Int J Pharm Bio Sci 2015 Oct; 6(4): (B) 616 - 628

Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

VIRTUAL SCREENING FOR IDENTIFYING A PUTATIVE INHIBITOR OF RMLC, A MAJOR TARGET PROTEIN IN TUBERCULOSIS DISEASE

SHEFIN B*[1] AND BINDU A. SUNILKUMAR[2]

[1]Department of Bioinformatics, National College, Kerala, India,
[2] Department of Applied Nutrition and Food Chemistry Lund University, Lund, Sweden

ABSTRACT

Tuberculosis, is caused by various strains of *Mycobacteria*, and specifically the one by *Mycobacterium tuberculosis* needs effective treatment. Administration of antibiotics fails owing to multi-drug resistant capability of Mycobacterium. Traditional methods for the identification of a potential drug take time and require huge investment The present study was focused on virtual screening to identify an effective drug candidate against tuberculosis. RmlC protein, an important enzyme of the rhamnose pathway for bacterial pathogenicity is selected, as the target molecule. Inhibition of the RmlC protein function was achieved by docking with 1200 ligand molecules using GOLD software. Finally, five leads were identified with a GOLD score above 90. Hydrogen bonding pattern, Lipinski's rule, and Drug likeliness test of the potential leads were determined. On the basis of screening it can be summarized that the identified compounds can be further studied for their potentiality as suitable drug candidates.

KEY WORDS: Tuberculosis, disease, drug design, pathogen, target, lead, docking

SHEFIN B
Department of Bioinformatics, National College, Kerala, India

INTRODUCTION

Tuberculosis (TB) is the second largest infectious disease after HIV which are grouped asneglected tropical diseases (NTDs) including Malaria and Leishmaniasis. NTDs, mainly affecting low income populations of the developed countries are characterized by excessive suffering till death[1]. In 2013, it was estimated that more than 1.5 million people died because of TB and 9 millions were infected[2]. The problem is worse in a situation where more than one-third of cases are not reported on time and are not getting proper medication and care. Our human body opts a variety of fights when a foreign body invades in. Macrophages are one among the specialized cells that form a key part of the front line of defense against invading pathogens in the human immune system. They are formed in response to an infection or an accumulation of damaged or dead cells. When a person is infected with the tubercle bacilli, the causative bacteria of TB, his body starts producing macrophages and majority of these bacilli are destroyed or inhibited and are ingested by alveolar macrophages. However, a small number of bacteria gets the chance to multiply at intra-cellular level and are released when the macrophages are dead. This person is infectious and could spread TB in the air. When another person breathes in these germs, chances are more that he/she will become infected with tuberculosis. If alive, these bacilli may spread by way of lymphatic channels or through the bloodstream to more distant tissues and organs (including areas of the body in which TB disease is most likely to develop: regional nodes, apex of the lung, kidneys, brain, and bone)[3]. Taking proper antibiotics and other medications are the best practice to treat TB. Currently, TB chemotherapy is widely preferred as it shows positive effect in treating TB. It uses a cocktail of first-line drugs isoniazid (INH), Rifampicin (RIF), Pyrazinamide (PZA) and Ethambutol (EMB) given for six months. If the treatment fails as a result of bacterial drug resistance, or intolerance to one or more drugs, second-line drugs such as para-amino salicylate (PAS), Kanamycin, Fluoroquinolones, Capreomycin, Ethionamide and Cycloserine, are used that are either less effective or more toxic with serious side effects[4].The existing new agents also stand a negligible chance of slowing the rising epidemics of TB types such as Multidrug-resistant tuberculosis (MDR-TB), Extensively drug-resistant tuberculosis (XDR-TB) and now Totally drug-resistant (TDR- TB)[5]. Besides these, another major hindrance is that Bacillus Calmette Guerin (BCG) is the only vaccine available for prevention of TB in humans[6]. The experimental techniques available for identification of *Mycobacterium tuberculosis* inhibitors such as plating, micro plate Alamar blue assay, BACTEC 460 etc. are very expensive, time-consuming and tedious those require sophisticated systems for controlling the risk of infection. Hence the theoretical techniques such as Virtual Screening or *In Silico* screening gaining its importance[7,8]. Structure Based Drug Design through Virtual Screening is used to discover new drug candidates from different chemical scaffolds by searching commercial, public, or private three-dimensional chemical structure databases. The goal is to enrich set of molecules with desirable properties (active, drug-like, lead-like) and eliminate compounds with undesirable properties (inactive, reactive, toxic, poor ADMET/PK)[9].The Structure Based Drug Design (SBDD) uses known three-dimensional geometric shapes or structure of proteins to assist in the development of new drug compounds. Technical advancements in X-ray crystallography and NMR (Nuclear Magnetic Resonance) Spectroscopy techniques have made it easier to obtain high-resolution structural data for many protein-ligand complexes[10]. This ability to work at high resolution with both proteins and drug compounds makes SBDD one of the most powerful methods in drug designing. Compared to conventional methods which were time consuming and less logical, new drug designing based on structure is rational, evidence based, faster and more scientific in nature[11]. In the era of modern medicine, where newer insights into the molecular level of disease processes are available, it is very essential that drug designing should be based on molecular mechanism of pathologic processes. Thus, the computational structure-based drug designing opens the door

