950 (yellow) complexes. (b) RMSD of heavy atoms of the ligands MolPort-050-872-358 (blue), MolPort-050-884-068 (red), MolPort-051-135-630 (green), and MCC950 (yellow)	106
Figure 5.12	Residue-wise protein RMSF of the C α of the protein– MolPort-050-872-358 (blue), protein– MolPort-050-884-068 (red), protein– MolPort-051-135-630 (green), and protein– MCC950 (yellow) complexes.....	107
Figure 5.13	(a) Residue-wise ligand RMSF of the molecules MolPort-050-872-358 (blue), MolPort-050-884-068 (red), MolPort-051-135-630 (green), and MCC950 (yellow). Atom identifier numbers are shown for the molecular structures of (b) MolPort-050-872-358 , (c) MolPort-050-884-068 , (d) MolPort-051-135-630 , and (e) MCC950	108
Figure 5.14	The number of intermolecular H-bonds formed between ligand and protein during MD simulation for MolPort-050-872-358 (blue), MolPort-050-884-068 (red), MolPort-051-135-630 (green), and MCC950 (yellow)	109
Figure 5.15	Rg of the ligands MolPort-050-872-358 (blue), MolPort-050-884-068 (red), and MolPort-051-135-630 (green), and MCC950 (yellow)	110
Figure 5.16	Graph of solvent accessible surface areas (SASAs) of protein– MolPort-050-872-358 (blue), protein– MolPort-050-884-068 (red), protein– MolPort-051-135-630 (green), and protein– MCC950 (yellow) complexes.....	112

Figure 6.1	Structural domains of human COX-2 dimer visualized by Chimera. N-terminal epidermal growth factor (EGF)-like domain (red), membrane-binding domain (yellow), C-terminal catalytic domain (cyan), cyclooxygenase active site (magenta), and peroxidase.....	120
Figure 6.2	Distribution of molecules in training set, internal test set, and external validation set.....	123
Figure 6.3	Optimization process to find the best parameters of each basic classifier (top); the 5-fold and 10-fold cross-validation processes for the basic classifiers and the ensemble models (bottom).....	126
Figure 6.4	Schematic representation of the majority voting strategy for building ensemble models.....	131
Figure 6.5	Pair plot of COX-2 inhibitor dataset. The diagonal represents the frequency of distribution.....	137
Figure 6.6	Pearson correlation between final features and pIC ₅₀	139
Figure 6.7	Heatmap of Pearson correlation coefficient among the features after performing feature selection.....	140
Figure 6.8	Data distribution of active and inactive molecules in the complete COX-2 inhibitor dataset, training set, internal test set, and external validation set.....	141
Figure 6.9	2D t-SNE plot (left) and 3D PCA plot (right) represent the molecule distribution of training, test, and validation set.....	142
Figure 6.10	Confusion matrix of the final ensemble model for the internal test set and external validation set.....	145
Figure 6.11	Balanced accuracy, sensitivity, specificity, MCC, and F1 scores obtained from the final ensemble model for the internal test set and external validation set.....	145
Figure 6.12	ROC curves for the internal test set (left) and the external validation set (right) obtained from the descriptor-based ensembles (top), algorithm-based ensembles (middle), and final ensemble (bottom).....	148
Figure 6.13	Clustering of compounds in 10 subsets along with their centroids.....	152
Figure 6.14	Structures of compounds closest to the centroids and their corresponding activities.....	152
Figure 6.15	COX-2 Inhibitor Prediction Workflow for Virtual Screening.....	156
Figure 6.16	Structural domains of human COX-2 dimer visualized by Chimera. N-terminal epidermal growth factor (EGF)-like domain (red), membrane-binding domain (yellow), C-terminal catalytic domain (cyan), cyclooxygenase active site (magenta), and peroxidase.....	162
Figure 6.17	The dual-model strategy to prioritize selective COX-2 inhibitors for further investigation.....	164

