BioGecko

Vol 12 Issue 02 2023
ISSN NO: 2230-5807

Identification of inhibitors for *Mycobacterium tuberculosis* PE16 Serine Hydrolase Domain using virtual Screening and MD Simulations studies

E. Hariprasad[1], Rafiya Sultana[2*]
1 Assistant Professor, Department of Chemistry, Vasavi College of engineering, Ibrahimbagh, Hyderabad
2 Assistant Professor, Department of Chemistry, Government Degree College for Women, Begumpet

*Corresponding author
Corresponding author Email ID: Rafiya.hani@gmail.com

Abstract
The importance of PE and PPE proteins in *Mycobacterium tuberculosis (Mtb)* along with the significant role played by cutinases/esterases/lipases in the lipid metabolism of pathogenic *mycobacteria* strongly implicates that the serine hydrolase, PE-PPE domain is a biological receptor to design new anti-tuberculosis potential drug candidates. With use of *in silico* screening of large databases of molecules, molecular docking, molecular dynamics simulations and SIE free energy calculations to select potential Rv1430 PE-PPE domain inhibitors. The best binding among the identified list of molecules, two molecules are ZINC13681668 and ZINC16052749 showing promising binding free energies. The results are also providing more insights for the design and identification of inhibitors to this drug target with high binding affinities thereby increasing the inhibitory activity.

Keywords: *Mycobacterium tuberculosis*; tuberculosis; PE proteins; PPE proteins; PE-PPE domain; serine hydrolase; esterase; Rv1430; molecular docking; molecular dynamics simulations; free energy calculations

Introduction

Mycobacterium tuberculosis (Mtb), a member of the closely related group of slow growing pathogenic *mycobacteria* called *M. tuberculosis* complex is responsible for causing deadliest infectious disease tuberculosis (TB) (Mcevoy *et al.*, 2012). TB remains a major health issue all over the world infecting over one third of the world's population. Although the BCG vaccination was introduced 90 years ago, the heterogeneity in the efficacy of protective effect of BCG vaccine varies due to several factors such as the differences in exposure to a typical *mycobacteria* in the environment, the genetic susceptibility of the population, differences in the virulence effect of *Mtb*, high risk of reinfection or reactivation, the availability of different BCG strains, nutritional differences and etc (Pereira *et al.*, 2007; Sterne *et al.*, 1998). The current first line drugs for anti-TB treatment are a multidrug course of therapy which consists of rifampin, isoniazid, pyrazinamide, and ethambutol. The second line of drugs consists of aminoglycosides, fluoroquinolones and *p*-aminosalicylic acid. The third line drugs consist of clofazimine, linezolid, amoxicillin, clavulanate, imipenem, cilastatin and clarithromycin.

The basic treatment for TB involves the first-line anti-TB drugs. If these drugs are misused or poorly managed, multi-drug-resistant tuberculosis (MDR-TB) can develop. MDR-TB refers to the strains of the TB bacillus that are resistant to at least the two most powerful drugs of first line anti-TB drugs (rifampin and isoniazid). Further, extensively drug-resistant TB (XDR-TB) can develop with the poor management of these second line drugs or a subset of MDR-TB that are resistant to nearly all current anti-TB drugs (Keshavjee *et al.*, 2008). *Mtb* can also become drug resistant due to mutations in genes encoding drug targets (Madania *et al.*, 2012). There is an increasing clinical occurrence of *Mtb* strains with XDR-TB that results in higher rate of mortality (Berry and Kon, 2009). The problem with the present anti-TB drugs is their side-effects that mainly cause liver damage and other adverse reactions like nausea, vomiting and anorexia leading to hospitalization (Ormerod and Horsfield, 1996; Shang *et al.*, 2011; Yee *et al.*, 2003). Therefore the hunt for novel TB drug targets is an ongoing pursuit to inhibit drug resistant bacteria that are unique to the pathogen and absent in the human host.

Two large gene families, PE and PPE were first reported during the complete genome sequencing of *Mtb* H37Rv (Cole *et al.*, 1998). Further investigations revealed that these multi-gene family members are highly expanded throughout the pathogenic *mycobacteria*, whereas nonpathogenic *mycobacteria* comprises relatively few PE and PPE genes (Ekiert and Cox, 2014). These gene families are specific to only *mycobacteria* and absent in the human host. These proteins therefore become good candidates for drug design studies. Cutinases, esterases and lipases are hydrolases required for lipid metabolism in both prokaryotes and eukaryotes. In *Mtb*, it was reported that there are at least 250 enzymes related to lipid metabolism which includes extracellular secreted enzymes, integrated cell wall enzymes and intracellular esterases/lipases (Camus *et al.*, 2002; Cole *et al.*, 1998). Further, it has been reported that most of the mycobacterial genes involved in lipid metabolism, cell division, chromosomal partitioning and secretion are required during infection in mouse model (Lamichhane *et al.*, 2005; Sassetti *et al.*, 2003). Several studies reported esterases/lipases as drug targets. Based on the analysis using sequence and structure prediction, and molecular docking studies, few potential inhibitors were proposed for Rv3802c, a secreted protein known for its virulence role in pathogenesis of *Mtb* (Saravanan *et al.*, 2012). The Rv3203, another lipase from *Mtb* H37Rv is known to up regulate during acidic stress (Singh *et al.*, 2014). The Rv2224c, belonging to the microbial esterase/lipase family was identified as a virulence gene, required for bacterial survival in mouse model and full virulence of *Mtb* (Lun and Bishai, 2007). Two novel proteins Rv1399c and Rv0045c which have putative hydrolase function are probably involved in the ester/lipid metabolism of *Mtb* (Canaan *et al*., 2004; Guo *et al*., 2010). In our earlier work, it was observed PE-PPE domain commonly present in some PE and PPE families of proteins (Adindla and Guruprasad, 2003). The PE or PPE domain is present in the N-terminus, while the PE-PPE domain is located towards the C-terminus of these proteins. Sequence analyses and structure modelling studies identified the presence of characteristic pentapeptide sequence motif and catalytic triad, and predicted that this 225 amino acid PE-PPE domain folds into a serine hydrolase domain (Sultana *et al.*, 2011). Subsequently, it has been revealed that the function of the PE-PPE domain is indeed an esterase that has specificity for the hydrolysis of short chain length esters (C4 and C6) (Sultana et al., 2013).

