# COMPUTATIONAL INTELLIGENCE–DRIVEN DRUG DISCOVERY APPROACH FOR TARGETING TAU PROTEIN IN ALZHEIMER'S DISEASE

**Vineeta Johri**
Subject: Biotechnology
Research scholar, Department of Biotechnology, NIMS Institute of Medical Science and Technology, NIMS University, Jaipur
Email: johrivineeta06@gmail.com)

## Abstract

Alzheimer's disease (AD) is a developing neurodegenerative disease identified by abnormal aggregation of tau protein, which leads to neuronal dysfunction and cognitive decline. Identifying effective tau-targeting therapeutics through conventional experimental approaches is time-consuming, costly, and resource-intensive. In this study, artificial intelligence-enabled computational methodologies have appeared as powerful research tools for increasing knowledge generation and drug discovery. The present study proposes a computational intelligence–driven in silico framework to identify potential natural inhibitors of the human tau protein, which is related to Alzheimer's disease.

We used the high-resolution crystal structure of the human tau protein (PDB ID: 8Y76) and a selected library of bioactive natural compounds. This library was computationally screened using an automated workflow executed with Bio-Python. Artificial intelligence and Machine learning assisted algorithms were used for ligand preparation, structural optimization, and automated virtual screening, so we can minimizing human bias and enhance reproducibility. Molecular docking was performed to investigate binding affinity and stability of ligand–protein complexes, followed by detailed interaction analysis to identify key hydrogen bonds, hydrophobic interactions, and active-site residues involved in tau inhibition.

Computational intelligence–based ranking and filtering strategies were applied to prioritise lead compounds with favourable binding scores and interaction profiles. The identified leads demonstrate strong binding affinity toward critical tau functional regions, suggesting their potential role in inhibiting tau aggregation. This study highlights how AI-assisted in silico techniques can transform traditional drug discovery by enabling rapid screening, systematic analysis, and data-driven decision-making.

Overall, the proposed computational intelligence framework illustrates the growing role of artificial intelligence as a research tool in neurodegenerative disease therapeutics and provides a scalable platform for future experimental validation and precision drug discovery in Alzheimer's disease

**Keywords:** Computational intelligence, Alzheimer's disease, tau protein, molecular docking, natural compounds, Bio-Python, In-silico drug discovery

## 1. Introduction

Alzheimer's disease is a progressive and critical neurodegenerative disorder. It is a common type of dementia which affects over millions of people worldwide, and scientific predictions indicate this data will nearly triple by 2050 if no effective interventions are introduced (WHO, 2023).[1] This disorder shows progressive memory loss, cognitive deterioration, and behavioural changes. So, it places a significant socio-economic burden on patients, carers, and healthcare services. Regardless many years of research and effective treatments that modify the disease, proper treatment is still not available, and ongoing medications such as cholinesterase inhibitors and NMDA receptor antagonists mainly offer symptomatic relief without stopping the disease from advancing. This therapeutic and diagnostic gap underscores the urgent need for innovative drug-discovery approaches that target the fundamental molecular mechanisms underlying AD. A primary neuropathological feature of Alzheimer's disease (AD) is the abnormal buildup of hyperphosphorylated tau protein. This protein forms aggregates called paired helical filaments and neurofibrillary tangles within neurons. Typically, tau is a microtubule-associated protein essential for stabilising the neuronal cytoskeleton. In pathological states, hyperphosphorylation hampers its ability to bind microtubules, resulting in microtubule destabilization, disrupted axonal transport, and neuronal death. Growing evidence indicates that tau pathology is more closely linked to cognitive decline than amyloid-β accumulation, making tau a vital target for AD therapy. [2]

