Journal of Innovations in Pharmaceuticals and Biological Sciences
www.jipbs.com
JIPBS
ISSN: 2349-2759

Research Article

Designing Novel Inhibitors for Tuberculosis (Tb) by Targeting InhA and KasA Using Ligand Based Drug Design

P. Sharma*[1], S. Chaudhary[1], T. Pardeshi[1], N. Sood[1], S. Sheikh[1], K. Deshmukh[1]

[1]RASA Life Science Informatics, 301, Dhanashree Apartments, opposite Chitaranjan Vatika, Model colony, Shivaji Nagar Pune- MH 411016, India

Abstract

Tuberculosis (TB) is a contagious disease caused by bacteria called Mycobacterium tuberculosis, infects an estimated nearly one third of the world population has latent tuberculosis infection, as it has been documented according to the World Health Organization. The emergence of multidrug resistant varieties of Mycobacterium tuberculosis has led to a search for novel drug targets. We have performed an in silico comparative analysis of metabolic pathways of the host Homo sapiens and the pathogen M. tuberculosis. Current therapy targets for TB treatment are based on the inhibiting of main proteins: the fatty-acid enoyl-acyl carrier protein reductase (InhA) and a complex of an acyl carrier protein (AcpM) and a β-ketoacyl-ACP synthase (KasA). In this study novel inhibitor was designed against the proteins responsible for mycolic acid synthesis found in Mycobacterium tuberculosis. The ligands were screened using integrated computational protocol that relies on methods such as docking, in house method of "loop docking" and ADMET analysis. The ADMET analysis of the ligand indicated that it is likely to be a drug candidate. It was observed that ligand with ID ZINC01757652 (Silybin) may prove to be a promising candidate drug for TB.

Key words: *Mycobacterium tuberculosis*, mycolic acid synthesis, Virtual Screening, Pharmacophore, Docking, Drug Designing.

***Corresponding Author: P. Sharma,** RASA Life Science Informatics, 301, Dhanashree Apartments, opposite Chitaranjan Vatika, Model colony, Shivaji Nagar Pune- MH 411016, India.

1. Introduction

Tuberculosis (TB) is a contagious disease caused by bacteria called *Mycobacterium tuberculosis*. Like the common cold, it is carried in airborne particles, called droplet nuclei of 1-5 microns in diameter. Infectious droplet nuclei are generated when persons who have pulmonary or laryngeal TB disease cough, sneeze, shout, or sing. Depending on the environment, these tiny particles can remain suspended in the air for several hours. Infected person have different symptoms

depending on the site of infection. Like in pulmonary TB, common symptoms are cough, shortness of breath, chest pain, loss of appetite, fever, etc. Tuberculosis is either latent or active. Latent TB means you have TB infection in your body but can't spread the infection to others. Active TB means that the TB bacteria are growing and causing symptoms, also easy to spread the disease to others. According to the World Health Organization, nearly one third of the world population has latent tuberculosis infection. After infection, a tuberculosis bacterium travels to the lungs and enters in to the aveoli, where they are recognized as foreign and are attacked by the body's macrophages. In the case of tuberculosis, not all of the bacteria cells will be destroyed no matter how excellent the host's immune system is, and the survivors infect and hijack macrophages feeding on them while increasing the bacteria population. Once macrophages are infected, they either kill the bacteria inside them or the bacteria multiply until they burst the macrophage, leading to further infection and extracellular bacilli. After that macrophages eliminate the bacteria by the process of phagocytosis. Phagocytes, the primary innate immune cells, engulf microorganism in phagosomes, which later combines with lysosome to form phagolysosomes. The acidic environment of the phagolysosome degrades the disease causing microorganism. They can't be destroyed by active acids produces by phagolysosomes because *M. Tuberculosis* has a thick, waxy mycolic acid capsule. The infected areas gradually transform into granuloma, a wall of macrophages intended to contain the infection. This also allows the *Mycobacterium Tuberculosis* to continue growing and overwhelm the cells it has infected until they die. Mycolic acids are major and specific lipid components of the mycobacterium cell envelope surrounding the cell membrane, providing extremely rigidity and protection to the cell wall. Such unique features in the cell envelope are important in the virulence and persistence of *M. Tuberculosis. M. Tuberculosis has* three main types: alpha-, methoxy-, and keto-mycolic acid.