The importance of PE and PPE proteins in *Mtb* compounded with significant role played by cutinases/esterases/lipases in the cell wall and lipid metabolism of pathogenic *mycobacteria* strongly implicates that the PE-PPE domain is a biological receptor to design new anti-TB potential drug candidates with high affinity and selectivity using various computer aided-drug design tools to combat TB. Therefore, *in silico* screening of large databases of molecules and molecular docking techniques were used to select new and potent PE-PPE domain inhibitors as better drug candidates. To confirm the stability of their binding and energies, molecular dynamics (MD) simulations of the protein-inhibitor complexes were carried out. These results explain the basis for inhibitor binding to Rv1430 PE-PPE domain and provide more precise directions in the design of new inhibitors.

Methods and Materials

Protein Modeling and Virtual Screening

The PE-PPE domain (amino acid residues from 108 to 337) region of Rv1430 corresponds to a serine hydrolase. The three-dimensional homology model of this domain was built using MODELLER software (Sali and Blundell, 1993) and was discussed in our earlier work (Sultana et al., 2011). For virtual screening Drug Bank (Wishart *et al.*, 2006) approved molecules from ZINC database were used, which is freely available at http://zinc.docking.org and contains commercially available molecules in 3D format for structure/target based virtual screening (Irwin and Shoichet, 2005; Irwin *et al.*, 2012). Initially all the 6447 drug molecules were converted into pdbqt format using MGLTools-1.5.6 for virtual screening into the Rv1430 serine hydrolase domain using AutoDock Vina (Trott and Olson, 2010). Hydrogen atoms were added to the structure model followed by the addition of Gasteiger–Marsili charges. The non-polar hydrogens were merged onto their respective heavy atoms and atom types were fixed using AutoDock Tools (Morris *et al.*, 2009). The active site of the enzyme was covered by a grid box with 20 x 22 x 22 points and 1 Å was given for grid spacing with 47.0 x 1.0 x 30.8 centre of the grid box. AutoDock Vina software uses gradient optimization method and multithreading with local optimization (Trott and Olson, 2010). The number of

orientations for each molecule was set to 10. The best conformers with highest binding energies were chosen on the basis of affinity score. These protein ligand docking conformations were analyzed using Discovery Studio 2.5 visualizer to understand the inhibitor binding interactions. The top hit molecules with highest affinity were subjected to MD simulations and binding energy calculations were carried out subsequently. Using similar molecular docking protocol, these selected inhibitors were also docked into the active site of esterase D (PDB_ID: 3FCX_A), to estimate their binding to the human serine hydrolases.

Molecular Dynamics Simulations

The Rv1430 serine hydrolase domain - high affinity inhibitor complexes were subjected to MD simulations studies. These complexes were placed in a 10 Å cubic water box with TIP3P water molecules. MD simulations studies were carried out using GROMACS 4.5.5 simulation package (Hess *et al.*, 2008; Van Der Spoel *et al.*, 2005) with the AMBER99SB as force field for protein. The inhibitor parameter files were generated with GAFF force fields in antechamber (Wang *et al.*, 2006; Wang *et al.*, 2004) module using ACPYPE script (Sousa Da Silva and Vranken, 2012). Then the protein-inhibitor complexes were allowed for energy minimization for 2000 steps of steepest descent, 2000 steps of conjugate gradient and 1 ns position-restrained dynamics for the distribution of water molecules throughout the system. MD simulations were performed on the whole system for 25 ns, using 0.002 ps time step. The particle-mesh Ewald (PME) summation method (Darden *et al*., 1993; Essmann *et al*., 1995) was employed for the electrostatic calculation with a real space cut-off of 10 Å, PME order of 6, and relative tolerance between long- and short-range energies were set to 10^{-6}. To evaluate the short-range interactions, a neighbour list of 10 Å was updated for every 10 steps, Lennard–Jones (LJ) interactions and the real space electrostatic interactions were truncated at 9 Å. Isothermal-isobaric ensemble (NPT) with 298 K temperature and 1 atmosphere pressure were used along with the periodic boundary conditions. The V-rescale thermostat method for temperature (Bussi *et al.*, 2007) and the Parrinello-Rahman algorithm (Parrinello M, 1981) was used for pressure maintenance and the hydrogen bonds were constrained using LINCS algorithm (Hess *et al.*, 1997). Binding free energy calculations were performed by using the trajectory file obtained from MD simulations. The RMSD of protein C-α atoms was calculated from *g_rms* program of GROMACS by least-square fitting the structure to the reference structure.

Binding Free Energy Calculations

The binding free energies of Rv1430 serine hydrolase domain with potential inhibitors were calculated using Solvated Interaction Energies (SIE) method (Naim *et al.*, 2007; Sulea *et al.*, 2011) implemented in Sietraj software. Sietraj (http://www2.bri.nrc.ca/ccb/pub/sietraj_main.php) is an alternative method to the MM-PBSA software, and is compatible also with AMBER distribution (Cui *et al.*, 2008; Naim *et al.*, 2007). This method was successfully utilized in our earlier work and the SIE equation, coeffiecients and constant values were used as in our previous work (Tanneeru *et al.*, 2015; Tanneeru and Guruprasad, 2013). The binding free energies (ΔG) of the Rv1430 serine hydrolase domain-inhibitor complexes were calculated for the snapshot structures obtained from the MD trajectory of the systems. The ΔG is the sum of the change in Coulomb interactions, intermolecular van der Waals (vdW), reaction-field energy (obtained from solving the Poisson-Boltzmann equation) and non-polar solvation energy (that is proportional to the solvent accessible surface area) (Naim *et al.*, 2007). Similar to MM-PBSA/GBSA, SIE method treats the protein-ligand system in atomistic detail and implicit solvent model. The free energy of binding of the protein-inhibitor complex was calculated as components including van der Waals, Coulombic, reaction field and molecular surface area with BRIBEM software (Purisima, 1998; Purisima and Nilar, 1995, Chen *et al.*, 2004; Naim *et al.*, 2007). Here, the free energy of binding ΔG of structures estimated from each of the selected 100 snapshots of the MD trajectory and averaging over the resulting free energies obtained from each snapshot. Then the contribution of each active site residue is calculated to the binding free energy of the inhibitors.

