

Identification of hotspot regions of MurB oxidoreductase enzyme using homology modeling, molecular dynamics and molecular docking techniques

Vivek Kumar · Parameswaran Saravanan ·
Akanksha Arvind · C. Gopi Mohan

Received: 13 April 2010 / Accepted: 18 June 2010 / Published online: 9 July 2010
© Springer-Verlag 2010

Abstract Despite the availability of effective chemotherapy and a moderately protective vaccine, new anti-tuberculosis agents are urgently needed to decrease the global incidence of tuberculosis (TB) disease. The *MurB* gene belongs to the bacterial cell wall biosynthesis pathway and is an essential drug target in *Mycobacterium tuberculosis* (Mtb) that has no mammalian counterparts. Here, we present an integrated approach involving homology modeling, molecular dynamics and molecular docking studies on Mtb-MurB oxidoreductase enzyme. A homology model of Mtb-MurB enzyme was built for the first time in order to carry out structure-based inhibitor design. The accuracy of the model was validated using different techniques. The molecular docking study on this enzyme was undertaken using different classes of well known MurB inhibitors. Estimation of binding free energy by docking analysis indicated the importance of Tyr155, Arg156, Ser237, Asn241 and His304 residues within the Mtb-MurB binding pocket. Our computational analysis is in good agreement with experimental results of site-directed mutagenesis. The present study should therefore play a guiding role in the experimental design of Mtb-MurB inhibitors for in vitro/in vivo analysis.

Keywords *Mycobacterium tuberculosis* · MurB · Homology modeling · Molecular dynamics · Molecular docking

Electronic supplementary material The online version of this article (doi:10.1007/s00894-010-0788-3) contains supplementary material, which is available to authorized users.

V. Kumar · P. Saravanan · A. Arvind · C. G. Mohan (✉)
Centre for Pharmacoinformatics,
National Institute of Pharmaceutical Education and Research,
Sector 67,
S.A.S. Nagar 160 062 Punjab, India
e-mail: cmohan@nipер.ac.in

Introduction

Tuberculosis (TB) is a contagious-infectious disease caused mainly by *Mycobacterium tuberculosis* (Mtb). The resumption of TB, which is due mainly to the emergence of multidrug-resistant (MDR) and extensively drug-resistant strains of Mtb as well as HIV epidemics, has led to an increased need to understand the molecular mechanisms of drug action and drug resistance. These complicated issues need to be tackled by discovering novel compounds for TB therapy.

The physiology of Mtb requires a high level of oxygen, and it usually establishes infection in the mammalian respiratory system [1]. In the lungs, Mtb are taken up by alveolar macrophages having high oxygen tension leading to accelerated growth of Mtb. It is quite difficult to stain such bacteria as they have an atypical waxy coating on the cell surface, comprised chiefly of mycolic acid, which makes the cells impervious to Gram staining.

In 1882, Robert Koch, an esteemed scientist of his time, isolated and cultured Mtb from crushed tubercles. His experimental work identified the bacterium as the etiological agent of TB. He also developed staining methods for identification of the bacillus, and these techniques were subsequently improved on by the German doctor and bacteriologist Paul Ehrlich, whose method for the detection of the bacillus provided the basis for the development of the Ziehl-Nielsen acid-fast staining, which still is an important tool in TB therapy [2].

Mtb is a tenacious and remarkably successful pathogen that has latently infected one-third of the world's population. Every year, 8 million new cases of TB appear and 2 million deaths occur [3]. The increasing emergence of drug-resistant TB and HIV infection, which compromises host defense and allows latent infection to reactivate or

render individuals more susceptible to TB, pose further challenges to the effective control of this disease [4].

Chemotherapy of TB has been divided into first-line and second-line drugs. First-line drugs include isoniazid, rifampin, pyrazinamide and ethambutol, while second-line drugs such as para-aminosalicylate, kanamycin, fluoroquinolones, capreomycin, ethionamide and cycloserine are used, but these are either less effective or more toxic with serious side effects [5]. Treatment is made quite difficult by the presence within host lesions of metabolically silent, persistent or dormant Mtb, which are not susceptible to the anti-TB drugs that usually kill growing bacteria but not the persistent bacteria. During the interaction between Mtb and host cells, a cyclic reinfection of host macrophages by tubercle bacillus can occur, allowing the prolonged survival and persistence of the bacilli. TB drugs currently in use were developed 40 years ago and showed some extent of resistance problem towards Mtb [6]. Thus, it is almost impossible to achieve complete sterilization of lesions. It is this unique characteristic of these bacilli to withstand chemotherapy and host immune attack, and its ability to survive for decades before reactivation that makes TB so difficult to treat and eradicate. Due to the heterogeneous bacterial populations in TB lesions and insufficient host immunity, treatment with a combination of drugs must be given for extended periods of time to prevent reactivation of disease by persisting bacilli. Much research effort focuses on understanding the biology of persistence and developing therapies that kill persistent bacilli more efficiently [7].

