|| ISSN(online): 2589-8698 || ISSN(print): 2589-868X ||
International Journal of Medical and Biomedical Studies
Available Online at www.ijmbs.info
NLM (National Library of Medicine ID: 101738825)
Index Copernicus Value 2020: 79.44
Original research article Volume 6, Issue 07; July: 2022; Page No. 37-47

In Silico Phytochemicals Screening for Tuberculosis

Krati Solanki[1], Radhika.B.S[2], Rida Fatimah[3]

Mohanlal Sukhadia University[1],

Sri Venkateswara College of Engineering[2],

Aligarh Muslim University[3]

Article Info: Received 05 January 2021; Accepted 30 June 2022
doi: https://doi.org/10.32553/ijmbs.v6i7.2587
Corresponding author: Krati Solanki
Conflict of interest: No conflict of interest.

Abstract

Tuberculosis (TB) is a highly contagious disease, and the increased mortality rate continues to be a cause for concern. Although anti-tuberculosis drugs are available, it still affects mankind, causing millions of deaths each year. By integrating biological data on gene sequences, gene expressions, protein structures, and drug target prioritization. The number of cheminformatics databases for small molecules and tools to assist drug discovery has also grown. Many researchers are investigating potential medication therapies that could assist in this battle along with the medical care system throughout the world. We present a docking-based screening using a quantum mechanical scoring of a library built from approved drugs and compounds that Bergenin, Laccaic acid, Swertiamarin, Magnolol, Curcumin, Emodin, Pasakbumin A, Umckalin, Plumbagin, Maritinone, Tetrahydro3,3□-biplumbagin and epigallocatechin with Proteins with PDB id's 4W4I, 3Q0V could display antiviral activity against Tuberculosis. Ideally, these compounds should undergo experimental assays and clinical trials to verify their effectiveness against the disease. We hope that these findings may contribute to the rational drug design against Tuberculosis.

Introduction

According to WHO data, as of 2021, Tuberculosis remains the leading infectious cause of death worldwide (Kramarska et al., 2021); despite mainly being a curable disease, it remains a significant cause of morbidity and mortality. It is an airborne bacterial disease caused by the rod-shaped, non-spore-forming, aerobic bacterium Mycobacterium Tuberculosis. Mycobacterium measure 0.5 mm by 3 mm, is classified as acid-fast bacteria, and has a unique cell wall structure essential for survival. (Knechel, 2009)

For its survival within the cell, it relies on host lipids as the primary source of energy and carbon. Furthermore, it transforms these lipid molecules into precursors for cell- membrane remodeling, cell-wall homeostasis, and ultimately pathogenesis (Asthana et al., 2021)

Its pathogenicity is attributed to the following mechanisms: (i) the ability of the bacteria to reprogramme the immune system after primary infection, preventing the pathogen's elimination; (ii) the development of granulomas, where the pathogen survives in equilibrium with the hosts' defenses; and (iii) the slow control of central

metabolism and replication of the bacteria and their resistance to host defenses and therapy (Miggiano et al., 2020)

The development of TB disease involves interactions between the environment, the host, and the pathogen, with known risk factors such as HIV infection, immunosuppression, diabetes mellitus, overcrowding, malnutrition, and general poverty (O'Garra et al., 2013). It is a chronic granulomatous infectious disease. The infection occurs through the inhalation of aerosols contaminated with M. Tuberculosis bacilli. Following infection, Tuberculosis progresses in two stages. The first stage is an asymptomatic state that may persist for many years, called latent TB. When the immune system is weak, the bacteria begin to reproduce, causing typical symptoms such as coughing, chest pain, fatigue, and unexplained weight loss. If left untreated, the disease can lead to death (Musasa, 2015)

The infection usually occurs in the lungs in 80% of cases, but signs can include cough, bleeding, and chest pain, shortness of breath, fever, weight loss, and drenching night sweats. Infection slowly develops; however, the organisms can spread to other organs, such as lymphatics, the pleura, the bones/joints, or the meninges, leading to an extrapulmonary case of Tuberculosis (Knechel, 2009)