to novel treatments in modern medicine. Therefore, the present study focuses on Docking Studies through Virtual Screening, using the GOLD (Genetic Optimisation for Ligand Docking) software that employs Genetic Algorithm (GA). This method allows a partial flexibility of ligand molecule. The two scoring functions available in the GOLD software to measure the affinity of ligand for protein binding site are GOLDScore fitness function and ChemScore fitness function. Based on the Gold Score of Docking process, further redocking, H-bonding investigation, Lipinski's rule evaluation and Drug likeliness tests were conducted to identify a potential inhibitor.

MATERIALS AND METHODS

(i) Preparation of Target protein

The X-ray crystallographic structure of the target protein RmlC complexed with Rhamnose (PDB ID: 2IXC) with a resolution of 1.79 A^0 was retrieved from RCSB PDB(Research Collaboratory for Structural Bioinformatics) database, which is a repository for the three-dimensional structural data for protein & nucleic acids (http://www.rcsb.org)[12]. The intact RmlC protein with active site was positioned within 4.000 A^0 vicinity of the ligand molecule by using the SWISS-PDB viewer software[13]. Preparation of active site was done by removing water molecules, ionizing key amino acid side chains, adding and optimization of hydrogen atoms in the protein. All the atoms making the active site was listed as a file using ArgusLab software (Argus lab 4.0.1)[14].The total number of atoms present in the active site was saved in a notepad.

(ii)Ligand Input File Preparation and Optimization

Ligand input structure of 1200 biologically annotated small molecules with structure were drawn from a database named Meta Database Ligand.Info(http://Ligand.Info)[15]. All the downloaded molecules were displayed in Swiss PDB viewer and are saved in SDF format for virtual high throughput screening.

(iii) Docking methodology

When the target protein and the ligand molecules were ready, molecular docking was performed using GOLD (Genetic Optimization for Ligand Docking) 5.0 program[16].Protein preparation including binding pocket information and ligand library preparation were carried out using the docking wizard of GOLD program. A few independent parameters were applied for fitness function (hydrogen bond energies, atom radii and polarizabilities, torsion potentials etc.), which were taken from the GOLD parameter file. Best-docked complexes were ranked based on their GOLD fitness score that ranges from 40 and above and non-bonded interaction analysis. Hydrogen bond contacts, lipophilic interactions and non-bonded contacts were calculated using LIGPLOT v.4.5.3[17]. Analysis of hydrogen bonding interaction pattern was conducted using Silver, a post processing docking search algorithm.

(iv)Re-docking of the top compounds with good Gold Fitness Score

In order to avoid the false positives and negative rates that arise from the methodological discrepancies, redocking was carried out independently. About 30 molecules with good Gold Fitness Score obtained from docking study is subjected for redocking. This was followed by strict visualization using Pymol software. Even though GOLD offers a variety of scoring functions, like Gold Score, Chem Score and ASP (Astex Statistical Potential), Gold Fitness Score being the original and default scoring function is preferred. The fitness score is taken as the negative of the sum of the component energy items so that the larger scores are better. Consequently, five compounds were identified with respect to the Gold Fitness Score.

(v) Hydrogen bonding investigation

Hydrogen bonds stipulate the specificity of a ligand binding to its target macromolecule. Hence, it is very significant to perform the hydrogen bond investigation. This study can be done using the PDBsum server. The PDB file of the docked complex for which the analysis has to be done is submitted to PDBsum server via Generate option in the EMBL EBI website along

with our email ID. Once the run is over, the result will be sent to our mail ID. The result shows a complete structural analysis along with the hydrogen bond details.

(vi)Lipinski's Rule Filtering and Drug likeliness Test

The Lipinski's rule filter necessitate the PDB format of the compounds identified through docking and redocking process, to execute the filtering course[18]. Hence, the 3D structures of the identified five compounds were generated using CORINA interface provided with SMILES format of the ligands. The 3D structures saved in the PDB format are submitted to the filtering tool available at http://www.scfbio-iitd.res.in/utility/LipinskiFilters.jsp. It predicts high probability of success rates of the identified five compounds in clinical trials. With respect to this rule, poor absorption is anticipated, if Molecular weight is greater than 500; log P greater than 5, Hydrogen bond donors greater than 5 and Hydrogen bond acceptors greater than 10.

