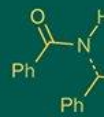


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
ISSN Online: 2617-4707
NAAS Rating (2026): 5.29
IJABR 2026; SP-10(8): 1179-1187
www.biochemjournal.com
Received: 25-05-2026
Accepted: 30-06-2026

Pawan Kumar Raghav
Project Trainee at
Bioinformatics Project and
Research Institute, Noida,
Uttar Pradesh, India

Uma Kumari
Senior Bioinformatics
Scientist, Bioinformatics
Project and Research Institute,
Noida, Uttar Pradesh, India

Harshita Anand
Project Trainee at
Bioinformatics Project and
Research Institute, Noida,
Uttar Pradesh, India

Abhinav Kumar Sinha
Project Trainee at
Bioinformatics Project and
Research Institute, Noida,
Uttar Pradesh, India

Corresponding Author:
Uma Kumari
Senior Bioinformatics
Scientist, Bioinformatics
Project and Research Institute,
Noida, Uttar Pradesh, India

Structure-guided discovery and cheminformatics optimization of NLRP3 inflammasome inhibitors for neurodegenerative disease therapeutics

Pawan Kumar Raghav, Uma Kumari, Harshita Anand and Abhinav Kumar Sinha

DOI: <https://www.doi.org/10.33545/26174693.2026.v10.i8Sn.9594>

Abstract

Neurodegenerative diseases, including Alzheimer's disease and Parkinson's disease, are a major public health issue and the subject of global interest and research efforts over the past few decades. These diseases, which are caused by sustained inflammation and neuronal loss, are characterized by relentless degeneration of neuronal function and are considered to be the major causes of mortality and severe disability. Recently, the involvement of the NLRP3 inflammasome has been recognized as a central mediator of the inflammatory response, making it an attractive therapeutic target for neurodegenerative diseases. In analysis used a structure-guided computational strategy to discover lead compounds that inhibit the NLRP3 NACHT domain (PDB ID: 7ALV). After analyzing the structure for appropriate binding pockets, we screened compounds using molecular docking to predict their binding affinity for the target protein. Selected compounds with high docking scores were then subjected to various cheminformatics approaches, including calculation of physicochemical properties, analysis of the chemical space, estimation of drug-likeness, evaluation of synthetic accessibility, prediction of PAINS (Pan-assay interference compounds) alerts, and estimation of blood-brain barrier (BBB) permeability.

Keywords: NLRP3 inflammasome, neurodegenerative diseases, molecular docking, structure-guided drug discovery, Cheminformatics, NACHT domain

1. Introduction

Neurodegenerative diseases, such as AD, PD, ALS and MS, constitute one of the greatest burdens facing modern medicine. Degradation occurs within the structural and functional integrity of the nervous system, culminating in cognitive decline and motor deficiency, which ultimately causes deterioration of quality of life. Worldwide demographic trends indicate neurodegenerative diseases are reaching epidemic proportions, with dementia-related syndromes repeatedly identified as a future priority for intervention by the WHO. Few disease-modifying therapies are currently available. In recent years there has been increasing evidence to suggest that chronic neuroinflammation is not only a bystander to neurodegeneration but may actually be an active contributor to it. When challenged by accumulated proteins such as amyloid, synuclein, or misfolded tau protein, the resident central nervous system immune cells, microglia, release reactive oxygen species, inflammatory cytokines, and other mediators. This prolonged inflammatory response damages neurons directly and synaptically, enhances further protein aggregation, and enhances the ongoing inflammatory process, creating a self-sustaining cycle of neuronal death and glial activation. Accordingly, elucidating these molecular mechanisms is a major goal for the future development of neuroprotective drugs.

1.1 The NLRP3 Inflammasome: Structure and Mechanism of Activation

Within the various molecular platforms mediating the innate immune inflammatory response, the NOD-like receptor family pyrin domain-containing 3 (NLRP3) inflammasome has been identified as particularly important. NLRP3 is an intra-cytosolic pattern-recognition receptor which, when activated by a broad spectrum of danger signals including extracellular ATP, reactive oxygen species, lysosomal insult, or protein aggregates, polymerizes with the adaptor molecule ASC (PYCARD) and pro-caspase-1 into a large multiprotein complex ^[1].

This induces the autoproteolytic activation of caspase-1, which cleaves pro-IL-1 β and pro-IL-18 into their secreted, biologically active form, and gasdermin D into a pore-forming domain to induce pyroptotic cell death [2]. NLRP3 is a multidomain protein with an N-terminal pyrin domain (PYD); a central nucleotide-binding and oligomerisation domain (also called the NACHT domain; includes a nucleotide-binding domain, helical domain 1, winged-helix domain, and helical domain 2); and a C-terminal leucine-rich repeat (LRR) domain. The NACHT domain is the catalytic domain that binds ATP/ADP and oligomerizes, and is the main region of interest for structure-based inhibitor design, as ligand binding in or proximal to it can hold NLRP3 inactive, preventing inflammasome formation. The availability of high-resolution cryo-electron microscopy structures of the NLRP3 NACHT domain, such as Protein Data Bank identifier 7ALV, the main structural template for this work, has enabled the visualisation of these ligand-binding pockets in atomic detail and the use of rational structure-guided inhibition.

1.2 NLRP3 in Neurodegenerative Disease Pathology

However, when considered within the scope of neurodegeneration, inflammasome activation has been causally linked to disease progression and pathology in several settings, ruling out a solely secondary role. In AD, A β fibrils are a particularly potent activator of microglial NLRP3 inflammasomes and isotype-specific inducers of inflammasome-derived IL-1 β that directly build on a body of work implicating pro-inflammatory responses and AD autoantibodies to A β in the pathogenesis [3]. Moreover, inflammasome-derived ASC specks released from activated microglia were shown to cross-seed tau histopathology. This study positions the NLRP3 inflammasome as a biologically relevant mechanistic bridge between the genesis of the two characteristic pathologies of AD, i.e. A β -cored fibrils and tau, activating the innate immune compartment responsible in this paradigm for culmination in a discovery-driven approach with the potential to inform the development of a disease-modifying therapeutic. By contrast, when neurodegeneration develops in the context of additional pathology, inflammasome activation is a disease-driving event in PD, as microglial NLRP3 inflammasome activation induced dopaminergic neurodegeneration in the substantia nigra; while in other neurodegenerative conditions, similar inflammasome-dependent inflammatory mechanisms drive axonal injury and demyelination [4].