Tuberculosis is classified based on these criteria: 1) Anatomical site, 2) Drug-resistance, 3) Treatment history, 4) HIV status (Sharma & Yadav, 2017). Specific risk groups are at greater risk of infection, including young adults (more commonly males), those in developing countries, healthcare workers, and people whose immune systems are weak, such as those who are HIV positive or who smoke (Fogel, 2015)

Proteins that serve essential roles in a pathogen's survival or growth are vital to its maintenance in a dormant stage, responsible for its pathogenesis, involved in the reactivation of the disease, or beneficial to the pathogen in other ways, could prove to be promising drug targets in the future (Sundaramurthi et al., 2012)

This study analyzed two proteins, 4W4I and 3Q0V, members of the PE-PPE and MCE protein families, respectively.

The genomes of Mycobacterium Tuberculosis and other pathogenic mycobacteria have demonstrated the prominence of the PE/PPE proteins. The PE and PPE proteins share a highly conserved N-terminal domain with a proline-glutamic acid (Pro-Glu) or proline-proline-glutamic acid (Pro-Pro-Glu) motif. PE–PPE proteins also share a variable C-terminal domain that contributes to structural and functional diversification within the [EspG3 structure (PDB code 4W4I)] (Chen et al., 2017)

The EthR structure contains synthetic inhibitors that occupy an allosteric pocket. The chain A and B chains are dimeric and contain 188 and 194 ordered residues, respectively, bound to inhibitor BDM31369. Based on the structure validation from PDBe, chain A with higher quality backbone and side chain geometry percentiles, was used in our docking calculations. According to studies into the resistance mechanism of ethionamide, an increase in the expression of HTH-Transcriptional regulator, EthR, results in the reduction of transcription and the level of EthA protein, which then results in resistance by *M.Tuberculosis (Mugumbate et al., 2017)*
The toxicity and adverse side effects of allopathic medicine have prompted an increase in the use of medicinal plants today. Medicinal plants can be used as medicines to treat TB and have been used since ancient times to treat diseases. Approximately 70% to 80% of the world's population trusts the traditional medical system, which extensively relies on herbal medicines (Shanley & Luz, 2003) Secondary metabolites,

including alkaloids, coumarins, flavonoids, polyphenols, terpenoids, quinones, and phytosteroid compounds, are available in all parts of plants, including leaves, stems, roots, and stem bark, and are widely used as medicines, including TB (Faron et al., 2004)

In order to control drug-resistant strains of Tuberculosis, new interventions are needed. There is a pressing need to create an intense enemy of this disease, to avoid bacterial contamination. So here certain phytochemical drugs were used against *M.Tuberculosis by means of Bioinformatics applications.*

Procedure

1. Ligand screening

For the Initial Ligand Screening purposes, a web based tool named SwissADME (http://www.swissadme.ch/) was used to eliminate a few compounds according to Lipinski's rule of five parameters. For a compound to qualify as ligand it should have <500Da molecular weight, a high lipophilicity i,e. Value of log P being less than 5, Hydrogen bond acceptors being less than 10 and H-bond donors less than 5. Any compound with more than 2 violations was ruled out for further study.

2. Protein Preparation and Active site Determination.

Required protein in pdb format was downloaded from the website **rcsb.org** commonly known as **Protein Data Bank**. 3D Conformers of the ligand were downloaded from Pubchem.

Using **PyMOL (Version 2.4.1)** software water molecules as well as native ligands from the proteins were removed, defined as cleaning/purification of the protein for further application. **Using a web server called Deep Site Active Pockets** of the proteins were calculated. The results calculated by the web server were in the form of different ids, centers and scores.

Scoring in Deep Site was using neural networking based on following instructions using DCNN architecture https://academic.oup.com/bioinformatics/article/33/19/3036/3859178. Center values for the grid were selected keeping score greater than 0.98.