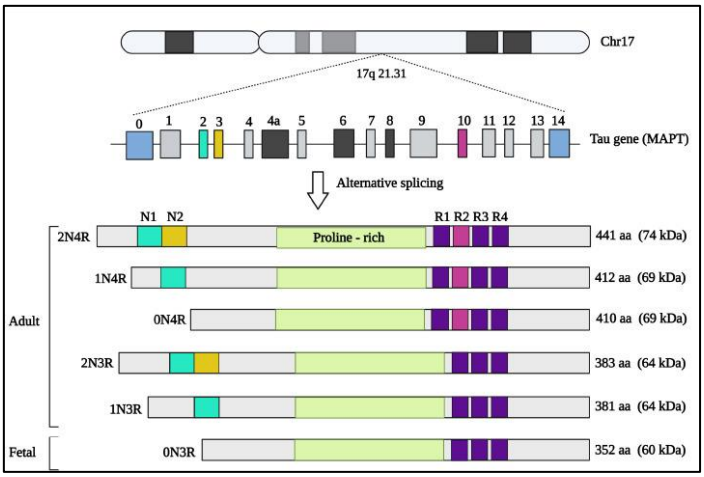


Figure 1- Tau protein gene structure

Traditional methodologies for developing tau-targeted drugs face challenges due to complex protein-protein interactions, their high clinical trial failure rates, and the significant time and cost of laboratory screening. Molecular docking and computer simulation have become an important method for accelerating early drug discovery, to predicting how small molecules bind to target proteins. This technique lowers the need for extensive in vitro testing and helps to identify promising lead compounds more efficiently.[3]

Natural compounds, such as phytochemicals, are a rich source of structurally diverse, biologically active molecules with scientifically proven neuroprotective and anti-inflammatory

properties. There are many plant-derived compounds like curcumin, resveratrol, quercetin, and epigallocatechin gallate that have demonstrated inhibitory activity against tau aggregation in experimental models. [3,4] These compounds mostly come with favourable safety profiles and may cross the blood–brain barrier, making them promising candidates for further research. However, the systematic identification and assessment of natural compounds that target the tau protein remain limited, particularly with respect to modern bioinformatics approaches and workflows. [5,6,7]

Computational intelligence integrates algorithm-driven automation, predictive modelling, and decision-support systems to reduce human bias and enhance reproducibility. In neurodegenerative disease research, these methods are particularly valuable due to the multifactorial nature of disease mechanisms and the limited availability of experimental models. The development of computational biology tools and programming-driven automation has made drug screening pipelines more efficient, reproducible, and scalable. Bio-Python is a computational biology tool that provides a flexible working pipeline for tasks such as retrieving protein structures, preparing ligands, running docking simulations, and analysing results. With the integration of Biopython and docking platforms such as CB-Dock2, AutoDock Vina. Researchers can streamline screening of large ligand datasets against specific protein targets, thereby accelerating the discovery of promising candidates. [8,9,10]

In this study, we used an in silico docking method, implemented in a BioPython-based workflow, to evaluate a library of natural compounds for their potential to inhibit human tau protein. We retrieved the three-dimensional structure of tau from the Protein Data Bank (PDB), prepared ligand datasets from plant-derived bioactive compounds, performed docking simulations using CB-Dock2 with BioPython scripting, and examined binding affinities along with protein-ligand interaction patterns. To enhance translational relevance, we further screened the top candidates for drug-likeness and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties, selecting compounds with the most favourable pharmacokinetic and safety profiles. [17,18,19]

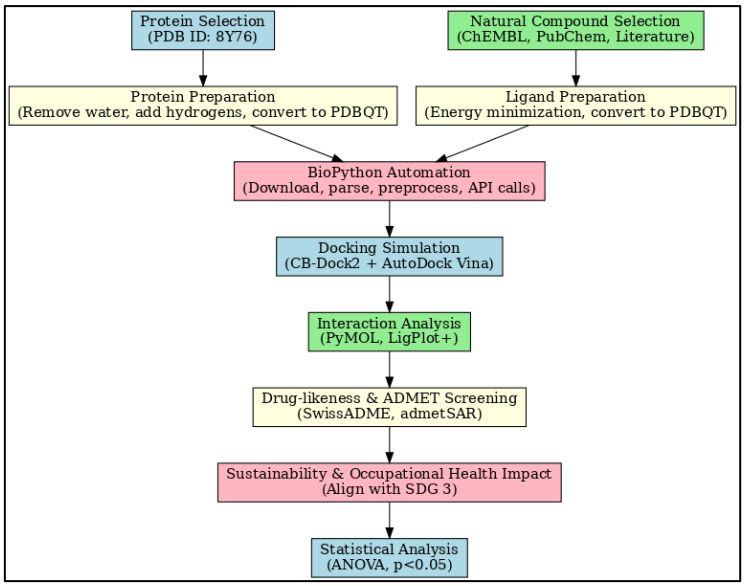


Figure 2- flowchart of workflow

Our research study also supports the United Nations Sustainable Development Goal (SDG) 3. The aim of this goal is to reduce premature mortality from non-communicable diseases (NCDs) through prevention, treatment, and mental health promotion. Alzheimer's disease is a part of NCDs with significant personal, societal, and economic impacts. By investigating tau-targeted drug discovery using natural compounds, this study helps develop affordable, potentially safer treatments, and also ensures equitable healthcare access and better NCD management.