(vii)Absorption, Distribution, Metabolism, Excretion, and toxicity (ADMET)

ADMET test serves as a prerogative of computational study analyses the disposition of drug molecule within organism. The kinetics and behavior of drug molecule depend on these criteria and need to be evaluated to scrutiny the performance and pharmacological activity of the compound as a drug. To select drug-like molecules, DruLiTo software was used essentially to screen the selected five compounds based on eight filters namely Lipinski's rule, MDDR-like rule, Veber rule, Ghose filter, BBB rule, CMC-50 like rule, weighted and unweighted Quantitative Estimate of Drug-likeness[19]. DruLiTo is an open-source virtual drug-likeness tool which can be downloaded freely from the internet. If we provide the SDF format of the compounds into the interface of DruLiTo software, and proceed to calculate properties, the molecular properties of the selected compound will be displayed in the window.

RESULTS AND DISCUSSION

According to World health Organisation, about 15 million (> 25 %) of 57 million annual deaths world wide are the direct result of infectious diseases[20.] It is generally accepted that the identification of causative organisms and the development of proper drug or antibiotics has changed the mortality rate recently, but it is not always true in the case of most infectious diseases especially Neglected Tropical Diseases. Therefore, a proper and powerful drugs become paramount that are need to be safe and effective[21]. Bioinformatics especially medical informatics focuses on the application of computer and information science in the field of medicine that have brought special algorithms, softwares and programs for the identification causative organism and specific protein providing immunogenicity in host cell. Structure based Virtual screening has been widely used in the area of medical informatics that helping to identify target molecules, hit to lead selection, structural annotation of selected lead molecules and validation. In addition absorption, distribution, metabolism, excretion and toxicity identification as well as safety issues can also identify using bioinformatics tools[22]. In silico drug target identification mainly relies on the principle "a good drug target is a gene essential for bacterial survival yet cannot be found in a host" [23]. In this study screening of 1200 molecules against target proteins has resulted 5 potential drug molecules which were later validated with different bioinformatics tools.

1.Target protein

RmlC protein (PDB ID: 2IXC) complexed with L-Rhamnose downloaded from PDB data bank is displayed in Figure 1.RmlC (dTDP-6-deoxy-d- xylo-4-hexulose 3,5-epimerase, EC 5.1.3.13) is one of the four enzymes in biosynthetic pathway of dTDP-L-rhamnose pathway, which catalyzes an unusual double epimerization reaction at positions C3 and C5. L-Rhamnose pathway commonly exists in bacterial cells, which is required for the synthesis of bacterial cell wall.

RmlC protein structure

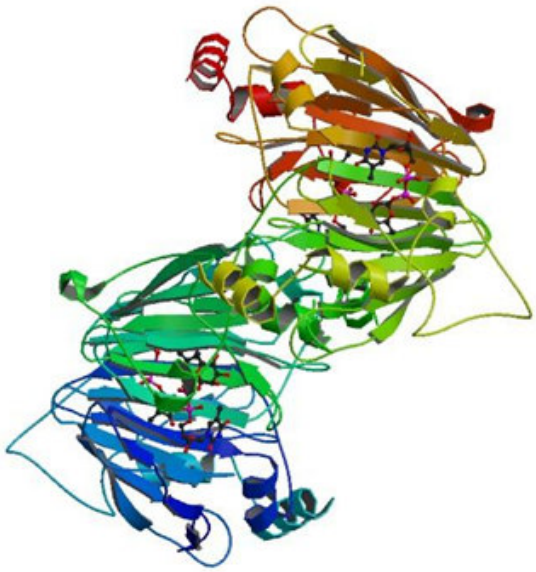

Figure 1
RmlC protein with natural ligand DTDP L Rhamnose

The natural ligand dTDP-L-rhamnose forms the disaccharide linker unit (α-l-rhamnosyl-(1→3)-α-d-N-acetyl-glucosaminosyl-1-phosphate), which connects the middle layer of bacterial cell wall (arabinogalactan) to the innermost layer peptidoglycan. dTDP-L-rhamnose is synthesized from glucose-1-phosphate which requires the enzyme RmlC . By inhibiting the RmlC protein, the dTDP-L-rhamnose biosynthetic pathway can be blocked, which in turn blocks the linker unit biosynthesis. Without the linker unit, cell wall cannot uphold which open-up an entirely new prospect of targeting RmlC as a potential drug candidate to address the infection caused by *Mycobacterium tuberculosis*. The downloaded structure was validated using Pro Check (Pro Check 2.3) program[24] and is shown in Figure 2. Ramachandran plot demonstrated the relationship between the phi and psi angles of RmlC protein structure and the position of amino acids (aa)[25]. From the Phi and Psi scatter plot stability of RmlC protein structure were confirmed.