1.3 Therapeutic Targeting of NLRP3: Current Landscape and Challenges

Pharmacological interest in NLRP3 inhibition has increased dramatically since the recent identification of compounds such as MCC950, which binds the NACHT domain of NLRP3 to potentially inhibit ATP hydrolysis required to induce conformational rearrangements of NLRP3 that activate inflammasome assembly [5]. However, MCC950 and other earlier paradigm compounds have been disappointing for clinical development due to off target hepatotoxicity and poor pharmacokinetic properties highlighting the challenge of completing the optimization paradigm for NLRP3-targeted drugs that are both sufficiently potent and safe [6]. For neurodegeneration indications, additional requirements specific to this route of administration include the ability to traverse the blood brain

barrier (BBB), achieve adequate central nervous system (CNS) exposure, and not induce non-specific immunosuppression. Therefore, forward NLRP3 inhibitor discovery for neurodegeneration cannot focus solely on binding; structure, pharmacokinetics and safety profile considerations must be integrated at the earliest points in the discovery pipeline.

1.4 Structure-Guided Drug Discovery and Chemoinformatics: Rationale

Structure-guided drug discovery exploits the three-dimensional structure of a validated target to select, screen and optimize candidate ligands computationally prior to the time-consuming and costly processes of *in vitro* synthesis and testing. Approaches are generally based on retrieval and quality control of the target crystal or cryo-EM structure; identification of druggable binding sites; virtual screening of candidate compounds by molecular docking, estimating binding affinities and interaction geometries; and chemoinformatic profiling of high scoring hits using a variety of criteria including Lipinski's Rule of Five, pan-assay interference (PAINS) filters synthetic accessibility and blood brain barrier permeability prediction [7, 8]. The biological rationale is strengthened by orthogonal bioinformatics, including sequence identity searches, structural alignment, and protein-protein interaction network analysis. This integrated structural bioinformatics and cheminformatics pipeline is particularly suitable for a target such as NLRP3, where the aim is not just to find any binder, but a chemically tractable, CNS-penetrant, and NACHT domain-selective inhibitor. By combining structural validation, molecular docking, descriptor-based filtering, and systems-level pathway analysis, this synthetic set of properties ultimately results in a small, mechanistically supported, computationally prioritized set of candidate compounds for experimental testing. This study has been conducted for a total of nine weeks duration systematically covering structural bioinformatics, cheminformatics, drug discovery process, and systems biology. The study was undertaken collaboratively by three student research fellows with the project structured in a phased manner that involves literature review, materials and methods, results and discussion sections. The computational work performed under this project is focused on structural data retrieve and validate, sequence and structural comparison, preparation of the receptor, molecular docking against both the selected candidates and the drug databases, chemical filtration of identified inhibitors and custom structural analysis of the protein, protein interaction network mapping, and pathway analysis of the protein, all targeted towards establishing NLRP3 NACHT domain (PDB ID: 7ALV) as a pharmacologically targetable domain and to screen candidate small molecule inhibitors for neurodegenerative diseases.

2. Materials and Methods

This section outlines the computational pipeline used to identify and evaluate candidate small-molecule inhibitors of the NLRP3 inflammasome as potential therapeutics for neurodegenerative disease. The pipeline combines structural bioinformatics, molecular docking, and cheminformatics to characterize the binding behavior and drug-like properties of the selected compounds. Every stage of the analysis was carried out *in silico*, drawing on public structural

repositories, sequence-analysis tools, and Python-based cheminformatics libraries.

2.1 Target Protein Selection and Structural Database Analysis

NLRP3 was chosen as the target protein for inhibitor discovery. High-resolution structures were located by querying the Protein Data Bank (PDB) [9] and the Molecular Modelling Database (MMDb) [10]. The crystal structure was selected on the basis of resolution and completeness, and its chain composition, resolution, and any bound ligands or cofactors were reviewed before it was carried forward for downstream analysis.

2.2 Structure Visualization

An initial visual inspection of the 7ALV structure was carried out in RasMol [11]. Command-line rendering was used to switch between wireframe, cartoon, and spacefill representations, apply residue- and chain-based colouring, and display hydrogen bonds, which together helped characterise the secondary structure and the key interactions stabilising the fold.

2.3 Sequence Retrieval and Homology Analysis

The FASTA sequence of the target protein was extracted from the PDB entry and was aligned to the non-redundant protein database using protein BLAST (BLASTP) [12]. We evaluated homology based on the conventional BLAST output parameters, including the E-value, percent identity, alignment length, and gap distribution, and retrieved the closest homologs from the alignment summary table. The top BLASTP hits were later plotted in a network graph using Python's NetworkX library [13], showing how they relate to each other.

2.4 Structural Validation

Model quality was evaluated with the SAVES server, where ERRAT indicated unusual non-bonded atomic packing in different regions [14]. For stereochemical quality analysis, we performed a Ramachandran plot that shows the distribution of backbone torsion angles (phi and psi) into favoured, allowed, and disallowed regions [15]. Top B-factors in PyMOL are visualised as a representation to differentiate the rigid and flexible parts of the protein at the residue level [16].

2.5 Comparative structural and sequence alignment

7ALV was structurally aligned to another NLRP3 protein structure (PDB id: 8ZEM) and RMSD between the two was calculated [17].