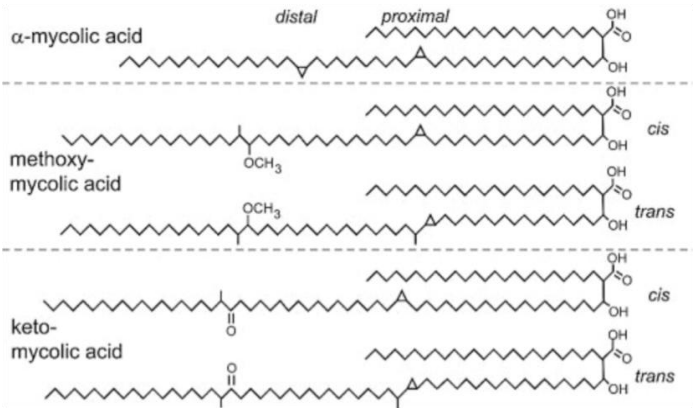


Figure 1: Three types of mycolic acid.

The three polymers in the cell wall, arabinogalactan-mycolate covalently linked with peptidoglycan and trehalose dimycolate, provide a thick layer that protects the tubercle bacillus from general antibiotics and the host's immune system. The enzymes involved in synthesis of mycolic acids are fatty-acid enoyl-acyl carrier protein reductase (InhA), and a complex of an acyl carrier protein (AcpM) and a β-ketoacyl-ACP synthase (KasA). The drugs Isoniazid, Ethionamide and all other play an important role in the inhibition of biosynthesis of mycolic acids. Specifically isoniazid inhibits InhA, the enoyl reductase from Mycobacterium tuberculosis, by forming a covalent adduct with the NAD cofactor. It is the INH-NAD adducts that acts as a slow, tight-binding competitive inhibitor of InhA. Treatment of TB will always involve a combination of many drugs at different times of the day. If people do not take their tuberculosis medications as recommended, the infection becomes much more difficult to

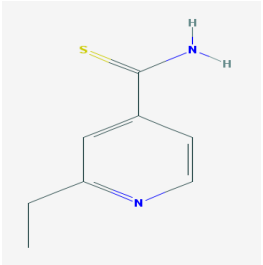
treat. Standard therapy for active TB consists of a six months regimen: two months with Rifater (Isoniazid, Rifampin, and Pyrazinamide), four months of Isoniazid and Rifampin (Rifamate, Rimactane) Ethambutol (Myambutol) or streptomycin added until your drug sensitivity is known [1]. Drug resistant tuberculosis, particularly that caused by strains resistant to Isoniazid and Rifampin, is much harder to treat and often is fatal. Medication side effects can be serious when they do occur, like highly toxic to liver. Rifampin can also cause severe flu-like signs and symptoms. When taking these medications, a person may experience nausea or vomiting, loss of appetite, a yellow colour to your skin, dark urine, etc. The increasing prevalence of MDR-TB has contributed to increase the difficulties in the treatment and control of TB. So, the development and design of new and potent anti-TB drugs without cross-resistance constitute a challenge for the scientific community. Nowadays, rational drug design can't be carried out without the use of important disciplines such as Chemoinformatics or Bioinformatics. This comes from the fact that from one side, Chemoinformatics includes design, creation, organization, management, analysis, dissemination, and visualization of chemical information [2]. On the other hand, Bioinformatics is focused on the creation and advancement of databases, algorithms, computational and statistical techniques and theory to solve formal and practical problems arising from the management and analysis of biological data [3].

In silico methods can help in identifying drug targets via bioinformatics tools. They can also be used to analyze the target structures for possible binding/ active sites, generate candidate molecules, check for their drug likeness, dock these molecules with the target, rank them according to their binding affinities, further optimize the molecules to improve binding characteristics. The use of computers and computational methods permeates all aspects of drug discovery today and forms the core of structure-based drug design. Our conclusion is that the in silico pharmacology paradigm is ongoing and presents a rich of opportunities that will assist in expatiating the discovery of new targets, and ultimately lead to compounds with predicted biological activity for these novel targets. The aim of this study was to design a lead compound against TB with the help of various in silico approaches against multiple protein targets followed by ADMET analysis for toxicity prediction.