Moreover, our study is relevant to occupational health, where workers exposed to heavy metals, pesticides, organic solvents, and other neurotoxic agents may face higher risks of developing neurodegenerative diseases, including AD. Identifying tau-targeted therapeutics with neuroprotective properties could help to shape occupational health strategies and reduce long-term cognitive decline among high-risk groups. Merging these therapeutic advances into workplace health programmes aligns with SDG 3 goals, especially those focused on prevention and early intervention. The choice of an In-silico approach is justified by its efficiency, scalability, and sustainability. Computational screening significantly lowers laboratory resource use, reduces experimental waste, and lowers the environmental impact of drug discovery, an increasingly important factor in sustainable research practices. Additionally, the design of the Bio-Python workflow ensures reproducibility and flexibility, and allows for the future integration of additional compound libraries, docking algorithms, and targets relevant to other neurodegenerative diseases. [11,12,13]

This research addresses a critical unmet need in Alzheimer's disease treatment by employing computational tools to identify natural compounds that inhibit tau protein aggregation. The combination of the BioPython library and molecular docking provides a high-throughput, cost-effective, and environmentally friendly platform for early-stage drug discovery. Aligning with global health priorities and occupational health concerns, the study enhances scientific understanding of tau-targeted therapeutics and supports the broader goal of sustainable health

innovation. These results are expected to set up a foundation for experimental validation and future translational research aimed at developing accessible, effective, and safe treatments for Alzheimer's disease. [14,15,16]

## 2. Materials and Methods

### 2.1 Collection of Data and Fingerprinting Data Analysis

We collect samples from the PDB database and download the 3D structure of the tau protein sample ID 8Y76. We obtain a sequence of samples using the Molecular Modelling Database (MMDB) tool. This structure was selected due to its high-resolution experimental validation and relevance to AD-associated tau pathology. For ligands, we use ChEMBL with Alzheimer's neuroprotective annotations. We utilised a fingerprint database analysis tool to identify protein signature patterns and profiles within protein families and domains, and to analyse their evolution, biological function, and structure.

### 2.2 Preparation for ligand and protein

A library of bioactive natural compounds was compiled from publicly available chemical databases. Ligand structures were optimised and converted into appropriate formats using automated scripts developed in Bio-Python, enabling standardised and reproducible ligand preparation. After collecting the data file, we convert the ligand file to a .pdbqt file using Python libraries such as RDKit and OpenBabel. We remove water and heteroatoms, then perform energy minimisation of the sample using Biopython and PyMOL.

### 2.3 Computational Intelligence–Assisted Screening

Artificial intelligence–assisted automation was implemented using Bio-Python to streamline ligand screening, structural validation, and filtering. Algorithm-driven rules were applied to evaluate molecular compatibility, physicochemical properties, and structural stability, allowing efficient elimination of unsuitable compounds.

### 2.4 Molecular Docking Analysis

Molecular docking was performed to assess the binding affinity and orientation of each ligand within the active regions of the tau protein. Docking scores were used as quantitative indicators of binding strength, and top-ranking ligand–protein complexes were selected for further analysis. In the docking setup, we define the active site using either known binding sites or a blind docking approach. For docking, we use the AutoDock Vina tool with the help of Bio-Python docking codes. Subsequently, we perform batch docking for all ChEMBL ligands.

### 2.4 Post-Docking Analysis

After completing docking for all ligands, we analyse docking scores and ligand ranks. We visualise the top ligand interaction using PyMOL, and then generate a graph using seaborn and matplotlib. Protein–ligand interaction analysis was conducted to identify hydrogen bonds, hydrophobic interactions, and key amino acid residues involved in binding. These interactions

were visualised using molecular visualisation tools to assess ligand binding stability and specificity. Computational intelligence–based ranking and prioritization strategies were applied to identify lead compounds with optimal binding affinity and interaction profiles. The final leads were selected based on docking scores, interaction stability, and biological relevance to tau inhibition.