2.6 Protein-Protein Interaction Network Analysis

The researchers generated the network of NLRP3 protein-protein interactions. They used the STRING database to find what proteins interact with NLRP3. This way, they tried to understand the role of NLRP3 protein-protein interaction in the development of inflammation and neurodegeneration. In order to verify the reliability of the interactions they discovered, researchers used the False Discovery Rate test to identify true interactions. In this way, they are able to distinguish true from false interactions. The NLRP3 protein plays a very important role in this mechanism and those researchers were interested in discovering how it interacts with other proteins leading to neurodegeneration and inflammation.

2.7 Molecular Docking Studies Docking

Five selected candidate inhibitors, Rasagiline, Oridonin, Fingolimod Hydrochloride, Tranilast, and Edaravone, were docked with the 7ALV structure using the CB-Dock server [18]. Cavity volume, docking center, docking score and interacting residues for each ligand were documented. Tranilast had the highest binding score among all five candidate ligands (docking score = 8.6 kcal/mol).

A custom, reproducible pipeline was built around the Bio Python library to extend beyond what the web-based servers could offer. Core Bio Python modules handled transcription, translation, and complement/reverse-complement operations on the relevant DNA sequences, while the Bio.PDB module was used to parse and programmatically represent three-dimensional protein and DNA structures [19]. Docking was also automated using Auto Dock Vina within the Python pipeline. Vina binding-affinity scores for each ligand were then extracted and compared programmatically across runs [20, 21].

2.7 Structural Feature and Hydrophobicity Analysis

A contact map of the structure was created to determine the spatial distribution of residues and identify regions that are tightly packed and may serve as binding sites for the 7ALV structure with the potential ligand. Additionally, 7ALV was examined in order to define which parts of the molecule are hydrophobic and which parts of the structure are hydrophilic based on the Kyte-Doolittle scale. In this case, the scale was used to differentiate the hydrophobic from hydrophilic areas and identify the regions where the ligand can bind on the 7ALV surface [22, 23].

3. Results

3.1 Structural Retrieval, Visualization, and Validation

The structure of the NLRP3 NACHT domain, Protein Structure was visualized with RasMol and PyMOL to determine the elements, secondary structure, and network of hydrogen bonds, in keeping with previous structural reports on NLRP [24, 25]. BLASTP was used on the FASTA sequence retrieved from the PDB database and confirmed the species and similarity of the identified molecule lending high confidence that the crystal structure 7ALV belongs to the human NLRP3 protein [26].

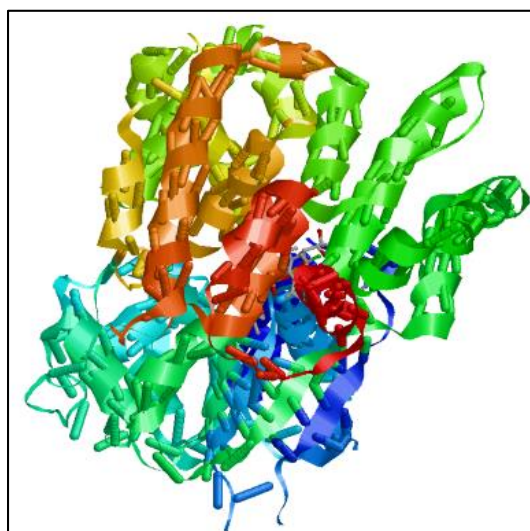


Fig 1: Cartoon representation of the NLRP3 NACHT domain (PDB ID: 7ALV), rainbow-colored from N-terminus (blue) to C-terminus (red), shown with a bound ligand at the docking pocket.

In total, the quality of the model based on protein structure validation achieved an ERRAT score of 92.689, which is clearly over the limit of 90 and hence “excellent” for a difficult calculation, as per Dock [27, 28]. The ERRAT detected only two areas with high error values in the model (residues 280-290 and 355-365), which most likely

correspond to flexible loop regions and are not indications of errors but of the inherent flexibility typical of NACHT domain-containing proteins. Almost all residues fell within allowed areas in the Ramachandran plot of backbone dihedral angles [29].

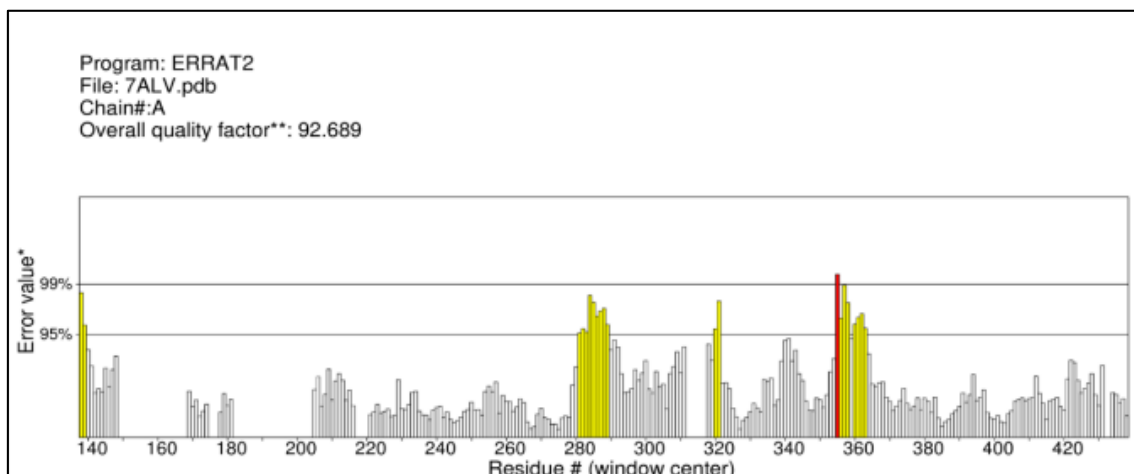


Fig 2: ERRAT2 quality assessment plot for 7ALV (overall quality factor: 92.689), showing per-residue error values across the sequence.