TB is curable but the therapy still takes at least 6–9 months. Treatment is laborious and lengthy, and often results in patient non-compliance, significant toxicity and multi-drug resistance tuberculosis (MDR-TB). The increasing problems caused by MDR-TB has focused attention on developing new drugs that are not only active against drug-resistant TB, but also shorten the lengthy therapy. This urgent need to develop novel and potent TB drugs to overcome this neglected disease is of national importance, as the discovery process in this area has been dormant for the past 40 years [8].

The cell walls of both Gram-positive and Gram-negative bacteria possess a unique biopolymer, peptidoglycan, critical for maintaining the osmotic integrity of the cell. Enzymes of the bacterial cell wall biosynthesis pathway are essential proteins without mammalian counterparts [9]. The synthesis of peptidoglycan begins in the cytoplasm, with the condensation of phosphoenolpyruvate and UDP-N-acetylglucosamine catalyzed by MurA. This is followed by a MurB-catalyzed reduction of the enolpyruvate moiety to D-lactate, yielding UDP-N-acetylmuramate. A series of ATP-dependent amino acid ligases (MurC, MurD, MurE and MurF) catalyze the

stepwise addition of the pentapeptide side-chain on the newly reduced D-lactyl group, resulting in the formation of UDP-N-acetylmuramyl pentapeptide [10]. These initial steps in the biosynthesis of the cytoplasmic peptidoglycan precursor are poorly exploited as anti-bacterial targets.

The genes encoding MurA to MurF are all essential in Mtb. In addition, Mur proteins are highly conserved among various bacterial species, and common structural motifs can be identified. For this reason, a potential Mur inhibitor would be expected to be bactericidal and to have a wide spectrum, which validates the choice of these important bacterial enzymes as targets for the development of new inhibitors [10–14].

Different computational studies have already been reported for several Mtb protein targets: the lysine/DAP pathway of Mtb, methionine S-adenosyltransferase enzyme, Asp kinase and DNA gyrase for the treatment of TB [15–18].

The main objective of the present work is to report the three-dimensional (3D) model of Mtb-MurB enzyme (ORF Rv0482) for the first time based on the crystal structure of the (*Escherichia coli*)-MurB. The model provides the geometry of hot spot regions, i.e., binding site residues, and therefore provides a clear insight into the importance of the active site residues in terms of their contribution to protein–ligand complexes. Moreover, the rational design of an inhibitor selective towards the Mtb-MurB enzyme could be more effective if the key residues and atomic level binding site interactions are known. A combined approach combining a molecular dynamics study and molecular docking analysis of 28 already reported inhibitors of (*E. coli*)-MurB protein was applied to the Mtb-MurB homology model to probe its active site residues, interactions, thereby revealing its mechanism of action.

Methods

Experimental section

All computations and molecular modeling of Mtb-MurB enzyme were carried out on a Silicon Graphics Fuel Workstation with IRIX 6.5 operating system and Linux workstation using MODELLER9v3, Autodock4.0 and SYBYL7.1 molecular modeling packages.

Homology modeling

The Mtb-MurB (ORF Rv0482) sequence was retrieved from UniProtKB/TrEMBL database (primary accession number P65460). In order to identify homologous sequences with known 3-D structure, a BLASTP search was carried out against the protein data bank (PDB) using Mtb-MurB as query sequence [19]. A number of sequences homologous to Mtb-

Molecular dynamics study

Molecular dynamics (MD) stimulations were carried out using 43A1 force field of Gromacs96 implemented in the Gromacs3.3.3 package [28]. A cubic box with the SPC water model was built and submitted to a maximum of 1,000 rounds of energy minimization using the steepest descent gradient algorithm [29]. The leap-frog algorithm was used for integrating Newton's equations in MD simulation.