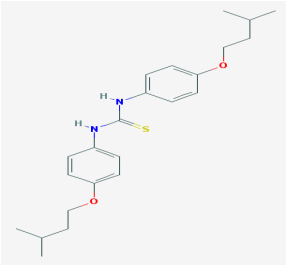
2. Materials and Methods

2.1 Screening the Targets for the gene responsible for the synthesis of mycolic acid:

Mycobacterium tuberculosis is known to synthesis alpha-, methoxy-, and keto-mycolic acids. The fatty-acid enoyl-acyl carrier protein reductase (InhA) and a complex of an acyl carrier protein (AcpM) and a beta-ketoacyl-ACP synthase (KasA) are the enzymes commonly involved in the synthesis of mycolic acid. The genes that encode these proteins are InhA, AcpM, and KasA. We are finalised six drugs that inhibit the pathway of mycolic acid synthesis: Ethionamide, Isoniazid, Triclosan, Isoxyl, Thiolactomycin, and Pyridomycin. (Figure 2)



Ethionamdie



Isoxyl

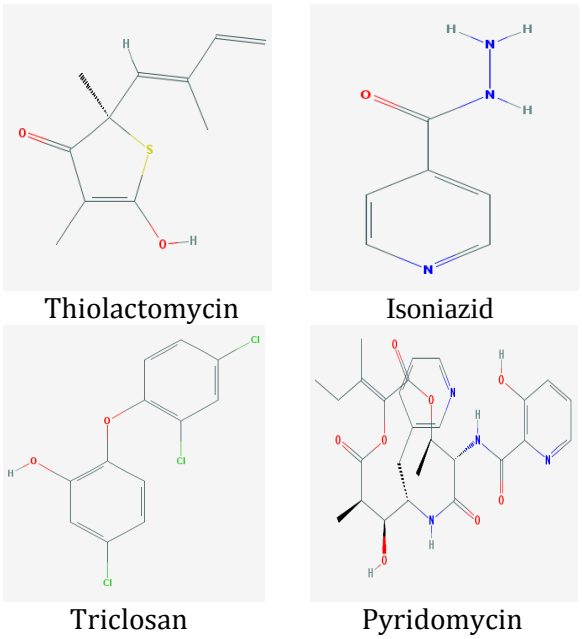


Figure 2: Anti-tuberculosis Drugs.

These drugs were downloaded from the server like Drug bank, Pubchem [4] in SDF file format and cleaned in Marvin View. Then combined into a single mol2 file in the Discovery Studio and submitted to PharmaGist server to generate pharmacophores, this uses ligand-based pharmacophore (http://bioinfo3d.cs.tau.ac.il/pharma/in ex.html) detection. It aligns a set of drug- like molecule that can bind to the receptors and pharmacophore detection of ligand. While submitting, the molecule was set as 'first input molecule' set a key key-molecule in an advanced option and a min number of pharmacophore feature was set to 5. Each of the six complex files were taken as pivot molecules with five features and the result was calculated each time. The pharmacophore alignment obtained from the uploaded complexes which showed highest score was chosen for further analysis. From six complex files, 24 pharmacophore were generated using the above mentioned combinations. Out of the pharmacophores generated only two were selected for further analysis.

These pharmacophores were then uploaded in the ZINCPharmer server (http://zincpharmer.csb.pitt.edu/). This server is free pharmacophore search software for screening the purchasable subset of the ZINC database it has features like it can identify pharmacophore features directly from structure and can identify the subset of ZINC database [5] and gives the result of structurally similar molecules to the uploaded pharmacophore. The pharmacophore was then analyzed using different subsets of ZINC database like Zinc Drug Database, Zinc in Man and Zinc Drug Database (Metabolites). Finally the ligands file for uploaded pharmacophores were downloaded. The total ligands obtained were around thirty thousand from both pharmacophore. But we used an in-house JAVA tool to remove the duplicate molecules which then drastically reduced the number of ligands to 753.

2.2 Binding Site Identification:

PDBsum (http://www.ebi.ac.uk/pdbsum/) database was used to find the native ligand binding site for the given receptor. The output of PDBsum is a colour, or black-and-white, PostScript file giving a simple and informative representation of the intermolecular interactions and their strengths, including hydrogen bonds, hydrophobic interactions and atom accessibilities. The proteins InhA and KasA were downloaded and checked in PDBsum for the ligplot. Using ligplot co-ordinates where the ligand can bind was estimated and noted down as X, Y and Z-axis. Then mean sum was calculated. Out of two proteins we have ligplot for 2X23 proteins only.