3.2 Structural and Sequence Conservation Analysis

Structural alignment of 7ALV against a related structure, 8ZEM, yielded an RMSD of 0.5 Å, indicating a very high degree of structural conservation and suggesting that the ligand-binding surface identified in 7ALV is not an artifact of a single structure determination. Clustal Omega multiple sequence alignment reinforced this by highlighting conserved residue stretches across homologous sequences, confirming that the residues involved in ligand binding occupy evolutionarily conserved and functionally important positions [30].

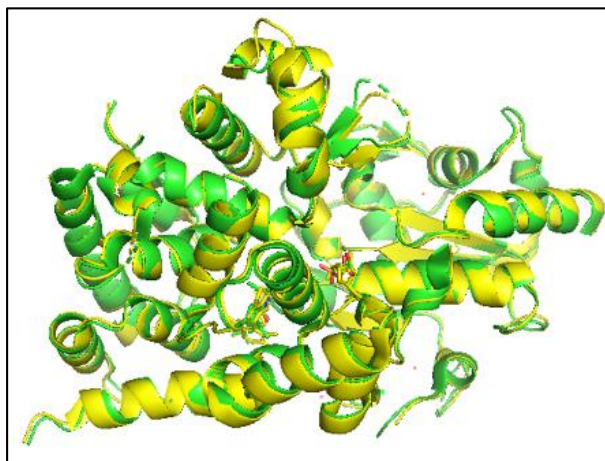


Fig 3: Structural superposition of 7ALV (green) and 8ZEM (yellow); (RMSD = 0.5 Å), illustrating the high degree of structural conservation between the two structures.

The B-factors of the single residues viewed to illustrate clearly that the core region is rigid while the loops on the surface are flexible [31].

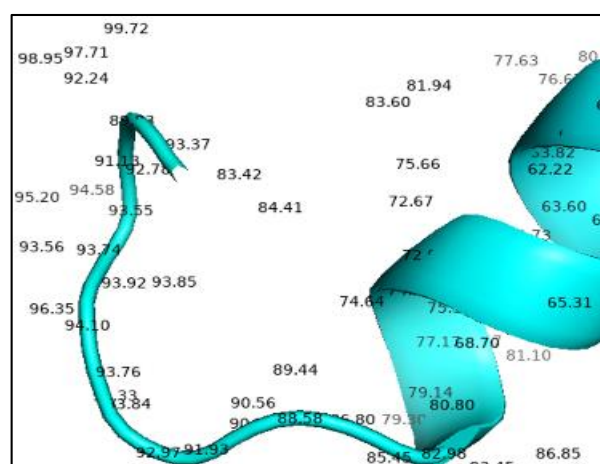


Fig 4: B-factor-labelled cartoon of a representative loop region of 7ALV in Python molecular object, illustrating relative atomic mobility across the backbone.

3.3 Protein-Protein Interaction Network and pathway analysis

An analysis of interactions in the STRING database showed NLRP3 as a highly connected hub protein with strong likelihood links to CASP1, PYCARD, NEK7, TXNIP, AIM2, and MAVS (scores 0.998-0.999); these data agreed with mechanisms of inflammasome assembly [32].

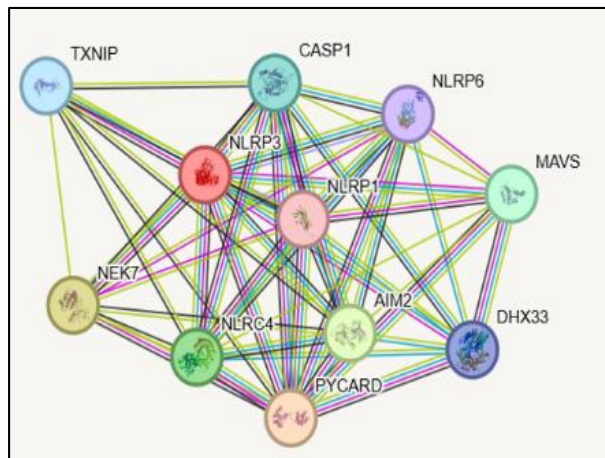


Fig 5: STRING protein-protein interaction network showing NLRP3 as a hub protein with high-confidence associations to CASP1, PYCARD, NEK7, TXNIP, AIM2, and MAVS

3.4 Overall Protein Structure

The NLRP3 NACHT domain is composed of a three-dimensional conformation. The structure of the NLRP3 NACHT domain was established by investigating its conformational form. A type of image, a three-dimensional scatter plot, was generated to observe the distribution of the amino acid residues in the protein structure [33]. The NLRP3

NACHT domain has a highly ordered structure without large void spaces, making the NLRP3 NACHT domain structural form fit for computer analysis. Therefore, the NLRP3 NACHT domain is an appropriate protein structure for computer analysis.