Ramachandran plot of RmlC protein

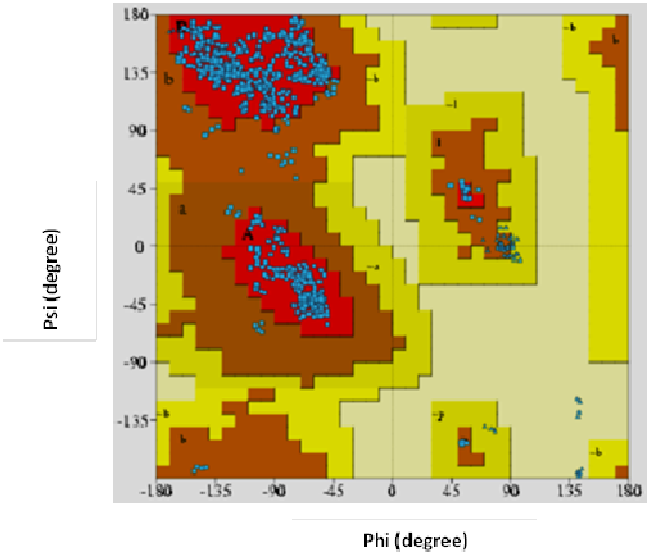

Figure 2
Ramachandran plot of RmlC protein generated using ProCheck 2.3 tool

2. Identification and preparation of active site

The active site (Figure 3) of the RmlC protein is located using the PDBsum server through which the active site residues of the natural ligand, DTDP L Rhamnose was selected for the study. Almost 20 amino acid residues within 4.000 A^0 vicinity of the ligand molecule is found and displayed through the Control panel of the Swiss PDB Viewer Stand alone software application.

The Active site of RmlC target protein

Figure 3
The active site of RmlC target protein with amino acid residues located using PDB sum server.

3. Virtual Screening

Docking of RmlC protein using virtual screening procedures was carried out using the software GOLD 5.0. At first, the target protein was docked with about 1200 ligand compounds. Of the 1200 molecules docked, the first top 30 compounds were selected based on GOLD score which ranges from 99.53 to 86.84 and is shown in Table 1.

Table 1
First 30 lead molecules with GOLD Score

Sl.No	IUPAC Names of Hit molecules	Gold Score
1	3-((4-O-(4-O-(3-O-acetyl-2,6-dideoxy-4-O-hexopyranosylhexopyranosyl)-2,6-dideoxy hexopyranosyl)-2,6-dideoxyhexopyranosyl)oxy)-12,14-dihydroxycard-20(22)-enolide	99.53
2	No IUPAC name available	99.30
3	No IUPAC name available	98.92
4	3-((4-O-(4-O-(3-O-acetyl-2,6-dideoxy-4-O-hexopyranosylhexopyranosyl)-2,6-dideoxyhexopyranosyl)-2,6-dideoxyhexopyranosyl)oxy)-14-hydroxycard-20(22)-enolide	96.22
5	1-deoxy-1-(7,8-dimethyl-2,4-dioxo-3,4-dihydrobenzo[g]pteridinn-10(2H)-yl)hexitol	95.70
6	3-((ethyl-4-((4-(ethyl(3-sulfobenzyl)-lambda~5~-azanylidene)-2,5-cyclohexadien-1-ylidene)(phenyl)methyl)aniline)methyl)benzenesulfonic acid	94.19
7	3-((4-((2-chlorophenyl)(4-(ethyl(3-sulfo)-lambda~5~-azanylidene)-2,5-cyclohexadien-1-ylidene))(ethyl)aniline)methyl)benzenesulfonic acid	94.19
8	29-hydroxy-11,29-dioxoolean-12-en-3-yl 2-O-hexopyranuronosylhexopyranosiduronic acid	93.86
9	Phosphoric acid compound with N~1~-(7-chloro-1,2,3,4-tetrahydro-9-acridinyl)-N~3~,N~3~-dipentyl-1,3-propanediamine(1:1)	93.64
10	3-((12-hydroxy-9-octadecenoyl)oxy)propyl-12-hydroxy-9-octadecenoate	93.25
11	No IUPAC name available	92.66
12	Phosphoric acid compound with N~1~,N~1~-diethyl-N~4~-(1,2,3,4-tetrahydro-9-acridinyl)-1,4-pentanediamine(1:1)	92.17
13	Complexon IV	92.09
14	Phosphoric acid compound with N~1~-(7-chloro-1,2,3,4-tetrahydro-9-acridinyl)-N~3~,N~3~-diisopentyl-1,3-propanediamine (1:1)	91.83
15	2-methyl-2-(1-pyrrolidinyl)propyl 2-cyclopentyl pentanoate 2-hydroxy-1,2,3-propanetricarboxylate	91.83
16	No IUPAC name available	90.93
17	Phosphoric acid compound with N~1~,N~1~-diisopentyl-N~3~-(1,2,3,4-tetrahydro-9-acridinyl)-1,3-propanediamine(1:1)	90.84
18	2-methyl-5-((5-(4-methyl-3-sulfoanilino)-9,10-dioxo-9,10-dihydro-1-anthracenyl)amino)benzene sulfonic acid	90.80
19	3-((2,6-dideoxy-4-O-(2,6-dideoxy-4-O-(2,6-dideoxyhexopyranosyl)hexopyranosyl)hexopyranosyl)oxy)-14-hydroxycard-20(22)-enolide	89.99
20	HEEDTA	89.98
21	1-deoxy-1-(6,7-dimethyl-2,4-dioxo-3,4-dihydrobenzo[g]pteridin-10(2H)-yl)pentitol	89.38
22	No IUPAC name available	89.25
23	Phosphoric acid compound with N~1~,N~1~-dibutyl-N~3~-(1,2,3,4-tetrahydro-9-acidinyl)-1,3-propanediamine(1:1)	89.02
24	No IUPAC name available	88.64
25	Octadecanoic acid compound with 1,2-propanediol	88.49
26	Phenyl 4-O-hexopyranosyl-1-thiohexopyranoside	87.94
27	1-methyl-2-(1-pyrrolidinyl)propyl 2-cyclopenten-1-yl(phenyl)acetate 2-hydroxy-1,2,3-propanetricarboxylate	87.59
28	1-(4-aminosulfonyl)aniline)3-phenyl-1,3-propanedisulfonic acid	87.54
29	2,3-dihydroxypropyl 12-hydroxy-9-octadecenoate	87.18
30	9,10,16-trihydroxyhexadecanoic acid	86.84

