Journal of Medical Science, Biology, and Chemistry (JMSBC)

ISSN: 3079-2576 (Online)

Volume 2 Issue 2, (2025)

https://doi.org/10.69739/jmsbc.v2i2.918

https://journals.stecab.com/jmsbc

Published by
Stecab Publishing

Research Article

In Silico Identification of Natural Compounds Targeting the Essential ClpP1P2 Protease in *Mycobacterium tuberculosis*

*[1]Omoyele Olatunji Alli, [1]Olayinka Afeez Bankole, [1]Al-islam Abimbola Balogun, [2]Moyinoluwa Tolulope Abioye

About Article

Article History

Submission: July 22, 2025
Acceptance : August 25, 2025
Publication : September 02, 2025

Keywords

ClpP1P2 Protease, Drug-Resistant Tuberculosis, Molecular Docking, Neodiospyrin, Pharmacokinetic Modeling

About Author

[1] Department of Chemical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria

[2] Department of Chemistry, University of Medical Sciences, Ondo, Ondo State, Nigeria

Contact @ Omoyele Olatunji Alli
omoyelealli254@gmail.com

ABSTRACT

Drug-resistant strains of tuberculosis (TB) is caused by *Mycobacterium tuberculosis* and are making the disease a global health emergency. This calls for new treatments that target unidentified bacterial pathways. In order to find natural inhibitors of the ClpP1P2 protease that is a crucial virulence factor for mycobacterial survival, our study used computational techniques. We used molecular docking to screen 100 compounds against the ClpP1P2 structure (PDB: 4U0G) while giving preference to ligands with higher binding affinities than reference medications (ethambutol, isoniazid). To evaluate therapeutic potential, we subjected top hits to target fishing, toxicity prediction and pharmacokinetic profiling. Our top two hit compounds were neodiospyrin and arbutin respectively. According to physiologically-based pharmacokinetic modelling, neodiospyrin was the best candidate due to its remarkable binding energy (−13.28 kcal/mol), ligand efficiency (0.474), and intracellular lung accumulation. Also, Arbutin showed good safety but restricted tissue penetration. Through Prediction of Activity Spectra of substances (PASS) analysis, both compounds demonstrated complementary biological activities like modulation of apoptosis and anti-inflammatory effects. Target fishing indicated possible human off-target interactions (GPCRs for arbutin, kinases for neodiospyrin) but this requires further experimental validation. Our findings presented and demonstrated the potential of ClpP1P2 inhibition for tuberculosis treatment, with neodiospyrin identified as a key candidate for further improvement against drug-resistant strains.

Citation Style:

Alli, O. O., Bankole, O. A., Balogun, A. A., & Abioye, M. T. (2025). In Silico Identification of Natural Compounds Targeting the Essential ClpP1P2 Protease in Mycobacterium tuberculosis. *Journal of Medical Science, Biology, and Chemistry, 2*(2), 125-144. https://doi.org/10.69739/jmsbc.v2i2.918

1. INTRODUCTION

Tuberculosis (TB) is caused by *Mycobacterium tuberculosis* and it persists as a formidable global health challenge. There's an estimated 10.8 million new cases and 1.25 million deaths in 2023, allowing it to reclaim its position as the leading cause of death from a single infectious agent, surpassing COVID-19 in the process (World Health Organization, 2024). Approximately 1.7 billion individuals, or say 23% of the global population are estimated to harbour latent TB infection (LTBI). This serves as a reservoir for future active cases, especially when considering immunocompromised populations (Houben & Dodd, 2016). The disease disproportionately burdens low- and middle-income countries where limited access to healthcare infrastructure and diagnostic resources hampers effective TB control (World Health Organization, 2024). Socioeconomic factors like poverty and malnutrition, further worsen TB transmission, with 30 high-burden countries; led by India (26%), Indonesia (10%), and China (6.8%), accounts for 87% of global cases (World Health Organization, 2024). The rise of multidrug-resistant (MDR) and extensively drug-resistant (XDR) TB strains has intensified these challenges. In 2023 alone, approximately 400,000 new cases of MDR or rifampicin-resistant (RR) TB were reported which constitutes 3.7% of all new TB cases, with resistance rates of 3.2% among new cases and 16% among previously treated cases (World Health Organization, 2024). Current treatment regimens for MDR and XDR TB are protracted, often exceeding 18 months, and are associated with significant toxicity, high costs, and suboptimal success rates frequently below 60% (Zumla *et al.*, 2013).

By all means, the ClpP1P2 protease complex has emerged as a promising target for developing new anti-TB drugs due to its essential role in *M. tuberculosis* protein homeostasis. Unlike most bacteria, which possess a single ClpP protease, *M. tuberculosis* encodes two subunits, ClpP1 and ClpP2, that form a heterotetradecameric complex comprising two heptameric rings (Raju *et al.*, 2012). Indeed, this complex in collaboration with ATP-dependent chaperones ClpX and ClpC1, degrades misfolded or damaged proteins and ensures bacterial survival under stress conditions such as host immune responses or antibiotic pressure (Ollinger *et al.*, 2012). Inhibition of ClpP1P2 disrupts this protein quality control that then leads to the accumulation of toxic polypeptides and subsequent bacterial cell death, offering a novel mechanism to combat both drug-susceptible and drug-resistant TB strains (Raju *et al.*, 2012). The structural asymmetry of the ClpP1P2 complex with distinct binding interfaces for ClpX and ClpC1, further distinguishes it from other bacterial Clp proteases and enhances its potential as a selective drug target (Leodolter *et al.*, 2015).

Even more important, the critical role of ClpP1P2 in *M. tuberculosis* viability has been confirmed through genetic and biochemical studies, which demonstrate that depletion of either ClpP1 or ClpP2 results in rapid bacterial death both in vitro and during infection (Raju *et al.*, 2012; Ollinger *et al.*, 2012). Notably, the acyldepsipeptide (ADEP) antibiotics, which bind to ClpP1P2 and dysregulate its activity, have shown potent anti-TB activity. This supports its therapeutic potential (Li *et al.*, 2010), but that is not to say there are no challenges. Challenges remain in developing ClpP1P2 inhibitors with favorable pharmacokinetic profiles and minimal toxicity which necessitates advanced screening approaches to identify viable candidates.

Computational methods like molecular docking have revolutionized drug discovery by enabling the rapid screening of large compound libraries against specific targets. Docking studies facilitate the prioritization of compounds for experimental validation, significantly reducing the time and cost associated with traditional drug development by predicting binding affinities and molecular interactions (Meng *et al.*, 2011). Recent applications of molecular docking in TB research have identified promising inhibitors for targets such as InhA and DprE1, demonstrating the power of this approach in addressing drug resistance (Ekins *et al.*, 2017). Nonetheless, the application of docking to ClpP1P2 remains underexplored and presents an opportunity to identify novel inhibitors tailored to this protease. This study employs molecular docking to screen a curated library of 100 compounds, including 97 literature-derived antitubercular agents and three controls (isoniazid, ethambutol and ADEP), against the ClpP1P2 protease complex (PDB ID: 4U0G). We aim to identify high-affinity inhibitors with potential anti-TB activity by evaluating binding affinities, ligand efficiencies, and key interactions within the catalytic pocket. We evaluate the pharmacokinetic and toxicological profiles of top candidates using ADMET analysis, predict their biological activity, carry out pharmacokinetic simulations and target fishing analyses. This study aims to identify high-affinity ClpP1P2 inhibitors through molecular docking, characterize their interactions with the ClpP1P2 catalytic pocket, and assess drug-likeness and stability to prioritize candidates. These efforts contribute to developing novel anti-TB therapeutics to combat MDR and XDR TB.

2. LITERATURE REVIEW

2.1. Disease of interest

Tuberculosis (TB) is driven by *Mycobacterium tuberculosis* and continues to challenge global health systems with its host-pathogen interactions and therapeutic hurdles. Recent scholarship has shifted focus toward understanding how the bacterium manipulates host immunity and how emerging strategies might counter its resilience. The ability of *M. tuberculosis* to subvert host defenses remains a cornerstone of its pathogenicity. Once inhaled, the bacilli exploit alveolar macrophages as a niche. It then employs mechanisms like the inhibition of autophagosome maturation to persist intracellularly (Ahmad *et al.*, 2022). Recent studies talk about the role of the bacterial protein kinase G (PknG) in disrupting host autophagy by phosphorylating key regulatory proteins thus shielding the pathogen from degradation (Ge *et al.*, 2022). This facilitates survival and even contribute to the establishment of latency, a state where the bacterium remains dormant yet viable. Latency poses a significant barrier to eradication as it allows *M. tuberculosis* to evade both immune surveillance and conventional drugs, with reactivation often occurring years later under conditions of immune stress (Chandra *et al.*, 2022).

Therapeutic innovation offers a counterpoint to these challenges as recent advances show good promise against both active and latent TB. Host-directed therapies (HDTs) (Figure 1), for example, aim to bolster immune clearance by targeting

pathways like the mTOR signaling axis which *M. tuberculosis* hijacks to suppress autophagy (Tian *et al.*, 2025). Early clinical trials of rapamycin analogs as adjuncts to standard regimens suggest improved bacterial clearance even as scalability remains a concern in resource-limited settings (Tasneen *et al.*, 2021). In parallel, the development of shorter regimens for drug-resistant TB like the BPaL combination (bedaquiline, pretomanid, and linezolid), has reduced treatment duration to six months for some patients (Putra *et al.*, 2023). This is a marked improvement over the traditional 18-24 months but there are challenges. The regimen's toxicity, including peripheral neuropathy from linezolid shows the need for safer alternatives (Putra & Adiwinoto, 2023).

Reflecting on these advances, it is clear that TB control hinges on integrating insights from host-pathogen biology with practical therapeutic solutions. The pathogen's adaptability necessitates an approach of combining improved diagnostics like next-generation sequencing for rapid resistance profiling, with treatments that address both active and latent states. Even though progress is evident, the field must grapple with translating these innovations into equitable, real-world impact, especially in high-burden regions where infrastructure lags.

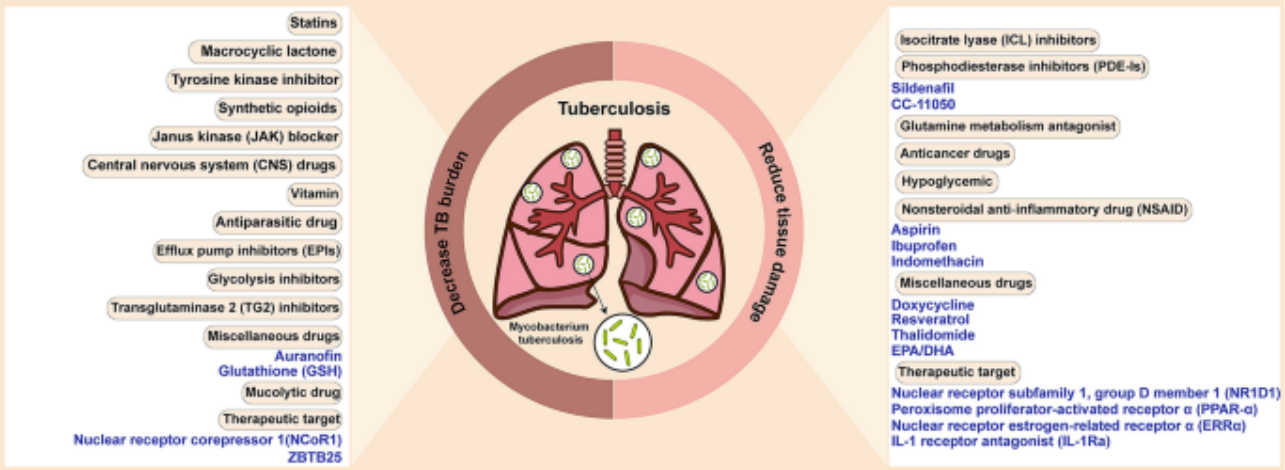

Figure 1. Main current HDTs used in tuberculosis (Tian *et al.*, 2025).

2.2. Life cycle of *mycobacterium tuberculosis* in humans

Mycobacterium tuberculosis initiates its life cycle in humans (Figure 2) through airborne transmission. Infectious droplets from individuals with active pulmonary tuberculosis are inhaled which then deposits bacilli in the lung alveoli (Coleman *et al.*, 2022). Once in the lungs, alveolar macrophages phagocytose the bacteria but *M. tuberculosis* employs mechanisms to evade intracellular killing, notably by inhibiting phagosome-lysosome fusion through the action of its lipid-rich cell wall and secreted effectors (Pieters, 2008). This allows the pathogen to survive and replicate within macrophages as it establishes an early infection that can progress to active disease or be contained by the host immune response.