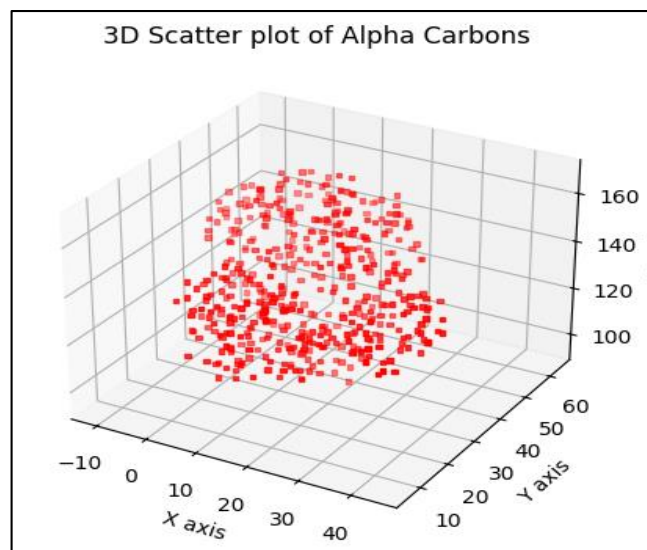


Fig 6: Spatial distribution of C α atoms of 7ALV

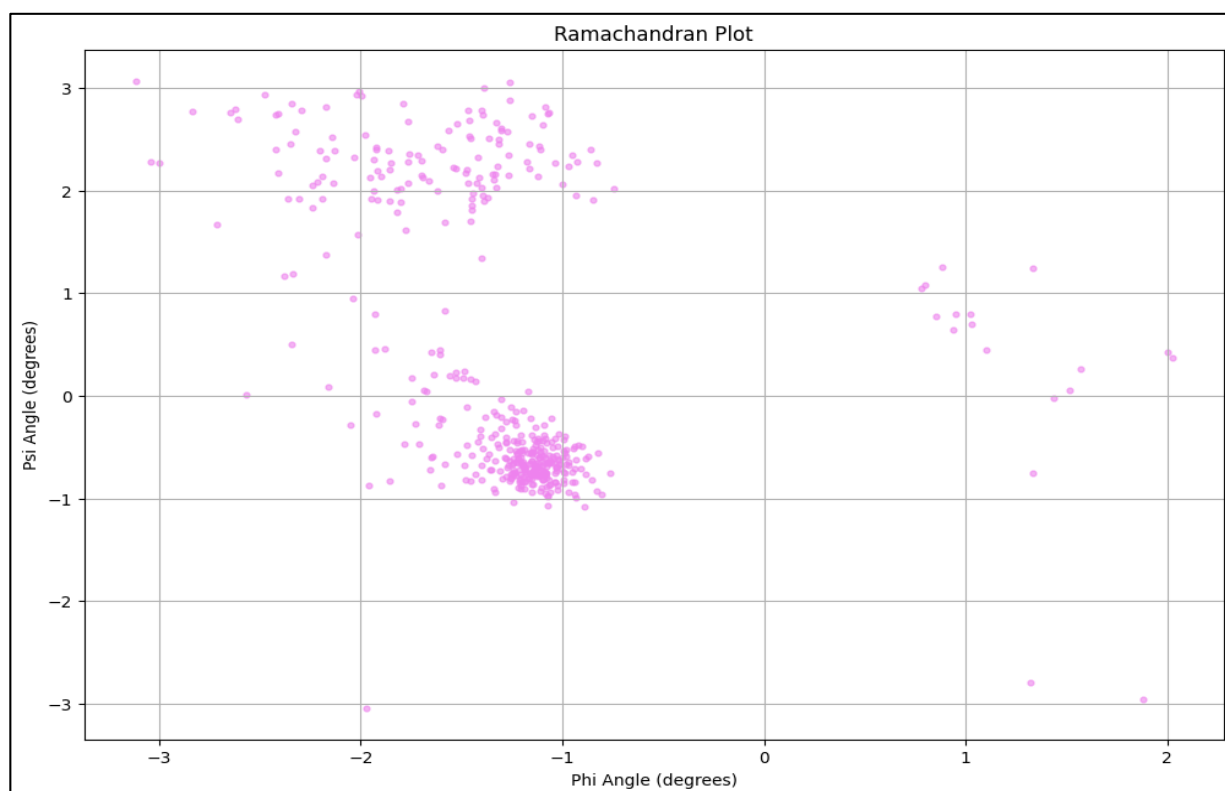


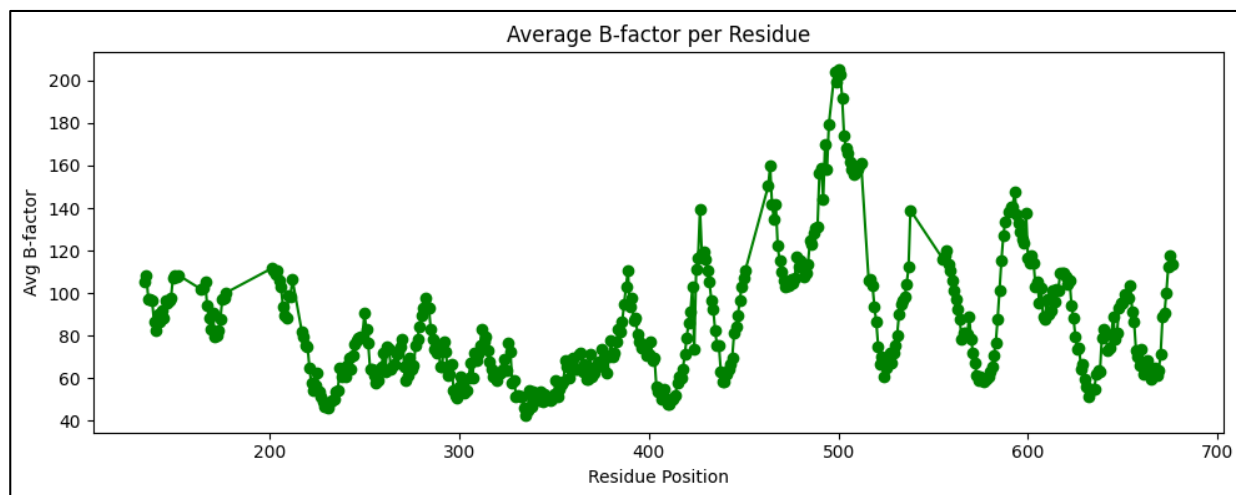
Figure 7. Ramachandran Plot of 7ALV

To further validate the stereochemical quality of the protein model, a Ramachandran plot was generated. The distribution of backbone ϕ (phi) and ψ (psi) dihedral angles revealed that the majority of amino acid residues occupied the most favoured regions, while only a limited number of residues were present within additionally allowed or generously allowed conformational spaces. Very few residues were observed within disallowed regions,

suggesting minimal stereochemical strain within the protein model [34].