Figure 6.18	Pair plot of COX-2 inhibitor dataset. The diagonal represents the frequency of distribution.....	168
Figure 6.19	Pearson correlation between final features and pIC ₅₀	169
Figure 6.20	Heatmap of Pearson correlation coefficient among the features after performing feature selection.....	170
Figure 6.21	Data distribution of active and inactive molecules in the complete COX-2 inhibitor dataset, training set, internal test set, and external validation set.....	171
Figure 6.22	2D t-SNE plot (left) and 3D PCA plot (right) represent the molecule distribution of training, test, and validation set.....	172
Figure 6.23	Confusion matrix of the final ensemble model for the internal test set and external validation set.....	176
Figure 6.24	Balanced accuracy, sensitivity, specificity, MCC, and F1 scores obtained from the final ensemble model for the internal test set and external validation set.....	177
Figure 6.25	ROC curves for the internal test set (left) and the external validation set (right) obtained from the final ensemble model.....	177
Figure 6.26	Clustering of compounds in 10 subsets along with their centroids.....	178
Figure 6.27	Structures of compounds closest to the centroids and their corresponding activity classes.....	179
Figure 6.28	Selective COX-2 inhibitor prediction workflow for virtual screening.....	182
Figure 7.1	Design rationale of compounds 33-57	190
Figure 7.2	Parallel artificial membrane permeation assay (PAMPA) setup.....	196
Figure 7.3	Schematic representation of the acute oral toxicity study procedure followed to evaluate the safety of the test drug.....	202
Figure 7.4	Schematic representation of the Y-maze apparatus and the assessment of spatial working memory in rats.....	204
Figure 7.5	Schematic representation of the experimental procedure for establishing the A β -induced AD phenotypic model, illustrating intraperitoneal (i.p.) injection, stereotaxic incision based on bregma and lambda landmarks, and Hamilton syringe-mediated intracerebroventricular (ICV) injection.....	206
Figure 7.6	Morris water maze test.....	207
Figure 7.7	Lineweaver Burk double reciprocal plot of compounds 34 (top) and 53 (bottom)	231
Figure 7.8	Dixon plot for compounds 34 (left) and 53 (right)	231
Figure 7.9	Effect of compound 34 and 53 on self-induced (left) and AChE-induced (right) A β ₁₋₄₂ aggregation. Bars are showing the NFI as	232

	the mean $\pm$ SEM of three separate experiments. ^a p < 0.05 vs. A β , ^b p < 0.05 vs. A β + DPZ (10:5), ^c p < 0.05 vs. A β + DPZ (10:10), ^d p < 0.05 vs. A β + DPZ (10:20), ^e p < 0.05 vs. A β + compound 34 (10:5), ^f p < 0.05 vs. A β + compound 34 (10:10), ^g p < 0.05 vs. A β + compound 34 (10:20), ^h p < 0.05 vs. A β + compound 53 (10:5), and ⁱ p < 0.05 vs. A β + compound 53 (10:10). One-way ANOVA followed by Newman - Keuls posthoc test.....	
Figure 7.10	Percent free radical scavenging (% FRS) activity of compound 34 , compound 53 , and ascorbic acid (AA). The individual plot points represent the % FRS value obtained from two independent experiments in triplicate, expressed as mean $\pm$ SEM.	233
Figure 7.11	2D interaction diagram of compound 34 with AChE (PDB ID 7E3H).....	235
Figure 7.12	3D interaction diagram of compound 34 with AChE (PDB ID 7E3H)	235
Figure 7.13	2D interaction diagram of compound 34 with BACE1 (PDB ID 6FGY)	235
Figure 7.14	3D interaction diagram of compound 34 with BACE1 (PDB ID 6FGY)	236
Figure 7.15	Effect of Compound 34 on organ toxicity: histology of kidney, liver, heart, and brain tissue of control (top) and drug-treated (bottom) rats.....	237
Figure 7.16	Estimation of learning and memory improvement effects of compound 34 (5, 10 and 20 mg/kg) in scopolamine-induced Y-maze test: % spontaneous alterations. ^a p < 0.05 vs. control; ^b p < 0.05 vs. scopolamine; ^c p < 0.05 vs DPZ (5 mg/kg); ^d p < 0.05 vs Compound 34 (5 mg/kg); ^e p < 0.05 vs Compound 34 (10 mg/kg); One-way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine]	237
Figure 7.17	Estimation of learning and memory improvement effects of compound 34 (5, 10 and 20 mg/kg) in scopolamine-induced Y-maze test: AChE activity-rate of substrate hydrolysed (left) and ACh levels (right). ^a p < 0.05 vs. control; ^b p < 0.05 vs. scopolamine; ^c p < 0.05 vs DPZ (5 mg/kg); ^d p < 0.05 vs Compound 34 (5 mg/kg); ^e p < 0.05 vs Compound 34 (10 mg/kg); One-way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine].....	238
Figure 7.18	Estimation of learning and memory improvement effects of compound 34 (5, 10 and 20 mg/kg) in scopolamine-induced Y-maze test: MDA levels (left) and CAT levels (right). ^a p < 0.05 vs. control; ^b p < 0.05 vs. scopolamine; ^c p < 0.05 vs DPZ (5 mg/kg); ^d p < 0.05 vs Compound 34 (5 mg/kg); ^e p < 0.05 vs Compound 34 (10 mg/kg); One-way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine]	239

