IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Homology Modeling And Virtual Screening Studies Of Hpt Protein To Target Tuberculosis By Identifying Lead Compounds

Rumandla Lavanya*[1], Pagudala Adavi Raju*[2]

[1] Rumandla Lavanya: Associate Professor, Department of Chemistry, Vivekananda GDC(A), Vidyanagar, Hyderabad-500044, Telangana, India.

[2] Pagudala Adavi Raju: Associate Professor, GDC, Malkajgiri, Medchal-500068, Telangana, India.

Abstract: Tuberculosis, caused by Mycobacterium Tuberculosis (*Mtb*), is a leading cause of death, with 10.7 million new cases and 1.23 million deaths expected worldwide by 2024. The Hypoxanthine guanine phosphoribosyltransferase protein (hpt) belongs to the purine/pyrimidine phosphoribosyl transferase (PRTase) family and is used to discover active ligand molecules that inhibit the purine salvage pathway. The homology simulated 3D model was created with modeler 9.13 software. The final structure was validated, and the active site was identified using pre-requisite predictitions. The protein's three-dimensional structure consists of nine sheets and twelve helices. The grid was built on the active site, followed by virtual screening using the drug bank ligand database. The majority of the lead compounds acquired contain imidazole moieties and derivatives as pharmacophore groups in their structures. The target protein amino acid residues identified by virtual screening are Val69, Asp74, Arg77, Ile79, and Pro80, which are critical for ligand binding. The chosen lead compounds have permitted ADME characteristics and can inhibit the purine salvage pathway in *Mtb*. The imidazole structural scaffolds can be found in the ligand molecules of docked complexes Lig1, Lig 2, and Lig 3.

KEYWORDS: *Mycobacterium tuberculosis*, Purine salvage pathway, Modeller, Virtual screening, ADME.

I. INTRODUCTION

The WHO Global TB Report 2025 shows mixed results. while TB is declining globally, it remains a leading cause of death, with 10.7 million new cases and 1.23 million deaths in 2024, owing to variables such as malnutrition and HIV. India leads the world in case numbers (25%), but has achieved a remarkable 21% incidence decline (2015-2024), increasing treatment coverage to 92% through technology and campaigns, although still facing high drug-resistant TB (MDR-TB) burdens[1]. In the current study, the target protein hpt, hypoxanthine guanine phosphoribosyltransferase protein, functionally enzymes of the purine phosphoribosyltransferase (PRTase) family, is a component of the purine salvage pathway and has been proposed as drug targets for the development of chemotherapeutic agents against infectious and parasitic diseases. Purine bases (adenine, guanine, hypoxanthine, or xanthine) and their counterparts to the corresponding nucleosides 5'-monophosphate and pyrophosphate are ribophosphorylated in a single step in the PRTase-catalyzed chemical reaction [2], [3]. Consequently, the htp protein was verified as a viable target for the development of anti-TB drugs[4].

II. Methods and materials

2.1. Homology modelling

Comprehending the biology of the target protein requires an understanding of its structure. Homology modeling is used to create a three-dimensional structure because "X-ray crystal and NMR structure of target protein" are not readily available. UniProtKB provides the target protein, htp of the Mycobacterium tuberculosis sequence ("UniProt: The Universal Protein Knowledgebase in 2021," 2021). According to the FASTA format. BLAST[5] (Basic Local Search Alignment engine) for structural similarity and Jpred 4[6] for secondary structure prediction are two template search tools used to choose templates. The homologous protein is used as a template to determine the hpt protein's three-dimensional structure.

2.2. Sequence Alignment and model building

A better 3D model will be produced if the target and template proteins have high sequence alignment. Using the ClustalW (Cambridge, UK)[7] server program, the generated alignment file provides information about identical, similar, and distinct residues. The automated homology modeling tool MODELLER was used to generate the hpt protein comparison model after the target and template sequences were aligned pairwise[8]."For validation and refinement studies, the final model with the lowest (MOF) model objective function was chosen." Protein secondary structure information is provided by the PDBsum server.