Results and Discussion:

Rv1430 Protein Sequence Analyses and Three-Dimensional Structure Modeling

Protein sequence analyses, fold prediction and homology modelling identified the GxSxG sequence motif, the catalytic triad (Ser199, Asp276 and His302) and an overall α/β hydrolase fold with central β-sheet, flanked by α-helices on either side for Rv1430 serine hydrolase domain (Fig. 1) and therefore protein was

predicted to function as lipase, esterase or cutinase. Biochemical and mutational studies of this protein confirmed the function of PE-PPE domain as esterase (Sultana *et al.*, 2013). The enzyme activities of the PE-PPE domain is comparable with the activities of the full-length protein indicating that the N-terminal PE domain in Rv1430 does not have role in the enzymatic activity.

In the serine α/β hydrolase proteins, a lid region covers the active site of the enzyme from the top of the catalytic site to render the catalytic Ser inaccessible to the solvent. The closed conformation of the lid is displaced to open and expose the catalytic triad during interfacial activation that allows the substrate to be activated. Therefore, the lid region will alter the conformation of the protein with opening and closing of the active site and is therefore responsible to regulate the activity of the protein (Schrag *et al.*, 1997; Miled *et al.*, 2003). In members of this structural fold that function as esterases and lipases, the lid region covers the active site and is shown to be highly flexible that allows the entry/exit of substrates and inhibitors (Anthonsen et al., 1995, Ramkrishnan et al., 2008).
The serine hydrolases have an oxyanion hole that constitutes a part of the active site and the main chain nitrogens of the amino acid residues of these oxyanion hole participate in the hydrogen bonding as donor atoms to the hydrolysed substrate, stabilizing the negative charge on the tetrahedral intermediate occurring from the nucleophilic attack of the catalytic Ser during activation (De Simone *et al.*, 2004). The location of the lid insertion region and the oxyanion hole are mapped in the PE-PPE domain model structure. In the close vicinity of the catalytic triad, amino acids, Ser120, Thr121, Tyr130, Met131, Phe157, Gln158, Pro159, Trp160, Thr161, Tyr168, Phe198, Gln200, Phe278, Ala294, Leu295, Leu296, Ile298, Tyr299 and Ser303 are located, thus identifying the larger substrate/inhibitor binding cleft. In the Rv1430 serine hydrolase domain, amino acid residues (285-315) are identified as the lid insertion region and the oxyanion hole is formed by the amino acid residues Gln200 and Thr121.

Virtual Screening
The Drug Bank approved database of inhibitors (6447 entries) was download from ZINC database and used for the virtual screening of the Rv1430 serine hydrolase domain. Virtual screening of the database was performed with an efficient docking program AutoDock Vina, which identified inhibitors that interacted well with the protein. The top 100 hits identified with binding affinity better than -9.0 kcal/mol and were further validated and confirmed by graphical visualization in Discovery Studio Visualizer 2.5. Observation from literature, hydrolases like lipase/esterase, the catalytic serine forms a covalent bond with many irreversible drug molecules, for example, JZL184 (Long *et al.*, 2009a; Long *et al.*, 2009b) and tetrahydrolipstatin (Saravanan *et al.*, 2012). By considering this concept of catalytic serine - inhibitor interaction, the orientation of the inhibitors in the active site that make hydrophilic interaction with the catalytic serine were observed. From this aspect, shortlisted 10 inhibitors that were again subjected to secondary screening by increasing the number of GA runs to 20 and finally four inhibitors (ZINC13681668, ZINC16052749, ZINC01547088, ZINC16052883) were identified those showing good interactions with the protein. The binding free energies and the conformation of binding are observed very similar to previously obtained conformations from virtual screening of the entire database. These selected 4 inhibitors were also docked in to the binding site of the human esterase D (PDB_ID: 3FCX_A) using AutoDock Vina with 20 GA runs. The binding energies of top 4 hits are shown in Table 1. The docking results assured us that the binding of these molecules is specific to Rv1430 and have only weak binding to the human esterase.

Molecular Docking
From the molecular docking results, four inhibitors that display good binding to the enzyme active site are i) ZINC13681668, the substituted xanthane is present as complexed with type II dehydroquinase from *H. pylori* (PDB_ID: 2C4W) (Kd = 20000 nM); ii) ZINC16052749 is present in complex with tyrosine kinase domain of the hepatocyte growth factor receptor c-met (PDB_ID: 3CTJ) (IC50= 100 nM); iii) ZINC01547088, known as glucose lowering drug is observed in the crystal structure complex of pyruvate dehydrogenase kinase (PDB_ID: 2Q8G) and iv) ZINC16052883, a benzimidazole derivative present in the crystal structure complexed with VEGFR2 kinase domain (PDB_ID: 2QU5). From the molecular docking results analyzed graphically in Discovery Studio 2.5 Visualizer, it was observed that the molecules were well resided in the active site of the protein and mediate several non-bonding intermolecular interactions.