3.5 Protein Flexibility Analysis

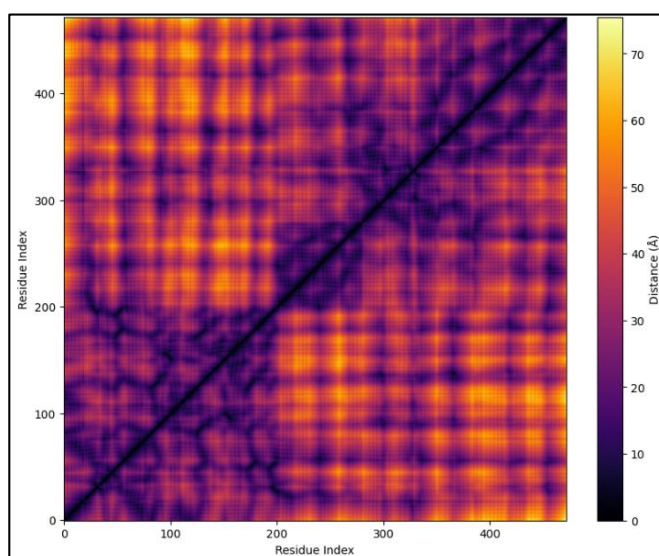
The structural flexibility of the NLRP3 NACHT domain was investigated by analysing the residue-wise average B-factor obtained from the crystal structure.

**Fig 8:** Average B-Factor per Residue

3.6 Residue Contact Analysis

To further investigate the structural organization of the NLRP3 NACHT domain, a residue contact map was

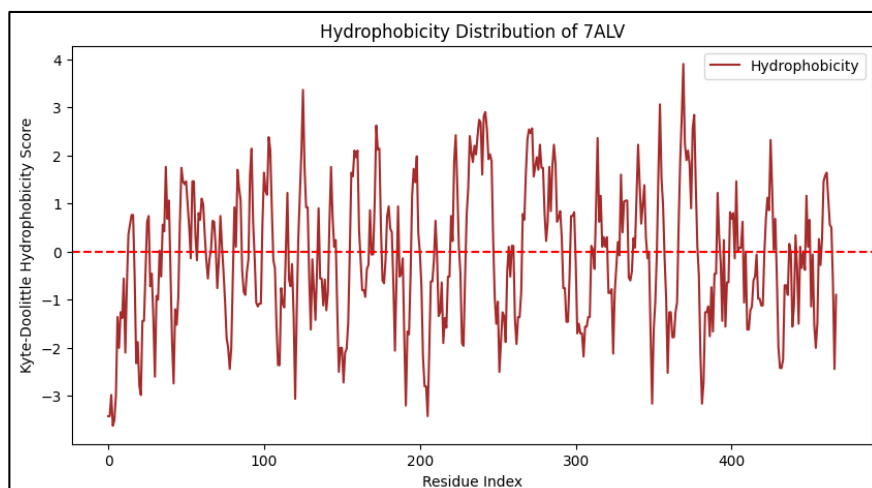
generated based on spatial proximity between amino acid residues.

**Fig 9:** Contact Map of 7ALV

3.7 Surface Hydrophobicity Analysis

Protein-ligand interactions are strongly influenced by the physicochemical properties of the protein surface. To

investigate the distribution of hydrophobic and hydrophilic regions within the NLRP3 NACHT domain, a surface hydrophobicity map was generated.

**Fig 10:** Hydrophobicity Distribution of 7ALV

3.8 Synthetic Accessibility and Blood-Brain Barrier Prediction

Synthetic accessibility (SA) scores ranged from approximately 3.27 to 3.78, with the majority of compounds clustered between 3.3 and 3.7. As lower SA scores

correspond to easier synthesis, these values indicate moderate synthetic complexity. No compound exhibited exceptionally high SA scores indicative of significant synthetic challenges^[35].

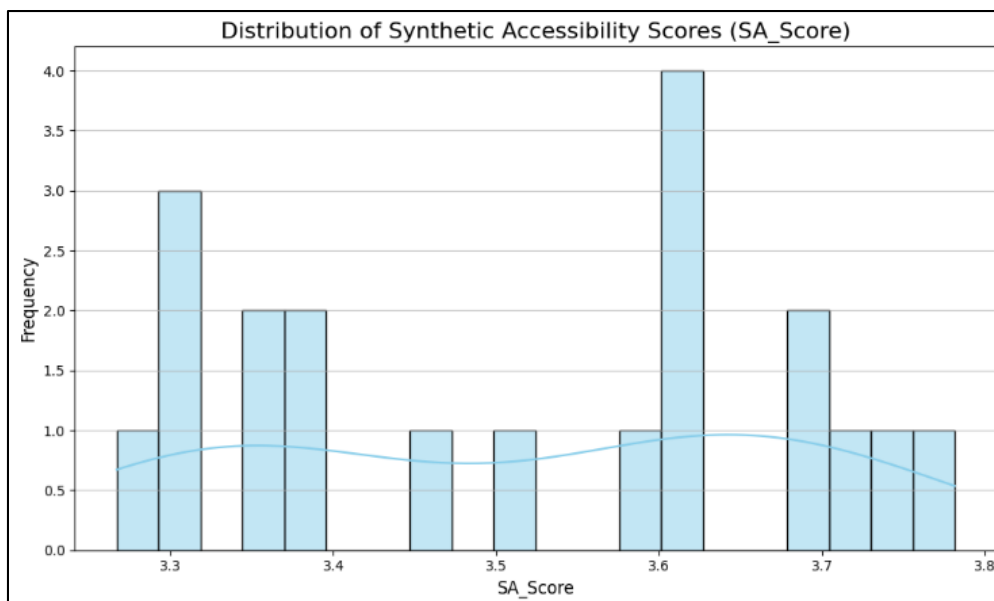


Fig 11: Distribution of synthetic accessibility scores among the screened compounds.