In most individuals, the immune system mounts a robust response. It recruits T cells and forms granulomas which are organized structures of immune cells that encase the bacteria to limit their spread (Rahman, 2024). This containment results in latent TB infection (LTBI), where the bacilli remain dormant, often for decades, without causing symptoms (Barry *et al.*, 2009). Reactivation to active TB disease can occur if immune control weakens. Sabotages such as during HIV infection, diabetes, or immunosuppressive therapy then leads to symptomatic disease and potential transmission through respiratory secretions (Ronacher *et al.*, 2015).

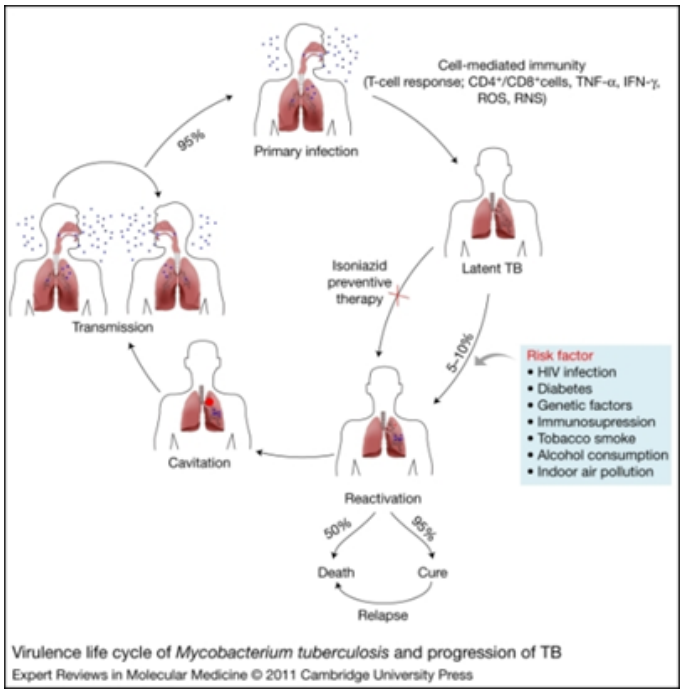

Figure 2. Life cycle of *mycobacterium tuberculosis* and progression of TB in humans (Kumar *et al.*, 2011).

2.3. *In Silico* studies in TB drug discovery

2.3.1. Molecular docking

Put simply, molecular docking is often described using the "lock and key" analogy, where a ligand (the key) is thought to fit into a specific binding site on a protein (the lock) based on shape complementarity. This analogy is an oversimplification however, because real-world protein-ligand interactions are more dynamic and even involve conformational changes in both the protein and ligand. A more suitable definition would be to describe it as a pivotal computational technique in structure-based drug design, predicting the optimal binding orientation of a small molecule (ligand) to a target protein (receptor) to form a stable complex (Meng *et al.*, 2011) (Figure 3). Docking estimates binding affinities and identifies key interaction sites, that then offers us insights into potential drug mechanisms by simulating ligand-receptor interactions. The process relies on high-resolution protein structures that are often sourced from the Protein Data Bank (PDB). It employs algorithms to explore ligand conformations within the receptor's binding pocket. Scoring functions, which assesses factors like hydrogen bonding, hydrophobic interactions, and shape complementarity, rank the resulting poses to predict binding strength (Kitchen *et al.*, 2004). Despite problems like accounting for solvent effects or protein flexibility, docking's ability to screen large compound libraries rapidly makes it important for tuberculosis (TB) drug discovery as novel inhibitors are urgently needed to combat *Mycobacterium tuberculosis* (Mtb) resistance (World Health Organization, 2024). Its integration with experimental validation enhances its utility in prioritizing compounds for further development.

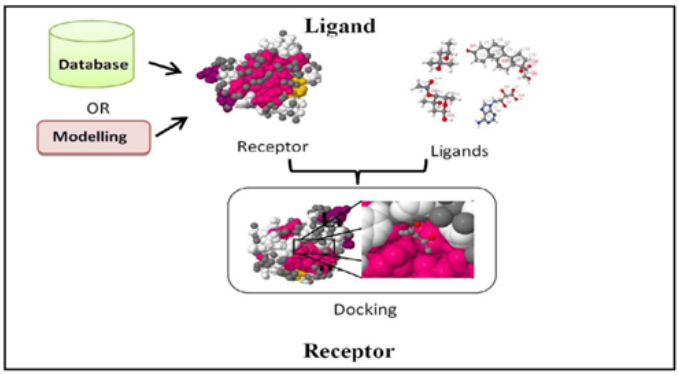

Figure 3. Molecular docking flow chart (Eweas *et al.*, 2014).

2.3.2. Types of docking

Molecular docking encompasses several approaches (Figure 4) that are distinguished by the degree of flexibility allowed for the ligand and receptor. Rigid docking treats both molecules as static and offers computational efficiency but limited accuracy for dynamic systems (Kitchen *et al.*, 2004). Flexible docking permits ligand conformational changes while keeping the receptor rigid. This captures the adaptability of small molecules in binding sites better. Induced-fit docking allows flexibility in both ligand and receptor, reflecting the conformational adjustments often observed in protein-ligand interactions, though it demands greater computational resources (Baptista *et al.*, 2021). Ensemble docking is a more advanced method that uses multiple receptor conformations often from molecular dynamics simulations to account for protein flexibility, improving prediction accuracy for complex targets. In TB research, flexible and induced-fit docking are preferred due to the conformational variability of Mtb proteins (such as ClpP1P2), ensuring more reliable binding predictions (Freitas de Freitas *et al.*, 2023).

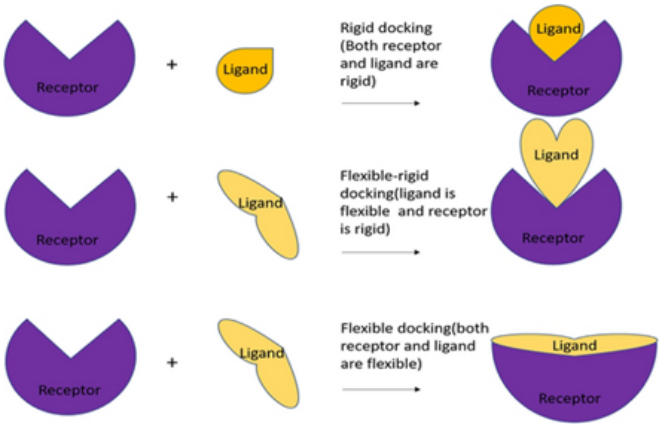

Figure 4. Different types of molecular docking approaches (Mohanty & Mohanty, 2023).

2.3.3. Applications of docking in TB research

Molecular docking has significantly advanced TB drug discovery by identifying potential inhibitors for critical Mtb targets. For example, Baptista *et al.* (2021) employed reverse docking to evaluate 53 anti-mycobacterial natural products against Mtb proteins, including ClpP1P2, InhA, and DprE1, identifying promising leads by comparing docking scores to known inhibitors. Similarly, docking studies targeting shikimate kinase (SK) that is an essential Mtb enzyme absent in humans, screened compound libraries to identify inhibitors with favorable binding affinities and low toxicity, validated through in vitro assays (Freitas de Freitas *et al.*, 2023). Docking has also elucidated resistance mechanisms; Shahbaaz *et al.* (2022) used it to study how KatG mutations reduce isoniazid binding, informing strategies to overcome resistance. Additionally, docking has explored host-pathogen interactions and guided host-directed therapies by modeling Mtb protein interactions with human immune components (Abreu *et al.*, 2020). These applications highlight docking's role in accelerating the discovery of novel anti-TB agents, especially for drug-resistant strains.

2.4. Gaps in the current research and unmet needs

Despite the promising validation of ClpP1P2 as a target, there are significant gaps in the translational path toward a clinical inhibitor. A primary challenge is in the structural complexity and asymmetry of the heterotetradecameric ClpP1P2 complex itself. As the crystal structure provides a foundational blueprint, the distinct binding interfaces for chaperones like ClpC1 and ClpX create dynamic binding sites that are not yet fully characterized for small-molecule inhibition (Schmitz *et al.*, 2014; Leodolter *et al.*, 2015). This complexity makes rational drug design especially challenging as effective inhibitors must precisely exploit these interfaces to achieve selectivity over human proteases.

Furthermore, although proof-of-concept molecules like

the acyldepsipeptides (ADEPs) demonstrate potent anti-mycobacterial activity by dysregulating ClpP1P2 function, their therapeutic potential is severely limited by unfavorable pharmacokinetic profiles and concerns over toxicity (Li *et al.*, 2010; Ollinger *et al.*, 2012). The development of lead compounds with improved drug-like properties like oral bioavailability, metabolic stability and a wider therapeutic window remains a critical unmet need. The field lacks a robust pipeline of chemically diverse scaffolds that can serve as starting points for medicinal chemistry optimization.

Finally, the pharmacodynamic understanding of ClpP1P2 inhibition within the context of an intracellular infection is still nascent. The ability of any inhibitor to penetrate the complex cell envelope of *M. tuberculosis* and subsequently achieve sufficient concentration within the bacterial cytosol to engage the target is significant but often unaddressed. The absence of comprehensive data on the tissue distribution and intracellular pharmacokinetics of proposed inhibitors creates a gap between in vitro activity and in vivo efficiency (Raju *et al.*, 2012). Bridging these gaps requires an integrated approach that combines advanced structural biology with rigorous *in silico* and pharmacokinetic profiling to identify novel, drug-like chemotypes capable of selectively targeting this crucial bacterial vulnerability.

3. METHODOLOGY

3.1. Retrieval and preparation of target protein structure

The ClpP1P2 protease complex of *Mycobacterium tuberculosis* is a critical component of the Clp proteolytic machinery essential for protein homeostasis and bacterial virulence (Schmitz *et al.*, 2014) and was selected as the target for inhibitor screening. We retrieved the three-dimensional crystal structure of the *M. tuberculosis* ClpP1P2 complex from the Protein Data Bank (Schmitz *et al.*, 2014) and saved it in legacy PDB format. Structural preparation was performed using Molecular Operating Environment (MOE) version 2019.0102 (Chemical Computing Group, Montreal, QC, Canada). Non-essential components like solvent molecules, bound ligands, and unwanted residues and chains were removed to isolate the ClpP1P2 protein in order to avoid interferences during the docking simulations. The structure was then processed using MOE's QuickPrep module which applied Protonate3D to assign protonation states, corrected structural errors and added hydrogens. This was followed by energy minimization using the Amber10: EHT forcefield with a root-mean-square (RMS) gradient of 0.1 kcal/mol/Å², and the default parameters were used to optimize the protein geometry for subsequent molecular docking studies.

3.2. Retrieval and preparation of ligands

The study curated a library of 100 ligands to screen for potential inhibitors of the *Mycobacterium tuberculosis* ClpP1P2 protease. Ninety-seven ligands were phytochemicals carefully sourced from literature and selected based on their reported anti-tubercular activity against *M. tuberculosis* (Kumar *et al.*, 2014). Additionally, two standard anti-tuberculosis drugs namely Isoniazid and Ethambutol were included as reference compounds and our basis for that was due to their established clinical efficiency in tuberculosis treatment (D'Ambrosio *et al.*, 2015). The final ligand was a modified acyldepsipeptide (ADEP-2B5Me) that was co-crystallized with the ClpP1P2 structure. The 99 non-co-crystallized ligands were retrieved from PubChem in 3D Spatial Data File (SDF) format using their respective compound identifiers (CIDs). The co-crystallized ADEP-2B5Me was extracted directly from the ClpP1P2 crystal structure. All ligands were imported into a molecular database using Molecular Operating Environment (MOE). Ligand preparation involved energy minimization using the MMFF94x forcefield with a root-mean-square (RMS) gradient of 0.1 kcal/mol/Å². The default dielectric conditions were employed to optimize molecular geometries for subsequent molecular docking studies.

3.3. Molecular docking and validation

Induced-fit Molecular docking was performed using the Molecular Operating Environment (MOE) software. Binding sites were identified using MOE's Site Finder tool by adding dummy atoms to map potential interaction regions based on geometric and chemical properties of the protein surface. The placement method utilized was the Triangle Matcher algorithm, which generates poses by aligning ligand conformations to triplets of receptor site points. Docking results were evaluated using the S-score. This score estimates binding affinity between the ligand and the receptor. Initial scoring was performed with the London dG scoring function and final pose refinement was conducted using the GBVI/WSA dG scoring function.

Although it is necessary to note that ligands with the lowest binding energy were typically prioritized but we still carried out visual inspection of each docked pose in MOE to ensure realistic binding conformations. This step ensured we didn't fall into the bias of selecting poses with low energy but unrealistic interactions like steric clashes or improbable hydrogen bonding geometries (Fischer *et al.*, 2021). Validation of docking results was performed using the PyMOL Molecular Graphics System, version 3.0 (Schrödinger, LLC, New York, NY, USA). The selected docked complex of the modified ADEP ligand and target was superimposed onto the original protein structure to derive the root-mean-square deviation (RMSD). It is noted that values of RMSD less than 2 Å is considered optimal.