2.3. Optimization and validation of the 3D model

In order to stabilize the generated 3D model, energy optimization is an essential step. The protein preparation module (Schrodinger Suite LLC, New York) is employed for structural optimization[9]. This phase involves removing unwanted water molecules, adding missing hydrogen atoms, and assigning bond ordering to the protein. The impact refinement module (Impref) and the optimal potential for liquid simulations (OPLS, 2005) are used to minimize energy with a cutoff of RMSD (Root Mean Square Deviation) set at 0.3 Å[10] power field. The PROCHECK and ProSA (Protein Structural Analysis) servers are used as validation tools for the revised 3D model. By examining the phi and psi angles of amino acid residues, the Ramachandran plot in PROCHECK[11] provides an explanation of the protein's stereochemical quality. The ProSA[12] server was used to assess the protein model's local compatibility. By taking advantage of energies based on amino acid sequence location, the ProSA energy graph illustrates the local quality of the protein model.

2.4. Prediction of Active site

The binding pockets at active sites are identified using the CASTp server application using the solvent-accessible surface model and the molecular surface of the protein[13]. The sitemap provides a very accurate indication of the overall site score. Protein binding areas, hydrophobicity, hydrogen bond donors, and interactions between the template and target proteins are all listed in the sitemap. The active site region can be identified using BLAST-Putative conserved domain[14].

2.5. Virtual Screening and ADME

The energy-minimized three-dimensional protein structure from the protein preparation wizard Maestro 9.1 is used for grid generation in virtual screening. The Schrödinger Suite of Grid gen module is used to generate a grid surrounding the energy-minimized protein structure's active site[15]. Using the LigPrep module of the Schrödinger Suite with the OPLS 2005[16] force field, the ligand molecules are chosen from the drug targets database in the structure data file (SDF) format. Schrodinger software's Glide (Grid-based Ligand Docking with Energetics) virtual screening workflow[17] features the following score protocols, each with a 10% output: XP stands for additional precision, SP for standard precision, and HTVS for high throughput virtual screening. The docked ligands with high Glide energy and Glide Score are assessed for additional research. The pharmacodynamic features of the docked complexes are revealed using the Schrodinger software's QikProp module[18], (Schrodinger LLC, 2010, New York). The new scaffold molecule's desirable (similar range) pharmacological properties and pharmacokinetic aspects, including absorption, distribution, metabolism, and excretion, are examined[19]. The Schrödinger Suite of the QikProp module evaluates the screened lead compounds.

III. RESULTS AND DISCUSSION

3.1. Structural assessment and validation

The purine/pyrimidine phosphoribosyl transferase (PRTase) family features the protein hpt, or hypoxanthine guanine phosphoribosyltransferase, which is involved in the purine salvage pathway. The target, the hpt protein, has 202 amino acids in total. The 3D structure of the hpt protein has not yet been disclosed to the Protein Data Bank by X-ray crystallography or NMR spectroscopy. In this context, a three-dimensional model was created using in-silico techniques. Homologous templates with the lowest e-values, sequence identity, and protein structural similarity prediction of secondary structure were found using BLAST and Jpred 4, respectively. The template protein with the lowest e-value is taken into consideration (PDB ID 4PFQ). The hpt protein has a lower e-value, a 56% identical sequence homology, and a 90% query coverage shown in Table 1.

Table 1. Template Selection-protein

No.	Server tool name	Parameter	PDB ID	E-Value
1.	BLAST	Protein structural similarity	4PFQ	3e-64
2.	Jpred 4	Secondary structure prediction	4PFQ	4e-41

Table 1. Template selection of hpt protein was done by using BLAST and Jpred 4 servers by the lowest e-value.

ClustalW aligned the hpt protein and 4PFQ template pair by pair; 89% of structurally identical residues were displayed as perfect alignment of template and protein pair by pair. Initially, 20 homogeneous models were produced using MODELLER 9.13. Schrodinger Suite's Protein Preparation Wizard eliminated steric conflicts and poor connections during the energy minimization phase, stabilizing the generated protein model. The findings show that using molecular dynamics considerably enhanced the hpt structure's stability features. To boost stereochemical quality, structural modifications were made and energy was optimized.