The docking results of ZINC13681668 in the active site of Rv1430 serine hydrolase were analyzed. The molecule was located in the active site and displays close interaction with catalytic Ser199 of the Rv1430 as shown in Fig. 2A. The inhibitor SO_2 group oxygen forms hydrogen bond with hydroxy group of Ser199 (OH...O, 2.725 Å). The nitrogen adjacent to tetrazine of the inhibitor forms hydrogen bond with hydroxy oxygen of Thr121(HO....HN, 2.796 Å). The Phe157 side chain forms pi-pi stacking interaction with the inhibitor aromatic ring in the active site. The side chain of Tyr299 on the lid region shields the tetrazole ring from the solvent accessible area to hold the molecule in the active site.

From the docking of ZINC16052749 into the protein active site reveals the carbonyl oxygen of the inhibitor forms hydrogen bond with the side chain hydroxy group of Ser199 (OH...O, 2.96 Å). The main chain carbonyl oxygen of Thr121 forms hydrogen bond with NH of the molecule (CO...HN, 3.0 Å). The *o*-substituted flourine forms hydrogen bond with main chain NH of the Met131 (NH...F, 2.9 Å). The *p*-flouro substituted phenyl ring of the inhibitor forms pi-pi stacking interaction with the side chain of Phe157. The inhibitor is located close to the lid region residues such as Tyr299 as shown in Fig. 2B.

From the observation of ZINC16052883 molecule docking conformations, the NH of inhibitor forms hydrogen bond with the Ser199 (HO..HN, 2.43 Å). The NH of the benzimidazole forms hydrogen bond with side chain hydroxy group of Thr121 (HO..HN, 2.47 Å). The substituted phenyl ring of the inhibitor forms pi-pi stacking interactions with the side chain phenyl ring of Phe157 as shown in Fig. 2C.

From the docking conformation ZINC01547088 molecule in the protein binding site, it was observed that one of the oxygens on the sulphur group of the inhibitor forms trifurcated hydrogen bonds with side chain hydroxy group of the Ser199 (OH...O , 2.7 Å), side chain hydroxy group of Thr121 (OH...O, 2.75 Å) and main chain NH group of Thr121 (NH...O, 2.85 Å). The hydroxy group of the inhibitor forms hydrogen bond with the main chain carbonyl oxygen of Phe157 (CO...HO, 2.60 Å). One of the flourine atoms of CF_3 group forms hydrogen bonding with the main chain NH of Trp160 (NH... F, 3.2 Å) as shown in Fig. 2D.

Molecular Dynamics Simulations

From the molecular docking studies observasions, the four inhibitors are making good non-bonding interactions with Rv1430 serine hydrolase domain. These protein-inhibitor complexes were subjected up to 25 ns of MD simulations to check their structural stability and binding. Among the four inhibitors screened from molecular docking, two inhibitors comparably showed weak interactions with active site residues during the MD simulations. In all cases, the lid region of the protein is actively involved in the movement of the inhibitor. Two molecules ZINC13681668 and ZINC16052749 show better interactions with the active site of Rv1430 serine hydrolase domain and retain non-bonding interactions with the enzyme and display continuous hydrogen bond with Ser199. From the trajectory of the MD simulations, the stability of the protein-inhibitor complexes observed and also analyzed the movement of the inhibitor in the enzyme active site. Low RMSD values of these inhibitors and only minor fluctuations in the active site residues were observed. The results of the MD simulations and binding free energies of two protein–inhibitor complexes are described below.

From the protein with ZINC13681668 complex MD trajectory analyses, it was observed that the inhibitor is stabilized in the active site throughout the MD simulations. The RMSD of the protein converged to 3 Å from 16 ns till the end of simulations, and the inhibitor RMSD is converged at 1 Å throughout the simulations as shown in Fig. 3A. A small movement of the inhibitor was made to enhance the binding with the protein active site and more number of hydrogen bonds were observed as shown in Fig. 4A. The tetrazine nitrogen forms new bifurcated hydrogen bonds with side chain OH (OH...N, 3.1 Å) and main chain NH (NH...N, 3.2 Å) of the Thr121, another tetrazine nitrogen forms hydrogen bond with side chain OH of Ser199 (OH...N, 2.162 Å). The SIE free energy calculations as shown in Table 3 also indicate that Thr121 has high coulombic energy value (-4.05 ± 0.86 kcal/mol) whereas Ser199 shows less coulombic interactions (0.68 ± 0.25 kcal/mol). One of the oxygens of inhibitor forms hydrogen bond with the side chain OH of Tyr168 (OH...O, 2.9 Å). One of the tetrazine nitrogen forms hydrogen bond with the side chain nitrogen of His302 (N...HN ,

2.9 Å). The side chains of Leu295 and Leu296 also exhibit good hydrophobic interactions and vdW contribution (-3.01 ± 0.67 kcal/mol and -3.43 ± 0.88 kcal/mol respectively) to the free energy of binding. The Tyr299 and Phe157 form pi-pi stacking interactions with the aromatic rings of the inhibitor to hold the molecule in the active site. The high vdW values of Phe157 (-3.43 ± 1.77 kcal/mol) and Tyr299 (-8.51 ± 1.28 kcal/mol) also indicate the hydrophobic nature of interactions in stabilizing the protein-inhibitor complex as shown in Table 2.

The ZINC16052749 complexed with protein MD simulations analysis reveals the protein RMSD fluctuated around 4 Å during the initial MD simulations but during the last 5 ns it converged near 3 Å. The inhibitor is stabilised in the enzyme active site and has low RMSD and makes continuous hydrogen bond with Ser199. The RMSD of the ligand initially converged and continued throughout the MD simulations at 1 Å of RMSD as shown in Fig. 3B. The side chain hydroxy group of Ser199 forms a bifurcated hydrogen bond with the two carbonyl oxygens of the inhibitor (OH...O, 3.3 Å ; OH...O, 2.7 Å) as shown in Fig. 4B.