Ligand efficiency (LE) and inhibition constant (Ki) were calculated to evaluate the binding potency and efficiency of the selected ligands. The calculations, statistical analysis, and scripting were performed using Python 3.12.7 within the Spyder integrated development environment. An acceptable threshold for drug-like molecules is typically LE ≥ 0.3 kcal/mol per heavy atom.

Ligand efficiency was determined using the formula (Hopkins *et al.*, 2014):

$$LE = \frac{-\Delta G}{N_{HA}} \quad \text{....(i)}$$

where ΔG is the binding free energy (in kcal/mol), and N_{HA} is the number of heavy atoms in the ligand (non-hydrogen atoms). The inhibition constant (Ki) was calculated using the standard thermodynamic relationship formula (Wu *et al.*, 2001):

$$K_i = e^{\frac{\Delta G \times 1000}{RT}} \quad \text{.....(2)}$$

Where,
ΔG is the binding free energy in kcal/mol, (R) is the gas constant (1.987 cal/mol·K), and (T) is the temperature in Kelvin (298 K was used).

3.4. ADMET and drug-likeness studies

Identified top hit ligands were subjected to absorption, distribution, metabolism, excretion, and toxicity (ADMET) evaluations along with drug-likeness assessments. These analyses were conducted using Swiss ADME and pkCSM online tools. These tools provide comprehensive predictions of physicochemical properties, pharmacokinetics and drug-like characteristics.

Drug-likeness was evaluated based on Lipinski's Rule of 5 and Veber's rules (Lipinski *et al.*, 1997; Veber *et al.*, 2002). These rules are established criteria for assessing oral bioavailability potential. Ligands with no violations of Lipinski's or Veber's rules were further screened for toxicity using ProTox 3.0. ProTox 3.0 is a web-based tool that predicts acute oral toxicity (LD_{50}) and organ-specific toxicities based on structural similarity to known toxicants which then aids in the identification of safe candidates for downstream studies.

3.5. Physiologically-based pharmacokinetic (PBPK) modeling

PBPK simulations were performed using PK-Sim (Open Systems Pharmacology Suite). We selected the lungs as the primary target organ because *Mycobacterium tuberculosis* establishes infection in the pulmonary alveoli and primarily resides within alveolar macrophages (Coleman *et al.*, 2022). ClpP1P2 is an intracellular protease in *M. tuberculosis* and the intracellular lung compartment was therefore selected as the primary compartment of interest.

A virtual healthy adult human subject (30 years, 70 kg) and a height of 176.00 cm (BMI 22.60 kg/m²) was used for simulations as this reflects the demographic with the highest tuberculosis burden (Suryanti & Ahmed, 2025). Key organ volumes used by the model for distribution calculations included liver (2.36 L), lungs (1.21 L), and kidneys (0.44 L) (Ann, 2002). For each compound, new entries were created in PK-Sim manually using SMILES-derived physicochemical parameters (molecular weight, logP, pKa, solubility) and ADME inputs (fraction unbound in plasma ($f_{u,p}$), blood-to-plasma ratio (B/P), total clearance, and Caco-2 permeability). Measured values were taken directly from pkCSM and SwissADME predictions and missing values were estimated *in silico*.

Tissue–plasma partition coefficients (Kp) were calculated using the Rodgers and Rowland method embedded in PK-Sim. This method accounts for tissue composition, ionization state, and lipophilicity (Rodgers & Rowland, 2007). The calculated Kp values were incorporated into a whole-body PBPK model.

A single intravenous (IV) bolus dose of 70 mg was simulated as an initial proof-of-concept to assess systemic and lung distribution. Three compartments were selected for output namely plasma (total concentration), lung interstitial fluid (ELF proxy), and lung intracellular (total concentration). Concentration-time profiles were generated for each compartment over 24 hours. Pharmacokinetic parameters were derived from these profiles using custom Python scripts with NumPy and SciPy libraries. Specifically, Cmax was determined as the maximum concentration in the dataset, and Tmax as the corresponding time point (first occurrence of multiple maxima). All equations used are according to Sujjavorakul *et al.* (2023) but with modifications. The area under the concentration-time curve from 0 to 24 hours (AUC_{0-24}) was calculated using the trapezoidal rule: for consecutive points (t_i, c_i) and (t_{i+1}, c_{i+1}), the area contribution is given by

$$Area = \frac{(t_{i+1} - t_i) \times (c_i + c_{i+1})}{2}, \quad \text{....(3)}$$

summed across all intervals. The terminal elimination rate constant (λ_z) was estimated by linear regression on the natural logarithm of concentrations, ln(c) versus time for the last 20 data points (~last 5 hours) where λ_z = -slope. This was applied provided the fit had $R^2 > 0.9$, the slope was negative, and concentrations were positive. The terminal half-life ($t_{1/2}$) was computed as

$$t_{1/2} = \frac{In(2)}{\lambda_z} \quad \text{....(4)}$$

The area under the curve extrapolated to infinity ($AUC_{0-\infty}$) was calculated as

$$AUC_{0-\infty} = AUC_{0-24} + \frac{c_{last}}{\lambda_z}, \quad \text{....(5)}$$

assuming mono-exponential decay beyond 24 hours. These parameters were computed for venous blood-plasma, lung-interstitial, and lung-intracellular compartments, though λ_z, $t_{1/2}$, and $AUC_{0-\infty}$ were not estimable for the intracellular compartments due to increasing concentrations (no observable elimination phase).

Unbound tissue-to-plasma partition coefficients ($K_{p,uu}$) were calculated to assess unbound drug exposure in lung compartments relative to plasma, using the formula

$$K_{p,uu} = \frac{AUC_{tissue,total} \times f_{u,tissue}}{AUC_{plasma,total} \times f_{u,p}} \quad \text{....(6)}$$

Where,
AUC values were from 0–24 hours, $f_{u,p}$ was 0.784 for Arbutin and 0.013 for Neodiospyrin (from pkCSM), and $f_{u,tissue}$ (fraction unbound in lung tissue) was assumed as 1 for baseline calculations (a sensitivity default for hydrophilic compounds like Arbutin; scaled linearly for other values like 0.01 – 0.5 for lipophilic compounds like Neodiospyrin). $K_{p,uu}$ was computed for lung-interstitial and lung-intracellular compartments providing us with insights into unbound intracellular lung concentrations as a proxy for potential efficiency against intracellular *M. tuberculosis*.

3.6. Biological activity prediction

The prediction of activity spectra for substances (PASS) was performed using the Way2Drug computational platform. This tool predicts potential biological activities of compounds based

on their chemical structures and utilizes a robust database and machine learning algorithms to estimate the probability of various pharmacological effects and mechanisms of action for our identified hits.

3.7. Target fishing

Target fishing was performed on identified hits using the SwissTargetPrediction web server. This computational tool predicts potential biological targets of small molecules by leveraging a ligand-based approach. It employs a combination of 2D and 3D similarity searches against a curated library of known protein-ligand interactions and integrates cheminformatics and machine learning to estimate the likelihood of the query compound interacting with specific protein targets across human and other species. This facilitates the identification of novel therapeutic targets.

4. RESULTS AND DISCUSSION

We present the key findings from our computational screening of 100 ligands against the *Mycobacterium tuberculosis* ClpP1P2 protein which integrated virtual screening, pharmacokinetic modeling and predictive analyses to identify anti-tuberculosis candidates. Results follow the methodological sequence of docking outcomes (binding affinities, poses, validation), ligand efficiency and inhibition constants, ADME/drug-likeness, ligand-receptor interactions analysis, toxicity, PBPK concentration-time profiles, biological activity predictions and target fishing.

We discuss these in the context of drug discovery and highlight top ligands, mechanistic insights, efficiency against intracellular pathogens, *in silico* assumptions, and literature comparisons. We conclude by selecting hits for experimental validation by emphasizing therapeutic potential and addressing challenges like bioavailability and toxicity, bridging predictions to practical applications.

4.1. Virtual screening and docking results

We conducted virtual screening of 97 natural product-derived ligands and three reference compounds against the prepared structure of *Mycobacterium tuberculosis* ClpP1P2 using induced-fit docking in MOE to identify potential inhibitors with superior binding affinity compared to the Isoniazid standard and co-crystallized modified ADEP analog. The Site Finder tool in MOE predicted multiple potential binding sites on the ClpP1P2 protein surface but the selected site was chosen based on prior literature evidence indicating its role as one of the primary ADEP-binding pockets (Schmitz *et al.*, 2014). The identified binding site included residues, that is, GLU39, ILE41, THR73, TYR75, GLN101, VAL103, LEU105, MET125, LEU127, LEU204 and ARG207. These residues collectively form a hydrophobic cleft with polar moieties conducive to ligand accommodation and are critical for ClpP1P2 function, as hydrophobic interactions mediated by ILE41, VAL103, LEU105, MET125, LEU127, and LEU204 stabilize ligand binding through van der Waals forces, while polar residues such as GLU39, THR73, TYR75, GLN101, and ARG207 facilitate hydrogen bonding and electrostatic interactions, enhancing specificity and affinity. Important active residues within this binding site are illustrated in Figure 5.

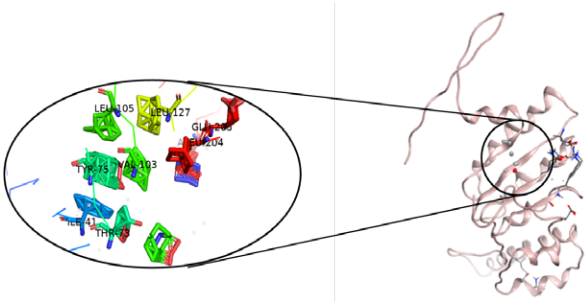

Figure 5. Important binding site residues of caseinolytic protease enzyme (ClpP1P2) (4U0G). The demonstrated binding site residues are involved in catalytic binding of target receptor.

Docking poses were generated to yield the final S score. The S score is an empirical estimate of binding free energy where more negative values indicate stronger predicted affinity. Although lower (more negative) S scores is generally believed to correlate with favorable binding thermodynamics, we performed visual inspection of all poses in MOE to ensure scoring function bias such as overestimation of entropic contributions or failure to penalize steric clashes is reduced (Fischer *et al.*, 2021). This ensures selection of biologically plausible conformations rather than just relying solely on numerical scores.

Of the 100 screened ligands, 25 exhibited S scores more negative than that of the modified ADEP reference (-6.3248 kcal/mol), with an additional 7 displaying affinities marginally less favorable than ADEP but superior to the standard drug isoniazid (-5.7111 kcal/mol). This broader selection was made using discretion to enable a comprehensive downstream analysis and incorporate ligands with potential advantages in efficiency despite not strictly outperforming ADEP, while using isoniazid as a secondary comparison for minimal binding viability. These 32 top-performing ligands are summarized in Table 1 and ranked by ascending S score.

For contextual comparisons, standard first-line anti-tuberculosis drugs isoniazid (-5.7111 kcal/mol) and ethambutol (-5.0209 kcal/mol) were included even though both demonstrated inferior binding affinities compared to ADEP and the screened hits. This is consistent with their distinct mechanisms targeting cell wall synthesis rather than ClpP1P2 (Vilchèze, 2020).

Table 1. Docking results for top ligands and reference compounds. Asterisks denote reference standards; modified ADEP lacks a standard PubChem CID as it is a customized analog.

Ligand Name	S Score (kcal/mol)	PubChem CID
Neodiospyrin	-13.2824	16072922
Silymarin	-12.0922	5213
Glisoflavanone	-10.409	480786
Hesperidin	-8.3375	10621
Rosmarinic acid	-7.8811	5281792
Arbutin	-7.9047	440936
Beilschmin A	-7.5559	22297613
Betulinic acid	-7.4479	64971

Patuletin	-6.991	5281678
Abietane	-6.962	6857485
Fisetin	-6.9083	5281614
Chelerythrine	-6.8353	2703
Glabridin	-6.7865	124052
alpha-Cubebin	-6.6773	25021463
Eupacunoxin	-6.6647	5281451
Andrographolide	-6.6763	5318517
Tiliacorinine	-6.7416	442369
Cyanidin	-6.5352	128861
Licochalcone A	-6.4264	5318998
Obtusifolin	-6.397	3083575
(+)-Eriodictyol	-6.3984	440735
(+)-Ursolic acid	-6.4506	64945
Eupatundin	-6.4831	5281464
Curcumin	-6.5228	969516
Hexahydrocurcumin	-6.3469	5318039
Capsaicin	-6.3203	1548943
Juglanin	-6.3191	5318717
Dehydrozingerone	-6.3097	5354238
Ferulenol	-6.216	54679300
Azorellanol	-6.2038	10383143
Aristolactam I	-6.1987	96710
Echinatin	-6.1556	643779
Isoniazid*	-5.7111	3767
Ethambutol*	-5.0209	14052
ADEP-2B5Me *	-6.3248	-

Notably, Neodiospyrin emerged as the top hit with an S score of -13.2824 kcal/mol. This reflects substantially stronger predicted binding than ADEP and may be due to its naphthoquinone moiety that enables extensive hydrophobic and π-π stacking interactions with residues like LEU105 and MET125. Silymarin (-12.0922 kcal/mol) and Glisoflavanone (-10.409 kcal/mol) followed, with their polyphenolic structures likely forming multiple hydrogen bonds with polar residues such as THR73 and GLN101, enhancing affinity. These results speak to the efficiency of our screening approach in prioritizing natural products with diverse structures for ClpP1P2 inhibition, outperforming or comparing favorably to conventional anti-TB agents.