Figure 7.19	Protective effect of compound 34 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, (C) The number of entries to platform zone during the probe trial. ^a p < 0.001 vs. control and ^b p < 0.05 vs. $A\beta_{1-42}$. One-way ANOVA followed by Newman - Keuls posthoc test.....	240
Figure 7.20	Protective effect of compound 34 and donepezil on $A\beta_{1-42}$ -induced memory deficits analyzed by the Morris water maze test: AChE activity (left) and ACh levels (right). ^a p < 0.001 vs. control and ^b p < 0.05 vs. $A\beta_{1-42}$. One-way ANOVA followed by Newman - Keuls posthoc test.....	242
Figure 7.21	Structure, IUPAC name, and biological activity of the identified lead.....	243
Figure 8.1	Design rationale of compounds 55-80	246
Figure 8.2	Lineweaver Burk double reciprocal plot of compounds 77 (top) and 78 (bottom).....	276
Figure 8.3	Dixon plot for compounds 77 (left) and 78 (right)	276
Figure 8.4	Effect of compound 77 and 78 on self-induced (left) and AChE-induced (right) $A\beta_{1-42}$ aggregation. Bars are showing the NFI as the mean $\pm$ SEM of three separate experiments ^a p < 0.05 vs. $A\beta$, ^b p < 0.05 vs. $A\beta$ + DPZ (10:5), ^c p < 0.05 vs. $A\beta$ + DPZ (10:10), ^d p < 0.05 vs. $A\beta$ + DPZ (10:20), ^e p < 0.05 vs. $A\beta$ + compound 77 (10:5), ^f p < 0.05 vs. $A\beta$ + compound 77 (10:10), ^g p < 0.05 vs. $A\beta$ + compound 77 (10:20), ^h p < 0.05 vs. $A\beta$ + compound 78 (10:5), and ⁱ p < 0.05 vs. $A\beta$ + compound 78 (10:10). One-way ANOVA followed by Newman - Keuls posthoc test.....	277
Figure 8.5	Percent free radical scavenging (% FRS) activity of compound 77 , compound 78 , and ascorbic acid (AA). The individual plot points represent the % FRS value obtained from two independent experiments in triplicate, expressed as mean $\pm$ SEM.....	278
Figure 8.6	2D interaction diagram of compound 78 with AChE (PDB ID 7E3H)	279
Figure 8.7	3D interaction diagram of compound 78 with AChE (PDB ID 7E3H)	280
Figure 8.8	2D interaction diagram of compound 78 with BACE1 (PDB ID 6FGY)	280
Figure 8.9	3D interaction diagram of compound 78 with BACE1 (PDB ID 6FGY)	280
Figure 8.10	Effect of Compound 78 on organ toxicity: histology of kidney, liver, heart, and brain tissue of control (top) and drug-treated (bottom) rats.....	282