SIE free energy calculations:

From the SIE free energy calculations as indicated in Table 3, the catalytic Ser199 has good coulombic interaction (-3.96 ± 0.92 kcal/mol) and less vdW interaction energy (-0.46 ± 1.04 kcal/mol) indicating the existence of continuous hydrogen bond. The NH of His302 forms a hydrogen bond with the carbonyl oxygen of the inhibitor (NH...O, 2.7 Å). The coulombic interactions (-2.47 ± 0.99 kcal/mol) and high vdW interaction (-4.19 ± 0.92 kcal/mol) energies indicate that the aromatic ring is involved in the hydrophobic interactions. The 1H-Pyrrolo[2,3-b]pyridine NH of inhibitor forms hydrogen bonds with the main chain carbonyl oxygen (CO...N_4, 1.8 Å) of Asp129. The side chain hydroxy group of Tyr299 forms hydrogen bond with the NH of the inhibitor (HO...HN_1, 2.4 Å). The Tyr299 shows high vdW (-4.75 ± 1.24 kcal/mol) and columbic (-1.24 ± 0.93 kcal/mol) interaction energies, and these values indicate that the phenyl ring of Tyr299 stabilizes the molecule in the active site. Based on the SIE calculations, the contribution from Phe157 to Coulombic and vdW energies are low, indicating its lesser participation in binding the inhibitor, and this is supported by the structural trajectories that indicate the loss of pi-pi stacking interactions from Phe157. The high vdW contribution from the Thr121 (-3.79 ± 0.79 kcal/mol), Tyr130 (-3.84 ± 0.68 kcal/mol) and Phe278 (-3.05 ± 0.66 kcal/mol) is an indicative that good hydrophobic interactions stabilize the inhibitor binding in this location.

Conclusions

The inhibitors from ZINC database were docked into the serine hydrolase domain of Rv1430, which identified four good inhibitors that have specific interactions and binding into the active site of the protein. From the MD simulations of the above four protein-inhibitor complexes, two molecules ZINC13681668 and ZINC16052749 complexes were identified with showing continuous hydrogen bonding with catalytic Ser199 and good interactions with amino acid residues in the active site of Rv1430 serine hydrolase domain. The lid region plays major role in holding the inhibitor in the active site of enzyme. These results identified two inhibitors to the serine hydrolase domain of Rv1430, a novel drug target to TB. The basis of this work provides more insights into the design of better inhibitors to Rv1430 PE-PPE domain that has increased the binding affinities there by increasing the inhibitory activity.

Acknowledgements:

Acknowledgements: Rafiya Sultana and E. Hariprasad is greatly acknowledged to Dr. Karunakar Tanneeru, for his constant support in MD data analysis.

References

Adindla, S., Guruprasad, L., 2003. Sequence analysis corresponding to the PPE and PE proteins in Mycobacterium tuberculosis and other genomes. J. Biosci. 28, 169-179.

Alonso, H., Bliznyuk, A.A., Gready, J.E., 2006. Combining docking and molecular dynamic simulations in drug design. Med. Res. Rev. 26, 531-568.

Anthonsen, H.W., Baptista, A., Drablø, F., et. al., 1995. Lipases and esterases: a review of their sequences, structure and evolution. Biotechnol. Annu. Rev. 1, 315-371.

Berry, M., Kon, O.M., 2009. Multidrug- and extensively drug-resistant tuberculosis: an emerging threat. Eur. Respir. Rev. 18, 195-197.
Bussi, G., Donadio, D., Parrinello, M., 2007. Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 014101.
Camus, J.C., Pryor, M.J., Medigue, C., Cole, S.T., 2002. Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv. Microbiology, 148, 2967-2973.
Canaan, S., Maurin, D., Chahinian, H., et al., 2004. Expression and characterization of the protein Rv1399c from Mycobacterium tuberculosis. A novel carboxyl esterase structurally related to the HSL family. Eur. J. Biochem. 271, 3953-3961.
Chen, W., Chang, C.E., Gilson, M.K., 2004. Calculation of cyclodextrin binding affinities: energy, entropy, and implications for drug design. Biophys. J. 87, 3035-3049.
Cole, S.T., Brosch, R., Parkhill, J., et al., 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature, 393, 537-544.
Cui, Q., Sulea, T., Schrag, J.D., et al., 2008. Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the MP1-p14 scaffolding complex. J. Mol. Biol. 379, 787-802.
Darden, T., York, D., Pedersen, L., 1993. Particle mesh Ewald: an Nlog (N) method for Ewald sums in large systems J. Chem. Phys. 98, 10089-10092.
De Simone, G., Mandrich, L., Menchise, V., et al., 2004. A substrate-induced switch in the reaction mechanism of a thermophilic esterase: kinetic evidences and structural basis. J. Biol. Chem. 279, 6815-6823.
Ekiert, D.C., Cox, J.S., 2014. Structure of a PE-PPE-EspG complex from Mycobacterium tuberculosis reveals molecular specificity of ESX protein secretion. Proceedings of the National Academy of Sciences of the United States of America 111, 14758-14763.
Essmann, U., Perera, L., Berkowitz, M.L., et al., 1995. A smooth particle meshes Ewald method J. Chem. Phys. 103, 8577-8593.

