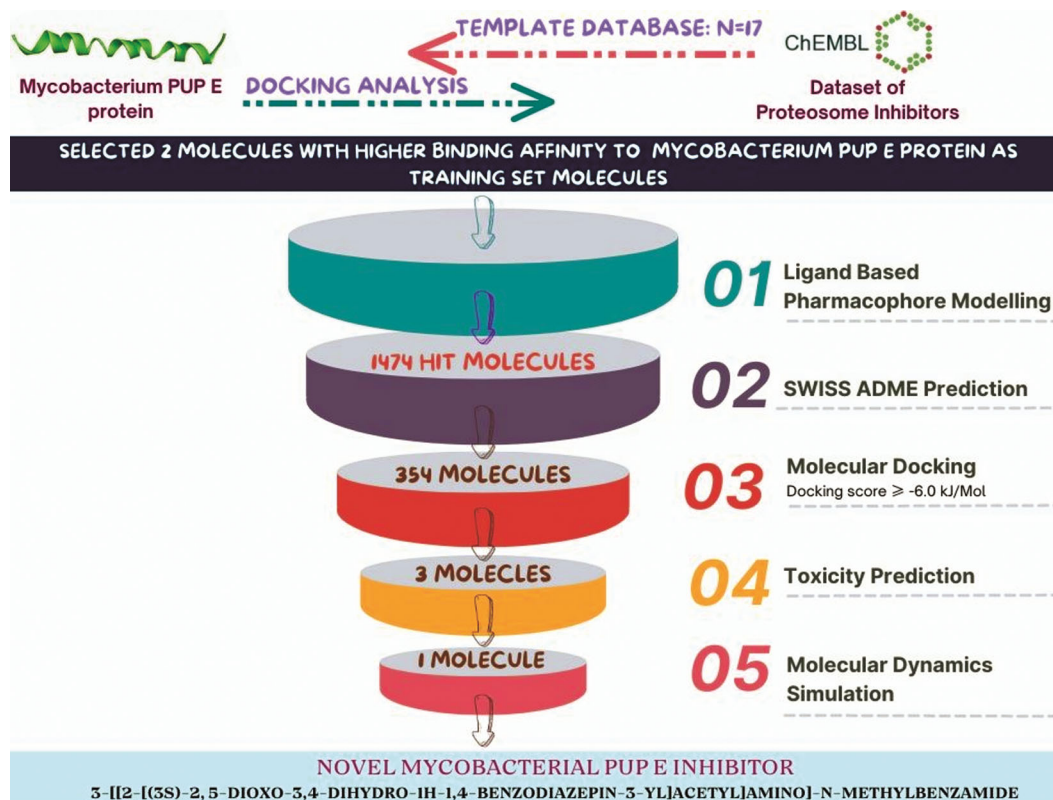




Graphical abstract

components of the PPS contribute to amino acid recycling and stress tolerance in mycobacteria. These findings established the PPS as a promising, species-specific target⁴. Direct inhibition of the *M. tuberculosis* 20S proteasome, however, is complicated by its structural and functional similarity to the human proteasome, creating a challenge in achieving sufficient pathogen selectivity while avoiding host-proteasome inhibition. Although selective *M. tuberculosis* proteasome inhibitors, including syringolin-derived scaffolds, have been identified, the potential for off-target inhibition of the human proteasome remains an important consideration. This limitation motivates exploration of upstream PPS components as alternative therapeutic targets^{5,6}.

In mycobacteria, proteasome substrates are earmarked for degradation by Pup, a 64-amino-acid prokaryotic ubiquitin-like protein that serves as a covalent degradation signal. Pup's C-terminal glutamine (PupQ) is first deamidated to glutamate by Dop (Deamidase of Pup), generating the active PupE form. PupE is then ligated to substrate proteins by PafA (Proteasome Accessory Factor A), forming an isopeptide bond between the Pup C-terminus and

lysine ϵ -amino groups in substrates. Pupylated proteins bind non-covalently to the proteasomal ATPase Mpa (Mycobacterial Proteasomal ATPase), which assembles into a hexameric ring and is thought to unfold and translocate substrates into the 20S core particle for degradation. Thus, Pup and its interactions with Dop, PafA and Mpa constitute the central recognition and targeting module of the PPS^{7,8}.

Given the critical role of Pup in survival, adaptation and virulence of *M. tuberculosis*, targeting Pup itself or its interfaces within the PPS represents an attractive, more selective strategy than directly inhibiting the 20S proteasome. Inhibitors that prevent Pup activation or ligation to substrates, or that disrupt Pup-Mpa engagement, could impair the bacterium's ability to clear damaged or unwanted proteins and increase its susceptibility to host defences and stress. Experimental studies on pupylation have shown that interfering with key steps of this pathway reduces *M. tuberculosis* survival under nutrient-deprivation and related stress conditions, further supporting Pup/PPS as a therapeutic target⁹.

Pup exists in two distinct forms: the inactive precursor PupQ and the active, glutamate-terminated form PupE, which alone is competent for interaction

with the proteasomal ATPase Mpa and for subsequent targeting of substrates to the mycobacterial proteasome. In this study, the structural model and docking calculations were explicitly performed on the PupE–Mpa complex². Accordingly, throughout the manuscript, the term “Pup” refers to the activated form PupE unless otherwise specified. Building on this definition and the central role of Pup in the PPS, our aim was to discover novel Pup-targeted inhibitors using computational methods, including 3D ligand-based pharmacophore querying, molecular docking and molecular dynamics (MD) simulations.

Materials and Methods

Identification of the structure of Prokaryotic Ubiquitin-like Protein (Pup) of *M. tuberculosis*

The current study commences by identifying the specific comprehensive identification of the three-dimensional configurations of the *M. tuberculosis* Prokaryotic Ubiquitin-like Protein (Pup), from different Pup-protein complexes found in RCSB PDB. The prokaryotic ubiquitin-like protein (Pup) of *M. tuberculosis* is encoded by the gene rv2111c⁷.

Preparation of protein target

The Pup–Mpa complex structure was obtained from the Protein Data Bank (PDB ID: 3M91), which contains the amino-terminal coiled-coil fragment of the Mpa ATPase (chains A and C) and the prokaryotic ubiquitin-like protein Pup (chains B and D). For docking, we used Pup chain B corresponding to residues 21–64 of the full-length Pup sequence (64 amino acids), as deposited in 3M91. To obtain the structurally defined Pup model used for the computational analysis, truncated construct comprising residues 21–51 were used. The docking analysis was designed to evaluate interactions involving the retained structural region rather than to model the native Glu64-dependent C-terminal recognition of full-length PupE. Further the binding site for docking was defined as the Pup–Mpa interface pocket surrounding the Pup-binding region observed in 3M91, and the AutoDock Vina grid box (25 × 25 × 25 Å) was centered on this pocket to encompass the experimentally observed interaction site¹⁰.

Evaluation of the prepared 3D structure

The conformational quality of the prepared protein structure was subsequently evaluated using The SAVES v6. 0 server (<https://saves.mbi.ucla.edu/>)

which is a Structure Validation Server that integrates different tools including ERRAT and PROCHECK for the prediction of different types of stereo chemical parameters of the protein structure. The resultant output was employed for analysing the geometry of the protein by considering both overall residues as well as residue by residue individually^{11,12}.