4.2. Docking validation

Docking validation is an important step in ascertaining the accuracy of the docking protocol. The docked complex of the modified ADEP ligand with ClpP1P2 was superimposed onto the original PDB structure (4U0G) using PyMOL. Pairwise alignment initially involved 202 atoms and this was followed by three cycles of refinement. Seven atoms were rejected in the first cycle, four in the second, and one in the third, bringing the final alignment to a total of 190 atoms (Figure 6).

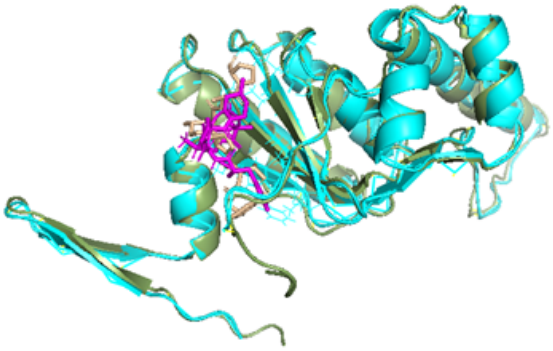

Figure 6. Superimposition of reference and docked complexes of modified ADEP and ClpP1P2 (4U0G) on PyMol Molecular Graphics System.

The root-mean-square deviation (RMSD) executed with a value of 0.746 Å. It is well established that the optimal threshold of RMSD is typically below 2 Å (Kufareva & Abagyan, 2012). As such, our value of RMSD confirms high structural congruence and reliability of the docked poses.

4.3. Ligand efficiency and inhibition constants

Ligand efficiency (LE) and inhibition constant (Ki) are key metrics in drug discovery. LE normalizes binding affinity by molecular size (heavy atoms) and enables fair comparisons across diverse ligands with typically LE ≥ 0.3 kcal/mol per heavy atom indicating drug-like potential (Hopkins *et al.*, 2014). Ki estimates the ligand concentration required for half-maximal inhibition, as lower values signify stronger binding affinity (classified as very strong: <0.1 μM; strong: 0.1–1 μM; good: 1–10 μM; moderate: 10–100 μM) (Wu *et al.*, 2001).

LE and Ki were computed for the 32 top-performing ligands from docking with the standards included making 35 compounds. Nineteen ligands achieved LE ≥ 0.3 or closely approximating this threshold. These 19 including ADEP included as reference, are presented in Table 2 and ranked according to LE.

Table 2. Ligand efficiency and inhibition constants for selected ligands and references (marked in asterisk).

Ligand Name	S Score	PubChem CID	Heavy Atoms	LE	Ki (μM)
Isoniazid*	-5.7111	3767	10	0.571	65.123
Echinatin	-6.1556	643779	11	0.560	30.755
Neodiospyrin	-13.2824	16072922	28	0.474	0.000

Dehydrozingerone	-6.3097	5354238	14	0.451	23.711
Arbutin	-7.9047	440936	19	0.416	1.606
Ethambutol*	-5.0209	14052	14	0.359	208.764
Abietane	-6.962	6857485	20	0.348	7.885
Silymarin	-12.0922	5213	35	0.345	0.001
Glisoflavanone	-10.409	480786	31	0.336	0.023
Fisetin	-6.9083	5281614	21	0.329	8.633
Cyanidin	-6.5352	128861	21	0.311	16.206
(+)-Eriodictyol	-6.3984	440735	21	0.305	20.415
Obtusifolin	-6.397	3083575	21	0.305	20.463
Rosmarinic acid	-7.8811	5281792	26	0.303	1.671
Patuletin	-6.991	5281678	24	0.291	7.509
Capsaicin	-6.3203	1548943	22	0.287	23.291
Glabridin	-6.7865	124052	24	0.283	10.604
Aristolactam I	-6.1987	96710	22	0.282	28.597
ADEP-2B5Me *	-6.3248	-	48	0.132	23.115

Whereas LE represents Ligand Efficiency & Ki represents Inhibition Constant.

Evaluation of LE and Ki values revealed that nearly two-thirds of the top-scoring ligands combined favorable size-normalized binding efficiency (LE ≥ 0.3) with meaningful inhibitory potential. This points towards their suitability for early-stage lead development. Low-molecular-weight ligands such as isoniazid and echinatin achieved the highest LE values (0.571 and 0.560), reflecting efficient target engagement despite only moderate Ki. This is quite consistent with the principle that smaller structures often bind more efficiently but not always with maximal potency (Reynolds *et al.*, 2008). In contrast, larger polyphenolic or terpenoid structures (silymarin, glisoflavanone, rosmarinic acid) exhibited strong to very strong binding affinities (sub-micromolar Ki) even when LE values were moderate (0.303–0.345). This indicates that increased molecular complexity can enhance potency, albeit at some cost to efficiency.

Neodiospyrin was a notable outlier as it combined high efficiency (LE = 0.474) with undetectably low Ki, suggesting optimal binding interactions across its heavy-atom framework. The clear performance gap between reference drugs and several of these ligands, especially ethambutol's low affinity (Ki = 208.764 μM) and modified ADEP's poor efficiency (LE = 0.132) highlights opportunities for hybrid design. Nine ligands here advanced to ADME and drug-likeness studies based on very strong, strong, or good binding (Ki <10 μM) and were supplemented by isoniazid, ethambutol, and ADEP for benchmarking.

4.4. Drug-Likeness and ADME Studies

The drug-likeness of the studied compounds was evaluated using Lipinski's Rule of Five (Ro5) and Veber's rules. These rules are widely accepted criteria for predicting oral bioavailability. Lipinski's Ro5 suggests that a molecule should have a molecular weight (MW) ≤ 500 Da, a consensus LogP ≤ 5, no more than 5 hydrogen bond donors (HBD), and no more than 10 hydrogen bond acceptors (HBA) if it is to be a drug-like molecule (Lipinski *et al.*, 1997). Veber's rules further refine these criteria by limiting the number of rotatable bonds to ≤ 10 and the topological polar surface area (TPSA) to ≤ 140 Å² (Veber *et al.*, 2002).

Our analysis revealed that most compounds (9 out of 11) fully comply with Lipinski's Ro5 (Table 3). Notable exceptions include Abietane that exceeds the logP threshold (6.14), and ADEP that violates both the MW (676.75 > 500) and HBA (10, at the upper limit) criteria. Interestingly, ADEP also exhibits an excessive number of rotatable bonds (16) that further diminishes its drug-likeness. These violations suggest that Abietane may face solubility challenges due to its high lipophilicity, while modified ADEP's large size and flexibility likely hinder oral absorption.

In terms of Veber's rules, Silymarin, Rosmarinic acid, and Patuletin exceed the TPSA threshold (>140 Å²). This indicates potential difficulties in membrane permeability. ADEP, in addition to its Lipinski violations, also fails Veber's criteria due to its high TPSA (182.45 Å²) and excessive rotatable bonds. These findings point to the fact that these compounds may possess bioactive potential but their pharmacokinetic profiles could limit their development as oral drugs.

The standard drugs, Isoniazid and Ethambutol fully comply with all drug-likeness rules and this aligns with their established clinical use and oral bioavailability. Smaller molecules like Fisetin and Arbutin also exhibit optimal drug-likeness. This makes them promising candidates for further study. Conversely, compounds with high TPSA or logP values may require structural modifications to improve their pharmacokinetic properties.

Table 3. Drug-likeness parameters of the evaluated compounds including violations of Lipinski's and Veber's rules. Reference compounds are shown in asterisk.

Compound	MW	LogP	HBD	HBA	nROT	TPSA (Å²)	LV	VV
Neodiospyrin	374.34	2.84	2	6	1	108.74	0	0
Silymarin	482.44	1.59	5	10	4	155.14	0	1
Glisoflavanone	424.49	4.42	4	6	5	107.22	0	0
Arbutin	272.25	-0.77	5	7	3	119.61	0	0
Abietane	276.50	6.14	0	0	1	0.00	1	0
Fisetin	286.24	1.55	4	6	1	111.13	0	0
Rosmarinic acid	360.31	1.58	5	8	7	144.52	0	1
Patuletin	332.26	1.38	5	8	2	140.59	0	1
Isoniazid*	137.14	-0.35	2	3	2	68.01	0	0
Ethambutol*	204.31	0.60	4	4	9	64.52	0	0
ADEP-2B5Me*	676.75	0.90	4	10	16	182.45	2	2

Whereas MW represents Molecular Weight; LogP represents Octanol-Water Partition Coefficient; HBD represents Number of Hydrogen Bond Donor; HBA represents Number of Hydrogen Bond Acceptor; nROT represents Number of Rotatable Bonds; TPSA represents Topological Polar Surface Area; LV represents Lipinski's Violations & VV represents Veber's Violations.

The ADME (Absorption, Distribution, Metabolism, and Excretion) analysis provided us key information about the pharmacokinetic behavior of the evaluated compounds with focus on four key parameters that determine their therapeutic potential (Table 4).

Table 4. Predicted ADME properties of the compounds. Reference compounds are shown in asterisk.

Compound	GI Absorption	BBB Permeant	Pgp Substrate	Bioavailability Score
Neodiospyrin	High	No	No	0.55
Silymarin	Low	No	No	0.55
Glisoflavanone	High	No	No	0.55
Arbutin	High	No	No	0.55
Abietane	Low	No	No	0.55
Fisetin	High	No	No	0.55
Rosmarinic acid	Low	No	No	0.56
Patuletin	Low	No	No	0.55
Isoniazid*	High	No	No	0.55
Ethambutol*	High	No	No	0.55
ADEP-2B5Me*	Low	No	No	0.17

Gastrointestinal (GI) Absorption predicts how well a compound is absorbed through the intestinal lining into systemic circulation, a critical factor for oral bioavailability. High absorption (>30% absorbed) is preferred for orally administered drugs (Zhang *et al.*, 2002). GI Absorption varied significantly among the compounds. Neodiospyrin, Glisoflavanone, Arbutin, Fisetin, Isoniazid, and Ethambutol were predicted to have high GI absorption, consistent with their favorable drug-likeness profiles. In contrast, Silymarin, Abietane, Rosmarinic acid, Patuletin, and modified ADEP exhibited low absorption. For Silymarin, Rosmarinic acid, and Patuletin, this is likely due to their high TPSA values as this reduces membrane permeability. Abietane's poor absorption can be attributed to its extreme lipophilicity (LogP = 6.14), while modified ADEP's large size and structural complexity further limit its bioavailability.

None of the compounds were predicted to permeate the BBB, which minimizes the risk of central nervous system (CNS)-related side effects. Blood-Brain Barrier (BBB) Permeation is a good indication of whether a compound can cross into the central nervous system (Banks, 2009). Even though it is desirable for CNS-targeting drugs, BBB permeation is typically unwanted for other therapeutics to avoid potential neurotoxicity. Additionally, none of the compounds were identified as Pgp substrates. This suggests that they are

unlikely to be effluxed by this transporter that is a common mechanism of drug resistance. Bioavailability Score provides a quantitative estimate (0-1) of the fraction of an orally administered dose that reaches systemic circulation intact and integrate absorption and also first-pass metabolism effects (Aungst, 2017). Scores that are less than 0.5 generally indicate good oral bioavailability potential. Most compounds scored 0.55 indicating moderate bioavailability. Modified ADEP scored significantly lower (0.17). This reinforces and even confirms its poor suitability for oral delivery. The standard drugs, Isoniazid and Ethambutol, both scored 0.55, which goes well with their known pharmacokinetic profiles and clinical efficiency.

4.5. Receptor ligand interaction analysis

After applying Lipinski's and Veber's rules, six compounds were selected for interaction analysis with the ClpP1P2 protease. This step was essential since these filters are widely used to identify molecules with acceptable oral bioavailability and pharmacokinetic behavior (Lipinski *et al.*, 1997; Veber *et al.*, 2002). The selected compounds included four natural molecules namely neodiospyrin, glisoflavanone, arbutin, and fisetin, then the reference drugs, isoniazid and ethambutol. Their binding interactions were analyzed in two dimensions on Molecular Operating Environment (MOE) and in three dimensions on PyMOL Molecular Graphics System. The results are consistent with their docking scores.