# 3. Result

**PDB ID 8Y76**

Crystal structure of the SAM domain of L3MBTL3

Classification: GENE REGULATION

Organism(s): Homo sapiens

Expression System: Escherichia coli

Mutation(s): Sequcence mutation

**MMDB DATABASE**

Chain A, Lethal (3) malignant brain tumour-like protein 3

PDB: 8Y76_A

>pdb|8Y76|A Chain A, Lethal (3) malignant brain tumour-like protein 3

SVSKWSTDEVSEFIQSLPGCEEHGKVFKDEQIDGEAFLLMTQTDIVKIMSIKLGPALKIFNSILM
FKAAE

**Fingerprint database analysis**

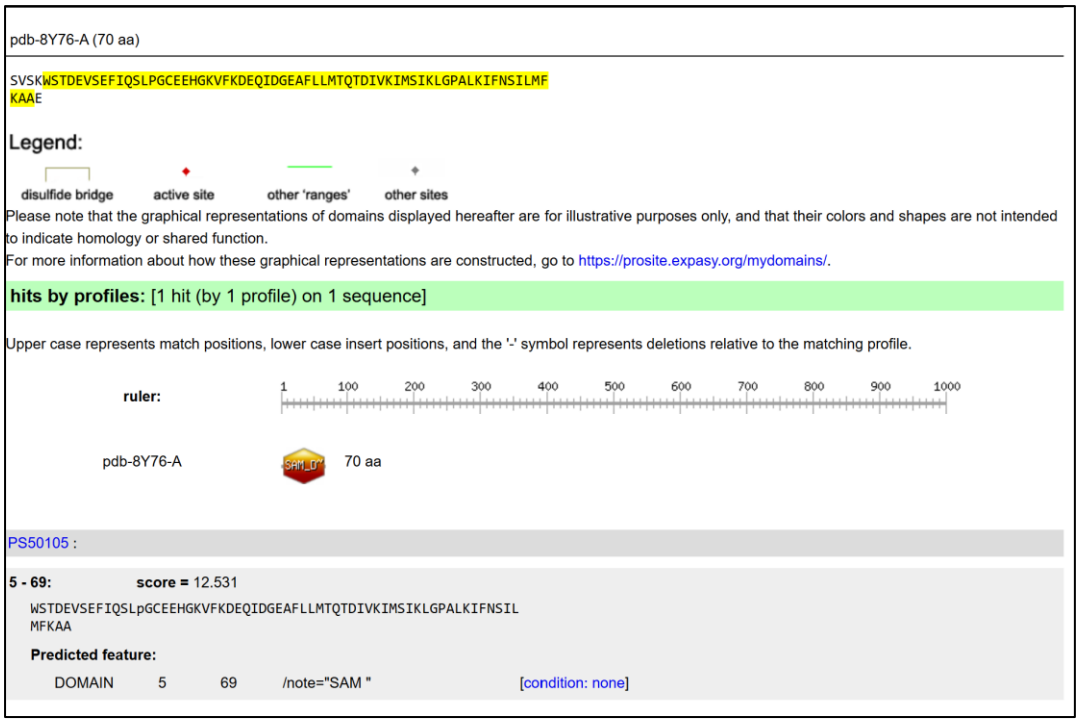


Figure 3: - Database analysis of sample 8Y76

Interpretation: In the domain prediction score, 12.531 indicates a strong match to the profile; the yellow-highlighted region shows the matched region relative to other families. In this analysis, a predicted domain, the SAM domain (Sterile Alpha Motif), is identified. This protein is involved in signal transduction, transcriptional regulation, and protein-protein interactions, and mediates binding to and oligomerisation of RNA/DNA. This domain also mediates tau protein phosphorylation by the scaffold kinase formation.

Structure of the sample 8Y76 by using PyMOL after the removal of water, heteroatoms and minimisation of energy.

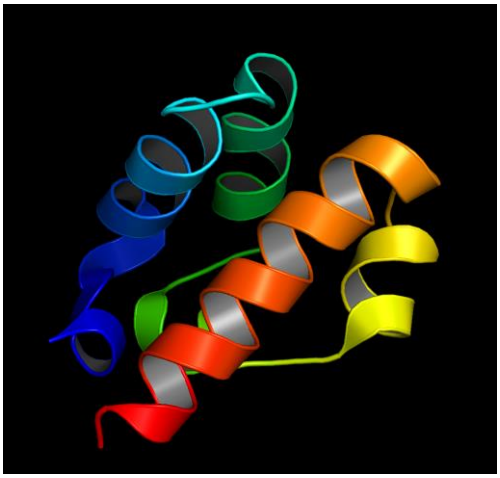

Figure 4: - 8y76 structure from Pymol

The structure highlights the overall conformation of the tau protein used for molecular docking and interaction analysis in Alzheimer's disease–related studies.