The Ramachandran plot of the protein amino acid residues' phi vs psi values is displayed in Figure 1. This figure is used to verify the model's stereochemistry quality. The results show that the strong stereo-chemical properties of the protein model were justified by 180 residues in the most desired zone, 14 in the extra permitted region, and 6 in the generously allowed portion out of 202 amino acids (Figure 1). Ninety percent of the amino acid residues were found in the most preferred location, according to the Ramachandran plot. This suggests that the stereochemical quality of the generated protein model is good. The ProSA server program assessed the structure using the obtained Z-score of -3.49. Figure 2 shows a high-quality hpt model structure with a negative value. Residues in amino acids that are below the baseline and have negative energies indicate that the generated hpt protein local model quality is more dependable.

3.2. Protein secondary structure and active site prediction

According to the PDBsum service, the secondary structure of the hpt protein consists of 12 helices and nine sheets. One important aspect of the protein structural stability seen in biological macromolecules is the non-bonding interactions. The potential active site was suggested by the CASTp server. The three-dimensional structure of the hpt protein was maintained by molecular dynamics simulation. Concave surface areas on the three-dimensional structure of the hpt protein are found and marked by the CASTp server. We identify binding pockets 1, 2, and 3. The optimized 3D model was displayed using PyMOL program. Active site of the hpt protein identified by using BLAST putative conserved domain in Figure 3. In the subsequent screening tests, the amino acids Ile62 to Val120 were found at the protein's active site.