Neodiospyrin demonstrated the strongest binding to ClpP1P2. With a docking score of −13.28 kcal/mol, the ligand engaged in four hydrogen-bond interactions with Gly122 at a distance of 3.06 Å with each contributing binding energies of −2.6 kcal/mol (Figure 7). Also, hydrophobic interactions were observed with Met125 at 3.12–3.46 Å but with lower energy contributions (−0.2 kcal/mol). These multiple stable hydrogen bonds likely explain its superior docking performance. This agrees with earlier reports showing that neodiospyrin exhibits potent antimycobacterial activity by targeting bacterial enzymes involved in redox and proteolytic regulation (Van der Kooy *et al.*, 2006). Its ability to form repeated hydrogen bonds with Gly122 spotlights its strong potential to disrupt ClpP1P2 activity. This is especially important for protein degradation and mycobacterial survival (Raju *et al.*, 2012).

Glisoflavanone also showed a strong docking score of −10.41 kcal/mol. It interacted with Gln101 at distance of 4.13 Å and Ala214 through repeated hydrogen bonds at distances of 3.96 Å (Figure 8). This was supported by an additional acceptor interaction with Leu204. And although each bond was individually weak (-0.1 kcal/mol), their combined effect stabilized the ligand within the binding pocket. Similar behavior has been observed for flavonoids in docking studies, where multiple moderate interactions were sufficient to secure strong overall binding to mycobacterial proteins (Baptista *et al.*, 2021). The visualization confirmed that the flavanone core was well accommodated within the receptor cavity.

Arbutin displayed a docking score of −7.90 kcal/mol and a particularly strong hydrogen bond with Glu39 at distance of 3.04 Å and binding energy of −4.0 kcal/mol. It displayed additional acceptor interactions with Tyr75 and Arg207 and a donor contact to Met125 at distances within the range of 3.13–3.29 Å and moderate binding energy of −0.3 kcal/mol (Figure 9). Previous studies have reported that glycosylated phenolics like arbutin often interact strongly with charged residues which possibly enhance their binding specificity (Qiu *et al.*, 2025). These results suggest that despite arbutin's moderate docking score, it could still be an effective modulator of ClpP1P2 activity due to the strength and diversity of its interactions.

Fisetin had a lower docking score of −6.91 kcal/mol. This reflects a less extensive interaction network. It mainly formed hydrogen-bond acceptor contacts with Gln101 at distances of 3.38 Å and 3.34 Å, alongside a pi–hydrogen interaction involving its aromatic ring at distance of 4.18 Å (Figure 10). Although this is modest compared to the top ligands, such interactions are common for polyphenols and contribute to moderate stability in the binding site (Baptista *et al.*, 2021).

The two reference drugs showed the weakest docking scores. This is consistent with the fact that ClpP1P2 is not their primary target. Isoniazid with docking score of −5.71 kcal/mol made only two pi–hydrogen contacts with Tyr75 and Thr73 at distances of 4.39 Å and 4.11 Å respectively (Figure 11), while ethambutol with a docking score of −5.02 kcal/mol interacted with Met125 through two weak hydrogen bonds at distances of 3.61 Å and 3.76 Å with energy of −0.4 kcal/mol apiece (Figure 12). Their limited interaction networks confirm their lower affinity for ClpP1P2 as this aligns with previous reports that their mechanism of action in tuberculosis involves enzymes other than the protease complex (Famulla *et al.*, 2016).

These results are consistent with earlier structural studies demonstrating that ClpP1P2 can be effectively targeted by small molecules forming multiple hydrogen bonds and polar contacts (Schmitz *et al.*, 2014; Famulla *et al.*, 2016). Overall, our findings support the potential of these natural compounds as promising candidates for further optimization.

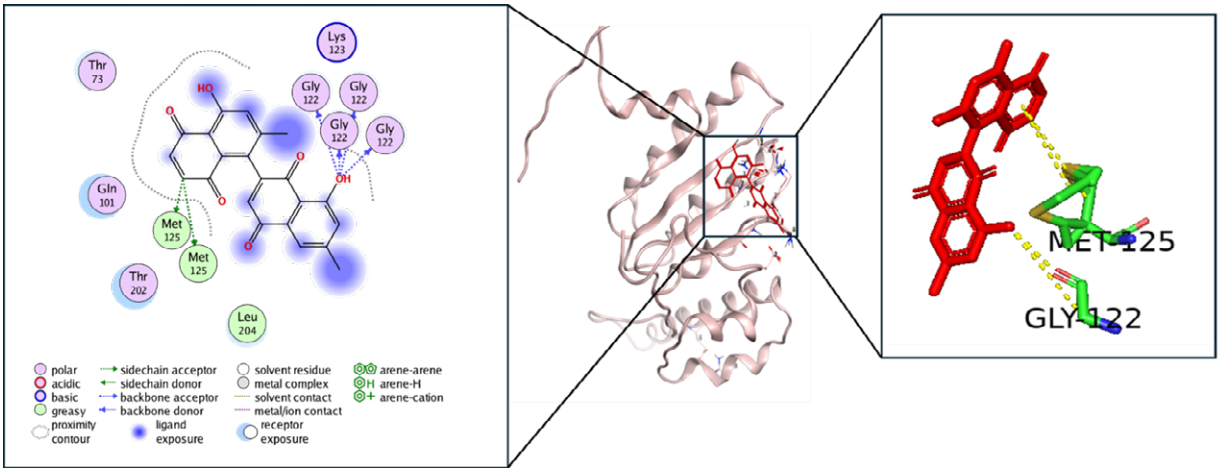

Figure 7. Interaction profile of Neodiospyrin with the potential binding site of ClpP1P2. The left side represents a 2D schematic representation of interactions by MOE, the central figure represents a 3D complex of protein and its respective ligand, and the right side represents interacting residues of protein with its respective ligand by PyMol

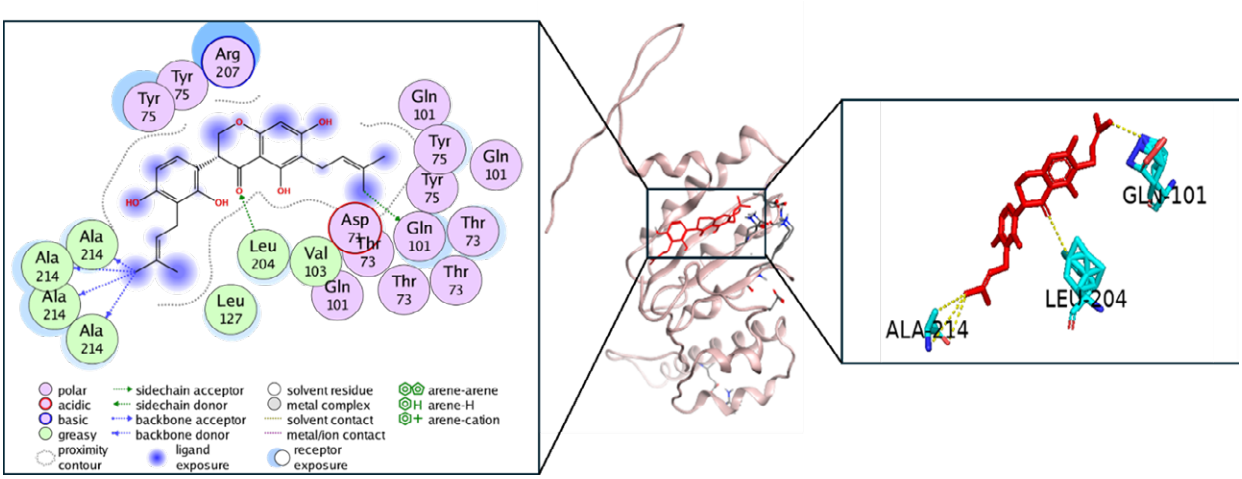

Figure 8. Interaction profile of Glisoflavanone with the potential binding site of ClpP1P2. The left side represents a 2D schematic representation of interactions by MOE, the central figure represents a 3D complex of protein and its respective ligand, and the right side represents interacting residues of protein with its respective ligand by PyMol.

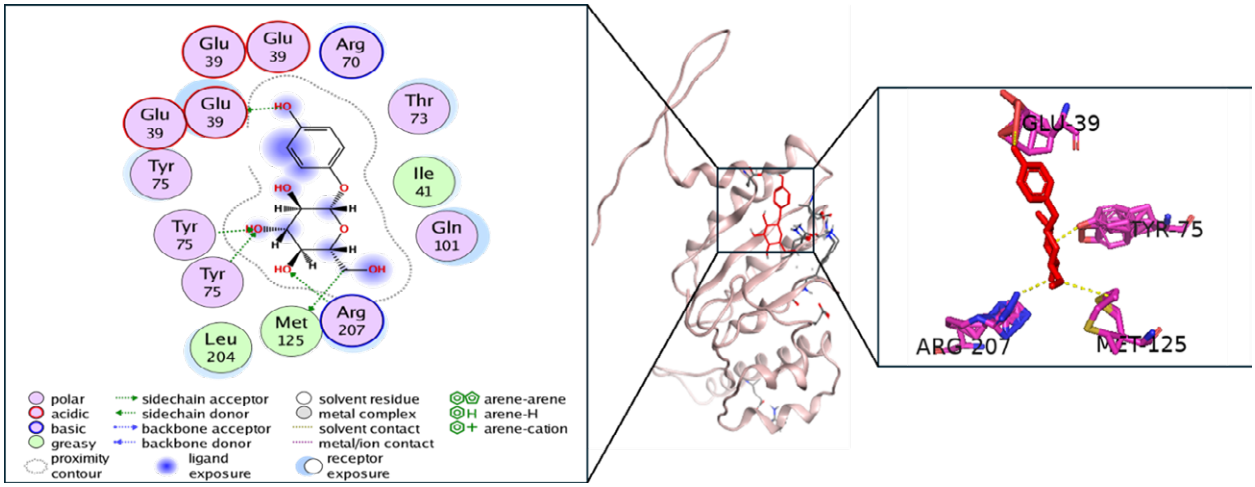

Figure 9. Interaction profile of Arbutin with the potential binding site of ClpP1P2. The left side represents a 2D schematic representation of interactions by MOE, the central figure represents a 3D complex of protein and its respective ligand, and the right side represents interacting residues of protein with its respective ligand by PyMol.

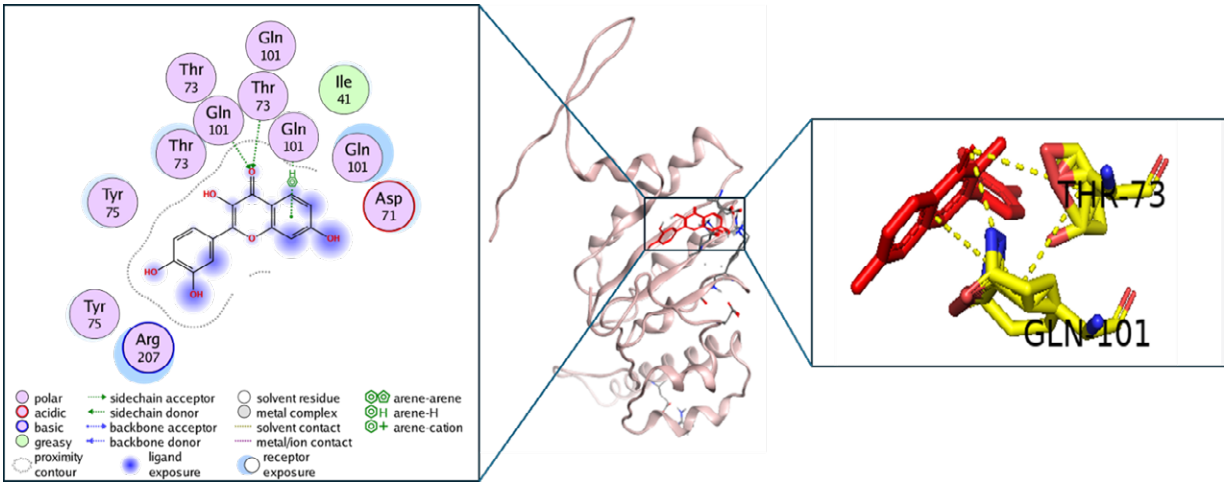

Figure 10. Interaction profile of Fisetin with the potential binding site of ClpP1P2. The left side represents a 2D schematic representation of interactions by MOE, the central figure represents a 3D complex of protein and its respective ligand, and the right side represents interacting residues of protein with its respective ligand by PyMol.