Figure 8.11	Estimation of learning and memory improvement effects of compound 78 (5, 10 and 20 mg/kg) in scopolamine-induced Y-maze test: % spontaneous alterations. ^a P < 0.05 vs. control; ^b P < 0.05 vs. scopolamine; ^c P < 0.05 vs DPZ (5 mg/kg); ^d P < 0.05 vs Compound 78 (5 mg/kg); ^e P < 0.05 vs Compound 78 (10 mg/kg); One-way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine].....	283
Figure 8.12	Estimation of learning and memory improvement effects of compound 78 (5, 10 and 20 mg/kg) in scopolamine-induced Y-maze test: AChE activity-rate of substrate hydrolysed (left) and ACh levels (right). ^a P < 0.05 vs. control; ^b P < 0.05 vs. scopolamine; ^c P < 0.05 vs DPZ (5mg/kg); ^d P < 0.05 vs Compound 78 (5 mg/kg); ^e P < 0.05 vs Compound 78 (10mg/kg); One-way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine].....	284
Figure 8.13	Estimation of learning and memory improvement effects of compound 78 (5, 10 and 20 mg/kg) in scopolamine-induced Y-maze test: CAT levels (left) and MDA levels (right). ^a P < 0.05 vs. control; ^b P < 0.05 vs. scopolamine; ^c P < 0.05 vs DPZ (5mg/kg); ^d P < 0.05 vs Compound 78 (5 mg/kg); ^e P < 0.05 vs Compound 78 (10mg/kg); One-way ANOVA followed by Newman - Keuls posthoc test. [SCO-Scopolamine].....	285
Figure 8.14	Protective effect of compound 78 and donepezil on A β 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, (C) The number of entries to platform zone during the probe trial. ^a p < 0.001 vs. control and ^b p < 0.05 vs. A β ₁₋₄₂ . One-way ANOVA followed by Newman - Keuls posthoc test.....	286
Figure 8.15	Protective effect of compound 78 and donepezil on A β ₁₋₄₂ -induced memory deficits analyzed by the Morris water maze test: AChE activity (left) and ACh levels (right). ^a p < 0.001 vs. control and ^b p < 0.05 vs. A β ₁₋₄₂ . One-way ANOVA followed by Newman - Keuls posthoc test.....	287
Figure 8.16	Structure, IUPAC name, and biological activity of the identified lead.....	289

List of Abbreviations

Abbreviation	Full Form
3D	Three-dimensional
5-CV	5-fold cross-validation accuracy
10-CV	10-fold cross-validation accuracy
A	Balanced accuracy
A β	Amyloid beta
Ach	Acetylcholine
AChE	Acetylcholinesterase
AD	Alzheimer's disease
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
AI	Artificial intelligence
ARIA	Amyloid-related imaging abnormalities
ASCII	American standard code for information interchange
AUC	Area under the curve
BOA	Bayesian optimization algorithm
CADD	Computer-aided drug design
CDSS	Clinical decision support systems
ChAT	Choline acetyltransferase
CNS	Central nervous system
COX-1	Cyclooxygenase-1
COX-2	Cyclooxygenase-2
CSF	Cerebrospinal fluid
DA	Dopamine
DAPK1	Death-associated protein kinase 1
DCM	Dichloromethane
DL	Deep learning
DT	Decision tree
EGF	Epidermal growth factor
EF	Enrichment factor
EEG	Electroencephalography