Figure 1. The hpt protein stereochemical analysis- Ramachandran Plot

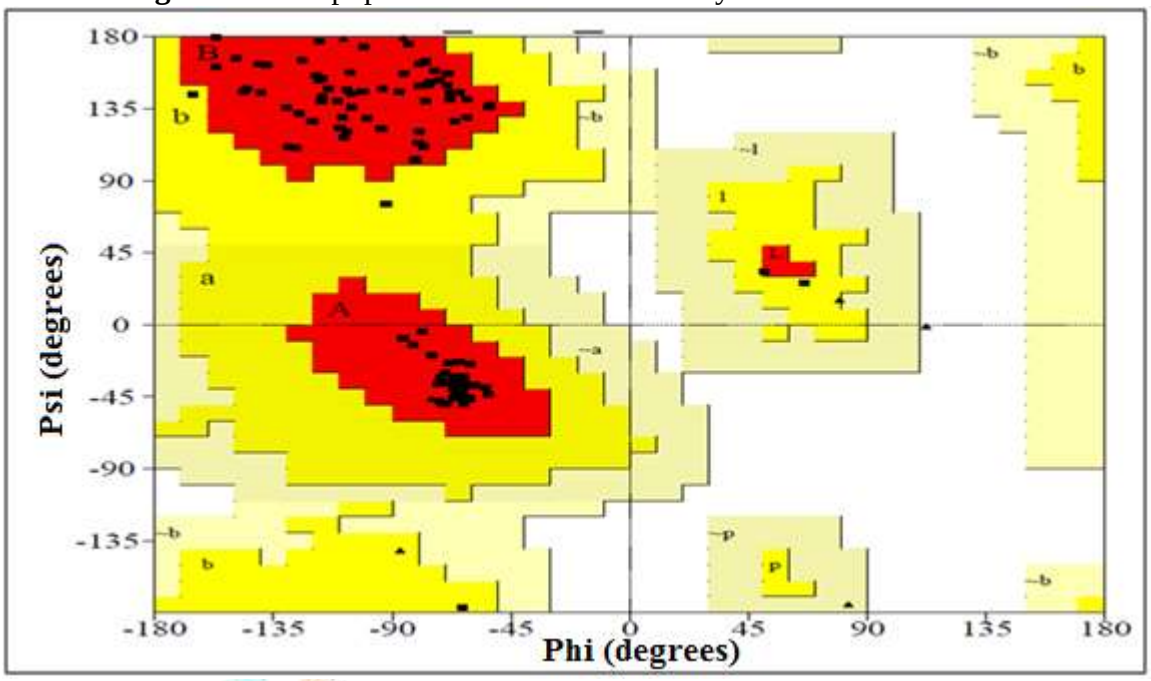


Figure 1. The stereo-chemical reliability of 3D model is confirmed by Ramachandran plot. The core zone is depicted in the colour red, with the extra permissible regions being brown, yellow, and light-yellow fields representing the banned regions.

Figure 2. ProSA analysis of the hpt protein 3D model

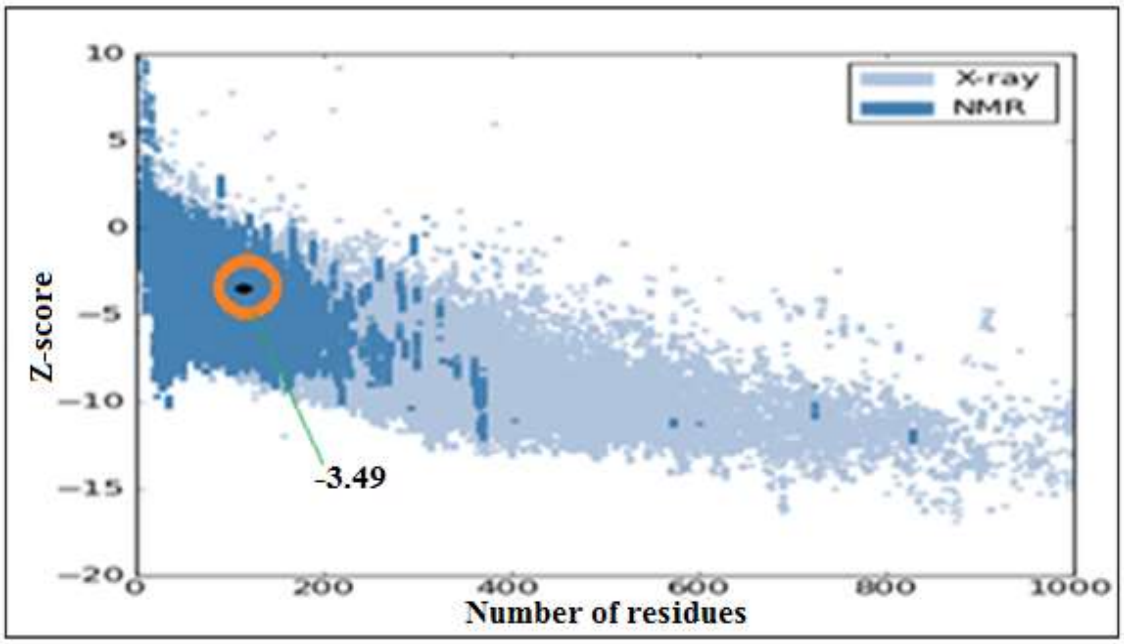


Figure 2. The hpt protein's energy profile is represented by the ProSA. The model's ProSA analysis revealed that the maximal residues had a negative energy zone. Analysis of the ProSA-web server reveals a -3.49 value of Z-score. The protein structure is of high quality if the Z-score is negative.