UCSF Chimera (Version1.14) was used to prepare the receptor using the **DockPrep** function. Dock Prep Prepared structures for Docking using these functions.

- Deleting water molecules
- Repairing truncated sidechains.
- Adding Hydrogens.
- Assigning partial charges
- Writing files in Mol2 format

1. In silico Docking Using Auto dock Vina

Auto dock Vina (Version 1.1.2)along with UCSF Chimera (Version 1.14) was used for molecular docking studies. Center values and size of the grid of different scores were used from **DEEPSITE** calculations done above.

Following parameters were set in the auto dock vina.

Receptor option-

- **Add Hydrogens in Chimera (true** /false) - whether to add hydrogens in Chimera before calling the script. The receptor prep script will check for hydrogens and add them if they are missing. AutoDock Vina needs the polar (potentially H- bonding) Hydrogens to identify atoms types for scoring purposes.

- **Merge charges and remove non -polar Hydrogens(true**/false)- note Autodock Vina does not use charges or nonpolar hydrogens, so this setting is not expected to affect the results except for the presence or absence of nonpolar hydrogens in the processed receptor.

- **Merge Charges and remove lone pairs(true**/false)- note Autodock vina does not use charges for lone pairs so this setting is not expected to affect results except for the presence or absence of lone pairs in the processed receptor (and there may not have beenany lone pairs to start with).

- **Ignore waters(true**/false)
- **Ignore chains of non- standard residues (true**/false)- ignore chains composed entirelyof residues other than 20 standard amino acids.
- **Ignore all non- standard residues (true**/false)-ignore all residues other than the 20 standard amino acids.

For Ligands

- **Merge charges and remove non- polar hydrogens(true**/false)- note Auto Dock Vina does not use charges or nonpolar hydrogens, so this setting is not expected to affectresults except for the presence or absence of nonpolar hydrogens in the ligand output files.
- **Merge charges and remove lone pairs (true**/false)- note AutoDockVina does not use charges or lone pairs, so this setting is not expected to affect results except for the presence or absence of lone pairs in the ligand output files(and there may not have been any lone pairs to start with).

Docking Parameters

- **Number of Binding Modes** (1-10, **10**) - Maximum number of binding modes to generate.
- **Exhaustiveness of search** (1-8, **8**) - thoroughness of search, roughly proportional to time.
- **Maximum Energy difference** (kcal/mol)(1-3,**3**)- maximum score range; binding modeswith scores not within this range of the best score will be discarded.

The Docking Results were calculated by Auto Dock Vina using its scoring function and results were displayed in the form of scores and RMSD values. Docking results with the highest value score accompanied by negative sign and least RMSD values were chosen for further studies.

4. Residue Analysis

PyMOL was used for visualization of interactions of the docked structure at the ligand sites. Discovery Studio 2020 was used to study the ligand interactions and total number of residues. It was also used to plot the 2D structure of the interactions and residues.

5. Statistical Analysis

Descriptive, estimation and Hypothesis testing with confidence interval 95% was appliedto data using formula1 given below

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}}$$

Where,
CI = confidence interval
$\bar{x}$ = sample mean
z = confidence level value
s = sample standard deviation
n = sample size
Formula 1 used for calculation of confidence interval

Results and Discussion:

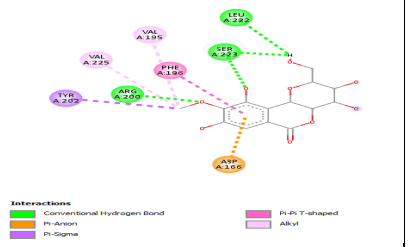
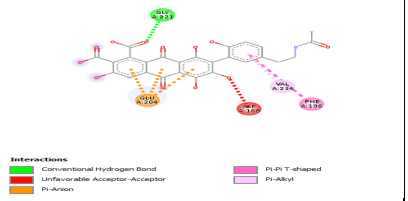
Molecular Docking
The docking result was obtained from AutoDock Vina in the form of Dock Score for all thethree proteins docked with above mentioned ligands.