Mtb-MurB oxidoreductase enzyme was subjected to a full MD simulation of 5,000 ps (or 5 ns) at 300 K, with no restrictions using 2 fs of integration time. Lincs constraints were used on all protein covalent bonds to maintain constant bond length [28]. Co-ordinates and energy terms (potential energy for the whole system) were saved each 20 ps. With the aim of evaluating the protein system stabilization throughout the MD period of 5,000 ps, its RMSD and potential energy were plotted against time.

Molecular docking

The ability of proteins to bind with different ligands in a highly specific manner is an important feature of many biological processes. The characterization of the structure and the energetics of molecular complexes is thus a key factor in understanding biological function.

Autodock4 program uses the Lamarckian genetic algorithm (GA) and is regarded as the best method in terms of its ability to find the lowest energy structure, and the accuracy of its structure predictions [30]. Polar hydrogens were added using the Autodock Tools interface. Grid maps were prepared using the AutoGrid utility with $58 \times 76 \times 48$ points and grid spacing set to 0.375 Å. Grid dimension was chosen to cover the 6.5 Å area around the ligand, so that all the active site residues can be accommodated. Docking parameters modified from the defaults were: number of individuals in the population (set to 150), maximum number of energy evaluations (set to 2,500,000), maximum number of generations (set to 2,700), and number of GA runs (set to 20).

To understand protein–ligand binding affinity from a thermodynamics viewpoint, this program provides binding energy (BE) and inhibition constants (K_i) for the docked ligands [30]. BE gives the extent of the binding affinity of the protein–ligand complexes, and is calculated as the sum of the intermolecular energy and torsional free-energy penalty. K_i is calculated by the equation given below.

$$K_i = \exp [(\Delta G \cdot 1000) / (R \cdot T)]$$

Where ΔG is docking energy, R (gas constant) is 1.98719 cal K⁻¹ mol⁻¹ and T (Temperature) is 298.15 K.

After setting the above mentioned docking protocol, we first extracted the ligand (naphthyl tetronic acid) from the

Mtb-MurB enzyme model and subjected it to energy minimization. the binding site for Mtb-MurB enzyme was defined as amino acids residues within a 6.5 Å region around the ligand. The geometry optimized extracted molecule was again docked to Mtb-MurB enzyme and showed an excellent RMSD of 0.60 Å, indicating that prediction ability of Autodock4 is acceptable and thus can be extended to docking of the other 28 Mtb-MurB inhibitors into the active site regions of the homology model.

A series of 28 inhibitors against Mtb-MurB enzyme belonging to 4-alkyl and 4,4'-dialkyl 1,2-bis(4-chlorophenyl) pyrazolidine-3,5-dione derivatives were reported by Kristinia et al. [31]. In the present study, we adopted the same numbering convention for these molecules. Docking analysis was carried out using the Autodock4 program with the aim of exposing the active site amino acid residues involving protein–ligand interactions in order to obtain information about the bioactive conformation of these Mtb-MurB inhibitors. All molecules were docked successfully in the active site of the homology model of the Mtb-MurB enzyme.

Results and discussion

Homology modeling of Mtb-MurB and its structure-function relationships

Mtb-MurB (EC 1.1.1.158) is an oxidoreductase enzyme involved in the cell wall biogenesis. Mtb-MurB has an amino acid sequence length of 369 and is located in the cytoplasm of the bacterial cell. To identify homologous sequences of Mtb-MurB that could be used as a template for model building, a NCBI BLASTP search against PDB using Mtb-MurB as query sequence was performed. The top hit from BLASTP, viz. (*E. coli*)-MurB (PDB code: 2Q85) has a crystal structure to 2.51 Å resolution, and exhibits 33% sequence identity and 50% sequence similarity with the Mtb-MurB enzyme (Fig. 1).

The crystal structure of (*E. coli*)-MurB enzyme complexed with the inhibitor naphthyl tetronic acid has structural similarity with known reported inhibitors of MurB enzyme. This analysis gave us more confidence in using this crystal structure (PDB code: 2Q85) for homology modeling of Mtb-MurB enzyme, with further molecular docking to identify the active binding site residues.