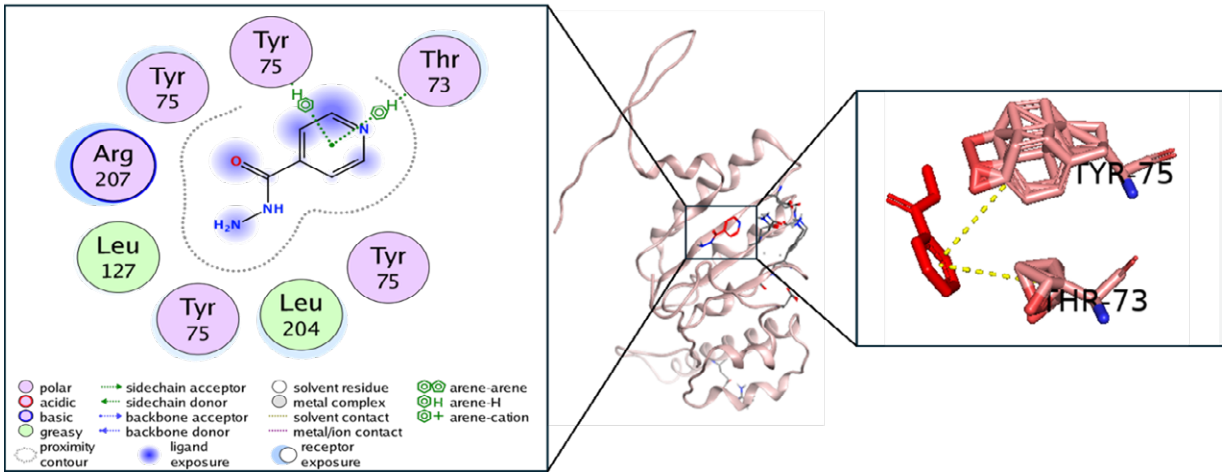

Figure 11. Interaction profile of Isoniazid with the potential binding site of ClpP1P2. The left side represents a 2D schematic representation of interactions by MOE, the central figure represents a 3D complex of protein and its respective ligand, and the right side represents interacting residues of protein with its respective ligand by PyMol.

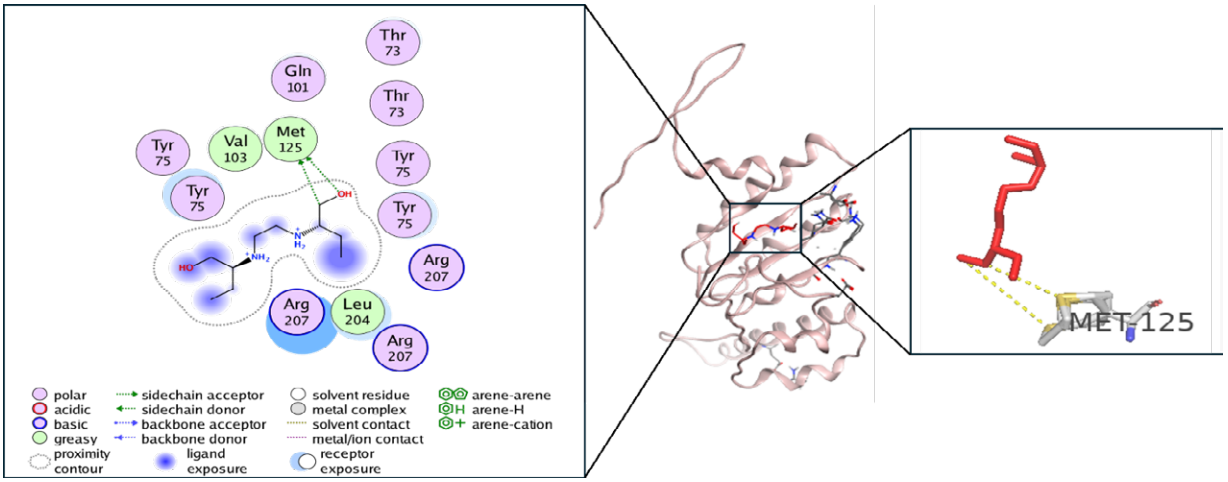

Figure 12. Interaction profile of Ethambutol with the potential binding site of ClpP1P2. The left side represents a 2D schematic representation of interactions by MOE, the central figure represents a 3D complex of protein and its respective ligand, and the right side represents interacting residues of protein with its respective ligand by PyMol.

4.6. Toxicity

Toxicity assessment is an essential stage in drug discovery since promising compounds can fail in late development due to adverse effects. The ProTox-III server was used to predict toxicity across six biologically relevant endpoints (Table 5) for the six selected ligands that passed drug likeness studies. These endpoints provide information about organ damage, genotoxicity, metabolic perturbations, and molecular initiating events (MOE).

Hepatotoxicity reflects the likelihood of liver injury which is a major cause of drug attrition since the liver is central to xenobiotic metabolism (Chen *et al.*, 2013). Neurotoxicity assesses risks to the nervous system as it is often linked with off-target effects on neurotransmission or mitochondrial function (Lein *et al.*, 2007). Mutagenicity indicates the potential of compounds to cause genetic mutations. These are often precursors to carcinogenicity (Eastmond *et al.*, 2009). In addition, nutritional toxicity models disruptions in nutrient metabolism. It can impair growth or metabolic health (Prentki *et al.*, 2020). At the cellular stress level, the Heat Shock Factor response element (HSE) is linked to protein damage and stress adaptation (Akerfelt *et al.*, 2010), while Pregnane X Receptor (PXR) activation reflects altered xenobiotic metabolism that may lead to drug–drug interactions and metabolic imbalances (Kliewer *et al.*, 2002).

From the predictions, arbutin emerged as the safest candidate (Class 5, LD_{50} = 2500 mg/kg) with no predicted activity across any endpoint. This goes well with its known use in dermatological and nutraceutical products, where it is generally regarded as safe at therapeutic doses (Maeda & Fukuda, 1991). Fisetin was placed in Class 3 (LD_{50} = 159 mg/kg) and flagged for nutritional toxicity. Although fisetin is a flavonoid with reported antioxidant and anticancer activities, its high doses have been associated with metabolic disturbances in some experimental models (Khan *et al.*, 2013).

Neodiospyrin (Class 4, LD_{50} = 1000 mg/kg) was largely inactive across endpoints but showed PXR activation. This suggests potential risks for drug–drug interactions via cytochrome P450 induction. Despite the fact that glisoflavanone showed potent docking scores, it still demonstrated significant limitations with a very low LD_{50} (10 mg/kg, Class 2) and predicted nutritional toxicity alongside PXR activation. This profile suggests a narrow therapeutic window and a higher risk of systemic toxicity.

The reference antituberculars showed expected toxicity profiles. Isoniazid (Class 3, LD_{50} = 133 mg/kg) was predicted to be hepatotoxic and neurotoxic in addition to nutritional toxicity. This is consistent with clinical observations of hepatotoxicity, peripheral neuropathy, and pyridoxine depletion in long-term therapy (Saukkonen *et al.*, 2006). Ethambutol (Class 4, LD_{50} ≈ 998 mg/kg) was flagged for hepatotoxicity but otherwise appeared less risky. This goes along with its established safety profile aside from dose-related optic neuritis (Leibold, 1966).

Taken together, it is worthy of note that these classes of toxicity do not tell the whole story. They are majorly predictive and serve as early flags but do not immediately condemn a potential drug based on toxicity predictions. These insights highlight the importance of integrating computational toxicity prediction with experimental validation in guiding lead optimization and preclinical prioritization.

Table 5. ProTox-III toxicity prediction across six endpoints for selected ligands. Standard drugs are indicated with asterisk.

Compound	LD_{50} (mg/kg)	TC	HT	NC	MT	NT	HSE	PR
Neodiospyrin	1000	4	–	–	–	–	–	+
Glisoflavanone	10	2	–	–	–	+	–	+
Arbutin	2500	5	–	–	–	–	–	–
Fisetin	159	3	–	–	–	+	–	–
Isoniazid*	133	3	+	+	–	+	–	–
Ethambutol*	998	4	+	–	–	–	–	–

Whereas LD_{50} represents Acute Toxicity; TC represents Toxicity Class; HT represents Hepatotoxicity; NC represents Neurotoxicity; MT represents Mutagenicity; NT represents Nutritional Toxicity; HSE represents Heat Shock Factor Response Element, PXR represents Pregnane X Receptor, + represents active & - represents inactive.

4.7. Physiologically-based Pharmacokinetic (PBPK) Modeling

Arbutin and neodiospyrin were promising candidates from the initial toxicity screening of the six ligands. These compounds were advanced to physiologically based pharmacokinetic (PBPK) modeling to evaluate their distribution into the lungs. This is the primary site of infection where the pathogen resides intracellularly within alveolar macrophages

For arbutin, the concentration-time profiles (Figure 13) revealed rapid distribution and minimal decline in plasma as peak concentrations (C_{max}) were achieved almost immediately post-dose. Plasma concentrations decreased minimally in micromolar range, indicative of efficient clearance. Lung interstitial levels were similar to this but at lower magnitudes. Intracellular lung concentrations remained minimal and increased gradually over time. This suggests limited penetration into the target compartment. This is reflected in the pharmacokinetic parameters (Table 6), where plasma AUC_{0-24} was substantially higher than in lung tissues. The unbound tissue-to-plasma partition coefficients ($K_{p,uu}$) were low (0.48 for interstitial, 0.01 for intracellular). This implies suboptimal unbound exposure at the site of action. The long terminal half-life (~228 hours) pointed to slow elimination, but without an observable decline in intracellular concentrations,

extrapolation to infinity was really not feasible for that compartment. Altogether, arbutin's poor intracellular accumulation may limit its efficiency against intracellular pathogens, despite favorable systemic kinetics.

In contrast, neodiospyrin that is a more lipophilic molecule exhibited distinct behavior (Figure 14). Plasma profiles showed quick peaking followed by a very slow decline. This is consistent with its extended half-life (>2500 hours) as this is likely due to high tissue binding and low clearance. Lung interstitial concentrations were comparable to plasma initially but stabilized, while intracellular levels rose progressively by surpassing interstitial by 24 hours. This suggests preferential accumulation in the intracellular space is necessary for targeting *M. tuberculosis*. Table 6 summarizes the endpoints and highlights higher $K_{p,uu}$ values (28.47 interstitial, 11.05 intracellular) which indicates superior unbound penetration relative to plasma. The extensive $AUC_{0-\infty}$ in plasma confirms prolonged exposure, although the increasing intracellular trend precludes terminal phase estimation there.

To compare, neodiospyrin demonstrated better lung tissue distribution and intracellular partitioning than arbutin. This could translate to greater potential against intracellular bacteria but its low unbound plasma fraction is concerning for bioavailability in oral formulations. Even though arbutin was safer in toxicity screens, it may require structural modifications for improved targeting. These analyses guide further optimization and emphasize the need for intracellular exposure in anti-TB drug design.

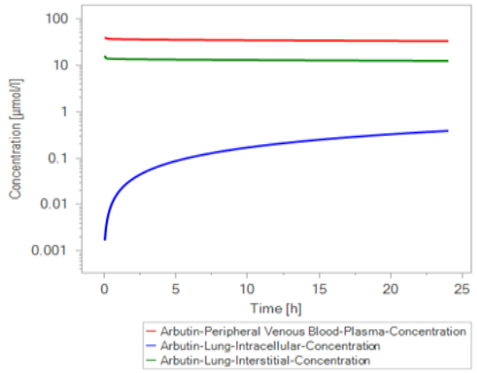

Figure 13. Simulated concentration-time profile of Arbutin in plasma, lung intracellular, and lung interstitial compartments.

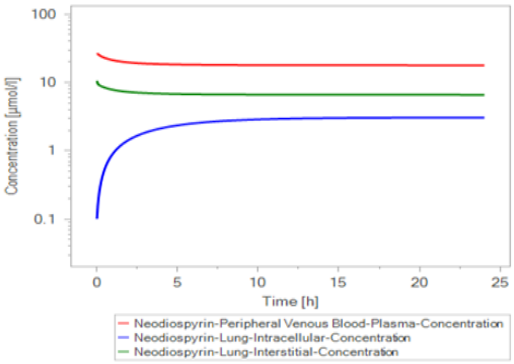

Figure 14. Simulated concentration-time profile of Neodiospyrin in plasma, lung intracellular, and lung interstitial compartments.

Table 6. Pharmacokinetic parameters for arbutin and neodiospyrin.

Compound	Compartment	C_{max} (µmol/L)	T_{max} (h)	AUC_{0-24} (µmol·h/L)	λ_z (h^{-1})	$T_{1/2}$ (h)	$AUC_{0-\infty}$ (µmol·h/L)	$K_{p,uu}$ ($f_{u,\ tissue}$ = 1)
Arbutin	Venous Blood-Plasma	42.09	0.05	819.98	0.00304	228.21	11603.31	-
	Lung-Interstitial	15.91	0.05	307.21	0.00304	228.20	4347.07	0.48
	Lung-Intracellular	0.39	24	4.75	N/A	N/A	N/A	0.01
Neodiospyrin	Venous Blood-Plasma	28.42	0.05	441.07	0.00027	2537.38	65855.09	-
	Lung-Interstitial	10.61	0.05	163.24	0.00027	2538.05	24377.86	28.47
	Lung-Intracellular	3.07	24	63.35	N/A	N/A	N/A	11.05

4.8. Biological activity predictions

The Prediction of Activity Spectra for Substances (PASS) tool is accessible via the Way2Drug computational platform and was employed to forecast the potential biological activities of arbutin and neodiospyrin. PASS uses structure-activity relationship (SAR) analysis derived from a vast training set of over 60,000 known compounds. This includes drugs and bioactive molecules to estimate the probability of a substance exhibiting specific pharmacological or toxicological effects (Filimonov *et al.*, 2014). For each predicted activity, PASS provides two metrics. Pa is the probability of activity (ranging from 0 to 1, where higher values indicate greater likelihood of activity), and Pi is the probability of inactivity (also 0 to 1, with lower values supporting potential activity). This *in silico* approach enables rapid screening of compounds for diverse therapeutic potentials which further guides experimental validation.