Identification of primary dataset

The current study initially considered the already known proteasome selective inhibitors¹³ available at ChEMBL database (Chemical Biology Database) (https://www.ebi.ac.uk/chembl/document/report_card/ChEMBL4402529/), which is a comprehensive and curated collection of bioactive chemicals and inhibitors, to prepare the dataset of molecules used to identify potential novel inhibitors targeted specifically against Pup protein. Further methodical strategies were applied for the systematic retrieval and screening of molecules which was eventually be combined to build the preliminary screening database.

Molecular docking

Molecular docking was performed using the PyRx platform (version 0.8), which implements AutoDock Vina (version 1.1.2) for virtual screening of the ligand database against the prepared *M. tuberculosis* Pup protein. The preparatory virtual screening of the database was accomplished via molecular docking studies, a computational technique that predicts the best spatial arrangement of one molecule with respect to another when they bind together to form a stable complex. The protein target was prepared as described in Section 2.2 by removing heteroatoms, including crystallographic water molecules, adding Kollman charges and polar hydrogens, and saving the structure in PDBQT format prior to docking. The library of ligand molecules was energy-minimized, protonated at physiological pH (7.4), had polar hydrogens added, and was converted to PDBQT format; the prepared protein structure and ligand library were then imported into the “Molecule” window of PyRx and defined as “Macromolecule” and “Ligands,” respectively. The search space was defined using the AutoDock Vina wizard as a cubic grid box of 25 × 25 × 25 Å with a grid spacing of 1.0 Å, centered on the Pup binding site, and docking was performed with an exhaustiveness value of 8, requesting 10 output poses (num_modes = 10) per ligand and an energy_range corresponding to binding affinities equal to or less than –6.0 kcal/mol. Binding

energies were obtained in terms of docking score kcal/mol and compared further, and the resulting protein–ligand interactions were analyzed using BIOVIA Discovery Studio Visualizer¹⁴.

Selection of template ligands

From preliminary screening of inhibitors, the best molecules having significant interaction with the Pup protein were selected based on the docking score. These ligands were studied further for their structural properties and IC₅₀ values against the proteasome using the Binding Database. The best two ligands were subsequently chosen as template compounds to construct a hypothesis-generating ligand-based pharmacophore query intended to enrich for potential Pup-binding inhibitors, rather than a fully validated Pup-specific pharmacophore model. Ligand-Based Pharmacophore Modelling was used, employing LigandScout Version 4.4 software for the construction of this pharmacophore query.

Hypothesis-generating pharmacophore query

A pharmacophore model is a spatial depiction of chemicals that have similar biological effects and is made up of elements such as hydrogen bond donors, acceptors, hydrophobic regions, and aromatic rings. Pharmacophore modelling is used to detect and define the inherent molecular constituents that contribute to the biological activity of a certain molecule. Here, LigandScout is used for the generation of pharmacophore models using the best screened template ligands (from the section 2.6) as the training set molecules. These template ligands were proteasome inhibitors docked to Pup and were employed to derive an exploratory, hypothesis-generating pharmacophore query, rather than a fully validated Pup-specific pharmacophore model. The pharmacophore generation was performed employing “Gaussian Shape Similarity” as the scoring function, with “merged and shared features” type of pharmacophore to generate a total of 10 pharmacophore queries. These 10 queries were evaluated using a set of positive controls (“actives”) and a set of negative controls (“decoys”). The best screened template molecules were grouped as actives, and decoys, molecules having similar structure to the actives but with no reported biological activity were generated using the DuD.E webserver (A Database for Useful Decoys, Enhanced; <https://dude.docking.org/>). The generated pharmacophore queries were screened against the actives and decoys databases,

and their ability to enrich actives over decoys was assessed using scoring factors including the area under the curve (AUC) and enrichment factor, which describe the sensitivity and specificity of the query within this training/decoy set. The pharmacophore query showing the most favourable enrichment behaviour was then used as a target-guided filter and screened against external databases to obtain hits enriched for compatibility with the hypothesised Pup-binding mode. These hit molecules were then classified and saved for further evaluations.

The basic requirement for this procedure is to have a set of training molecules. By entering the required parameters as mentioned in Section 2.7, a total of 10 pharmacophore queries were generated by the software, which were further assessed using the Screening window of LigandScout. Assessment of the queries with the two databases (3 actives and 204 decoys) verifies their ability to preferentially retrieve actives within this limited set; however, this enrichment is interpreted as support for the hypothesis-generating query rather than as formal validation of a Pup-specific pharmacophore model.

The best-performing pharmacophore query was then uploaded to the Pharmit web tool (<https://pharmit.csb.pitt.edu/>), an open web-based platform that allows virtual screening of chemical compounds using pharmacophore features, molecular shape and energy minimization to search large compound databases for molecules compatible with a given feature pattern¹⁵. The Pharmit server was used to screen this pharmacophore query against the MolProt database, which initially consisted of 64,824,361 conformers of 4,742,020 molecules (updated as of 05 September 2024). MolProt database screening was carried out using “Hit screening filter” constraints with the following cut-offs: molecular weight 196–500 g/mol, rotatable bonds 1–10, LogP –2 to 5, polar surface area (PSA) 1–150 Å², aromatic ring count 1–3, hydrogen bond acceptors (HBA) 1–10 and hydrogen bond donors (HBD) 1–5.

Prediction of ADME and toxicity of hit molecules

The parameters depicting pharmacokinetic properties like absorption, distribution, metabolism, and excretion of each hit molecule were further predicted using the SWISS-ADME web tool maintained by the Swiss Institute of Bioinformatics (<http://www.swissadme.ch/>). The values for different ADME parameters including molecular weight, solubility, polar surface area, number of aromatic

rings, lipophilicity etc. were predicted and the suitability of hit molecules to be developed as a prospective drug molecule was evaluated on the basis of desired cut-off values. Further the toxicity of each hit molecule was predicted using ProTox 3.0 webtool which is an artificial platform for the prediction of toxicities of small chemical or bioactive molecules (<https://tox.charite.de/protox3/>). Protox 3 is used to classify the molecule into different toxicity classes on the basis of predicted LD₅₀ values. The prediction is done along with a prediction accuracy parameter which enabled the step to be systematic to logically screen in the best molecule. The current study primarily identifies potential template ligands using which a pharmacophore is generated which was screened against external databases to get hit molecules. They were further screened in terms of toxicity and pharmacokinetic parameters. Finally, their inhibitory potential against *M. tuberculosis* Pup protein was validated.

Interactional analysis using molecular docking studies

After the initial screenings the best hit molecules were subsequently sorted and employed for checking the molecular interactional stability when complexed with the Pup protein. The 3D structure of the respective protein prepared as mentioned in section 2.2. The Multi-Ligand approach was utilized for docking and it was carried out with PyRx software using the method same as Section 2.5. The binding energy of best poses in which each candidate molecules bound to target protein were subsequently obtained in terms of docking score, which were further categorized and sorted to determine a cut-off value to screen out least interacting molecules. The 2D interaction diagram of complexes was correspondingly drawn using BIOVIA Discovery Studio Visualizer to define the key interacting residues as well as types of interactions involved. The molecules having best drug like interactions as well as docking score below the cut-off value were selected and validated by determining the interactional behaviour with the target protein in a virtual dynamic system mimicking molecular physiological environment.