EOAD	Early-onset Alzheimer's disease
FAD	Familial Alzheimer's disease
FPR	False positive rate
FP	False positive
FRET	Fluorescence resonance energy transfer
GANs	Generative adversarial networks
GABA	Gamma-aminobutyric acid
GH	Goodness of hit
GMQE	Global model quality estimation
GNNs	Graph neural networks
HBAs	Hydrogen bond acceptors
HBDs	Hydrogen bond donors
HTVS	High-throughput virtual screening
ICV	Intracerebroventricular
InChI	IUPAC international chemical identifier
IRLS	Iteratively reweighted least squares
LBDD	Ligand-based drug design
LE	Ligand efficiency
LGA	Lamarckian genetic algorithm
LOAD	Late-onset Alzheimer's disease
LR	Logistic regression
LRR	Leucine-rich repeat
MBD	Membrane-binding domain
MAE	Mean absolute error
MAP	Microtubule-associated protein
MCC	Matthews correlation coefficient
MD	Molecular dynamics
MDD	Major depressive disorder
MDL	Minimal description length
MM-GBSA	Molecular mechanics–Generalised Born surface area
MM-PBSA	Molecular mechanics-Poisson-Boltzmann surface area
MMSE	Mini-mental state examination

MTLDs	Multi-target-directed ligands
NA	Noradrenaline
NIA-AA	National institute on aging and Alzheimer's association
NFTs	Neurofibrillary tangles
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NOD	The nucleotide-binding oligomerization domain
OA	Overall accuracy
PAINS	Pan-assay interference compounds
PBS	Phosphate buffer saline
PI	Propidium iodide
PK/PD	Pharmacokinetic and pharmacodynamic
PR	Precision
QMEAN	Qualitative model energy analysis
QSAR	Quantitative structure-activity relationships
QSPR	Quantitative structure--property relationship
R _g	Radius of gyration
REOS	Rapid Elimination of Swill
RF	Random Forest
RC	Recall
RDF	Radial distribution function
ROC	Receiver Operating Characteristic
RoF	Rule of Five
RNNs	Recurrent neural networks
RMSD	Root-mean-square deviation
RMSF	Root mean square fluctuation
SBDD	Structure-based drug design
SCO	Scopolamine hydrobromide
SCMC	Sodium carboxymethylcellulose
SL	Supervised learning
TBA	Thiobarbituric acid
TBARS	Thiobarbituric acid reactive substance

TDP	Transactive response DNA-binding protein
TEA	triethylamine
TMS	transcranial magnetic stimulation
TN	true negative
TP	true positive
tDCS	Transcranial direct current stimulation
TPSA	topological polar surface area
UFF	universal force field
UMAP	uniform manifold approximation and projection
VAEs	variational autoencoders
WHO	World Health Organization
XGB	Extreme Gradient Boosting

List of Symbols

Symbol	Meaning
α	Alpha
β	Beta
δ	Delta
Δ	Delta
ε	Epsilon
λ	Lambda
$^{\circ}\text{C}$	Degree Celsius
$\text{\AA}$	Angstrom
©	Copyright
g	Gram
mg	Milligram
μg	Microgram
ng	Nanogram
mmol	Millimole
μM	Micromolar
nM	Nanomolar
mM	Millimolar
mL	Millilitre
μL	Microlitre
μm	Micrometer
mm	Millimetre
cm	Centimetre
mV	Millivolt
h	Hour
s	Second; Singlet
nm	Nanometre
ppm	Parts per million
rpm	Revolutions per minute
Kcal	Kilocalories
Hz	Hertz
MHz	Megahertz
J	Coupling constant
d	Doublet
t	Triplet

Symbol	Meaning
m	Multiplet
dd	Doublet of doublet
m/z	Mass-to-charge ratio
%	Percent
pH	Potential of hydrogen
$\leq$	Less than or equal
$<$	Less than
$>$	More than
$\pm$	Plus or minus