From the comprehensive PASS outputs for arbutin and neodiospyrin, 21 common activities were selected based on their Pa scores. Activities that are relevant to anti-tuberculosis drug development like anti-inflammatory, antineoplastic, and apoptosis-modulating effects were focused on (Figure 15). In total, both compounds displayed promising profiles, with average Pa values exceeding 0.5 for most activities. Arbutin exhibited higher Pa scores in several key areas. This include anticarcinogenic (Pa = 0.829), vasoprotector (Pa = 0.934), and membrane permeability inhibitor (Pa = 0.946), going well with its known hydrophilic nature and reported protective effects on cellular integrity (Bhalla *et al.*, 2022). In contrast, neodiospyrin

showed stronger predictions for CYP2C12 substrate activity (Pa = 0.896) and HIF1A expression inhibition (Pa = 0.769), indicative of its lipophilic structure potentially enhancing metabolic interactions and hypoxic stress modulation in infected tissues (Van der Kooy *et al.*, 2006).

Notably, both ligands scored highly in antineoplastic (average Pa = 0.785) and caspase-3 stimulant activities (average Pa = 0.7895). These are important for targeting *Mycobacterium tuberculosis* by inducing programmed cell death in infected macrophages. Arbutin's superior anti-inflammatory prediction (Pa = 0.673 as against 0.473 for neodiospyrin) is in line with experimental evidence of its efficiency in reducing pro-inflammatory cytokines as demonstrated in UVB-induced skin damage models (Shu *et al.*, 2024). Meanwhile, neodiospyrin's antimutagenic and antiseptic potentials that is inferred from related naphthoquinones like diospyrin, support its broader antimicrobial spectrum as this includes activity against Gram-positive bacteria (Adeniyi *et al.*, 2000). These predictions confirm to us that neodiospyrin has edge in apoptosis agonist activity (Pa = 0.707). This potentially agrees with its PBPK-modeled intracellular accumulation for enhanced efficiency against intracellular pathogens.

Although PASS offers valuable insights, there are discrepancies between predictions and in vivo outcomes. This has necessitated the need for empirical testing. Nonetheless, the shared high Pa for TP53 expression enhancement (average Pa = 0.7965) and NAD(P)+- arginine ADP-ribosyltransferase inhibition (average Pa = 0.859) position both compounds as candidates for adjunctive TB therapy. It won't be farfetched to say this warrants further investigation into their ClpP1P2-specific interactions.

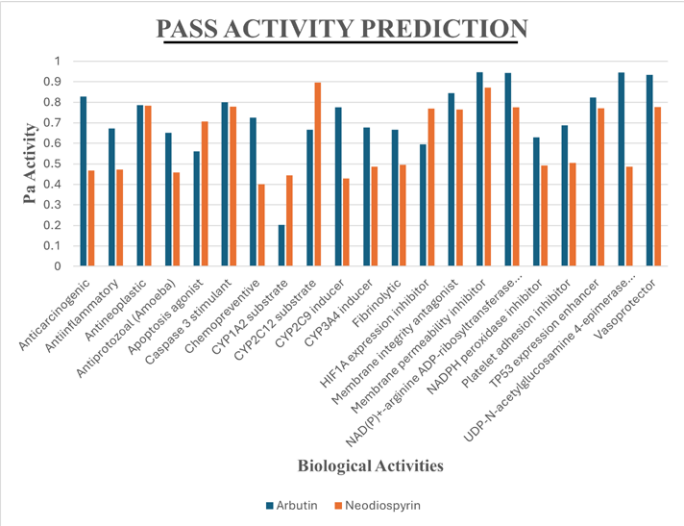

Figure 15. Predicted Biological Activity Spectra of Arbutin and Neodiospyrin, Highlighting Top 21 Shared Activities for Anti-Tuberculosis Drug Development.

4.9. Target fishing analysis

We used computational target fishing to identify potential human protein targets for Arbutin and Neodiospyrin, our selected hit compounds of interest. This approach is powered by SwissTargetPrediction and it compares compound structures against databases of known drug-target interactions to assign probability scores (0-1) for human target engagement. As Gfeller *et al.* (2014) note, such analysis is important for early pharmacodynamic profiling especially for anti-infectives like these targeting Mtb's ClpP1P2 protease. Critically, we sought to map both therapeutic and adverse interaction risks upfront, as human off-target effects directly impact drug safety and efficiency which potentially alters disease pathways or even introduce toxicity during treatment.

In the case of Arbutin, three human target classes showed 26.7% probability. These are Enzymes, Hydrolases, and Family A GPCRs (Figure 16a). Although hydrolase engagement might support desirable anti-protease effects against Mtb ClpP1P2, the GPCR signal raises concerning red flags. Human GPCR interactions commonly trigger side effects like cardiovascular or metabolic disturbances so it demands careful scrutiny. Lower-probability hits (Lyase at 6.7%) suggest minimal risk but warrant monitoring.

Neodiospyrin presented a broader but weaker profile, with human Protease as its top prediction (13.3%) (Figure 16b). Even though this is potentially advantageous for ClpP1P2 inhibition, the scattered low-probability hits (Kinase, Nuclear Receptor at 6.7% each) imply wider off-target risks. Such promiscuity could really undermine therapeutic utility. Kinase interference might disrupt immune responses to Mtb, while nuclear receptor binding could cause endocrine toxicity.

Ultimately it is worthy to note that these probabilities aren't confirmatory but highlight key human-specific risks and opportunities. As Yuan *et al.* (2021) emphasize, such computational insights must guide targeted experimental validation especially for Mtb drugs where human off-target effects could compromise treatment safety without careful pharmacodynamic optimization.

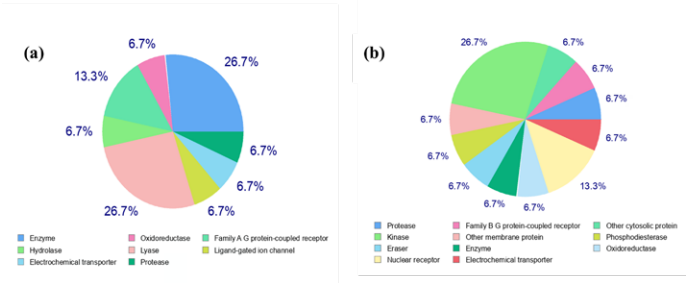

Figure 16. Computational target fishing results for (a) Arbutin and (b) Neodiospyrin using SwissTargetPrediction. Probabilities (0–1 scale) indicate predicted engagement likelihood with human target classes.

5. CONCLUSION

Our study computationally identified Neodiospyrin and Arbutin as promising anti-tubercular candidates targeting the *Mycobacterium tuberculosis* ClpP1P2 protease from a set of 97 literature derived compounds. Neodiospyrin demonstrated exceptional binding affinity (-13.28 kcal/mol), efficient ligand efficiency (0.474), and also favorable intracellular accumulation in lung macrophages that are key for eradicating the pathogen. Arbutin was less potent but exhibited superior safety profiles

and compliance with drug-likeness rules. Both compounds showed complementary biological activities, including anti-inflammatory and apoptosis-modulating effects. Though these findings are inherently limited by their *in silico* nature. Neodiospyrin's predicted hepatotoxicity risk and Arbutin's poor intracellular penetration require experimental validation. The target fishing analysis also flagged potential human off-target interactions that could impact clinical safety. Our work confirms the value of computational approaches in accelerating TB drug discovery. We recommend prioritizing Neodiospyrin for lead optimization to reduce toxicity while using its exceptional target engagement, and also looking at Arbutin in combination therapies to enhance its intracellular delivery. Advancing these candidates to *in vitro* and *in vivo* studies is essential to confirm their potential as next-generation therapeutics against drug-resistant TB.

REFERENCES

Abreu, R., Giri, P., & Quinn, F. (2020). Host-pathogen interaction as a novel target for host-directed therapies in tuberculosis. *Frontiers in immunology, 11*, 1553. https://doi.org/10.3389/fimmu.2020.01553

Adeniyi, B. A., Fong, H. H. S., Pezzuto, J. M., Luyengi, L., & Odelola, H. A. (2000). Antibacterial activity of diospyrin, isodiospyrin and bisisodiospyrin from the root of Diospyros piscatoria (Gurke)(Ebenaceae). *Phytotherapy Research: An International Journal Devoted to Pharmacological and Toxicological Evaluation of Natural Product Derivatives, 14*(2), 112-117. https://doi.org/10.1002/(SICI)1099-1573(200003)14:2%3C112::AID-PTR488%3E3.0.CO;2-T

Ahmad, F., Rani, A., Alam, A., Zarin, S., Pandey, S., Singh, H., ... & Ehtesham, N. Z. (2022). Macrophage: a cell with many faces and functions in tuberculosis. *Frontiers in immunology, 13*, 747799. https://doi.org/10.3389/fcimb.2020.618414

Åkerfelt, M., Morimoto, R. I., & Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. *Nature reviews Molecular cell biology, 11*(8), 545-555. https://doi.org/10.1038/nrm2938

Ann, I. C. R. P. (2002). Basic anatomical and physiological data for use in radiological protection: reference values. *Ann ICRP, 32*(3), 5. https://doi.org/10.1016/0146-6453(79)90123-4

Aungst, B. J. (2017). Optimizing oral bioavailability in drug discovery: an overview of design and testing strategies and formulation options. *Journal of pharmaceutical sciences, 106*(4), 921-929. https://doi.org/10.1016/j.xphs.2016.12.002

Banks, W. A. (2009). Characteristics of compounds that cross the blood-brain barrier. *BMC neurology, 9*(Suppl 1), S3. https://doi.org/10.1186/1471-2377-9-S1-S3

Baptista, R., Bhowmick, S., Shen, J., & Mur, L. A. (2021). Molecular docking suggests the targets of anti-mycobacterial natural products. *Molecules, 26*(2), 475. https://doi.org/10.3390/molecules26020475

Barry 3rd, C. E., Boshoff, H. I., Dartois, V., Dick, T., Ehrt, S., Flynn, J., ... & Young, D. (2009). The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nature Reviews Microbiology, 7*(12), 845-855. https://doi.org/10.1038/nrmicro2236

Bhalla, M., Mittal, R., Kumar, M., & Kushwah, A. S. (2022). Pharmacological Aspects of a Bioactive Compound Arbutin: A Comprehensive Review. *Biointerface Research in Applied Chemistry, 13*, 119. https://doi.org/10.33263/BRIAC132.119

Chandra, P., Grigsby, S. J., & Philips, J. A. (2022). Immune evasion and provocation by *Mycobacterium tuberculosis*. *Nature Reviews Microbiology, 20*(12), 750-766. https://doi.org/10.1038/s41579-022-00763-4

Chen, M., Borlak, J., & Tong, W. (2013). High lipophilicity and high daily dose of oral medications are associated with significant risk for drug-induced liver injury. *Hepatology, 58*(1), 388-396.