Interactional behaviour analysis using Molecular Dynamics Simulation studies

The interactional behaviour of the best screened molecule with the target protein was ultimately validated using molecular dynamics (MD) simulation studies conducted with Desmond (Maestro,

Schrödinger Inc.). Simulation analysis was carried out for the apo protein and for the protein–candidate molecule complex separately, and their behaviours were compared quantitatively. The respective 3D structures in “.pdb” format were initially imported into the simulation platform and prepared using the Protein Preparation Wizard to optimize the structures, add hydrogens and correct any structural issues. A virtual system was then constructed using the System Builder, employing the OPLS force field and an explicit TIP3P water model in a cubic simulation box with a minimum distance of 1.0 nm between any protein atom and the box edge; the system was solvated and neutralized by adding water molecules and 0.15 M NaCl. Energy minimization was performed using steepest descent for 5,000 steps to relax the system and remove steric clashes. Subsequently, the system was equilibrated first under NVT conditions for 100 ps at 300 K and then under NPT conditions for 100 ps at 1 bar, with temperature controlled by the V-rescale thermostat and pressure by the Parrinello–Rahman barostat. Production MD simulations were then run for 100 ns with a timestep of 2 fs, and each system was simulated in two independent replicates. After the simulations, the Simulation Event Analysis Wizard was used to evaluate the dynamical behaviour of the structures over 100 ns in terms of root mean square deviation (RMSD), root mean square fluctuation (RMSF) and radius of gyration (Rg), which were compared and evaluated to assess the interactional behaviour of the candidate molecules with the target protein.

Results

The current study focused on identifying novel and specific inhibitor molecules against prokaryotic ubiquitin-like protein (Pup) from *M. tuberculosis*. For the docking workflow, the Pup–Mpa complex structure (PDB ID: 3M91), which represents Pup complexed with the *M. tuberculosis* proteasomal ATPase Mpa and has a reported molecular weight of 21.65 kDa¹⁶, was used as the source of Pup coordinates. In this structure, the Pup portion was identified as chain B; chain B was extracted, stripped of all crystallographic water molecules, and supplemented with polar hydrogen atoms prior to use in the virtual screening pipeline (Fig. 1a).

The quality of the prepared protein structure was evaluated using the SAVES v6.1 server, yielding an ERRAT score of 95.65, indicating good overall structural quality¹⁰. Stereochemical assessment using

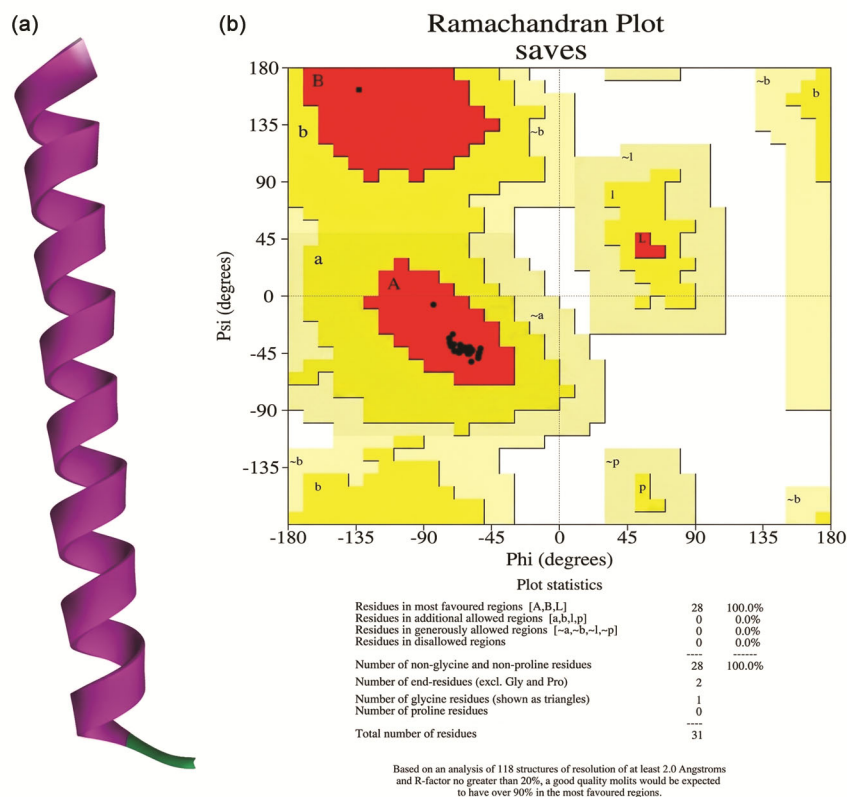


Fig. 1 — The 3D structure of PUP protein with PDB id: 3M91 (a) 3D structure of B chain 3M91, where B chain represents the Prokaryotic ubiquitin-like protein (Pup) crystal structure; and (b) Pup protein Ramachandran plot indicating the structural quality, i.e with no residues in the disallowed region, the plot shows that the 100% residues are in most favourable region indicating the stereochemical quality of the prepares 3D structure of the protein

PROCHECK (Ramachandran plot) showed no residues in disallowed regions, supporting acceptable geometry of the model¹⁷ (Fig. 1b). SAVES/ERRAT/ PROCHECK were used solely to assess the stereochemical and geometric quality of the PupE–Mpa structure and are not intended as proof that this conformation is uniquely suited for docking; accordingly, all docking-based conclusions are considered predictive and subject to experimental validation.

With the objective to find a specific inhibitor against protein Pup a primary database to start with was fundamentally required. For that objective this study considered the ChEMBL database. From the data available in the report card ChEMBL4402529, a comprehensive collection of molecules, which were the experimentally proved inhibitors of proteasomes along with their inhibitory concentration (IC₅₀) value was obtained¹³. From the obtained data the duplicates and those molecules with IC₅₀ value greater than 100nM were removed to get a more efficient set of molecules. And finally, a database of 17 best proteasome inhibitors was obtained.