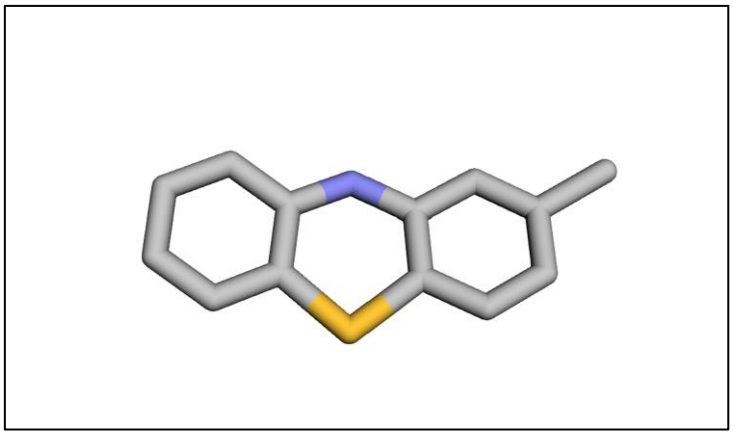

Figure 5: ligand structure by using Python Colab scripted py3mol

The ligand structure was optimized and visualized to ensure proper geometry prior to molecular docking analysis.

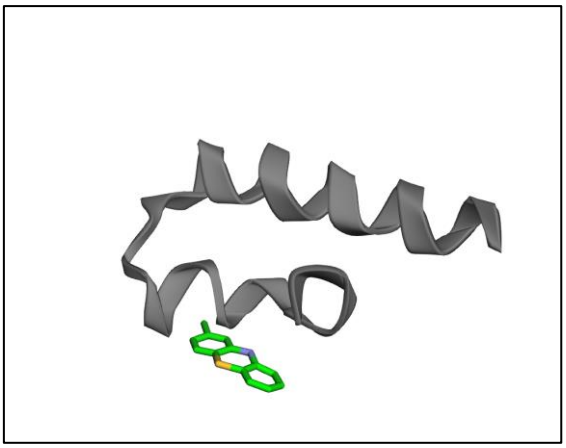

Figure 6: Ligand-Receptor docking result using AutoDock in Python Colab

The docked complex illustrates the binding orientation and interaction of the ligand within the functional region of the tau protein.

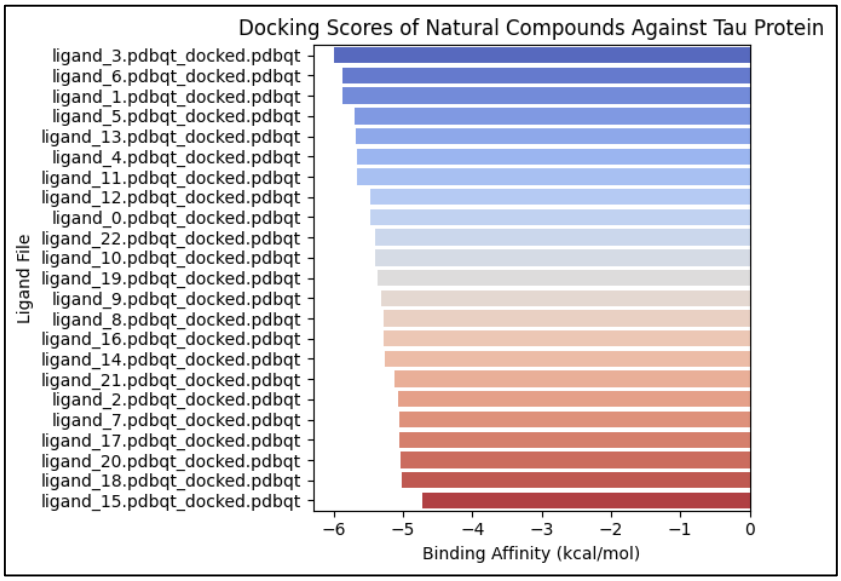


Figure 7: Horizontal bar graph displaying the docking scores (binding affinities) of various natural compounds (ligands) against the Tau protein, which is involved in Alzheimer's disease.