So for another protein 4C72, we have to check it out the pockets with the help of online pocket finder servers because it does not have ligplot. The different online servers for the prediction of the active site

are: **CASTp** (Computer Atlas of Surface Topography of Proteins) is a web server [6] that aims to provide a detailed quantitative characterization of interior cavities and surface pockets of proteins, which are prominent concave regions of proteins that are frequently associated with binding events (http:// sts.bioe.uic.edu/castp/calculation.php).
MetaPocket (http://projects.biotec.tu-dresden.de/metapocket/) is a comprehensive method in which the predicted site from eight methods: LIGSITE, PASS, Q-Site-Finder, SURFNET, Fpocket, GHECOM, ConCavity and POCASA are combined together to improve the protein-ligand binding prediction success rate. These eight methods are chosen because their binding provides source codes or executable binary and web-server available freely [7] [8].

2.3 Docking study:
The first phase of docking was performed on the molecules obtained from ZINCPharmer, PDB ID 2X23 and 4C72. As the ligands were large in number, multiple docking was done using AutoDockVina. The required script and parameters were taken from the (http://autodock. scripps.edu/) website. AutoDockVina is multiple docking open source software. AutoDock 4.2 software was used for the docking study combined with the Lamarckian genetic algorithm (LGA) to search for the globally optimized conformation. AutoDock 4.2 uses one of several conformational search algorithms to explore the conformational states of a flexible ligand, using the maps generated by AutoGrid to evaluate the ligand-protein interaction at each point in the docking simulation. The grid spacing was constant in each dimension, and each grid map consisted of a 60 x 60 x 60 grid point and the centre was calculated as per different receptors. At the end of a docking experiment with multiple runs, a cluster analysis was performed. Docking solutions with a ligand all-atom root mean square deviation (RMSD) within 0.1 nm of each other were clustered together and ranked by the lowest docking energy.

2.4 Loop Docking:
Loop docking is an in-house method used to validate the docking results. In most cases the best-docked structure cannot be a docking artefact and does not represent the best docking orientation. Hence we opt to perform the docking calculation using the best-docked structure from initial docking as starting structure for a second docking run. Few scripts were used to allow this process to be automated.
This automated ‘‘loop docking’’ will continue until a threshold value (d) is reached. The threshold value (d) is the difference between the docking binding energy of the last run and the preceding one. A threshold value of 0.05 was found to be appropriate. When this imposed value reached, the docking was stopped and the best-docked structure is selected. AutoDock software [9] is used for docking calculations.

2.5 ADMET Analysis:
ADMET analysis includes Absorption, Distribution, Metabolism, Excretion and Toxicity of a drug before its come under use. ADMET analysis done by the help of a comprehensive source and free tool for evaluating chemical ADMET properties like AMES Toxicity, Carcinogenicity, Carcinogens and other more properties by online server (http://lmmd.ecust.edu. cn:8000/predict/). These all properties follow the Lipinski’s Rule states that, in general an orally active drug has no more than one violation of the following criteria: 1. not more than 5 hydrogen bond donors (nitrogen or oxygen atoms with

one or more hydrogen atoms). 2. Not more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms). 3. A molecular mass less than 500 Daltons. 4. An octanol-water partition co-efficient logP not greater than 5. To find out the value of logP we have to check it out by using Pubchem server (http:// pubchem.ncbi.nlm.nih.gov).

3. Results and Discussion

In the present study, molecular targets for mycolic acid synthesis were identified using in silico technique. The molecules screened by forming the Pharmacophores of six molecules and further ligands obtained from ZINCPharmer were rigorously cleansed and then docking was performed on them with the following results:

3.1 Binding Site Identification:

There are two target proteins for the mycolic acid synthesis: 2X23 and 4C72 were downloaded and checked in PDBsum for ligplot. Out of these two proteins, only 2X23 has ligplot (Figure: 3). The ligplot co-ordinates where the ligand was shown interacting were selected and noted down as X, Y and Z-axis. Then mean sum was calculated for the protein 2X23 which was further used in grid preparation for docking.