This database of ligand molecules was docked against the prepared structure of Pup using the PyRx platform. The docking search space was defined as a cubic grid box centred on the Pup–Mpa interaction region observed in the Pup–Mpa complex (PDB ID: 3M91), with grid centre coordinates $x = -2.7250$, $y = -7.8782$ and $z = -8.2984$ and dimensions of $25 \times 25 \times 25$ Å. All ligand molecules were docked within this same grid box to evaluate their ability to bind at the Pup interface region and to enable a consistent comparison of their predicted binding affinities to Pup. The docking scores for the selected molecules are provided in (Table 1). Docking scores presented in Table 1 identified Mol-15 and Mol-14 as the best inhibitors from the initial ChEMBL dataset. Their interactions with key Pup residues suggest that these molecules can effectively bind to the protein, making them promising candidates for further optimization. Their 2D interaction diagram showed that the major interacting residue of Pup protein when it was docked with Mol 15 and Mol 14 were ASP 41, ASP 44, ASP 45, Glu 42 and Glu 48 (Fig. 2).

mol-id	ChEMBL ID	Molecular Weight (g/mol)	Experiment IC ₅₀ value against proteasome in nM	Docking Score (kcal/mol)
mol-1	CHEMBL4465456	604.79	97	-6.02836
mol-2	CHEMBL4473920	544.7	57	-5.86962
mol-3	CHEMBL4563115	566.72	14	-6.11259
mol-4	CHEMBL2141296	361.03	3.4	-4.5484
mol-5	CHEMBL270515	413.28	31	-5.10193
mol-6	CHEMBL4297468	586.69	0.623	-5.9654
mol-7	CHEMBL2103884	532.62	36	-5.28965
mol-8	CHEMBL371405	313.78	3.5	-4.08521
mol-9	CHEMBL4521524	600.69	4	-5.71018
mol-10	CHEMBL4443505	570.69	83	-5.09853
mol-11	CHEMBL4469397	566.72	5.5	-5.48221
mol-12	CHEMBL4592435	550.63	15	-5.81914
mol-13	CHEMBL325041	384.25	7	-4.58984
mol-14	CHEMBL451887	719.92	33	-6.30044
mol-15	CHEMBL4441524	569.7	34	-7.30622
mol-16	CHEMBL4591725	344.41	13	-4.59708
mol-17	CHEMBL4465266	445.43	9	-5.09276

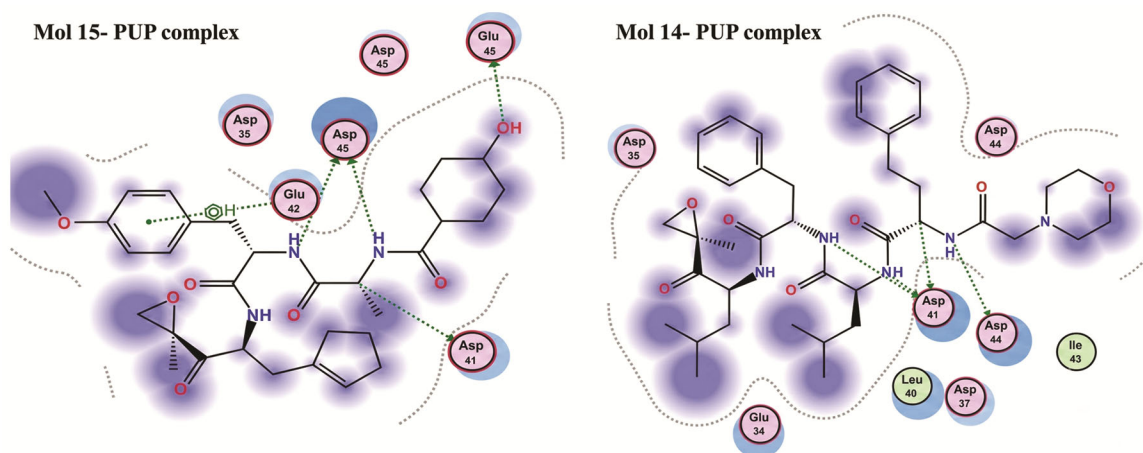


Fig. 2 — 2D interaction diagram of PUP protein complex with Mol 15 and Mol 14 respectively. (Biovia Discovery Studio Visualizing Software)

These molecules i.e., Mol 15 and Mol 14 with chemical name (IUPAC) N-[(2R)-1-[[[(2S)-1-[(2S)-3-(cyclopenten-1-yl)-1-[(2R)-2-methyloxiran-2-yl]-1-oxopropan-2-yl]amino]-3-(4-methoxyphenyl)-1-oxopropan-2-yl]amino]-1-oxopropan-2-yl]-4-hydroxycyclohexane-1-carboxamide and (2S)-4-methyl-N-[(2S)-1-[[[(2S)-4-methyl-1-[(2R)-2-methyloxiran-2-yl]-1-oxopentan-2-yl]amino]-1-oxo-3-phenylpropan-2-yl]-2-[[[(2S)-2-[(2-morpholin-4-yl)acetyl]amino]-4-phenylbutanoyl]amino]pentanamide, respectively¹⁸ were hence entitled as the template ligands in the current study, which were further employed as training set molecules in the subsequent step which was Ligand Based Pharmacophore query.

Pharmacophore modeling was performed, identifying Model 2 as the optimal model with an

AUC of 1.00 and EF of 69% (Fig. 3a). Which are considered suitable threshold values for reliable screening performance as discussed by various authors^{19,20}. This model featured seven key electronic characteristics, including three hydrogen bond acceptor sites, two hydrogen bond donor sites, and two hydrophobic interaction sites (Fig. 3b).

Finally, the best pharmacophore model, Model 2 was used as a pharmacophore query and screened against external database MolProt using Pharmit server and as a result a total of 1474 hit molecules were obtained. These 1474 molecules were examined, evaluated, and analysed for quality and 3D structure availability before being taken into consideration for a follow-up screening investigation that verifies the pharmacokinetic characteristics of the molecules. The

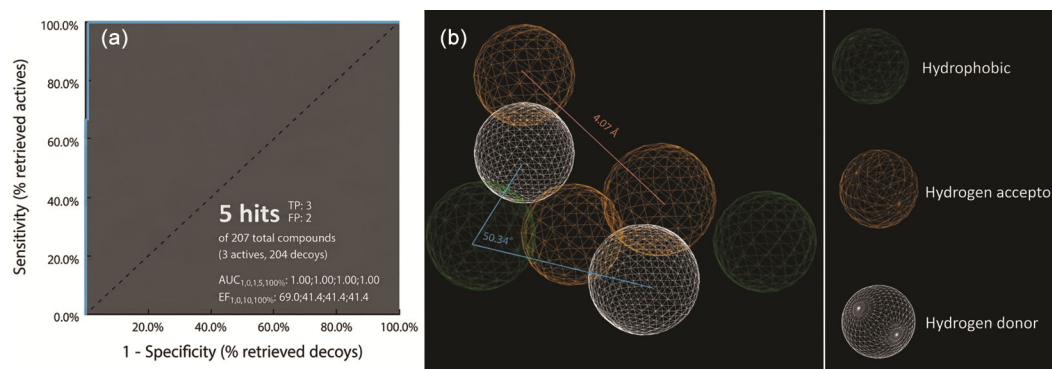


Fig. 3 — The generation of Pharmacophore model using Ligand Scout, (a) Receiver operating curve showing the specificity and sensitivity of the best pharmacophore model; and (b) the 3-D representation of the Best Pharmacophore Model (Model 2) visualized using the “Load Features” option at Pharmit Search Engine having Hydrophobic, HBA and HBD as prime characteristic features

pharmacokinetic predictions were conducted using SWISS-ADME-SWISS Drug Design Online platform (<http://www.swissadme.ch/>), which is operated and maintained by the Molecular Modelling Group of the University of Lausanne and the SIB Swiss Institute of Bioinformatics. For the screening of molecules based on SWISS ADME the current study prioritized the suitability of molecules based on certain parameters including, GI absorption, BBB permeability, water solubility, Lipinski's rule, Ghose-Veber filter, Egan filter, Muegge filter and CYP inhibition²⁰. The study considered those molecules which have moderate or high GI absorption, no BBB permeability and moderate water solubility to be screened in to next step. The screening technique further rejected compounds that inhibited more than 3 isoforms of CYP enzyme and those that violated the five filters previously described. As a result, a total of 354 molecules were selected after pharmacokinetic predictions. The detailed data for all 354 selected molecules are provided in the Supplementary Material (Suppl. Table 1). SwissADME provides computational predictions rather than experimental validation of pharmacokinetic behavior or oral efficacy. Therefore, these predicted properties require further validation through appropriate *in vitro* and *in vivo* studies.