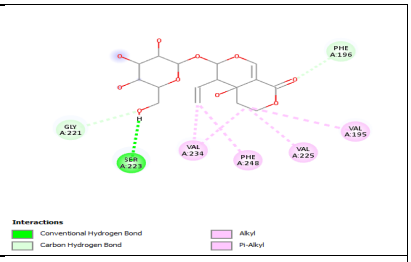
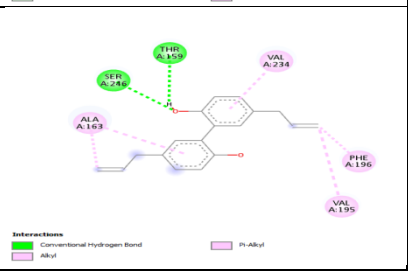
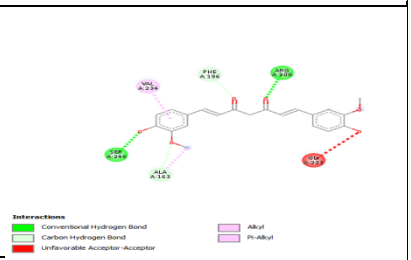
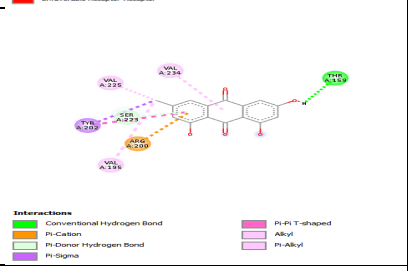
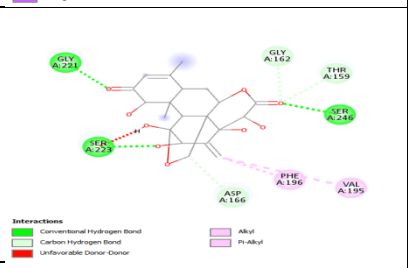
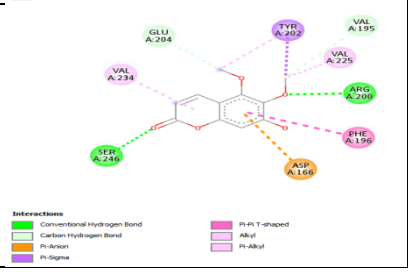
Tuberculosis Protein Docking Result
PDB-ID 4W4I
For 4W4I, Three active sites were selected out of which 0th active site was selected with a Deepsite score of 0.996. Table 1 the selection was made on the basis of the highest binding energy of the ligand- receptor. The docking results before statistics are shown in Table1 and Table 2 shows the post-statistical docking scores with Ligand Protein Interactions.

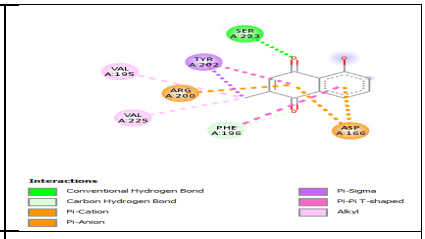
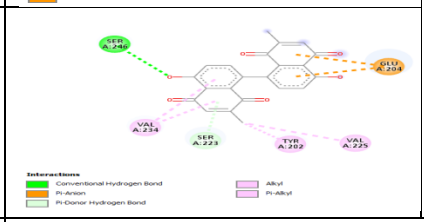
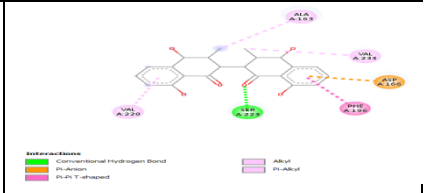
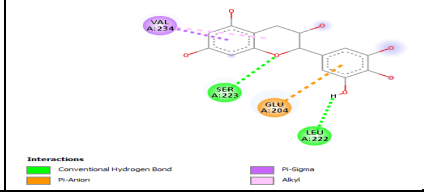
Sites	score	Selected/Not Selected
0	0.996	
1	0.986	
2	0.983	