Figure 3. BLAST Putative Conserved Domain

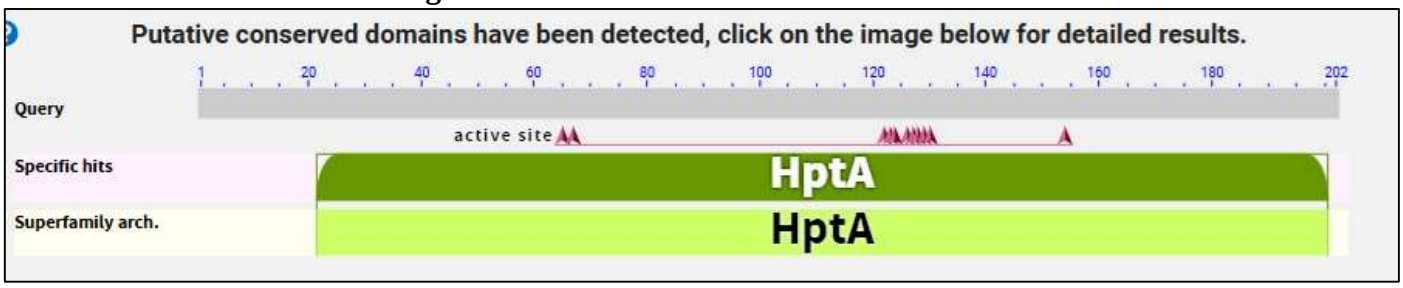


Figure 3. BLAST putative conserved domain for hpt protein has shown the active site region as shown in the figure**.**

3.3. Virtual screening

At the protein's binding site area, a grid measuring 32 x 32 x 32 Å was made. The ligand structures were prepared for virtual screening using the LigPrep tool of the Schrodinger suite. Ten thousand six hundred and ninety-two (10692) molecules were imported from the drug bank database into the ligand preparation module of the Schrödinger suite. With a 10% functional output by default, this module applies the filters High-throughput Virtual Screening (HTVS), Standard Precision (SP), and Extra Precision (XP) in a hierarchical manner. Out of the ten docked complexes, the top three are virtually screened based on their glide energy and glide score.

Table 2 shows that the top 3 ligand molecules from Lig1 to Lig3 were grouped based on non-bonding interactions. Glide scores range from -9.5 to -9.93, and Glide energy fluctuates between -48.45 and -51.56 K. Cal/mol. Val69, Asp74, Arg77, Ile79, and Pro80 are the amino acids with which the majority of protein-ligand docked complexes have demonstrated notable binding interactions. The majority of ligand molecules have pharmacophore groups in their structures that are imidazole moieties or their derivatives. The ligand molecules in the Lig1, Lig2, and Lig3 docking complexes include these imidazole structural scaffolds.

3.4. ADME Properties

In order to increase the success rate of drug discovery, the evaluation of the ADME study is crucial for later pre-clinical drug development research. For the docked complexes from Lig1 to Lig 3, the in-silico prediction of ADME characteristics is shown in Table 3. The QikProp module is used to compare the pharmacokinetic characteristics of existing medications in order to generate the pharmacokinetic parameters (Schrodinger LLC, 2010, New York). The investigation's findings suggest that because the found ligands correspond to the permitted range, they have drug-like outcomes. The ADME characteristics for Lig 1 through Lig 3 are all comparable to those of the majority of other authorized medication compounds. In contrast, by blocking the purine salvage pathway, the recently chosen ligand compounds are being explored as a potential treatment for tuberculosis.

Table 2. The docked complexes with the best Glide scores and Glide energy results.

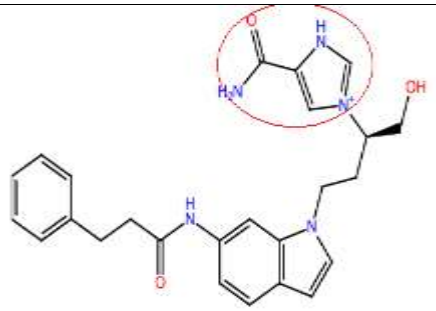
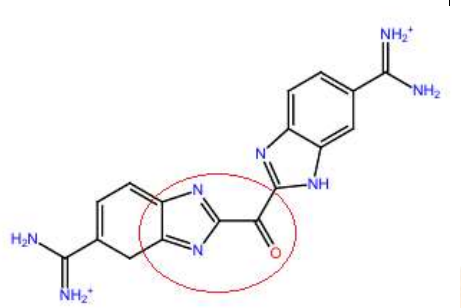
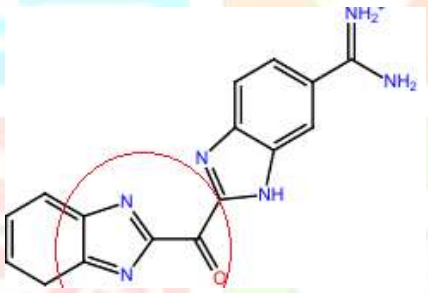
Binding ligand Number	Ligand structure	Glide energy	XP-glide score	Non-bonding interactions: H-bonds
Lig 1		-51.56	-9.93	ARG77:HH22-Lig 1:H27 ASP77:O-Lig 1:H34 PRO80:O-Lig 1:H32
Lig 2		-48.61	-9.75	VAL69:H-Lig 2: N2 VAL69:O-Lig 2:H15
Lig 3		-48.45	-9.5	VAL69:H-Lig 3: N2 ILE79:H-Lig 3: N5 PRO80:0-Lig 3:H16