The final model was validated using seven different tools:

- (1) Ramachandran plot showed that the backbone dihedral angle distribution of all the amino acids residues represents: 86.2% in core region, 12.1% in allowed

region and 1.7% in generously allowed region, as presented in Table 1 and Fig. S1 respectively (see electronic supplementary material). This indicated that the backbone dihedral angles ψ and ϕ , in the Mtb-MurB model were reasonably accurate.

- (2) An ERRAT plot calculated non-bonded interactions between different atom types of amino acid. We can determine the ‘structure error’ at each residue in the 3-D structure of the protein using this plot. This provides an indication for loop refinement at those regions where the plot bar exceeds the 95% error value in ERRAT plot, shown in Fig. S2. Loop refinement was performed in several cycles, taking a few amino acids (4–12) in each step and validating through ERRAT plot. ERRAT plot of the loop refined model showed an increase in the overall score for structural quality to 93%.
- (3) To support the evaluation of protein fold quality, Mtb-MurB model was subjected to ProsaII folding energy analysis. This program calculates the knowledge-based mean fields to judge the quality of the protein folds and has been used widely to measure the stability of a protein conformation. Folding energy of protein generally has negative values, and these values correspond to the stability and nativity of the molecules. Thus, protein folds with high negative value are considered more stable. The folding energy of both the (*E. coli*)-MurB crystal structure and the Mtb-MurB model, and its comparative energy profile is shown in Fig. S3. We note that the ProsaII folding energy of Mtb-MurB remains negative for all the amino acids, indicating the acceptability of the Mtb-MurB model structure (Fig. S3).
- (4) G-factor provides a measure of how unusual or out of the ordinary a property is on the basis of phi, psi, chi1, chi2 and chi3 angles. G-factor for Mtb-MurB model was -0.20 , which indicates the usual distribution of properties (Table 1). It has been established that a G-factor value between -0.5 and -1.0 indicates that the property is unusual, while below than -1.0 is highly unusual.
- (5) Verify 3D was used to assess whether a primary sequence is compatible with the current 3D structural

model; 92.51% amino acids have average 3D-1D compatibility scores >2 for the Mtb-MurB model structure (Table 1). The compatibility scores for all the residues in the developed model are also shown to be above zero and are presented in Fig. S4.

- (6) Evaluation of Mtb-MurB homology with ProSA-web reveals that the ProSA Z-score value is -8.33 in the range of native conformations of the crystal structures, as shown in Table 1 and Fig. S5.
- (7) RMSD validation criteria include the superimposition of the residues within the active site regions surrounding the naphthyl tetronic acid inhibitor, in both Mtb-MurB and (*E. coli*)-MurB protein. We obtained RMSD values of 0.78 \AA and 1.52 \AA for 6.5 \AA and 8.0 \AA regions surrounding the naphthyl tetronic acid inhibitor. RMSD of the C_{α} -backbone for the whole protein was 1.79 \AA (Table 2, Fig. 2a). Superimposition of the active site residues within the 5.5 \AA region surrounding naphthyl tetronic acid inhibitor (ball and stick) in Mtb-MurB (green) and (*E. coli*)-MurB (cyan) showed that most of the active site residues were conserved and superimposed in both these proteins, except that Asp270 and Gln288 of Mtb-MurB were replaced with Glu282 and His304 in (*E. coli*)-MurB, respectively (Fig. 3).

MurB is an oxidoreductase enzyme that belongs to the flavin adenine dinucleotide (FAD) family—a cofactor binding protein super family—that share a characteristic FAD-binding fold. MurB is located in the aminosugar pathway, a sub-pathway that provides precursor for peptidoglycan biosynthesis [12, 32]. It catalyzes the reduction of UDP-N-acetylglucosamine enolpyruvate to UDP-N-acetylmuramic acid by involving a sequence of two half-reactions. First, bound FAD is reduced by a two electron transfer from NADPH, and then it transfers these two electrons to the C-3 of the enol pyruvyl group of glucosamine.