3.9 Molecular Docking of Candidate Ligands

The prepared ligand library was docked against the predicted binding pocket of the NLRP3 NACHT domain (PDB ID: 7ALV) using AutoDock Vina. Docking scores were used as an initial measure of predicted binding affinity, with lower binding energies indicating stronger protein-

ligand interactions^[36, 37]. Docking of five candidate ligands yielded binding affinities of Rasagiline (-7.9 kcal/mol), Oridonin (-8.3 kcal/mol), Tranilast (-8.9 kcal/mol), Fingolimod hydrochloride (-7.2 kcal/mol), and Edaravone (-7.4 kcal/mol), with Tranilast showing the strongest predicted interaction with the NLRP3 NACHT domain.

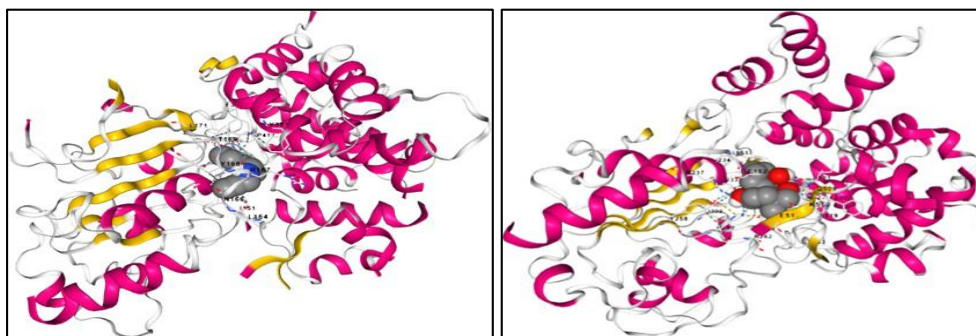


Fig 12: Rasagiline interaction in secondary structure style with a score of -7.9 kcal/mol; Figure 13. Oridonin interaction in secondary structure style with a score of -8.3 kcal/mol



Fig 14: Tranilast interaction in secondary structure style with a score of -8.9 kcal/mol; Figure 15. Edaravone interaction in secondary structure style with score -7.4 kcal/mol

Table 1: AutoDock Vina docking results for the five candidate ligands screened against the NLRP3 NACHT domain (PDB ID: 7ALV), including predicted binding affinity, search-space cavity volume, and grid centre coordinates.

PubChem CID	Drug name	AutoDock Vina score (kcal/mol)	Cavity volume (Å ³)	Center (x, y, z)	Remarks
3052776	Rasagiline	-7.8	1904	17, 31, 142	Average
5321010	Oridonin	-8.3	1904	17, 37, 142	Strong
107969	Fingolimod hydrochloride	-7.2	1904	17, 37, 142	Average
5282230	Tranilast	-8.6	1904	17, 37, 142	Strongest predicted binder
4021	Edaravone	-7.4	1904	17, 37, 142	Average

A core group of pocket residues, TYR381, PRO412, LEU413, TRP416, ILE521, and HIS522, was observed in nearly all of the docked complexes. This provides good evidence that this group represents the center of a true, target-binding pocket, rather than simply a random surface contact. Tranilast also interacted with an additional extended group of residues (ASN165, ARG167, LYS232, THR233, ARG237, HIS260, ARG262, GLN509). This likely reflects Tranilast's increased binding energy with respect to a larger contact surface area, and with a larger number of polar/charged interactions corresponding to its functional groups of amide, carboxylic acid and methoxy-aromatic, with a binding energy that is the highest among all the tested molecules. The poorest scoring molecules, Rasagiline and Edaravone, bind to nearly the exact same smaller subset of residues which correlates to their smaller size, and lower binding energies, indicating good consistency of the results. The scores, however, are just approximate binding energies that would need to be validated experimentally [38, 39].

5. Conclusion

Overall, the application of an integrated approach involving structural bioinformatics and chemoinformatics proved to be quite effective in finding prospective drug candidates for NLRP3 inflammasome inhibitors for the treatment of neurodegenerative diseases. Structural validation showed that NLRP3 (NACHT domain) could be used as a drug discovery target via computational approaches. From the results of molecular docking, pocket analysis and chemoinformatic evaluation, lead compounds with high scores also had desirable drug-like features, demonstrating the effectiveness of integrating molecular docking and chemoinformatics for selecting drug leads. Additionally, structural studies, sequence conservation analysis and protein-protein interaction network analysis pointed out the biological relevance of NLRP3 as a potential target in the treatment of neurodegenerative diseases. The results obtained using computational methods form a strong basis for future research, including molecular dynamics simulations and *in vitro* experiments for the study of NLRP3 as a therapeutic target.