Guo. J., Zheng. X., Xu, L., et al., 2010. Characterization of a novel esterase Rv0045c from Mycobacterium tuberculosis. PLoS One, 5, e13143
Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., 1997. LINCS: A linear constraint solver for molecular simulations. J. Comput. Chem. 18, 1463-1472.
Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E., 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput. 4:435-447.
Irwin, J.J., Shoichet, B.K., 2005. ZINC- a free database of commercially available compounds for virtual screening. J. Chem. Inf. Model. 45, 177-182.
Irwin, J.J., Sterling, T., Mysinger, M.M., et al., 2012. ZINC: a free tool to discover chemistry for biology. J. Chem. Inf. Model. 52, 1757-1768.
Keshavjee, S., Gelmanova, I.Y., Farmer, P.E., et al., 2008. Treatment of extensively drug-resistant tuberculosis in Tomsk, Russia: a retrospective cohort study. Lancet, 372, 1403-1409.
Lamichhane, G., Tyagi, S., Bishai, W.R., 2005. Designer arrays for defined mutant analysis to detect genes essential for survival of Mycobacterium tuberculosis in mouse lungs. Infect. Immun. 73, 2533-2540.
Long, J.Z., Li, W., Booker, L., et al., 2009a. Selective blockade of 2-arachidonoylglycerol hydrolysis produces cannabinoid behavioral effects. Nat. Chem. Biol. 5, 37-44.
Long, J.Z., Nomura, D.K., Cravatt, B.F., 2009b. Characterization of monoacylglycerol lipase inhibition reveals differences in central and peripheral endocannabinoid metabolism. Chem. Biol. 16, 744-753.
Lun, S., Bishai, W.R., 2007. Characterization of a novel cell wall-anchored protein with carboxylesterase activity required for virulence in Mycobacterium tuberculosis. J. Biol. Chem. 282, 18348-18356.
Luthy, R., Bowie, J.U., Eisenberg, D., 1992. Assessment of protein models with three-dimensional profiles. Nature, 356:83–85.
Madania, A., Habous, M., Zarzour, H., et al., 2012. Characterization of mutations causing rifampicin and isoniazid resistance of Mycobacterium tuberculosis in Syria. Pol. J. Microbiol. 61, 23-32.
Mc Evoy, C.R., Cloete, R., Muller, B., et al., 2012. Comparative analysis of Mycobacterium tuberculosis pe and ppe genes reveals high sequence variation and an apparent absence of selective constraints. PLoS One, 7, e30593.
Miled, N., Bussetta, C., De caro, A., et al., 2003. Importance of the lid and cap domains for the catalytic

activity of gastric lipases. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 136, 131–138.
Morris, G.M., Huey, R., Lindstrom, W., et al., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785-2791.
Naim, M., Bhat, S., Rankin, K.N., et al., 2007. Solvated interaction energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space. J. Chem. Inf. Model. 47, 122-133.
Ormerod, L.P., Horsfield, N., 1996. Frequency and type of reactions to antituberculosis drugs: observations in routine treatment. Tuber. Lung Dis. 77, 37-42.
Parrinello, M., Rahman, A., 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys 52, 7182-7190.
Pereira, S.M., Dantas, O.M., Ximenes, R., Barreto, M.L., 2007. [BCG vaccine against tuberculosis: its protective effect and vaccination policies]. Rev. Saude Publica. 41, Suppl, 1, 59-66
Perryman, A.L., Santiago, D.N., Forli, S., et al., 2014. Virtual screening with AutoDock Vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of HIV integrase: participation in the SAMPL4 protein-ligand binding challenge. J. Comput. Aided Mol. Des. 28(4), 429-41.
Purisima, E.O., 1998. Fast summation boundary element method for calculating solvation free energies of macromolecules. J. Comput. Chem. 19, 1494-1504.
Purisima, E.O., Nilar, S.H., 1995. A Simple yet Accurate Boundary-Element Method for Continuum Dielectric Calculations. J. Comput. Chem. 16, 681-689.
Ramakrishnan, S.K., Krishna, V., Kumar, K.S.V., et. al., 2008. Molecular dynamics simulation of lipases. Int. J. Integr. Biol. 2, 204–213.
Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779-815.
Saravanan, P., Avinash, H., Dubey. V.K., Patra, S., 2012. Targeting essential cell wall lipase Rv3802c for potential therapeutics against tuberculosis. J. Mol. Graph. Model. 38, 235-242.
Sassetti, C.M., Boyd, D.H., Rubin, E.J., 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol. Microbiol. 48, 77-84.

Schrag, J.D., Li, Y., Cygler, M., Lang, D., et al., 1997. The open conformation of a Pseudomonas lipase. Structure, 5, 187-202.
Shang, P., Xia, Y., Liu, F., Wang, X., et al., 2011. Incidence, clinical features and impact on anti-tuberculosis treatment of anti-tuberculosis drug induced liver injury (ATLI) in China. PLoS One, 6, e21836.
Shi, J., Blundell, T.L., Mizuguchi, K., 2001. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. 310, 243-257.
Singh, G., Arya, S., Narang, D., et al., 2014. Characterization of an acid inducible lipase Rv3203 from Mycobacterium tuberculosis H37Rv. Mol. Biol. Rep. 41, 285-296.
Sousa da Silva, A,W,, Vranken, W.F., 2012. ACPYPE - AnteChamber PYthon Parser interfacE. BMC. Res. Notes, 5, 367.
Sterne, J.A., Rodrigues, L.C., Guedes, I.N., 1998. Does the efficacy of BCG decline with time since vaccination? Int. J. Tuberc. Lung Dis. 2, 200-207.
Sulea, T., Cui, Q.Z., Purisima, E.O., 2011. Solvated Interaction Energy (SIE) for Scoring Protein-Ligand Binding Affinities. 2. Benchmark in the CSAR-2010 Scoring Exercise. J. Chem. Inf. Model. 51, 2066-2081.
Sultana, R., Tanneeru, K., Guruprasad, L., 2011 The PE-PPE domain in mycobacterium reveals a serine alpha/beta hydrolase fold and function: an in-silico analysis. PLoS One, 6, e16745.
Sultana, R., Vemula, M.H., Banerjee, S., Guruprasad, L., 2013. The PE16 (Rv1430) of Mycobacterium tuberculosis is an esterase belonging to serine hydrolase superfamily of proteins. PLoS One, 8, e55320.
Tanneeru, K., Balla, A.R., Guruprasad, L., 2015. In silico 3D structure modeling and inhibitor binding studies of human male germ cell-associated kinase. J. Biomol. Struct. Dyn. 33, 1710-1719.
Tanneeru, K., Guruprasad, L., 2013. Ponatinib is a pan-BCR-ABL kinase inhibitor: MD simulations and SIE study. PLoS One, 8, e78556.
Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455-461.
Van Der Spoel, D., Lindah, E., Hess, B., et al., 2005. GROMACS: fast, flexible, and free. J. Comput. Chem.