Ligands	Dock score
Bergenin	-7.5
Laccaic acid	-8
Swertiamarin	-7.3
Magnolol	-7.1
Curcumin	-7.8
Emodin	-8.1
Pasakbumin A	-8.9
Umckalin	-6.6
Plumbagin	-6.6
Maritinone	-8.3
Tetrahydro3,3□-biplumbagin	-8.1
Epigallocatechin	-7.2

4W4I

Ligands	Dockscore	Interactions
Bergenin	-7.5	
Laccaic acid	-8	

Swertiamarin	-7.3	
Magnolol	-7.1	
Curcumin	-7.8	
Emodin	-8.1	
Pasakbumin A	-8.9	
Umckalin	-6.6	

Plumbagin	-6.6	
Maritinone	-8.3	
Tetrahydro3,3□□biplumbagin	-8.1	
Epigallocatechin	-7.2	

PDB-ID 3Q0V

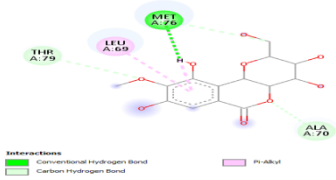
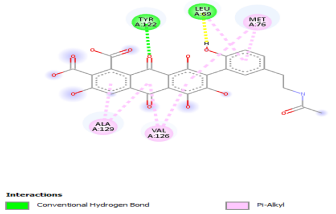
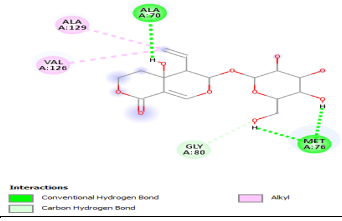
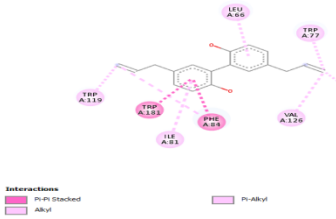
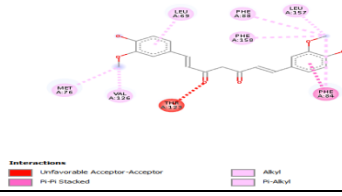
For 3Q0V, Two active sites were selected out of which 0th active site was selected with aDeep site score of 0.99. Table 1 the selection was made on the basis of the highest bindingenergy of the ligand- receptor. The docking results before statistics are shown in Table1 andTable 2 shows the post-statistical docking scores with Ligand Protein Interactions.

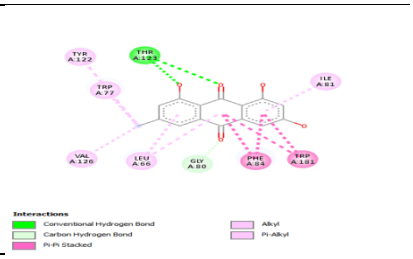
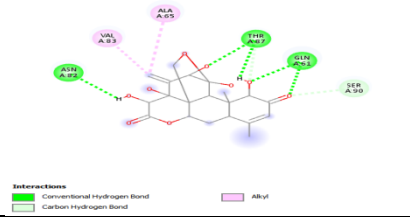
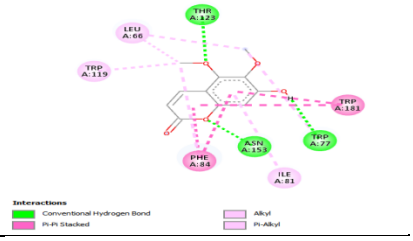
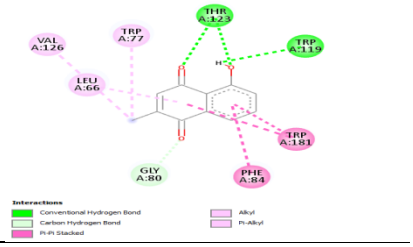
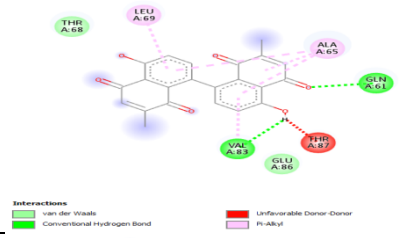
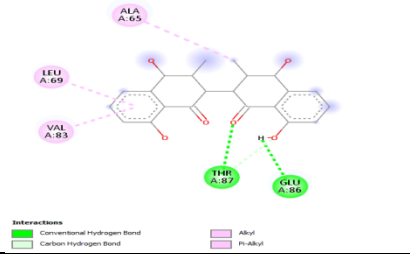
Sites	Score	Selected/Not selected
0	0.99	
1	0.87	