Table 2. 22 lead molecules docked at the hpt protein active region, out of which the top 3 ligands were prioritized. hpt protein, lead molecules with a common scaffold of the pharmacophore were shown, which are involved in the non-bonding interactions.

Table 3. The qikprop features of the lead molecules (Lig 1 – Lig 3) are prioritized.

S. No.	CNS	DonorHB	Accepted HB	Q Plog Po/w	QPlog BB	oral absorption in humans %	Rule of 3	Rule of 5
Lig 1	-1	5.0	8.5	0.636	-1.518	79.990	0	0
Lig 2	-2	2.0	4.9	0.786	-1.305	79.540	0	0
Lig 3	-2	3.0	5.95	-0.793	-3.306	78.99	**1**	1

Table 3. The ADME attributes are predicted by the Schrödinger suite's QikProp programme. The following allowed ADME value ranges are: Oral absorption in humans %: <25%low and >80% high;
Mol wt. (130 to 725); CNS: -2 (inactive) +2 (active): HB of the donor: (0.0 to 6.0); QPlog Po/w: (-2.0 to 6.5); QPlog BB: (-3.0 to 1.2); Accept HB (2.0 to 20).
ADME – Absorption distribution metabolism excretion, HB – Hydrogen bond,
Po/w – Partition coefficient of Octanol of Water, BB – Brain blood barrier.

IV. CONCLUSION

The three-dimension structure of the hpt protein was studied, and the protein's structural features were determined via energy minimization. Virtual screening experiments were conducted to find and report the drug-like lead inhibitor compounds of hpt protein to arrest the purine salvage pathway. The amino acids Val69, Asp74, Arg77, Ile79, and Pro80 have a significant impact on protein-ligand binding. The lead compounds with acceptable ADME characteristics and a good docking score were found and reported. According to the findings, the imidazole structural scaffolds were identified as drug-like lead potential molecules, which further helps in the development of anti-tuberculosis drugs.

V. ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Head of the Department of Chemistry and the Principal of the University College of Science at Osmania University for facilitating the research work.

REFERENCES

[1] *Global tuberculosis report 2025*. 2025.

[2] G. Biazus, C. Z. Schneider, M. S. Palma, L. A. Basso, and D. S. Santos, "Hypoxanthine–guanine phosphoribosyltransferase from Mycobacterium tuberculosis H37Rv: Cloning, expression, and biochemical characterization," *Protein Expr. Purif.*, vol. 66, no. 2, pp. 185–190, 2009, doi: https://doi.org/10.1016/j.pep.2009.04.001.

[3] W. S. Eng *et al.*, "First Crystal Structures of Mycobacterium tuberculosis 6-Oxopurine Phosphoribosyltransferase: Complexes with GMP and Pyrophosphate and with Acyclic Nucleoside Phosphonates Whose Prodrugs Have Antituberculosis Activity," *J. Med. Chem.*, vol. 58, no. 11, pp. 4822–4838, Jun. 2015, doi: 10.1021/acs.jmedchem.5b00611.