The homology model of Mtb-MurB oxidoreductase enzyme along with the ligand naphthyl tetronic acid and FAD cofactor is presented in Fig. 2b. Structurally Mtb-MurB is a mixed ($\alpha + \beta$) protein, composed of three

Table 1 Validation statistics comparison of predicted Mtb-MurB model with the template (*Escherichia coli*)-MurB crystal structure

Ramachandran plot statistics	(<i>E. coli</i>)-MurB	Mtb-MurB
% Amino acids in most favored regions	91.2	86.2
% Amino acids in additional allowed regions	8.5	12.1
% Amino acids in generously allowed regions	0.3	1.7
% Amino acids in disallowed regions	0.0	0.0
G-factor	-0.01	-0.20
Verify 3D: Average 3D-1D compatibility score for % amino acid >0.2	100	92.51
ProSA Z-score	-10.17	-8.33

Table 2 Three dimensional (3D) structure comparison root mean squares deviation (RMSD) of the Mtb-MurB homology model with the template (*E. coli*)-MurB enzyme

C α -backbone	6.5Å region around ligand	8Å region around ligand
1.79 Å	0.78 Å	1.52 Å

domains. Domain 1 contains five β -strands and one α -helix, which includes residues 3–67 and residues 327–346; Domain 2 contains seven β -strands and six α -helices, which includes residues 68–201; and Domain 3 contains five β -strands and five α -helices, which include residues 202–326. The topology of the Mtb-MurB model with the above mentioned three domains shown as rectangular boxes is presented in Fig. S6 (see electronic supplementary material). It can be seen that the ligand binding site is surrounded by domains 2 and 3, and is quite accessible from the protein surface, while cofactor FAD is surrounded by all three domains and is buried in the protein core.

The MurB protein family is divided into type I and type II, based on the presence or absence of various secondary structural elements, and on how these structural elements interact with its substrates. (*E. coli*)-MurB is classified as type I and contains a Tyr loop and split $\beta\alpha\beta\beta$ fold, whereas *Staphylococcus aureus* (*S. aureus*)-MurB is classified as type II and lacks these secondary elements.

X-ray model and homology model structural analyses have indicated that the *E. coli* and Mtb type I MurB proteins and *S. aureus* type II MurB protein have similar overall folds, although there are differences between the substrate binding regions due to amino acid deletions, as presented in Figs. 1 and 2, respectively [12, 32–34]. (*S. aureus*)-MurB lacks both a loop structure containing Tyr190 and the single split fold found in the substrate binding domain of type I (*E. coli*)-MurB [12], and the same

is found in the Mtb-MurB model. Similarities between type I and type II catalytic sites include: Ser237 of Mtb-MurB (Ser229 in (*E. coli*)-MurB, and Ser226 in (*S. aureus*)-MurB) is a catalytic residue and participates in the proton transfer to the enol intermediate formed during the second reduction step [35].

Experimental evidence for important amino acid residues and knowledge of their roles in enzyme structure–function relationships are crucial for developing enzyme-specific anti-bacterial drugs. Using the crystal structure of (*E. coli*)-MurB (PDB code: 2Q85) and (*S. aureus*)-MurB (PDB code: 1HSK) proteins, we identified key amino acid residues making hydrogen bonding interactions with FAD and the ligand in the Mtb-MurB model. A multiple sequence alignment of Mtb-MurB with (*E. coli*)-MurB and (*S. aureus*)-MurB protein is presented in Fig. 1. The observations generated from these sequence–structure relationships are described below.

Residues Gly47, Gly49, Ser50, Asn51, and Arg218 in Mtb-MurB; (*E. coli*)-MurB corresponding residues include: Gly47, Gly49, Ser50, Asn51 and Arg214; (*S. aureus*)-MurB corresponding residues include: Gly67, Gly69, Ser70, Asn71 and Arg213) are observed close to the FAD binding site. Also, residues Tyr175, Arg176, Arg213 Ser226, Asn230, His259, and Glu296 [corresponding residues in (*E. coli*)-MurB are Tyr158, Arg159, Arg214, Ser229, Asn233, Gln288 and Glu325] are located close to the UDP-N-acetylglucosamine enolpyruvate binding site in (*S. aureus*)-MurB protein [34]. Residue Ser226 of (*S. aureus*)-MurB [Ser229 of (*E. coli*)-MurB and Ser237 of Mtb-MurB] was reported as a catalytic residue whose hydroxyl group is thought to transfer a proton for the reduction of UDP-N-acetylglucosamine enolpyruvate [32, 35]. It is well established that the Gly-X-Gly motif, i.e., Gly67-Asn68-Gly69 in (*S. aureus*)-MurB and Gly47-Glu48-Gly49 in (*E. coli*)-MurB protein, corresponds to the FAD-binding domain of FAD-binding proteins (Fig. 1). The corresponding motif does not exist in Mtb-MurB due