Interpretation: - In this horizontal graph, blue bars show stronger binding and red bars show weaker binding affinities. The most negative docking scores indicate stronger binding affinity for the Tau protein, with ligand_3 (-6 kcal/mol), ligand_6 (-5.5 kcal/mol), and ligand_1 (-5 kcal/mol) exhibiting the strongest binding. These ligands could be potential Tau-targeting therapeutics for further investigation. Ligands that have less negative scores indicate weaker interactions, such as ligand_18(-1 kcal/mol) and ligand_15 (-0.5 kcal/mol). Tau protein aggregation is a hallmark of Alzheimer's disease. Identifying natural compounds with high binding affinity could assist in Inhibiting Tau aggregation, designing novel therapeutic agents, and exploring neuroprotective mechanisms. The docking score distribution demonstrates the effectiveness of the proposed AI-assisted in silico screening framework in differentiating

potential tau inhibitors from a natural compound library. Several ligands exhibited significantly lower binding energies, indicating strong and stable interactions with the tau protein.

This quantitative ranking of compounds directly supports the proposal's objective of lead compound identification, enabling data-driven prioritization before experimental validation. The use of automated docking analysis and visualization reflects the role of computational intelligence as a research tool, reducing manual screening bias and accelerating early-stage drug discovery for Alzheimer's disease.

## 4. Discussion

This study highlights the potential of computational intelligence methods as powerful tools in Alzheimer's disease drug discovery. By combining AI-driven automation, molecular docking, and interaction analysis, the framework effectively finds candidates targeting tau, reducing both experimental costs and time. The use of natural compounds increases therapeutic relevance because of their structural diversity and biological compatibility. Choosing PDB ID 8Y76 provided a reliable structure for precise docking and interaction studies. Although this study emphasizes in silico validation, experimental and clinical testing will be necessary to verify the therapeutic potential of the lead compounds identified.

## 5. Conclusion

This study presents a computational intelligence–driven in silico drug discovery framework for targeting tau protein in Alzheimer's disease. The integration of AI-assisted screening, molecular docking, and interaction analysis demonstrates how artificial intelligence can serve as an effective research tool for transforming knowledge creation in neurodegenerative disease therapeutics. The identified lead compounds provide a foundation for future experimental validation and precision medicine approaches in Alzheimer's disease.

The SAM domain prediction in pdb-8Y76-A (residues 5–69) indicates a structurally conserved region likely involved in protein–protein interactions, potentially relevant to signalling pathways. Simultaneously, docking analysis shows Ligand_3 as the most promising Tau binder (–6 kcal/mol), suggesting therapeutic potential for neurodegenerative diseases. Combining domain annotation with docking insights supports a targeted approach to modulate Tau interactions through SAM-containing proteins. Future work should validate these findings via molecular dynamics and mutational studies, explore SAM domain-mediated Tau regulation, and assess ligand specificity and bioactivity in cellular models to advance translational strategies for Alzheimer's disease and related tauopathies. This integrated analysis supports UN Sustainable Development Goal 3: Good Health and Well-being, focusing on decreasing mortality from non-communicable diseases and enhancing mental health. Applying molecular insights to translational research helps achieve SDG 3's aim of strengthening R&D for diseases that mainly impact ageing populations. Future validation and drug development can lessen dementia's impact, improving quality of life and long-term neurological health. Future work

will focus on integrating machine learning models to improve binding affinity prediction, performing ADMET and toxicity analyses, and experimentally validating the identified lead compounds.