For 4C72 protein we referred online pocket finding servers. The online servers used for prediction of the active site. CASTp web server showed results which had 126 pockets with different amino acids. Similarly MetaPocket had results in three different sites. By analyzing the results of these three different servers by docking we concluded that residue HIS in chain A at 345 positions is the appropriate binding site for our protein ligand docking. (Table No. 1)

3.2 Docking Analysis:

Initial docking was done on all the chemically diverse ligands from both the pharmacophores with the target proteins 2X23 and 4C72. AutoDock Vina results were analysed in AutoDock. Results were screened for hydrogen bond formation and lowest binding energy. This was done using the script to obtain top results and we analysed only top 10% to 20% of total ligands docked. These results were short-listed by lowest binding energy and top first results were selected as given in Table No. 2.

3.3 Loop Docking Results:

To carry out the loop docking procedure two ligand molecules were selected on the basis of the result obtained from the AutoDock vina. The ligands having zinc id ZINC01757652 with receptor 2X23 and ZINC01530603 with receptor 4C72 were selected for their lowest binding energy. Loop docking was done to validate the results obtained from AutoDock vina. On performing the loop docking and analysing the result both ligands show an average binding energy as shown in Figure No. 4 and Table No. 3.

3.4. ADMET Analysis:

ADMET analysis [10] had done by the help of a comprehensive source and free tool for evaluating chemical ADMET properties like AMES Toxicity, Carcinogenicity, Carcinogens and other more properties by online server (http**://**lmmd**.**ecust**.** edu**.**cn**:** 8000/predict/). Carcinogens denote a chemical substance or a mixture of chemical substances which induce cancer or increase its incidence (UNECE, 2004, p. 167). AMES Toxicity used for determining if a chemical is a mutagen or not. Ligands follow the Lipinski Rule and results are obtained by submitting the smiles are mention in the tabular form in Table No. 4 and 5.

Online Servers	Site No.	Pocket No.	Residue Name	Chain Name
CASTp		89	HIS	A
MetaPocket	Site 2		HIS	A
Ft site	Site 2		HIS	A

Table No. 1

Protein Name	Zinc Id	Binding Energy	Interacting amino acid	Conformation
2X23	ZINC01757652	-12.6	VAL65:HN1	1
	ZINC03831448	-12.7	VAL65:HN1	1
4C72	ZINC01530603	-10.7	HIS311:HE21	1
	ZINC01532344	-10.8	HIS311:HE21	1

Table No. 2

Protein Name	Zinc Id	Binding Energy	Interacting amino acid	Conformation
2X23	ZINC01757652	-12.6	VAL65:HN1	1
4C72	ZINC01530603	-10.7	HIS311:HE21	1