Coleman, M., Martinez, L., Theron, G., Wood, R., & Marais, B. (2022). *Mycobacterium tuberculosis* transmission in high-incidence settings—new paradigms and insights. *Pathogens, 11*(11), 1228. https://doi.org/10.3390/pathogens11111228

D'Ambrosio, L., Centis, R., Sotgiu, G., Pontali, E., Spanevello, A., & Migliori, G. B. (2015). New anti-tuberculosis drugs and regimens: 2015 update. *ERJ Open Research, 1*(1). https://doi.org/10.1183/23120541.00010-2015

Eastmond, D. A., Hartwig, A., Anderson, D., Anwar, W. A., Cimino, M. C., Dobrev, I., ... & Vickers, C. (2009). Mutagenicity testing for chemical risk assessment: update of the WHO/IPCS Harmonized Scheme. *Mutagenesis, 24*(4), 341-349. https://doi.org/10.1093/mutage/gep014

Ekins, S., Spektor, A. C., Clark, A. M., Dole, K., & Bunin, B. A. (2017). Collaborative drug discovery for More Medicines for Tuberculosis (MM4TB). *Drug discovery today, 22*(3), 555-565. https://doi.org/10.1016/j.drudis.2016.10.009

Eweas, A. F., Maghrabi, I. A., & Namarneh, A. I. (2014). Advances in molecular modeling and docking as a tool for modern drug discovery. *Der Pharma Chemica, 6*(6), 211-228. http://derpharmachemica.com/archive.html

Famulla, K., Sass, P., Malik, I., Akopian, T., Kandror, O., Alber, M., ... & Brötz-Oesterhelt, H. (2016). Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Molecular microbiology, 101*(2), 194-209. https://doi.org/10.1111/mmi.13362

Filimonov, D. A., Lagunin, A. A., Gloriozova, T. A., Rudik, A. V., Druzhilovskii, D. S., Pogodin, P. V., & Poroikov, V. V. (2014). Prediction of the biological activity spectra of organic

compounds using the PASS online web resource. *Chemistry of Heterocyclic Compounds, 50*(3), 444-457. https://doi.org/10.1007/s10593-014-1496-1

Fischer, A., Smiesko, M., Sellner, M., & Lill, M. A. (2021). Decision making in structure-based drug discovery: visual inspection of docking results. *Journal of Medicinal Chemistry, 64*(5), 2489-2500. https://doi.org/10.1021/acs.jmedchem.0c02227

Freitas de Freitas, T., Roth, C. D., Abbadi, B. L., Hopf, F. S. M., Perelló, M. A., de Matos Czeczot, A., ... & Timmers, L. F. S. M. (2023). Identification of potential inhibitors of *Mycobacterium tuberculosis* shikimate kinase: molecular docking, *in silico* toxicity and in vitro experiments. *Journal of Computer-Aided Molecular Design, 37*(3), 117-128. https://doi.org/10.1007/s10822-022-00495-w

Ge, P., Lei, Z., Yu, Y., Lu, Z., Qiang, L., Chai, Q., ... & Wang, J. (2022). *M. tuberculosis* PknG manipulates host autophagy flux to promote pathogen intracellular survival. *Autophagy, 18*(3), 576-594. https://doi.org/10.1080/15548627.2021.1938912

Gfeller, D., Grosdidier, A., Wirth, M., Daina, A., Michielin, O., & Zoete, V. (2014). SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic acids research, 42*(W1), W32-W38. https://doi.org/10.1093/nar/gku293

Hopkins, A. L., Keserü, G. M., Leeson, P. D., Rees, D. C., & Reynolds, C. H. (2014). The role of ligand efficiency metrics in drug discovery. *Nature reviews Drug discovery, 13*(2), 105-121. https://doi.org/10.1038/nrd4163

Houben, R. M., & Dodd, P. J. (2016). The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. *PLoS medicine, 13*(10), e1002152. https://doi.org/10.1371/journal.pmed.1002152

Khan, N., Syed, D. N., Ahmad, N., & Mukhtar, H. (2013). Fisetin: a dietary antioxidant for health promotion. *Antioxidants & redox signaling, 19*(2), 151-162. https://doi.org/10.1089/ars.2012.4901

Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery, 3*(11), 935-949. https://doi.org/10.1038/nrd1549

Kliewer, S. A., Goodwin, B., & Willson, T. M. (2002). The nuclear pregnane X receptor: a key regulator of xenobiotic metabolism. *Endocrine reviews, 23*(5), 687-702.

Kufareva, I., & Abagyan, R. (2012). Methods of protein structure comparison. In *Homology modeling: Methods and protocols* (pp. 231-257). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-61779-588-6_10

Kumar, A., Farhana, A., Guidry, L., Saini, V., Hondalus, M., & Steyn, A. J. (2011). *Redox homeostasis in mycobacteria: the key to tuberculosis control?*. Expert reviews in molecular medicine, 13, e39. https://doi.org/10.1017/S1462399411002079

Kumar, J. K., Prasad, A. D., & Chaturvedi, V. (2014). Phytochemical screening of five medicinal legumes and their evaluation for in vitro anti-tubercular activity. *AYU (An International Quarterly Journal of Research in Ayurveda), 35*(1), 98-102. https://doi.org/10.4103/0974-8520.141952

Leibold, J. E. (1966). The ocular toxicity of ethambutol and its relation to dose. *Annals of the New York Academy of Sciences, 135*(2), 904-909. https://doi.org/10.1111/j.1749-6632.1966.tb45532.x

Lein, P., Locke, P., & Goldberg, A. (2007). Meeting report: alternatives for developmental neurotoxicity testing. *Environmental health perspectives, 115*(5), 764-768. https://doi.org/10.1289/ehp.9841

Leodolter, J., Warweg, J., & Weber-Ban, E. (2015). The *Mycobacterium tuberculosis* ClpP1P2 protease interacts asymmetrically with its ATPase partners ClpX and ClpC1. *PloS one, 10*(5), e0125345. https://doi.org/10.1371/journal.pone.0125345

Li, D. H. S., Chung, Y. S., Gloyd, M., Joseph, E., Ghirlando, R., Wright, G. D., ... & Ortega, J. (2010). Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chemistry & biology, 17*(9), 959-969. https://doi.org/10.1016/j.chembiol.2010.07.008

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews, 23*(1-3), 3-25. https://doi.org/10.1016/S0169-409X(96)00423-1

Maeda, K. A., & Fukuda, M. (1991). In vitro effectiveness of several whitening cosmetic components in human melanocytes. *Journal of the Society of Cosmetic Chemists, 42*(6), 361-368. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=5508354

Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design, 7*(2), 146-157. https://doi.org/10.2174/157340911795677602

Mohanty, M., & Mohanty, P. S. (2023). Molecular docking in organic, inorganic, and hybrid systems: a tutorial review. *Monatshefte für Chemie-Chemical Monthly, 154*(7), 683-707. https://doi.org/10.1007/s00706-023-03076-1

Ollinger, J., O'Malley, T., Kesicki, E. A., Odingo, J., & Parish, T. (2012). Validation of the essential ClpP protease in *Mycobacterium tuberculosis* as a novel drug target. *Journal of bacteriology, 194*(3), 663-668. https://doi.org/10.1128/jb.06142-11

Pieters, J. (2008). *Mycobacterium tuberculosis* and the macrophage: maintaining a balance. *Cell host & microbe,*

3(6), 399-407. https://doi.org/10.1016/j.chom.2008.05.001

Prentki, M., Peyot, M. L., Masiello, P., & Madiraju, S. M. (2020). Nutrient-induced metabolic stress, adaptation, detoxification, and toxicity in the pancreatic β-cell. *Diabetes, 69*(3), 279-290. https://doi.org/10.2337/dbi19-0014

Putra, O. N., & Adiwinoto, R. P. (2023). Linezolid-associated neurologic toxicity in patients with drug-resistant tuberculosis in a bedaquiline-based regimen: A scoping review. *Journal of Preventive, Diagnostic and Treatment Strategies in Medicine, 2*(4), 194-201. https://doi.org/10.4103/jpdtsm.jpdtsm_117_23

Putra, O. N., Faizah, A. K., & DN, N. W. (2023). Six Months of Bedaquiline-Pretomanid-Linezolid (BPaL) Regimen in Patients with Drug-Resistant Tuberculosis: A Narrative Review. *Journal of Endocrinology, Tropical Medicine, and Infectious Disease (JETROMI), 5*(2), 83-95. https://doi.org/10.32734/jetromi.v5i2.12373

Qiu, K. H., Wang, Y. J., Cheng, K. L., Jiang, L. Q., Li, X., & Zhang, J. L. (2025). Preparation, characterization and analysis of anthocyanin arbutin co-amorphous complexes and evaluation of the inhibition of tyrosinase. I*nternational Journal of Biological Macromolecules*, 143600. https://doi.org/10.1016/j.ijbiomac.2025.143600

Rahman, F. (2024). Characterizing the immune response to *Mycobacterium tuberculosis*: a comprehensive narrative review and implications in disease relapse. *Frontiers in Immunology, 15*, 1437901. https://doi.org/10.3389/fimmu.2024.1437901

Raju, R. M., Unnikrishnan, M., Rubin, D. H., Krishnamoorthy, V., Kandror, O., Akopian, T. N., ... & Rubin, E. J. (2012). *Mycobacterium tuberculosis* ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS pathogens, 8*(2), e1002511. https://doi.org/10.1371/journal.ppat.1002511

Reynolds, C. H., Tounge, B. A., & Bembenek, S. D. (2008). Ligand binding efficiency: trends, physical basis, and implications. *Journal of medicinal chemistry, 51*(8), 2432-2438. https://doi.org/10.1021/jm701255b

Rodgers, T., & Rowland, M. (2007). Mechanistic approaches to volume of distribution predictions: understanding the processes. *Pharmaceutical research, 24*(5), 918-933. https://doi.org/10.1007/s11095-006-9210-3

Ronacher, K., Joosten, S. A., van Crevel, R., Dockrell, H. M., Walzl, G., & Ottenhoff, T. H. (2015). Acquired immunodeficiencies and tuberculosis: focus on HIV/AIDS and diabetes mellitus. *Immunological reviews, 264*(1), 121-137. https://doi.org/10.1111/imr.12257

Saukkonen, J. J., Cohn, D. L., Jasmer, R. M., Schenker, S., Jereb, J. A., Nolan, C. M., ... & Sterling, T. R. (2006). An official ATS statement: hepatotoxicity of antituberculosis therapy. *American journal of respiratory and critical care medicine, 174*(8), 935-952. https://doi.org/10.1164/rccm.200510-1666ST

Schmitz, K. R., Carney, D. W., Sello, J. K., & Sauer, R. T. (2014). Crystal structure of *Mycobacterium tuberculosis* ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proceedings of the National Academy of Sciences, 111*(43), E4587-E4595. https://doi.org/10.1073/pnas.1417120111

Shahbaaz, M., Qari, S. H., Abdellattif, M. H., & Hussien, M. A. (2022). Structural analyses and classification of novel isoniazid resistance coupled mutational landscapes in *Mycobacterium tuberculosis*: a combined molecular docking and MD simulation study. *Journal of Biomolecular Structure and Dynamics, 40*(11), 4791-4800. https://doi.org/10.1080/07391102.2020.1861986

Shu, P., Wang, Y., & Zhang, L. (2024). The effect of α-arbutin on UVB-induced damage and its underlying mechanism. *Molecules, 29*(9), 1921. https://doi.org/10.3390/molecules29091921

Sujjavorakul, K., Katip, W., Kerr, S. J., Wacharachaisurapol, N., & Puthanakit, T. (2023). Predicting the area under the plasma concentration-time curve (AUC) for first dose vancomycin using first-order pharmacokinetic equations. *Antibiotics, 12*(4), 630. https://doi.org/10.3390/antibiotics12040630

Suryanti, S., & Ahmed, I. A. (2025). Correlation Between Demographic Factors (Age, Gender, and Living Area) and Tuberculosis Notification Rates in Private Healthcare: a Cross-sectional Study. *International Journal of Health Literacy and Science, 3*(1), 33-39. https://doi.org/10.60074/ihelis.v3i1.68

Tasneen, R., Mortensen, D. S., Converse, P. J., Urbanowski, M. E., Upton, A., Fotouhi, N., ... & Hawryluk, N. (2021). Dual mTORC1/mTORC2 inhibition as a host-directed therapeutic target in pathologically distinct mouse models of tuberculosis. *Antimicrobial agents and chemotherapy, 65*(7), 10-1128. https://doi.org/10.1128/aac.00253-21

Tian, N., Chu, H., Li, Q., Sun, H., Zhang, J., Chu, N., & Sun, Z. (2025). Host-directed therapy for tuberculosis. *European Journal of Medical Research, 30*(1), 267. https://doi.org/10.1186/s40001-025-02443-4

Van der Kooy, F., Meyer, J. J. M., & Lall, N. (2006). Antimycobacterial activity and possible mode of action of newly isolated neodiospyrin and other naphthoquinones from Euclea natalensis. *South African Journal of Botany, 72*(3), 349-352. https://doi.org/10.1016/j.sajb.2005.09.009

Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of medicinal chemistry, 45*(12), 2615-2623. https://doi.org/10.1021/jm020017n

Vilchèze, C. (2020). Mycobacterial cell wall: a source of successful targets for old and new drugs. *Applied Sciences, 10*(7), 2278. https://doi.org/10.3390/app10072278

World Health Organization. (2024). *Global tuberculosis report 2024*. World Health Organization. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2024

Wu, S. Y., Dornan, J., Kontopidis, G., Taylor, P., & Walkinshaw, M. D. (2001). The first direct determination of a ligand binding constant in protein crystals. *Angewandte Chemie, 113*(3), 602-606. https://doi.org/10.1002/1521-3773(20010202)40:3%3C582::AID-ANIE582%3E3.0.CO;2-0

Yuan, T., Werman, J. M., & Sampson, N. S. (2021). The pursuit of mechanism of action: uncovering drug complexity in TB drug discovery. *RSC chemical biology, 2*(2), 423-440. https://doi.org/10.1039/D0CB00226G

Zhang, H., Zhang, J., & Streisand, J. B. (2002). Oral mucosal drug delivery: clinical pharmacokinetics and therapeutic applications. *Clinical pharmacokinetics, 41*(9), 661-680. https://doi.org/10.2165/00003088-200241090-00003

Zumla, A., Nahid, P., & Cole, S. T. (2013). Advances in the development of new tuberculosis drugs and treatment regimens. *Nature reviews Drug discovery, 12*(5), 388-404. https://doi.org/10.1038/nrd4001