The selected 30 lead molecules with corresponding GOLD Score after docking with 1200 ligands separately in RmlC protein active site.

4. Re-docking

Top 30 compounds with highest Gold Score obtained from the docking study was selected for the re-docking. Re-docking helps to stay away from selecting false positives. After redocking, five lead compounds were selected

as the best leads with high Gold Fitness Score ranging from 89.92-95.89. Redocking results are recorded in Table II and the structure of docked complex were shown in the figure 4. Also the hydrogen bonding pattern of the five top hits were necessarily studied as it indicates the binding affinity of the ligand to its target. Higher the number of hydrogen bond, better the affinity. The hydrogen bond investigation results are displayed in Table III.

Table II
The top 5 lead compounds with GOLD Score

Hits	IUPAC Names of hit molecules	GOLD Score
1	No IUPAC Name available	95.89
2	No IUPAC Name available	93.65
3	9,10,16-trihydroxyhexadecanoic acid	91.74
4	Complexon IV	90.24
5	3-((4-((2-chlorophenyl)(4-ethyl(3-sulfobenzyl)-lambda~5~-azanylidene)-2,5-cyclohexadien-1-ylidene)methyl)(ethyl)aniline)methyl)benzenesulfonic acid	89.92

Selected 5 lead compounds with IUPAC name and GOLD Score after redocking

Table III
Hydrogen bonding interaction details

Ligands	H Bonding Residues	Distance	Inhibitor
No IUPAC Name available	ARG170	2.906	O1
	HIS62	2.848	O1
	ARG23	2.815	O21
	PHE26	3.302	S12
	HIS19	2.988	S14
No IUPAC Name available	GLN47	2.925	O1
	GLU143	2.928	O11
	LYS72	2.412	O42
	ARG59	2.979	O27
	HIS19	2.293	O28
9,10,16-trihydroxyhexadecanoic acid	LYS72	2.752	O21
	HIS19	2.843	O21
	ARG23	2.855	O1
	GLN47	2.289	S14
	PHE26	2.968	O24
Complexon IV	HIS119	2.666	O24
	LYS72	2.621	O24
	TYR132	2.886	O19
	HIS62	2.869	N5
	HIS63	2.621	O23
	ARG23	2.764	O1
	ARG170	2.925	O8
	ARG170	2.869	O42
3-((4-((2-chlorophenyl)(4-ethyl(3-sulfobenzyl)-lambda~5~-azanylidene)-2,5-cyclohexadien-1-ylidene)methyl)(ethyl)aniline)methyl)benzenesulfonic acid	ARG59	2.718	O40
	ARG23	2.750	O37
	ARG59	3.047	O12
	ARG23	3.002	O14
	ARG170	3.008	O8
	HIS62	2.459	O21