## REFERENCES

1. Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, Cummings J, van der Flier WM. Alzheimer's disease. *Lancet*. 2021;397(10284):1577–1590.
2. Khan S, Barve KH, Kumar MS. Recent advancements in the pathogenesis, diagnostics, and treatment of Alzheimer's disease. *Curr Neuropharmacol*. 2020;18(11):1106–1125.
3. Gueorguieva I, Willis BA, Chua L, Chow K, Ernest CS, Shcherbinin S, Ardayfio P, Mullins GR, Sims JR. Donanemab population pharmacokinetics, amyloid plaque reduction, and safety in participants with Alzheimer's disease. *Clin Pharmacol Ther*. 2023;113(6):1258–1267.
4. Henley DB, Sundell KL, Sethuraman G, Dowsett SA, May PC. Safety profile of semagacestat, a gamma-secretase inhibitor: IDENTITY trial findings. *Curr Med Res Opin*. 2014;30(10):2021–2032.
5. Long HZ, Cheng Y, Zhou ZW, Luo HY, Wen DD, Gao LC. PI3K/AKT signalling pathway: A target of natural products in the prevention and treatment of Alzheimer's disease and Parkinson's disease. *Front Pharmacol*. 2021; 12:648636.
6. Jones WP, Kinghorn AD. Extraction of plant secondary metabolites. *Methods Mol Biol*. 2012; 864:341–366.
7. Niewiadomska G, Niewiadomski W, Steczkowska M, Gasiorowska A. Tau oligomers neurotoxicity. *Life (Basel)*. 2021;11(1):28.
8. Uma Kumari, Vineeta Johri, Swarali Dhopate, Tijil Jha. Structure Based Drug Designing for the prediction of epitope for targeting Malignant brain Tumor, 2024 JETIR July 2024, Volume 11, Issue 7
9. Vineeta Johri, Ekta Tyagi, Uma Kumari. (2024). Next Generation Sequencing Analysis of NSP5 NSP6 and NSP7 SARS CoV-2 Non-structural Protein with Bio-Python. South Eastern European Journal of Public Health, 1653–1666. https://doi.org/10.70135/seejph.vi.2936 Dec 2024
10. Kumari U, Kaur G, et al. Bio-python/network of protein identification and NGS analysis of glioma cancer ATP competitive type III C-MET inhibitor. *JETIR*. 2024;11(2):41–51. Published: 27 Jun 2024
11. Kumari U, Chaudhary S. Next-generation sequencing, molecular docking, and network pharmacology reveal potent inhibitors for the treatment of lung cancer. *JETIR*. 2024;11(9).
12. Uma Kumari, Adya A P, Shruthi Satheesan, Drug Discovery and Biopython Analysis MBP-MCL1 in Myeloid Cell Leukaemia. Journal of Emerging Technologies and Innovative Research (JETIR) (Jan, 2025). Volume 12, Issue 1. Pp f178-f187.
13. Uma Kumari, Shruti Gupta, NGS and Sequence Analysis with Biopython for Prospective Brain Cancer Therapeutic Studies. https://doi.org/10.22214/ijraset.2023.50885

14. Uma Kumari, Gurpreet Kaur *et al*,2024," Biopython/Network Of Protein Identification And NGS Analysis Of Glioma Cancer ATP Competitive Type III C-MET Inhibitor: 7.367 (Calculated by Google Scholar): Volume 11, Issue 2: 27-Jun-2024: pp 41-51: http://doi.one/10.1729/Journal.40229
15. Tanmay Bandbe, Juri Saikia, Uma Kumari. (2025). NGS Analysis Human Papillomavirus Type 18 E2 DNA-Binding Domain Bound to its DNA Target with Biopython. *South Eastern European Journal of Public Health*, 3781–3792. https://doi.org/10.70135/seejph.vi.5856,SEEJPH Volume XXVI, S2
16. Uma Kumari, Renu et al "Structure Analysis and molecular Docking of Mesothelin-207 fragment in human cancer", Jetir, Vol.12, Issue 3, page no.g98-g106, http://doi.one/10.1729/Journal.44261
17. Uma Kumari, Karren Mehrotra,"NGS and Proteomic gene expression analysis in Studies of NUDT5 silence hormone signaling in breast cancer", International Journal of Emerging Technologies and Innovative Research, Vol.12, Issue 2, page no. ppc630-c638, http://doi.one/10.1729/Journal.43563
18. Shipra Chaudhary Uma Kumari, "NGS, MOLECULAR DOCKING AND NETWORK PHARMACOLOGY REVEAL POTENT INHIBITOR FOR THE TREATMENT OF LUNG CANCER", International Journal of Emerging Technologies and Innovative Research, Vol.11, Issue 9, page no. ppf116-f126, https://doi.one/10.1729/Journal.41696
19. Uma Kumari, Pallavi Belokar[1], Amolika Esompalli[2], Shreya Deshpande[2], Manasi Kumkar[3], Ankita Tripathi[3] Next-Generation Sequencing To Investigate The P53 Cancer Mutant Y234C For Targeted Cancer Therapies, *IOSR Journal Of Pharmacy and Biological Sciences (IOSR-JPBS). Volume 20, Issue 4 Ser. 1 (July – August 2025), PP 63-72*