To efficiently screen this database of molecules, molecular docking studies were conducted prior to toxicity analysis. 3D structures of 354 molecules were obtained in “.sdf” format from the Pubchem database. It was discovered that 33 compounds lacked a well-defined 3D structure in the PubChem database and were therefore excluded from subsequent studies. Molecular docking was performed using PyRx software for multi-ligand docking against the

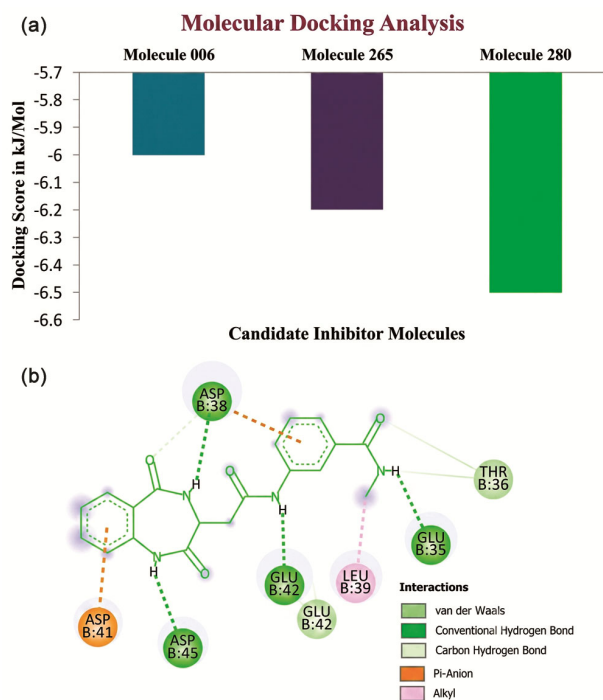


Fig. 4 — The prediction of molecular interactional behaviour of best molecules with PUP protein, (a) The Docking score of hit molecules with best molecule; Mol280 having the lowest docking score; and (b) 2-Dimensional interaction diagram of PUP-Mol280 in a static environment that includes van der Waals interactions, hydrogen bonding, pi bonding and alkyl bonding, respectively

prepared *M. tuberculosis* Pup protein. The molecules were sorted based on the docking score and the study picked the molecules with a docking score equal to or less than -6.0 kcal/mol. to be considered for further studies and hence, a total of 3 molecules encoded as Molecule 280, Molecule 265 and Molecule 006 were selected as the best molecules (Fig. 4a). To further prioritize these hits, their predicted physicochemical and ADME properties were evaluated using

Table 2 — Predicted physicochemical, ADME/drug-likeness and toxicity parameters (SwissADME and ProTox-III) of the three final hit molecules (006, 280 and 265) obtained from docking against *M. tuberculosis* Pup

Parameter	Molecule 6	Molecule 280	Molecule 265
Molecule ID	Molecule 6	Molecule 280	Molecule 265
Formula	C ₁₄ H ₂₂ N ₄ O ₃ S	C ₁₉ H ₁₈ N ₄ O ₄	C ₂₁ H ₂₃ ClN ₄ O ₄ S
IUPAC Name	N-[(Z)-(3,4-dihydroxyphenyl)methylideneamino]-4-methylbenzenesulfonamide	3-[[2-[(3S)-2,5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl]acetyl]amino]-N-methylbenzamide	3-chloro-N-[(2R)-1-[2-[2-(1H-indol-3-yl)acetyl]hydrazinyl]-3-methyl-1-oxobutan-2-yl]benzenesulfonamide
PubChem CID	135641595	91637624	26009832
MW (g/mol)	326.41	366.37	462.95
# H-bond donors	3	4	4
# H-bond acceptors	4	4	5
Rotatable bonds	10	6	10
TPSA (Å ²)	128.43	116.40	128.54
Fraction Csp ³	0.57	0.16	0.24
Consensus LogP	1.16	0.57	2.70
ESOL solubility class	Soluble	Soluble	Moderately soluble
GI absorption	High	High	High
BBB permeant	No	No	No
P-gp substrate	Yes	Yes	Yes
CYP inhibition profile	No inhibition was predicted for any of the five CYP isoforms.	No inhibition was predicted for any of the five CYP isoforms.	2C9, 2D6, 3A4 were predicted to be inhibited out of five CYP isoforms.
Lipinski violations	0	0	0
Veber violations	0	0	0
Bioavailability score	0.55	0.55	0.55
Lead-likeness violations	1	1	2
Synthetic accessibility	3.68	3.25	3.91
Tox class / LD ₅₀ (mg/kg)	3 / 250	4 / 1500	3 / 300

SwissADME. The key parameters relevant to oral drug-likeness are summarized in (Table 2).

The 2-D interaction diagram after docking of the best inhibitor Mol 280 was consequently analysed to reveal ample amount of binding with the interacting amino acids of the protein the major ones included Asp B:38, Asp B: 45, Glu B:42 and Glu B:35 with conventional Hydrogen bond , Asp B:38, Asp B: 41, Pi- Anion bond and Leu B: 39 with Alkyl bond interaction in a static environment contributing to a binding energy of -6.5 kcal/mol (Fig. 4b). These interacting forces are stable and normally seen in any kind of drug target protein complex hence it is discussed that probability of this complex to behave as a drug-protein complex is highly likely as interpreted by other authors²¹.

The predicted best inhibitor molecule, which was encoded as Molecule 280 and had a chemical nomenclature of “3-[[2-[(3S)-2, 5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl] acetyl]amino]-N-methylbenzamide,” was hence finalized (Fig. 5).