[4] W. B. Parker and M. C. Long, "Purine Metabolism in Mycobacterium tuberculosis as a Target for Drug Development," 2007. doi: http://dx.doi.org/10.2174/138161207780162863.

[5] R. Zaru and S. E. Orchard, "UniProt Tools: BLAST, Align, Peptide Search, and ID Mapping," *Curr. Protoc.*, vol. 3, 2023, [Online]. Available: https://api.semanticscholar.org/CorpusID:257637623

[6] A. Drozdetskiy, C. Cole, J. Procter, and G. Barton, "JPred4: A protein secondary structure prediction server," *Nucleic Acids Res.*, vol. 43, Apr. 2015, doi: 10.1093/nar/gkv332.

[7] M. A. Larkin *et al.*, "Clustal W and Clustal X version 2.0," vol. 23, no. 21, pp. 2947–2948, 2007, doi: 10.1093/bioinformatics/btm404.

[8] Y. Peng *et al.*, "P-9.13: Design and Optimization of Triple-focal Microlens Array for Integral Imaging 3D Display," *SID Symp. Dig. Tech. Pap.*, vol. 56, pp. 1359–1364, Jul. 2025, doi: 10.1002/sdtp.19085.

[9] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, and W. Sherman, "Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments," *J. Comput. Aided. Mol. Des.*, vol. 27, no. 3, pp. 221–234, 2013, doi: 10.1007/s10822-013-9644-8.

[10] I. Kufareva and R. Abagyan, "Methods of protein structure comparison," *Methods Mol. Biol.*, vol. 857, pp. 231–257, 2012, doi: 10.1007/978-1-61779-588-6_10.

[11] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J. Appl. Crystallogr.*, vol. 26, no. 2, pp. 283–291, Apr. 1993, doi: https://doi.org/10.1107/S0021889892009944.

[12] M. Wiederstein and M. J. Sippl, "ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins," *Nucleic Acids Res.*, vol. 35, no. SUPPL.2, pp. 407–410, 2007, doi: 10.1093/nar/gkm290.

[13] W. Tian, C. Chen, X. Lei, J. Zhao, and J. Liang, "CASTp 3.0: computed atlas of surface topography of proteins," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W363–W367, Jul. 2018, doi: 10.1093/nar/gky473.

[14] L. Rumandla, B. Mounika, R. Malikanti, R. Vadija, K. K. Mustyala, and V. Malkhed, "Virtual Screening Technique to Identify Inhibitors of Mycobacterium tuberculosis Rv3032 Protein Involved in MGLP Biosynthesis," *Russ. J. Bioorganic Chem.*, vol. 50, pp. 1067–1081, Jun. 2024, doi: 10.1134/S1068162024030300.

[15] R. A. Friesner *et al.*, "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy," *J. Med. Chem.*, vol. 47, no. 7, pp. 1739–1749, 2004, doi: 10.1021/jm0306430.

[16] "Schrödinger Release 2021-3: LigPrep, Schrödinger, LLC, New York, NY.," *Schrödinger Release 2021-3 LigPrep, Schrödinger, LLC, New York, NY.*.

[17] L. E. Bultum, G. B. Tolossa, G. Kim, O. Kwon, and D. Lee, "In silico activity and ADMET profiling of phytochemicals from Ethiopian indigenous aloes using pharmacophore models," *Sci. Rep.*, vol. 12, no. 1, pp. 1–19, 2022, doi: 10.1038/s41598-022-26446-x.

[18] A. J. Lucas, J. L. Sproston, P. Barton, and R. J. Riley, "Estimating human ADME properties, pharmacokinetic parameters and likely clinical dose in drug discovery," *Expert Opin. Drug Discov.*, vol. 14, no. 12, pp. 1313–1327, 2019, doi: 10.1080/17460441.2019.1660642.

[19] D. Lagorce, D. Douguet, M. A. Miteva, and B. O. Villoutreix, "Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors," *Sci. Rep.*, vol. 7, no. April, pp. 1–15, 2017, doi: 10.1038/srep46277.