Fig. 2 **a** Superimposition of C α -backbone of Mtb-MurB (yellow) and template (*E. coli*)-MurB (blue). **b** Homology model of Mtb-MurB showing three domains: domain 1 (yellow), domain 2 (pink) and domain 3 (cyan). Cofactor FAD and ligand naphthyl tetronic acid is shown in ball and stick model

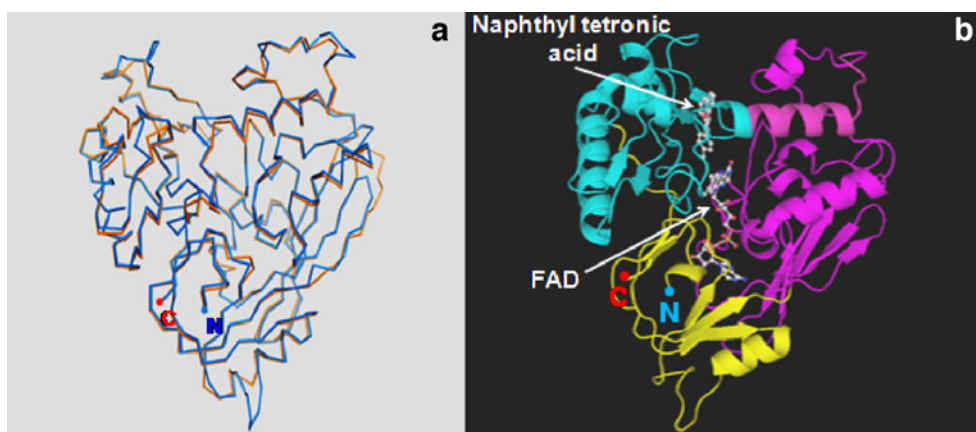
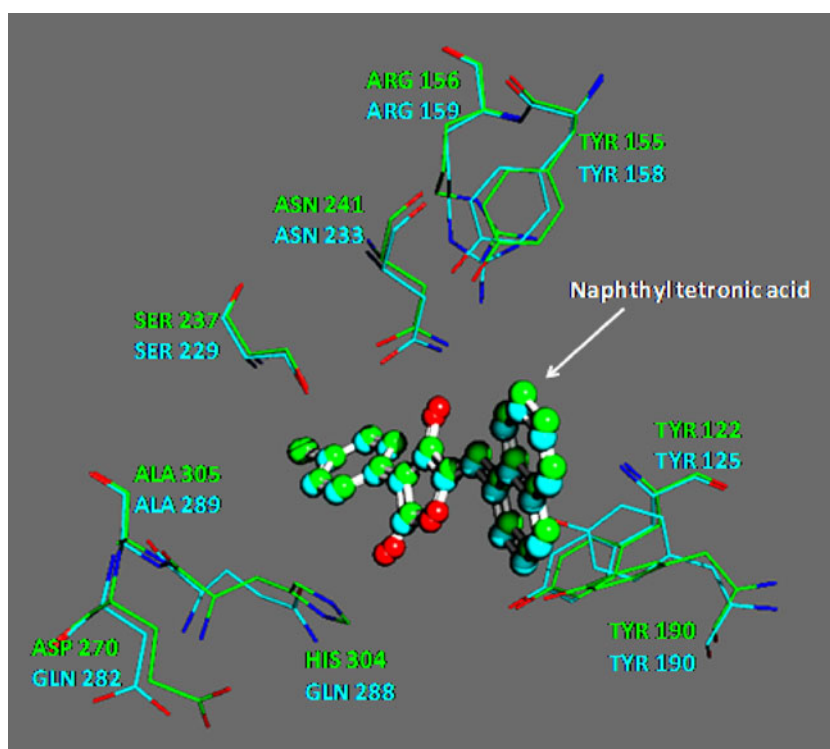