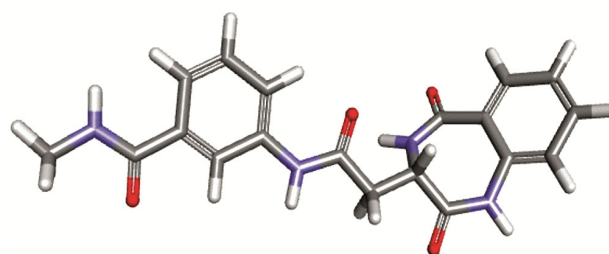


Fig. 5 — 3D structure of the best identified PUP inhibitor, 3-[[2-[(3S)-2,5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl]acetyl]amino]-N-methylbenzamide of *M. tuberculosis*

This was proceeded by checking the toxicity analysis using ProTox III webserver, which predicted the molecule to be in a Class 4 for Oral Toxicity which is defined as non-fatal and non-toxic, slightly harmful if swallowed according to the Globally Harmonized System of classification and labeling of chemicals²² (GHS) with an LD₅₀ value of 1500 mg/kg. The extensive analysis employing molecular similarity, fragment propensities, and

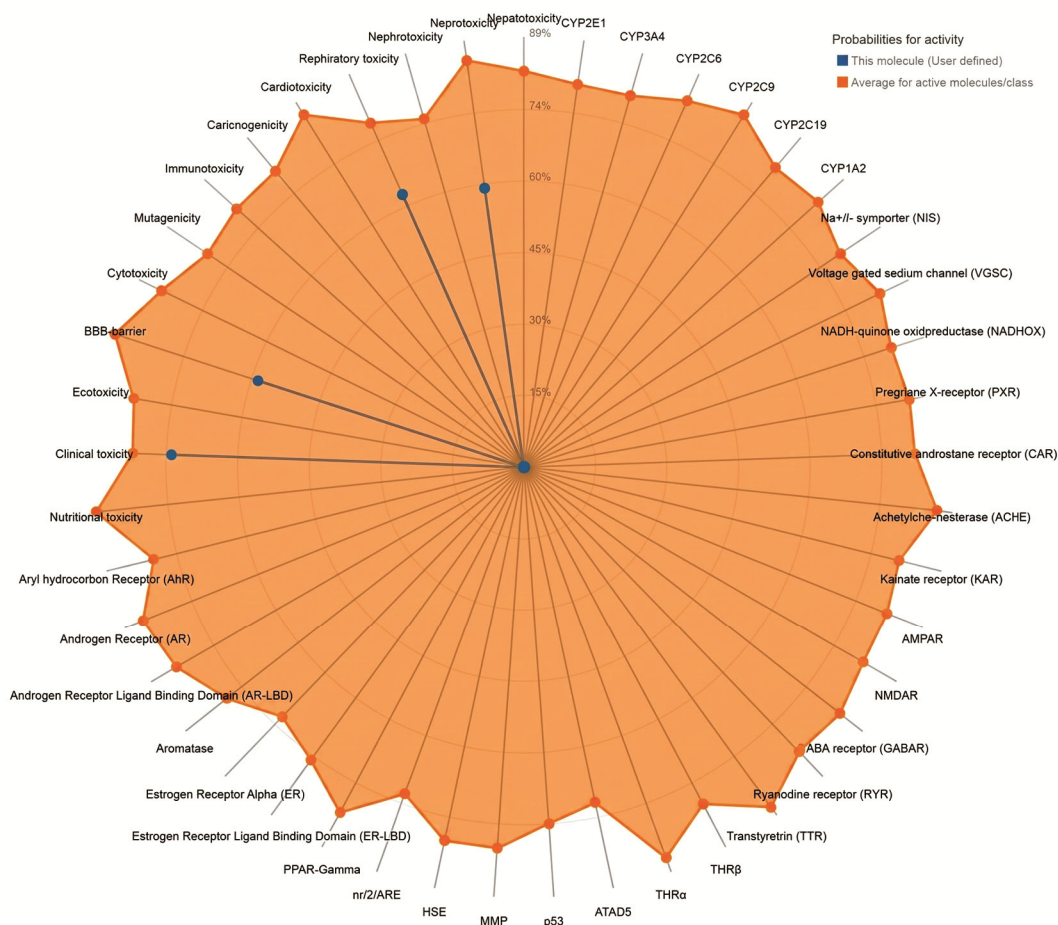


Fig. 6 — Predicted toxicity radar chart of 3-[[2-[(3S)-2,5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl]acetyl]amino]-N-methylbenzamide, showing inactive toxicity range of all predicted toxicity end points in the radar, implicating its safer application as a drug molecule

machine learning based predictions revealed that the molecule is inactive for toxicity endpoints including; clinical toxicity, BBB barrier, respiratory toxicity and neurotoxicity (Fig. 6).

The final validation of the identified inhibitor was conducted using molecular dynamics simulation studies using the Desmond Maestro tool provided by Schrödinger Software²³ by importing the 3D structure of the protein alone and the Protein-Molecule 280 complex into this platform, which is compatible with the PDB format. The whole simulation was performed in a virtual system similar to a physiological system, built under the OPLS force field, with ions and atoms inserted. The minimization run was performed prior to the beginning of the production MD run to achieve stable environment for dynamic simulation of both Pup protein alone and Pup-Molecule 280 complex (Fig. 7a). The MD simulation trajectories were subsequently analysed to examine the dynamic

behaviour of protein and protein complex exhibited for the entire simulation of 100ns. RMSD time series plot revealed that the overall behaviour of backbone atoms in both apoprotein (protein alone, shown in green color) and holoprotein (protein complex, shown in purple color) were adopting a stable behaviour with the average RMSD value for apoprotein for 100 ns as 3.98 Å which was decreased to 3.55 Å for holoprotein. The holoprotein showed a significant decrease in RMSD after 80 ns of simulation run, which clearly indicated the attainment of structural stability with the progression of time (Fig. 7b).

The RMSF plot also revealed a similar trend where holoprotein (purple) showed a steep peak at the 40th residue indicating structural non-favourability but other residues were having comparatively stable trend signifying lower fluctuation in the structure of apoprotein even with binding of the inhibitor molecule (Fig. 7c)