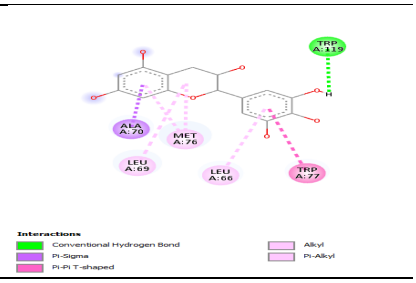
Ligands	Dock score
Bergenin	-5.4
Laccaic acid	-7.3
Swertiamarin	-5.6
Magnolol	-9.9
Curcumin	-9.9
Emodin	-9.3
Pasakbumin A	-6.1

Umckalin	-7.7
Plumbagin	-8.7
Maritinone	-6.6
Tetrahydro3,3□-biplumbagin	-6.1
Epigallocatechin	-8.2

3Q0V

Ligands	Dockscore	Interactions
Bergenin	-5.4	
Laccaic acid	-7.3	
Swertiamarin	-5.6	
Magnolol	-9.9	
Curcumin	-9.9	

Emodin	-9.3	
Pasakbumin A	-6.1	
Umckalin	-7.7	
Plumbagin	-8.7	
Maritinone	-6.6	
Tetrahydro3,3□-biplumbagin	-6.1	

Epigallocatechin	-8.2	

Ligands	**Acceptance**	
	PDB ID- 4W4I	PDB ID-3Q0V
Bergenin	Accepted	Strongly Accepted
Laccaic acid	Accepted	Accepted
Swertiamarin	Accepted	Strongly Accepted
Magnolol	Accepted	Accepted
Curcumin	Accepted	Accepted
Emodin	Accepted	Accepted
Pasakbumin A	Accepted	Strongly Accepted
Umckalin	Strongly Accepted	Accepted
Plumbagin	Strongly Accepted	Accepted
Maritinone	Accepted	Strongly Accepted
Tetrahydro3,3□-biplumbagin	Accepted	Strongly Accepted
Epigallocatechin	Accepted	Accepted

Conclusion:

All the twelve Ligands were studied using Bioavailability Radar. Our Result Proposed Umckalin and Plumbagin showed the best docking result for Tuberculosis Proteins with PDB Ids 4W4I. While Bergenin, Swertiamarin, Pasakbumin A, Maritinone, Tetrahydro3,3-biplumbagin showed the best docking result for Tuberculosis proteins with PDB IDs 3Q0V.To find the effectiveness and to propose the exact mechanism In-Vitro Studies can be encouraged on Umckalin, Plumbagin, Bergenin, Swertiamarin, Pasakbumin A, Maritinone, Tetrahydro3,3-biplumbagin targeting respective disease that are discussed above to understand the mechanism and potential cure for Tuberculosis.

Acknowledgement:

We would like to thank our supervisor, Bharat Kwatra, from Bharat Kwatra's Lab. whose expertise was invaluable in formulating the research questions, methodology and drawing Conclusions. His insightful feedback and guidance pushed us to sharpen our thinking and brought our work to a higher level.