Table No. 3

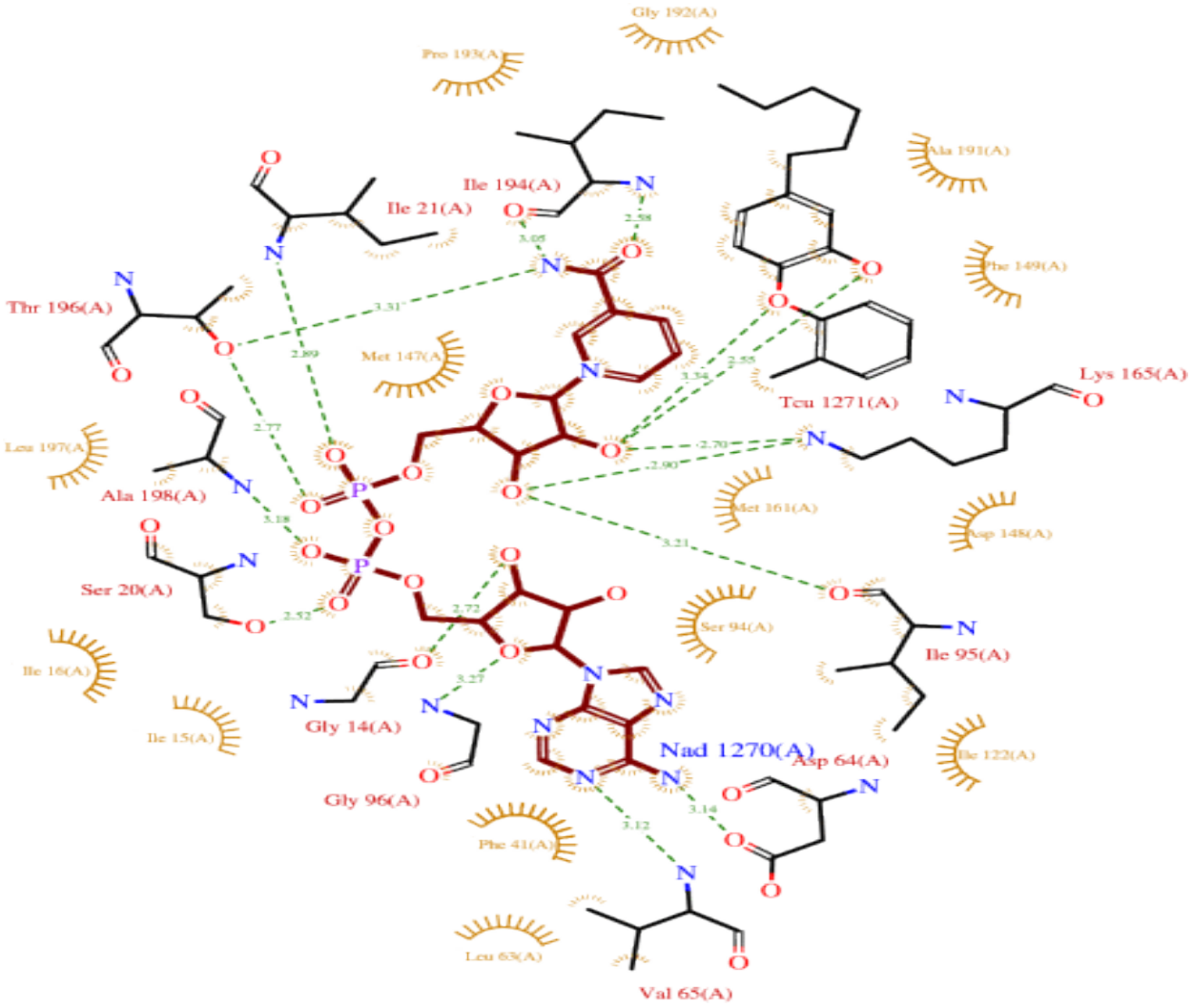


Figure 3: 2x23, Ligplot of interactions with ligplot.

ZINC ID	ZINC01757652	ZINC01530603
Protein Name	2X23	4C72
Molecular Weight	482.43618	434.87338
Molecular Formula	$C_{25}H_{22}O_{10}$	$C_{19}H_{17}ClN_3O_6S$
Hydrogen Acceptors	10	8
Hydrogen Donors	5	1
XLogP3	2.4	3.1
Carcinogen	Non-carcinogen	Carcinogen
Biodegradation	Not ready biodegradable	Not ready biodegradable

Table No. 4

Table No. 5

Toxicity		
Human Ether-a-go-go-Related Gene Inhibition	Weak inhibitor	0.9644
	Non-inhibitor	0.8101
AMES Toxicity	Non AMES toxic	0.9133
Carcinogens	Non-carcinogens	0.9385
Fish Toxicity	High FHMT	0.5506
Tetrahymena Pyriformis Toxicity	High TPT	0.9792
Honey Bee Toxicity	High HBT	0.6303
Biodegradation	Not ready biodegradable	0.8849
Acute Oral Toxicity	III	0.7236
Carcinogenicity (Three-class)	Non-required	0.7023

A. Toxicity prediction of 2X23 Protein with respective ligand having ZINC01757652.

Toxicity		
Human Ether-a-go-go-Related Gene Inhibition	Weak inhibitor	0.9994
	Non-inhibitor	0.8826
AMES Toxicity	Non AMES toxic	0.6492
Carcinogens	Carcinogens	0.5174
Fish Toxicity	High FHMT	0.9959
Tetrahymena Pyriformis Toxicity	High TPT	0.9226
Honey Bee Toxicity	Low HBT	0.7698
Biodegradation	Not ready biodegradable	0.9752
Acute Oral Toxicity	III	0.5001
Carcinogenicity (Three-class)	Non-required	0.5453

B. Toxicity prediction of 4C72 Protein with respective ligand having ZINC01530603.

After toxicity analysis we inferred that the ligand ZINC01757652 with a popular name Silybin (http://zinc.docking. Org/ substance/1757652) followed all the good characteristics of an ideal drug candidate and it should be used for further study whereas other ZINC01530603 with a popular name Sodium cloxacillin (http://zinc.docking.org/substance/1530 603) had followed Lipinski rule but was showed to be carcinogenic, hence this ligand would need more modifications to identify it as ideal drug candidate.