Hydrogen bonding investigation results of five lead compounds

RmlC *active site with lead molecules (docked complex)*

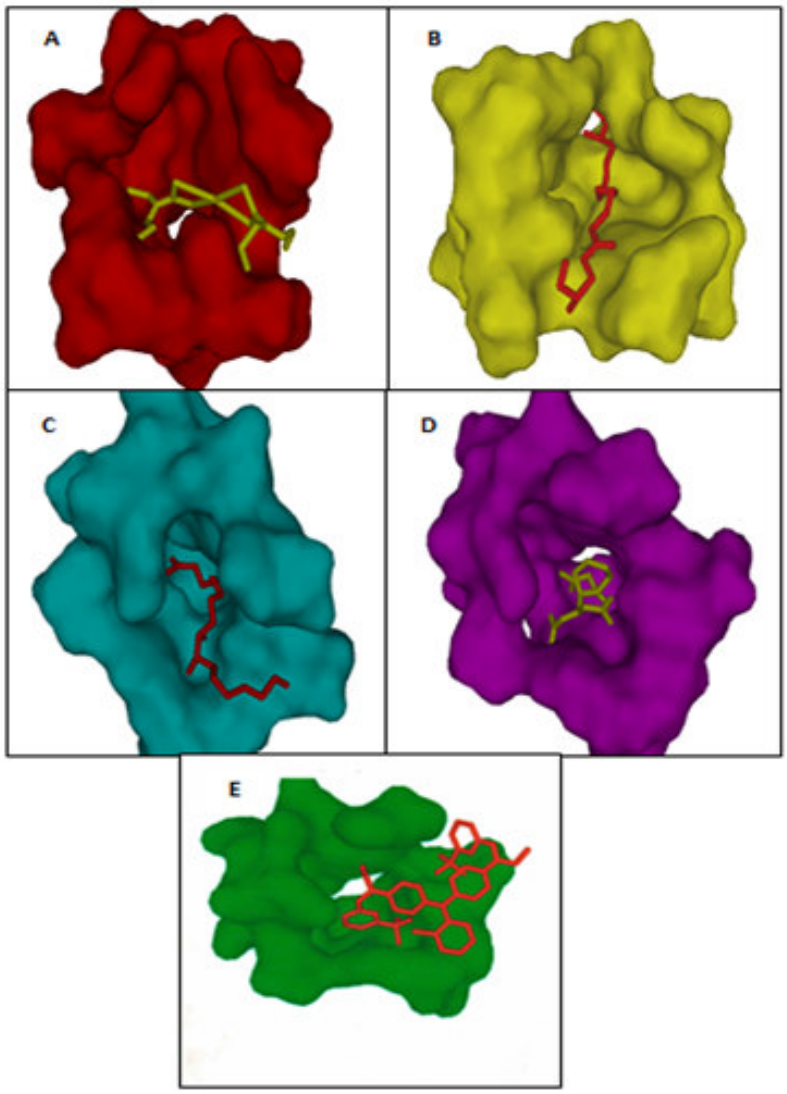

Figure 4
Structure of docked complexes

(A) RmlC active site with first hit (no IUPAC name) having Gold Score 95.89,
(B) RmlC active site with second hit (no IUPAC name) having Gold Score 93.65
(C) RmlC active site with 9,10,16-trihydroxyhexadecanoic acid having Gold Score 91.74
(D) RmlC active site with Complexon IV) having Gold Score 90.24
(E) RmlC active site with 3-((4-((2-chlorophenyl)(4-ethyl(3-sulfobenzyl)-lambda~5~-azanylidene cyclohexadieylidene)methyl)(ethyl)aniline)ethyl)benzenesulfonic acid having Gold Score 89.92.

5.Lipinski's Rule Filtering and Drug likeliness Test

Lipinski's Rule is widely followed in computational biology for screening the drug likeliness of the candidate molecules. The rules are based on the 90-percentile values of the drugs property distributions, applicable only to absorption by passive diffusion of compounds through cell membranes; compounds that are actively transported through cell membranes by transporter proteins are exceptions to the rule. Due in no small part to their simplicity, the Lipinski criteria are widely used in medicinal chemistry to predict not only the absorption of compounds, as Lipinski originally intended, but also overall drug-likeliness. Two hits out of five obtained from re-docking results

satisfied the Lipinski's rule of five. The results are shown in Table IV. Drug likeliness rules describe the guidelines for the structural properties of a compound and it helps to identify whether the compounds identified through docking and re-docking satisfies the structural properties so that they can be suggested as a lead molecule. Also the filters for drug-likeness are not enough to state that the compounds will be successful in further drug discovery stages, but identification of such 'warhead agents', 'frequent hitter' and 'promiscuous inhibitors' are important to evaluate the chances of success for a particular scaffold in a drug discovery process[21].The Drug likeliness test was performed using the DruLito software. Table.V shows the DruLito results.

Table IV
Over all drug –likeliness - Lipinskis rule of five

Hits	IUPAC name	Molecular Formula	MolWt	H bond Acceptors	H bond donors	Log P
1	Nil	C10H20HgN2O4S4	562	8	2	-0.6
2	Nil	C20H10 CuN2O16S4	726	5	3	-5.658
3	9,10,16-trihydroxy hexadecanoic acid	C16H32O5	304	5	4	-0.603
4	Complexon IV	C14H22N2O8	346	6	3	-4.883
5	3-((4-((2-chlorophenyl)(4-ethyl(3-sulfobenzyl)-lambda~5~-azanylidene)-2,5-cyclohexadien-1-ylidene)methyl)(ethyl)aniline)methyl)benzenesulfonic acid	C37H35N20O6S2	919	6	3	-4.883

The physicochemical properties of selected lead compounds after testing the Lipinskis rule of five.