26, 1701-1718.
Wang, J., Wang, W., Kollman, P.A., Case, D.A., 2006. Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model. 25, 247-260.
Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A., 2004. Development and testing of a general amber force field. J. Comput. Chem. 25, 1157-1174.
Wishart, D.S., Knox, C., Guo, A.C., Shrivastava, S., et al., 2006. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res, 34, D668-72.
Yee, D., Valiquette, C., Pelletier, M., Parisien, I., Rocher, I., Menzies, D., 2003. Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis. Am. J. Respir. Crit. Care Med. 167, 1472-1477.

Legends of Tables

Table 1: Free energy of binding of the top 5 inhibitors from ZINC database selected by virtual screening performed using Autodock Vina docking program into the active site of Rv1430 PE-PPE domain.

Table 2: The SIE free energy (kcal/mol) calculations of ZINC16052749 in the active site of the Rv1430 PE-PPE domain with contribution from Coulombic and vdW interaction energies.

Table 3: The SIE free energy (kcal/mol) calculations of ZINC13681668 in the active site of the Rv1430 PE-PPE domain with contribution from Coulombic and vdW interaction energies.

Legends of Figures

Fig. 1. Homology model of Rv1430 constructed using MODELLER. The catalytic triad is shown in ball and stick model (colour according to atom type). The lid insertion region is indicated in blue and the catalytic triad is indicated.

Fig. 2. A) The interaction of top hit ZINC13681668 with the active site of the Rv1430 PE-PPE domain. **B)** The interaction of top hit ZINC16052749 with the active site of the Rv1430 PE-PPE domain. **C)** The interaction of top hit ZINC16052883 with the active site of the Rv1430 PE-PPE domain. **D)** The interaction of top hit ZINC01547088 with the active site of the Rv1430 PE-PPE domain. Hydrogen bonding interactions are indicated in yellow dashed lines.

Fig. 3. A) The RMSD plot of Rv1430 PE-PPE domain (black) in complex with the inhibitor ZINC13681668 (red). **B)** The RMSD plot of Rv1430 PE-PPE domain (black) in complex with the inhibitor ZINC16052749 (red).

Fig. 4. A) The interaction of the best hit ZINC13681668 with the active site of the Rv1430 PE-PPE domain from MD simulations. **B)** The interaction of the best hit ZINC16052749 with the active site of the Rv1430 PE-PPE domain from MD simulations. Hydrogen bonding interactions are indicated in yellow dashed lines.

Table 1

S.No.	Molecule	Free energy of binding (kcal/mol)		ΔG difference	SIE of Rv1430
		Rv1430 ΔG	3FCX_A ΔG		
1	ZINC01547088	-9.7	-6.6	-3.1	-7.39 ± 0.65
2	ZINC13681668	-9.9	-7.2	-2.7	-7.51 ± 0.47
3	ZINC16052749	-10.2	-7.1	-3.1	-9.16 ± 0.44
4	ZINC16052883	-10.0	-7.4	-2.6	-7.93 ±0.49

Table 2

S.No	Inter vdW+ stdev.	Inter Coulomb + stdev.
Phe157	-3.43 ± 1.77	0.06 ± 0.13
Ser120	-1.75 ± 0.34	-1.02 ± 0.26
Ser199	-1.14 ± 0.21	0.68 ± 0.25
Gln200	-2.23 ± 0.68	-0.91 ± 0.71
Tyr299	-8.51 ± 1.28	-0.18 ± 0.32
Thr121	-1.90 ± 1.51	-4.05 ± 0.86
Tyr168	-1.59 ± 1.20	-2.74 ± 1.37
His302	-1.03 ± 1.00	-2.77 ± 1.67
Leu295	-3.01 ± 0.67	0.08 ± 0.31
Leu296	-3.43 ± 0.88	0.15 ± 0.11
protein with ZINC13681668	**-36.18 ± 3.87**	**-11.46 ± 2.43**

Table 3

S.No	Inter vdW+ stdev.	Inter Coulomb + stdev.
Phe157	-0.51 ± 0.31	0.31 ± 0.11
Phe198	-2.84 ± 0.62	-0.41 ± 0.27
Ser199	-0.46 ± 1.04	-3.96 ± 0.92
Gln200	-2.52 ± 0.72	-2.28 ± 0.52
Tyr299	-4.75 ± 1.24	-1.24 ± 0.93
Thr121	-3.79 ± 0.79	-0.04 ± 0.52
Tyr130	-3.84 ± 0.68	-1.25 ± 0.28
His302	-4.19 ± 0.92	-2.47 ± 0.99
Pro262	-1.38 ± 0.39	-0.18 ± 0.10
Phe278	-3.05 ± 0.66	-0.47 ± 0.29
protein with ZINC16052749	***-54.10 ± 3.11***	***-17.57 ± 2.19***

Fig. 1.

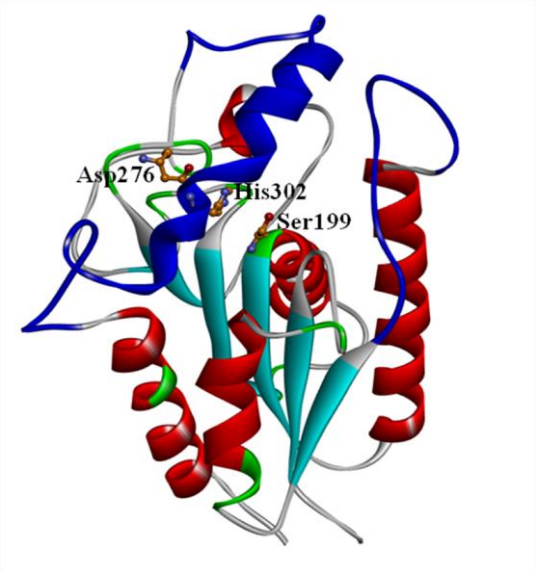
Asp276
His302
Ser199

Fig. 2 A-D.