Reference

- Swanson KV, Deng M, Ting JPY. The NLRP3 inflammasome: Molecular activation and regulation to therapeutics. *Nature Reviews Immunology*. 2019;19:477-489. doi:10.1038/s41577-019-0165-0.
- Guo H, Callaway JB, Ting JPY. Inflammasomes: Mechanism of action, role in disease, and therapeutics. *Nature Medicine*. 2015;21(7):677-687.
- Heneka MT, Kummer MP, Stutz A, Delekate A, Schwartz S, Vieira-Saecker A, *et al.* NLRP3 is activated in Alzheimer's disease and contributes to pathology in APP/PS1 mice. *Nature*. 2013;493:674-678. doi:10.1038/nature11729.
- Ising C, Venegas C, Zhang S, Scheiblich H, Schmidt SV, Vieira-Saecker A, *et al.* NLRP3 inflammasome activation drives tau pathology. *Nature*. 2019;575:669-673.
- Coll RC, Robertson AAB, Chae JJ, Higgins SC, Muñoz-Planillo R, Inserra MC, *et al.* A small-molecule inhibitor of the NLRP3 inflammasome for the treatment of inflammatory diseases. *Nature Medicine*. 2015;21(3):248-255. doi:10.1038/nm.3806.
- Mangan MSJ, Olhava EJ, Roush WR, Seidel HM, Glick GD, Latz E. Targeting the NLRP3 inflammasome in inflammatory diseases. *Nature Reviews Drug Discovery*. 2018;17:588-606. doi:10.1038/nrd.2018.97.
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*. 2001;46(1-3):3-26. doi:10.1016/S0169-409X(00)00129-0.
- Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry*. 2010;53(7):2719-2740. doi:10.1021/jm901137j.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, *et al.* The Protein Data Bank. *Nucleic Acids Research*. 2000;28(1):235-242. doi:10.1093/nar/28.1.235.
- Madej T, Lanczycki CJ, Zhang D, Thiessen PA, Geer RC, Marchler-Bauer A, *et al.* MMDB and VAST + : Tracking structural similarities between macromolecular complexes. *Nucleic Acids Research*. 2014;42(D1):D297-D303.
- Sayle RA, Milner-White EJ. RasMol: Biomolecular graphics for all. *Trends in Biochemical Sciences*. 1995;20(9):374-376. doi:10.1016/S0968-0004(00)89080-5.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. 1990;215(3):403-410. doi:10.1016/S0022-2836(05)80360-2.
- Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: *Proceedings of the 7th Python in Science Conference (SciPy2008)*. 2008. p.11-15.
- Colovos C, Yeates TO. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*. 1993;2(9):1511-1519. doi:10.1002/pro.5560020916.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*. 1993;26(2):283-291. doi:10.1107/S0021889892009944.

16. DeLano WL. The PyMOL molecular graphics system. Schrödinger, LLC; 2002.
17. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*. 2011;7:539. doi:10.1038/msb.2011.75.
18. Liu Y, Yang X, Gan J, Chen S, Xiao ZX, Cao Y. CB-Dock2: Improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Research*. 2022;50(W1):W159-W164. doi:10.1093/nar/gkac394.
19. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, *et al.* Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*. 2009;25(11):1422-1423. doi:10.1093/bioinformatics/btp163.
20. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*. 2010;31(2):455-461. doi:10.1002/jcc.21334.
21. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *Journal of Cheminformatics*. 2011;3:33. doi:10.1186/1758-2946-3-33.
22. Rogers D, Hahn M. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*. 2010;50(5):742-754. doi:10.1021/ci100050t.
23. McInnes L, Healy J, Melville J. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*. 2018.
24. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*. 1982;157(1):105-132.
25. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, *et al.* The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*. 2021;49(D1):D605-D612. doi:10.1093/nar/gkaa1074.
26. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 2000;28(1):27-30. doi:10.1093/nar/28.1.27.
27. Blum M, Chang HY, Chuguransky S, Grego T, Kandasamy S, Mitchell A, *et al.* The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*. 2021;49(D1):D344-D354. doi:10.1093/nar/gkaa977.
28. Kumari U, Mukherjee S, Pradhan M, Chakrabarti S. NGS analysis approach for neurodegenerative disease with Biopython. *Journal of Emerging Technologies and Innovative Research*. 2023;10(9). doi:10.1729/Journal.36043.
29. Kumari U, Pacholee K. In-silico drug discovery-based approach to treat impairments in patients of Alzheimer's disease. *Journal of Emerging Technologies and Innovative Research*. 2023;10(12):B236-B245. doi:10.1729/Journal.37001.
30. Kumari U, Gunjan, Basak S. NGS contribution in genetic diagnosis for brain disorder. *Journal of Clinical Otorhinolaryngology, Head, and Neck Surgery*. 2023;27(1):5525-5536.
31. Kumari U, Gaur D, Gupta A. Development of Parkinson's disease by genetic mutation analysis. *Journal of Clinical Otorhinolaryngology, Head, and Neck Surgery*. 2023;27(1):5225-5252.
32. Nagpal G, Kumari U. Next generation sequence analysis of amyloid precursor-like protein 2 (APLP2) E2 domain in Alzheimer's disease. *Journal of Emerging Technologies and Innovative Research*. 2025;12(1):ppd72-ppd79.
33. Uniyal U, Kumari U. Next generation sequencing data analysis of A25T transthyretin structure associated with Parkinson's disease. *Journal of Emerging Technologies and Innovative Research*. 2025;12(3):B612-B622. doi:10.1729/Journal.44079.
34. Kokad M, Kumari U. Next-generation sequencing data analysis in DYRK1A associated with neurodegenerative disease. *IOSR Journal of Biotechnology and Biochemistry*. 2025;11(6):40-48. doi:10.9790/264X-1106014048.
35. Mehrotra K, Kumari U, Pande S. Integrative workflow of Biopython and molecular docking to explore novel therapeutics targeting PARK7-DJ1. *IOSR Journal of Biotechnology and Biochemistry*. 2025;11(6):1-10. doi:10.9790/264X-1106010110.
36. Singh N, Kumari U. Multi-omics-based drug discovery pipelines incorporating machine learning, AI, and BioPython in Parkinson's disease: A case study of 5A polymorph of alpha-synuclein. *Journal of Emerging Technologies and Innovative Research*. 2026;13(2). doi:10.56975/.
37. Mandra D, Kumari U, Singh N, Marouthu A. Integrated multi-omics and Biopython approaches for target identification of α -synuclein polymorph in Parkinson's disease. *IOSR Journal of Biotechnology and Biochemistry*. 2026;12(2):12-21.
38. Johri V, Kumari U. Sustainable neuroprotection via proteomics and integrative NGS-CADD analysis of 6CVP variant to improve future neurodegenerative health outcomes. *IOSR Journal of Pharmacy and Biological Sciences*. 2026;21(3):1-10. doi:10.9790/3008-2103010110.
39. Singh N, Kumari U, Mandra D, Marouthu A. Integrated machine learning-based toxicity prediction with molecular docking for safer drug candidate screening in Parkinson's disease. *Computational Biology and Bioinformatics*. 2026;14(1):41-53. doi:10.11648/j.cbb.20261401.14.