Table V
Drug –likeness – DruLito Results

Hits	Clogp*	TPSA*	AMR*	nRB*	nAtom*	nAcidicGroup*	RC*	nRB*	nAR*	nHB*	SAlerts*
1	0	87.4	0	9	41	0	2	9	0	10	3
2	-3.1	160.7	92.2	5	61	0	3	24	0	19	0
3	-4.8	98	63	15	53	1	0	5	0	9	1
4	-2.9	155.7	71.4	10	46	4	1	14	0	14	0
5	-1.9	87.4	157.3	11	108	0	5	36	0	10	0

DruLiTo Results. *clogP:compound's hydrophilicity, TPSA:The Polar Surface Area Prediction, AMR: molculear refractivity, nRB-number of Rotatable Bonds, nAtom- number of Atom, nAcidicGroup-number of acidic groups, RC:Rotatable bond count, nRigidB:number of rigid bond, nAtomRing:number of Atom Ring, nHB:number of Hydrogen Bond, SAlerts-Structure alerts.

Ever since the advent of tuberculosis disease, *Mycobacterium tuberculosis* has been subjected to bottom-up research work. The unique cell wall structure that helps the bacteria to invade the immune mechanism of humans is well characterized. Mycobacterium species are provided with unique cell wall components, consisting of a covalently linked complex of mycolic acids, D-Araban and D-Galactan (mycolylarabinogalactan, mAG), which in turn is linked to peptidoglycan via a special linkage unit, -α-L-Rha*p*(1→3)-D-GlcNAc-P-.This unique cell wall structure makes it resilient to the autophagolysosome complex and thus survives as a successful pathogen[26].The traditional methodology employed in designing and discovery of drugs were time consuming and demanded high throughput screening procedures. It involves processes like *in vivo* biological screens, pharmacokinetic properties and potential toxicity studies, which resulted in poor efficacy, animal toxicity, and adverse human effects. Thus, drug discovery process has been revolutionized in many aspects. The molecular docking methods are widely used to reduce the cost and time involved in the process of drug discovery[23]. The method used in the current study has previously been confirmed to be competent in identifying novel

inhibitors from the wide library of compounds. The study identified five putative small molecules as inhibitors that bind well to the active site of the target molecule chosen for the study. All these five hits predicted to "dock" well as inferred from the GOLD score into the active site of RmlC protein. Detailed study of the redocking results with Lipiniski's filter showed that only two compounds 9,10,16-trihydroxy hexadecanoic acid, Complexon IV, out of five satisfied the rule. Among the lead compounds, ComplexonIV shows eight hydrogen bonds, which is essential for the structure and function of biological molecules, whereas the natural ligand shows only four hydrogen bonds. The hydrogen bonding indicates that all lead molecules are capable of modulating the target activity. In this, Histidine, Lysine, and Phenylalanine show hydrogen bond with hit molecules and the same hydrogen-bonding pattern with existing inhibitor Oxygen, Nitrogen and Sulphur. Lipinski rule test and Drug likeliness test were also performed. It is interesting that the GOLD score calculated for "hits" are significantly higher than that of the GOLD Score calculated for the existing inhibitor (Gold Score of natural ligand – 78). From all these findings, these molecules are suggested to be interesting candidates for further testing in the laboratory. Therefore, it is worth testing the actual binding affinities of these small molecules to the target protein to see whether the computer predictions reflect the biological characteristics. Thus, the study has its significance in identifying chemical compounds as safer drugs devoid of side effects. Hence, the virtual screening and associated analysis explained here can be used as a superior approach to identify potential lead molecules as well as to study the pharmacological model of the compound. Finally, this *in silico* study strongly underscores the importance of computational approaches in drug discovery, supplementing classical methods, thus saving enormous amount of time and money[27].

CONCLUSION

In the present study, the target RmlC protein was docked with 1,200 molecules which were virtually screened against its active site by employing molecular docking. The computational experiment with the help of different softwares has resulted in the identification of few small molecules, which docked well into the active site of the target. The overall drug-likeness was tested with DruLito and Lipinski's rule of five. Hydrogen bonding investigation also showed well affinity pattern of target protein active site with lead molecules. Apart from these, RmlC active sites with its inhibitor molecules and with its natural ligand are also compared. The study identified five putative small molecules as inhibitors that bind well to the active site of the target molecule chosen for the study. All these five hits predicted to "dock" well as inferred from the GOLD score into the active site of RmlC should be considered as interesting molecules that are needed to be further tested in the laboratory.

ACKNOWLEDGEMENT

I am deeply indebted to Dr. Achuthsankar S Nair, Director of Department of Computational Biology, and Bioinformatics for giving permission to do the project in DCB. I wish to express my deep sense of gratitude to Amjesh. R for providing the invaluable guidance. I am happy to acknowledge my deep sense of gratitude to Dr. Suseela Gomathi K, Head of the Department of Bioinformatics, National College for providing suggestions and help throughout this work.