References:

1. Asthana, P., Singh, D., Pedersen, J. S., Hynönen, M. J., Sulu, R., Murthy, A. V., Laitaoja, M., Jänis, J., Riley, L. W., & Venkatesan, R. (2021). Structural insights into the substrate-binding proteins Mce1A and Mce4A from Mycobacterium tubercu losis. IUCrJ, 8, 757–774. https://doi.org/10.1107/S2052252521006199

2. Chen, X., Cheng, H. F., Zhou, J., Chan, C. Y., Lau, K. F., Tsui, S. K. W., & Au, S. W. ngor. (2017). Structural basis of the PE–PPE protein interaction in Mycobacterium tuberculosis. Journal of Biological Chemistry, 292(41), 16880–16890. https://doi.org/10.1074/jbc.M117.802645

3. Faron, M. L. B., Perecin, M. B., Lago, A. A. do, Bovi, O. A., & Maia, N. B. (2004). Temperatura, nitrato de potássio e fotoperíodo na germinação de sementes de Hypericum perforatum L. e H. Brasiliense Choisy. Bragantia, 63(2), 193–199. https://doi.org/10.1590/s00068705200400 0200004

4. Fogel, N. (2015). Tuberculosis: A disease without boundaries. Tuberculosis, 95(5), 527–531. https://doi.org/10.1016/j.tube.2015.05.017

5. Knechel, N. A. (2009). Tuberculosis: Pathophysiology, clinical features, and diagnosis. Critical Care Nurse, 29(2), 34–43. https://doi.org/10.4037/ccn2009968

6. Kramarska, E., Squeglia, F., De Maio, F., Delogu, G., & Berisio, R. (2021). Pe_pgrs33, an important virulence factor of mycobacterium tuberculosis and potential target of host humoral immune response. In Cells (Vol. 10, Issue 1, pp. 1–20). https://doi.org/10.3390/cells10010161

7. Miggiano, R., Rizzi, M., & Ferraris, D. M. (2020). Mycobacterium tuberculosis pathogenesis, infection prevention and treatment. Pathogens, 9(5), 10–13. https://doi.org/10.3390/pathogens9050385

8. Mugumbate, G., Mendes, V., Blaszczyk, M., Sabbah, M., Papadatos, G., Lelievre, J., Ballell, L., Barros, D., Abell, C., Blundell, T. L., & Overington, J. P. (2017). Target identification of Mycobacterium tuberculosis phenotypic hits using a concerted chemogenomic, biophysical, and structural approach. Frontiers in Pharmacology, 8(SEP), 1–13. https://doi.org/10.3389/fphar.2017.00681

9. Musasa, J. I. (2015). Knowledge, attitude and practice with regard to tuberculosis and human immunodeficiency virus co-infection among patients with tuberculosis in Walvis Bay District, Namibia. The International Journal of Tuberculosis and Lung Disease, 144(27), 1399–1405.

10. O'Garra, A., Redford, P. S., McNab, F. W., Bloom, C. I., Wilkinson, R. J., & Berry, M. P. R. (2013). The immune response in tuberculosis. In Annual Review of Immunology (Vol. 31). https://doi.org/10.1146/annurev-immunol-032712-095939

11. Shanley, P., & Luz, L. (2003). The impacts of forest degradation on medicinal plant use and implications for health care in eastern Amazonia. BioScience, 53(6), 573–584. https://doi.org/10.1641/0006-3568(2003)053[0573:TIOFDO]2.0.CO;2

12. Sharma, D., & Yadav, J. P. (2017). Pe N rs ot on fo al r D U is se tri O bu n tio ly n. March. https://doi.org/10.2174/1389 5575166661605051146

13. Sundaramurthi, J. C., Brindha, S., Reddy, T. B. K., & Hanna, L. E. (2012). Informatics resources for tuberculosis - Towards drug discovery. Tuberculosis, 92(2), 133–138. https://doi.org/10.1016 /j.tube.2011.08.006