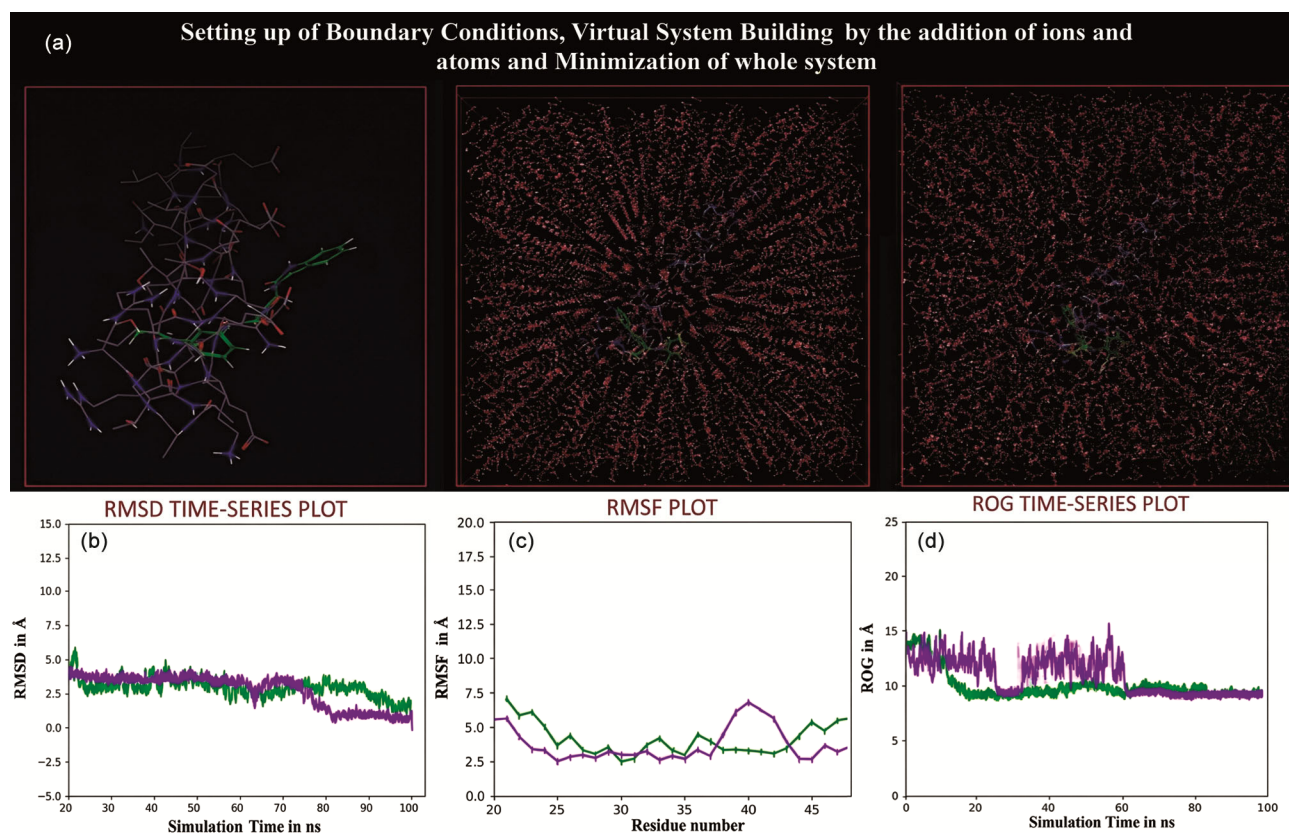


Fig. 7 — The Virtual Molecular Dynamics Simulation Studies (a) Setting up of boundary conditions, virtual system building by the addition of ions and atoms and final minimization of whole system which mimics a physiological system; (b) RMSD Time series Plot showing the much stabled and aligned trend of RMSD value with in a time limit of 100 ns implicating the stability of complex; (c) RMSF plot showing a steep peak at 40th residue in the complex, while other residues maintain a stable pattern; and (d) Radius of Gyration plot, showing a very fluctuating plot up to 60 ns which gets stabilized afterwards without any deviation till 100 ns. Green line represents apoprotein and purple colour represents holoprotein

A similar observation was made in the ROG, where the holoprotein (purple) showed a higher rate of fluctuations at the start of simulation, but the fluctuation adopted a rather stable plot after a time period of 60 ns, and it was maintained till the end of the simulation (Fig. 7d).

Discussion

Scientists around the globe are continuously searching for novel biomolecular targets and drugs to deal with infectious diseases, especially in the face of emerging resistance of pathogens to existing antibiotics^{17,24,25}. Computational studies and analysis play crucial roles in modern drug design, in exploring novel drugs from natural resources and in repurposing existing drugs for various human diseases^{26,27}. In the context of tuberculosis, the mycobacterial proteasome and its Pup-proteasome system (PPS) have emerged as attractive targets; however, directly inhibiting the 20S proteasome with agents such as bortezomib or

carfilzomib raises concerns about toxicity to the human proteasome, motivating strategies that instead focus on enzymes and interfaces unique to *M. tuberculosis*. In this study we aimed to identify small-molecule inhibitors that target the prokaryotic ubiquitin-like protein Pup and its engagement with the Mpa ATPase, thereby disrupting PPS-mediated protein quality control as a potential anti-tubercular mechanism. To do this, we first retrieved and prepared the chain B of Pup-Mpa complex structure (PDB ID 3M91). The present study focused only on the structured Pup region available in the crystal structure and therefore does not represent the complete-length Pup sequence. This limitation will be addressed in future studies by incorporating the missing residues and evaluating the complete PupE structure, particularly its C-terminal region and recognition mechanism and used ChEMBL-derived proteasome inhibitors as guide ligands for docking and ligand-based pharmacophore generation, creating

a hypothesis-generating pharmacophore query that captures key hydrogen-bonding and hydrophobic features at the Pup interface. The best-performing pharmacophore model (Model 2), which showed favourable enrichment of actives over decoys (AUC and EF values), was then applied to the MolProt database and combined with SwissADME filters (GI absorption, BBB permeability, Lipinski's rule) to narrow 1,474 hits down to 354 candidates with acceptable predicted drug-likeness, after which AutoDock Vina docking highlighted Molecules 280, 265 and 006 as top-scoring Pup binders, with Molecule 280 showing the lowest docking score and forming consistent interactions with acidic residues such as Asp B:38, Asp B:45, Glu B:42 and Glu B:35 at the Pup-Mpa interface. Subsequent MD simulations of apo Pup versus the Pup-Mol 280 complex were interpreted using RMSD, RMSF and radius of gyration, demonstrating that both systems remained structurally stable over 100 ns, that interface residues were relatively rigid compared with terminal and loop regions, and that the global compactness of the complex was preserved, all of which supports the structural plausibility and dynamic stability of the predicted Pup-Mol 280 binding mode²⁸. *In silico* toxicity assessment with ProTox III further suggested that Molecule 280 has a relatively favourable predicted safety profile (toxicity Class 4, LD₅₀ $\approx$ 1500 mg/kg), although experimental toxicity and efficacy studies will be required to confirm these predictions. Taken together, these pharmacophore-guided screening, docking, MD and toxicity results do not yet prove therapeutic efficacy but do justify prioritizing 3-[[2-[(3S)-2,5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl]acetyl]amino]-N-methylbenzamide (Molecule 280) as a promising Pup-interface binder for future biochemical and biological validation against the Pup-proteasome system of *M. tuberculosis*.

Despite the valuable insights provided by the MD simulations, some limitations should be acknowledged. Although two independent 100-ns simulation replicates were performed, the finite simulation time may not capture slow conformational changes, rare ligand-unbinding events, or all relevant binding modes. The results also depend on the selected force field, ligand parameters, protonation states, solvent model, and starting docking pose. Therefore, the RMSD, RMSF, and Rg results support the predicted complex stability under the simulated conditions but do not confirm experimental binding

affinity, proteasome inhibition, selectivity, or antimycobacterial activity. Experimental validation is required to confirm these computational findings.

It is important to note that the compounds identified in this study, including Mol280, represent computationally prioritized candidates rather than experimentally validated inhibitors. Their predicted binding modes and potential to disrupt the Pup-Mpa interaction are based on virtual screening, molecular docking, pharmacophore mapping, and molecular-dynamics simulations. Experimental validation is therefore required to confirm direct target binding, disruption of the Pup-Mpa interaction, target selectivity, antimycobacterial activity, and cytotoxicity before these compounds can be considered as potential anti-tubercular leads.