REFERENCES

1. Ekins S., Williams AJ., Krasowski MD., Freundlich JS. In silico repositioning of approved drugs for rare and neglected diseases. Drug Discovery Today, 16(7-8): 298-310, (2011)
2. WHO. World Health Organisations Global Tuberculosis Report Geneva

(2013).http://www.who.int/tb/publications/global_report/en.
3. Bates JH. Transmission and pathogenesis of tuberculosis.Clinics in chest medicine, 1(2);167-174,(1980)
4. Blumberg HMea et al. American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America: treatment of tuberculosis. American journal of respiratory and critical care medicine,167(4):603-662,(2003)
5. Institute of Medicine (US). Addressing the Threat of Drug-Resistant Tuberculosis: A Realistic Assessment of the Challenge: Workshop Summary, The National Academies Press, Washington DC(2009). http://www.ncbi.nlm.nih.gov/books/NBK45006/
6. Barry CE., Blanchard JS. The Chemical Biology of New Drugs in Development for Tuberculosis. Current opinion in chemical biology,14(4) : 456-466, (2010)
7. Ekins S., Freundlich JS., Choi I., Sarker M, Talcott C. Computational databases, pathway and cheminformatics tools for tuberculosis drug discovery. Trends in microbiology,19(2): 65-74(2011)
8. Valasani KR., Vangavaragu JR., Day VW., Yan SS. Structure Based Design, Synthesis, Pharmacophore Modeling, Virtual Screening, and Molecular Docking Studies for Identification of Novel Cyclophilin D Inhibitors. Journal of Chemical Information and Modeling, 54(3) :902-912, (2014)
9. Kapetanovic IM. Computer Aided Drug Discovery and Development (CADDD): In silico-chemico-biological approach.Chemico-biological interactions. 171(2):165-176, (2008).
10. Hornak V., Dvorsky R., Sturdik E. Receptor-ligand interaction and molecular modelling. General physiology and biophysics, 18(3):231-248 (1999).
11. Vijayakrishnan R. Structure-based drug design and modern medicine. Journal of postgraduate medicine, 55(4):301-304, (2009)
12. Rose PW, Bi C, Bluhm WF, et al.The RCSB Protein Data Bank: new resources for research and education. Nucleic acids research, 41(Database issue):D475-482, (2013)
13. Kaplan W, Littlejohn TG. Swiss-PDB Viewer (Deep View).Briefings in bioinformatics, 2(2):195-197, (2001)
14. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis, 18(15): 2714-2723, (1997)
15. Von Grotthuss M, Koczyk G, Pas J, Wyrwicz LS, Rychlewski L. Ligand.Info small-molecule Meta-Database. Combinatorial chemistry & high throughput screening, 7(8):757-761, (2004)
16. Morris GM, Green LG, Radić Z, et al. Automated Docking with Protein Flexibility in the Design of Femtomolar “Click Chemistry” Inhibitors of Acetylcholinesterase. Journal of Chemical Information and Modeling, 53(4):898-906, (2013)
17. Wallace AC, Laskowski RA, Thornton JM.LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein engineering, 8(2):127-134, (1995)
18. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. Drug discovery today, 1(4):337-341, (2004)
19. http://www.niper.ac.in/pi_dev_tools/DruLiToWeb/DruLiTo_index.html (Accessed on 5th April 2015).
20. World Health Organization (2002) Deaths by cause, sex and mortality stratum in WHO Regions, estimates for 2001.The World Health Report.
21. Reichman LB., Felton CP., Edsall JR. Drug Dependence, A Possible New Risk Factor for Tuberculosis Disease. Arch Intern Med,139(3):337-339, (1979).
22. Kapetanovic.Computer-aided drug discovery and development (caddd): In silicochemico-biological approach. Chemico-Biological Interactions, 171:165-176, (2008)
23. Allsop AE., Brooks G., Bruton G., Coulton S., Edwards PD., HattonIK., Kaura AC., McLean SD., Pearson ND., Smale TC.,

Southgate R. Penem inhibitors of bacterial signal peptidase. Bioorganic and Medicinal Chemistry Letters, 5:443-448, (1995)
24. Laskowski RA., Macarthur MW., Moss DS., Thornton JM. {PROCHECK}: a program to check the stereochemical quality of protein structures. J. Appl. Cryst, 26:283-291, (1993)
25. RamachandranGN., Ramakrishnan CT., Sasisekharan V. Stereochemistry of polypeptide chain configurations. Journal of Molecular Biology, 7:95-9,(1963)
26. J.E. Korf., G. Pynaert., K. Tournoy., T. Boonefaes., A. Van Oosterhout., D. Ginneberge., A. Haegeman., J.A. Verschoor., P. De Baetselier., J. Grooten. Macrophage reprogramming by mycolic acid promotes a tolerogenic response in experimental asthma. American Journal of Respiratory and Critical Care, 174:152-160, (2006)
27. Ghose A.K., Wendoloski J.J. In Perspective in Drug Discovery and Design. Kluwer/Escom,9/10/11, 253, (1998)