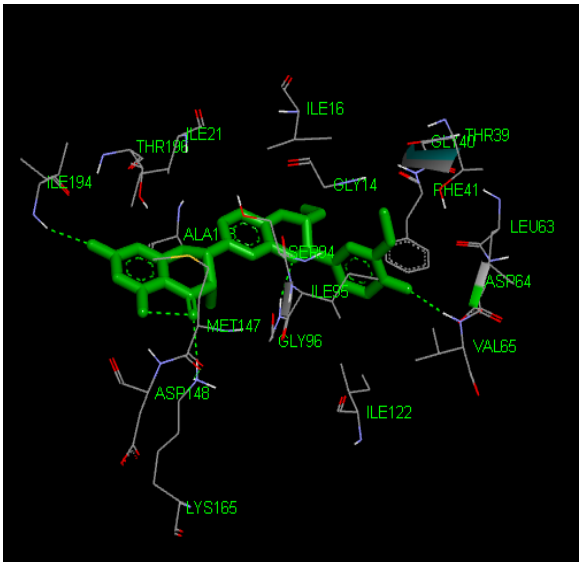


Figure 4: (A) shows the protein ligand interaction of ZINC01757652

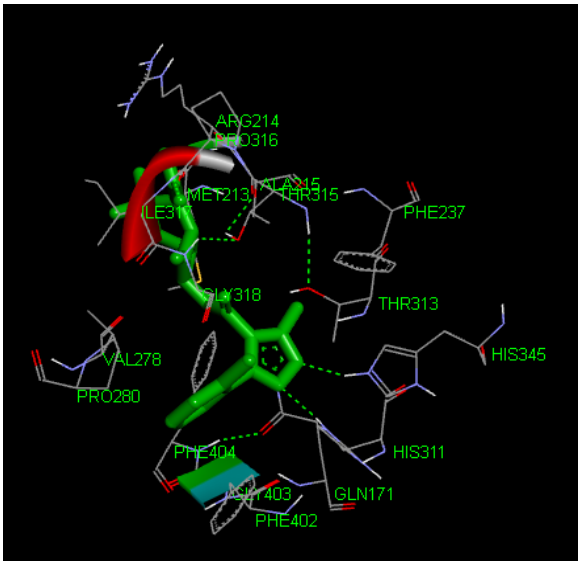


Figure 4: (B) shows the Protein ligand interaction of ZINC01530603

4. Conclusion

The present study uses in-silico screening approach using PharmaGist, AutoDock and Loop Docking to find out the potential inhibitor against mycolic acid synthesis. This study has revealed from pharmacophore virtual screening and docking studies that compound Silybin proposed as a potential to inhibit the synthesis of mycolic acid in the wall of *Mycobacterium tuberculosis* and a good start point for further research and optimization in laboratory.

References

1. Bhowmik D, Chiranjib, Chandira R. M, Jayakar B, and Kumar K.P.S: Recent trends of drug used treatment of tuberculosis. Journal of Chemical and Pharmaceutical Research 2009; 1:113-133.
2. Gasteiger, J. Handbook of Chemoinformatics, WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, 2003.
3. Larson, R.S. Bioinformatics and Drug Discovery, Humana Press Inc: Totowa, New Jersey, 2006.
4. Ihlenfeldt WD, et al: J Cheminform. 2009; 1:20.
5. Irwin JJ, Shoichet BK: J Chem Inf Model. 2005; 45:117.
6. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J: CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 2006; 34:116-118.
7. Konieczna I R: Identification of Ligand Binding Site and Protein-Protein Interaction Area. Springer Science & Business Media 2012.
8. Zhang1 Z, Li1 Y, Lin1 B, Schroeder2 M and Huang1, 2 B*: Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction.
9. Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J.: Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. J. Computational Chemistry, 1998; (19): 1639-1662.
10. Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, W. Lee P, Tang Y: admetSAR: a comprehensive source and free tool for evaluating chemical ADMET properties. J. Chem. Inf. Model., 2012; 52(11): 3099-3105