Conclusion

The rising prevalence of drug-resistant tuberculosis, particularly in high-burden nations, highlights the critical need for improved diagnostic, treatment and management strategies. Although several therapeutic development approaches exist and some proteasome-targeting agents have shown promising activity *in vitro* and in animal models, their significant side effects and lack of species selectivity limit their suitability for routine use against *M. tuberculosis*. In this context, we focused on the prokaryotic ubiquitin-like protein Pup, which is unique to bacteria, and employed diverse computational methods (pharmacophore querying, molecular docking, molecular dynamics simulations and *in silico* ADME/toxicity assessment) to generate and prioritize candidate Pup-targeted compounds. Among the screened molecules, 3-[[2-[(3S)-2,5-dioxo-3,4-dihydro-1H-1,4-benzodiazepin-3-yl]acetyl]amino]-N-methylbenzamide (Molecule 280) emerged as a preliminary computational hit that is predicted to bind at the Pup-Mpa interface with favourable docking scores, stable MD behaviour and acceptable pharmacokinetic and toxicity profiles. However, molecular docking and MD simulations are inherently hypothesis-generating, and docking scores alone are not reliable evidence of binding affinity or inhibitory activity; therefore, Molecule 280 should be regarded as a computationally predicted lead rather than a validated Pup inhibitor. Its proposed activity within the mycobacterial Pup-proteasome pathway remains to be tested, and further biochemical assays, cellular studies and *in vivo* infection models will be required

to establish whether targeting Pup with this chemotype can translate into safe and effective anti-tubercular therapy.

Acknowledgement

This work was supported by a major research project No. 37(1714)/18/EMR-II, awarded to CRP by Council of Scientific and Industrial Research (CSIR), India.

Conflict of interest

All authors declare no conflict of interest.

References

- World Health Organization, *Global tuberculosis report 2025* (World Health Organization, Geneva) 2025, available at <https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2025>.
- Bode NJ & Darwin KH, The Pup-proteasome system of mycobacteria. *Microbiol Spectr*, 2 (2014) MGM2-0008-2013.
- Cerda-Maira FA & Darwin KH, The *Mycobacterium tuberculosis* proteasome: More than just a barrel-shaped protease. *Microbes Infect*, 11 (2009) 1150.
- Gandotra S, Schnappinger D, Monteleone M, Hillen W & Ehrt S, *In vivo* gene silencing identifies the *Mycobacterium tuberculosis* proteasome as essential for the bacteria to persist in mice. *Nat Med*, 13 (2007) 1515.
- Darwin KH, Ehrt S, Gutierrez-Ramos JC, Weich N & Nathan CF, The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science*, 302 (2003) 1963.
- Totaro KA, Barthelme D, Simpson PT, Jiang X, Lin G, Nathan CF, Sauer RT & Sello JK, Rational design of selective and bioactive inhibitors of the *Mycobacterium tuberculosis* proteasome. *ACS Infect Dis*, 3 (2017) 176.
- Samanovic MI, Li H & Darwin KH, The Pup-proteasome system of *Mycobacterium tuberculosis*. In: *Prokaryotic Ubiquitin-Like Proteins*, edited by KH Darwin (Springer, Dordrecht) 2013, 267.
- Striebel F, Imkamp F, Sutter M, Steiner M, Mamedov A & Weber-Ban E, Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat Struct Mol Biol*, 16 (2009) 647.
- DeMott CM, Girardin R, Cobbert J, Reverdatto S, Burz DS, McDonough K & Shekhtman A, Potent inhibitors of *Mycobacterium tuberculosis* growth identified by using in-cell NMR-based screening. *ACS Chem Biol*, 13 (2018) 733.
- Wang T, Darwin KH & Li H, Binding-induced folding of prokaryotic ubiquitin-like protein on the mycobacterial proteasomal ATPase targets substrates for degradation. *Nat Struct Mol Biol*, 17 (2010) 1352.
- Colovos C & Yeates TO, Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci*, 2 (1993) 1511.
- Laskowski RA, MacArthur MW & Thornton JM, PROCHECK: Validation of protein-structure coordinates. In: *International Tables for Crystallography*, edited by E Arnold, DM Himmel & MG Rossmann (International Union of Crystallography, Chester) 2012, 684.
- Xi J, Zhuang R, Kong L, He R, Zhu H & Zhang J, Immunoproteasome-selective inhibitors: An overview of recent developments as potential drugs for hematologic malignancies and autoimmune diseases. *Eur J Med Chem*, 182 (2019) 111646.
- Dallakyan S & Olson AJ, Small-molecule library screening by docking with PyRx. *Methods Mol Biol*, 1263 (2015) 243.
- Sunseri J & Koes DR, Pharmit: Interactive exploration of chemical space. *Nucleic Acids Res*, 44 (2016) W442.
- Wang T, Darwin KH & Li H, Binding-induced folding of prokaryotic ubiquitin-like protein on the mycobacterial proteasomal ATPase targets substrates for degradation. *Nat Struct Mol Biol*, 17 (2010) 1352.
- Laskowski RA, MacArthur MW, Moss DS & Thornton JM, PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr*, 26 (1993) 283.
- Khan ML & Stewart AK, Carfilzomib: A novel second-generation proteasome inhibitor. *Future Oncol*, 7 (2011) 607.
- Jade DD, Pandey R, Kumar R & Gupta D, Ligand-based pharmacophore modeling of TNF- α to design novel inhibitors using virtual screening and molecular dynamics. *J Biomol Struct Dyn*, 40 (2022) 1702.
- Nivya RM & Joy A, *In silico* discovery of novel calcineurin inhibitors using ligand-based 3-D pharmacophore modelling and molecular dynamics simulation. *J Biomol Struct Dyn*, 42 (2024) 12608.
- Rashdan H & Abdelmonsef A, *In silico* study to identify novel potential thiadiazole-based molecules as anti-COVID-19 candidates by hierarchical virtual screening and molecular dynamics simulations. *Struct Chem*, 34 (2023) 1351.
- Banerjee P, Kemmler E, Dunkel M & Preissner R, ProTox 3.0: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res*, 52 (2024) W513.
- Bhachoo J & Beuming T, Investigating protein-peptide interactions using the Schrödinger computational suite. *Methods Mol Biol*, 1561 (2017) 235.
- Srivastava J, Singh S & Sinha R, *In silico* evaluation of potential plant-based tyrosinase inhibitors for cosmetic and pharmaceutical applications. *Biotechnol Lett*, 47 (2025) 26.
- Upert G, Luther A, Obrecht D & Ermer P, Emerging peptide antibiotics with therapeutic potential. *Med Drug Discov*, 9 (2021) 100078.
- Álvarez-Pérez A, Prieto-Santamaría L, Casas AI, Loscalzo J & Rodríguez-González A, Navigating the computational landscape for drug repurposing. *Annu Rev Pharmacol Toxicol*, 66 (2026) 149.
- Rapicavoli RV, Alaimo S, Ferro A & Pulvirenti A, Computational methods for drug repurposing. *Adv Exp Med Biol*, 1361 (2022) 119.
- Shah M, Sarfraz A, Shehroz M, Perveen A, Jaan S, Zaman A, Nishan U, Moura AA, Ullah R, Iqbal Z & Ibrahim MA, Computer-aided rational design of a mRNA vaccine against Guanarito mammarenavirus. *Biotechnol Lett*, 47 (2025) 2.