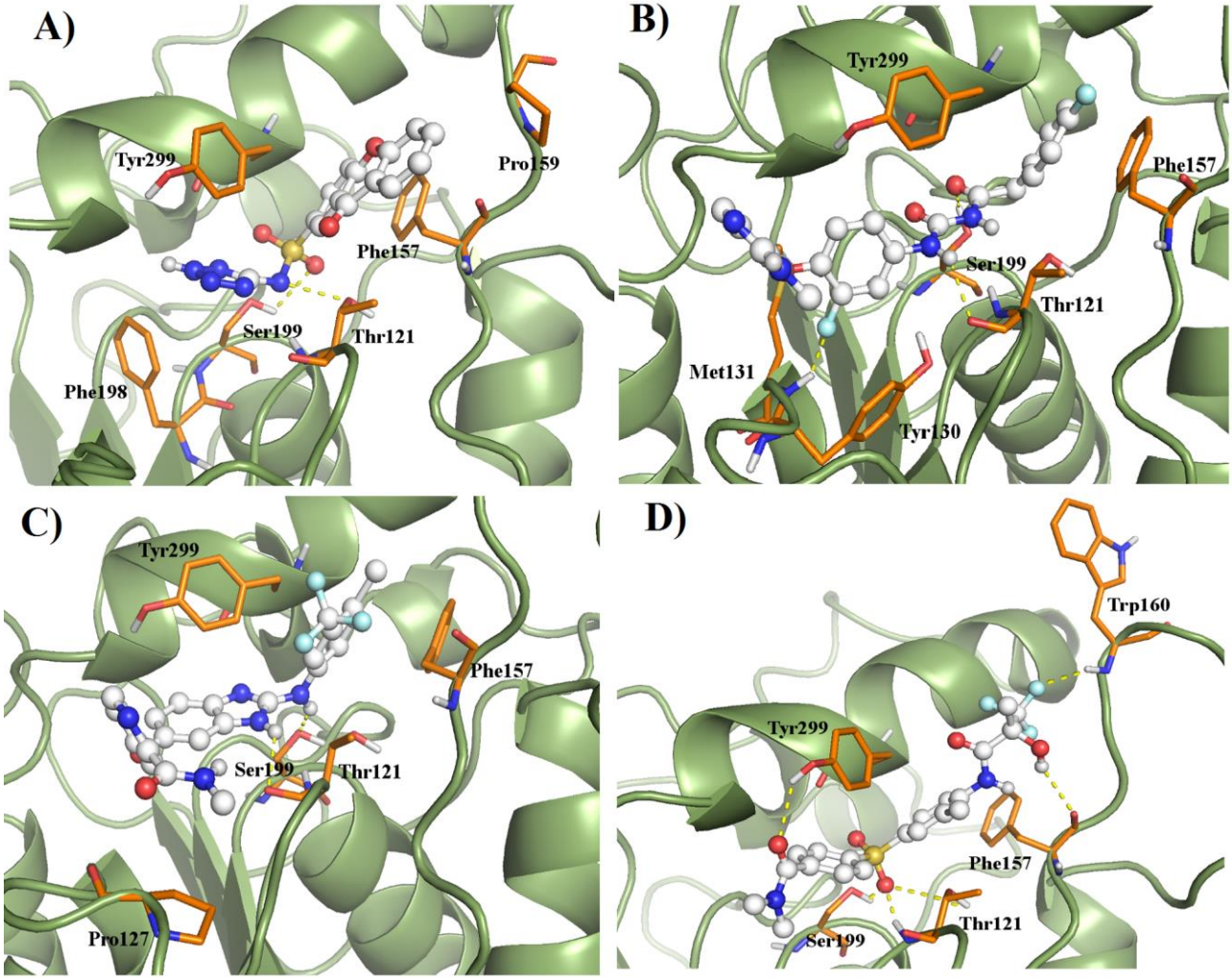
A)
Tyr299
Pro159
Phe157
Ser199
Thr121
Phe198
B)
Tyr299
Phe157
Ser199
Thr121
Met131
Tyr130
C)
Tyr299
Phe157
Ser199
Thr121
Pro127
D)
Trp160
Tyr299
Phe157
Thr121
Ser199

Fig. 3A-B.

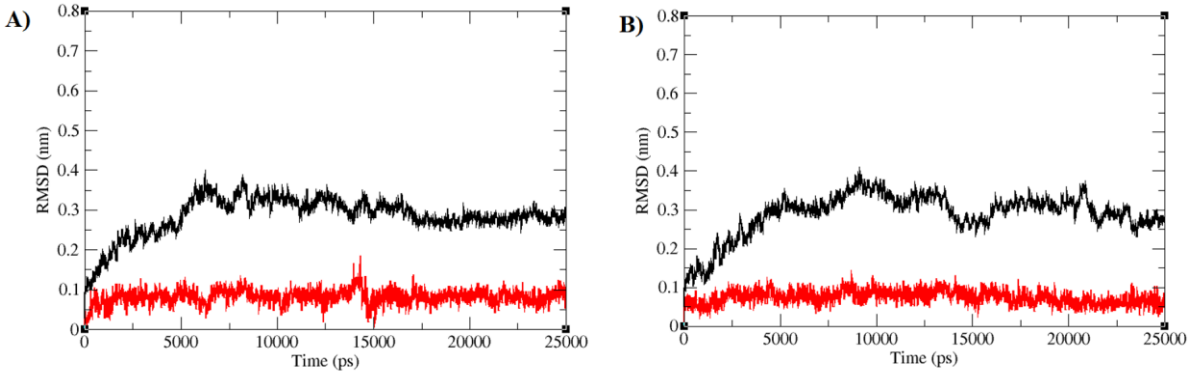

Fig. 4A-B.