Fig. 3 Superimposition of active site residues within 5.5 Å region surrounding naphthyl tetronic acid inhibitor in Mtb-MurB (green) and (*E. coli*)-MurB (cyan)



to deletion of the residue (possibly due to occurrence of a point mutation during protein evolution) between these residues (Fig. 1). Interestingly, we noted that in Mtb-MurB, Gly47 and Gly49 made three hydrogen bonds with both the α -phosphoryl and the β -phosphoryl moiety in FAD, while in (*S. aureus*)-MurB, Gly67 and Gly69 made single hydrogen bonds with the α -phosphoryl moiety of FAD. This analysis clearly showed that FAD is more tightly bound within Mtb-MurB as compared to (*S. aureus*)-MurB protein. This unique feature of the Mtb-MurB protein might be important in the MurB family of proteins. This motif in (*S. aureus*)-MurB interacts with the α -phosphoryl moiety of FAD via the main-chain nitrogen, and the Ser70 residue (Ser50 residue in (*E. coli*)-MurB and Mtb-MurB) interacts with the β -phosphoryl moiety of FAD via both its main-chain nitrogen and side-chain hydroxyl group; while Arg213 and His259 [Arg214 and Gln288 in (*E. coli*)-MurB and Mtb-MurB] help in maintaining the electronic state of the FAD cofactor [33].

Nishida et al. [34] performed site-directed mutagenesis of (*S. aureus*)-MurB genes and tested the ability of the resulting constructs to complement a temperature-sensitive MurB mutant of *S. aureus*. Residues Asn71, Tyr175, Arg176, Arg213, Ser226, His259, and Glu296 of (*S. aureus*)-MurB, were found to be essential for its activity. Sequence alignment provided evidence of conservation of these amino acid residues among 15 bacterial species other than Mtb [34]. We examined in detail the amino acid

residues of MurB protein in three different species—*S. aureus*, *E. coli* and Mtb—using multiple sequence alignment methodology (Fig. 1).

Residues Phe46, Gly48, Gly49, Ser50, Asn51, Gln117, Val119 and Val168 are observed to form hydrogen bonds with FAD, while Asn241 bonds with naphthyl tetronic acid in the Mtb-MurB homology model (Fig. 4). Residues Ile45, Gly47, Gly49, Ser50, Asn51, Pro111, Cys113, Ser116, Gly123 and Arg214 are observed to form hydrogen bonds with FAD, whereas Asn233 bonds with naphthyl tetronic acid in the crystal structure of (*E. coli*)-MurB [32, 33]. Hydrogen bonding residues of Mtb-MurB model that interact with FAD and the ligand (naphthyl tetronic acid) are presented in Fig. 4, to give a clearer view of the structure–function relationships of this enzyme.

Further, we employed molecular electrostatic potential surface (MEPS) maps as a tool to understand the difference in the charge distribution of the Mtb-MurB model compared with that of the (*E. coli*)-MurB crystal structure. It is a well established fact that molecular electrostatics form the main part of the non-bonded interaction energy between ligands and proteins, governing the strength of non-bonded interactions as well as molecular reactivity. Different physico-chemical environments, such as electrostatic, steric, and hydrophobic forces experienced by the ligand at the active site of the protein could contribute to ligand selectivity/specificity [36]. Thus, comparing the MEP of (*E. coli*)-MurB and Mtb-

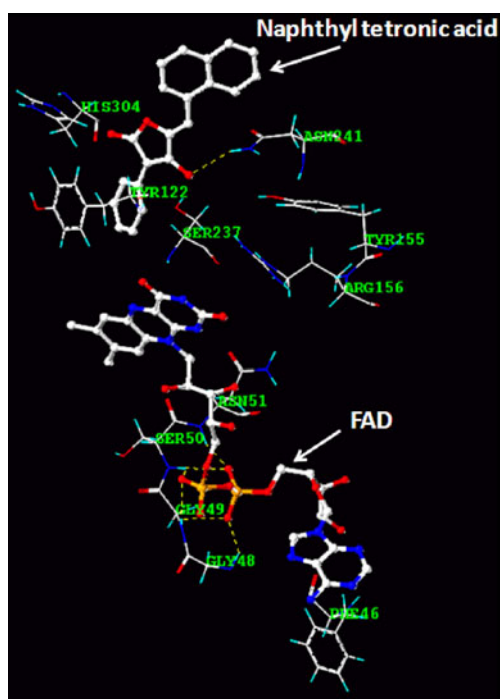


Fig. 4 FAD along with the key residues of the Mtb-MurB model. FAD and naphthyltetronic acid is represented by ball and stick model. Yellow dotted lines indicate hydrogen bonds

MurB proteins at the inhibitor binding pocket could provide an effective means of understanding the selective/specific inhibition of Mtb-MurB protein.

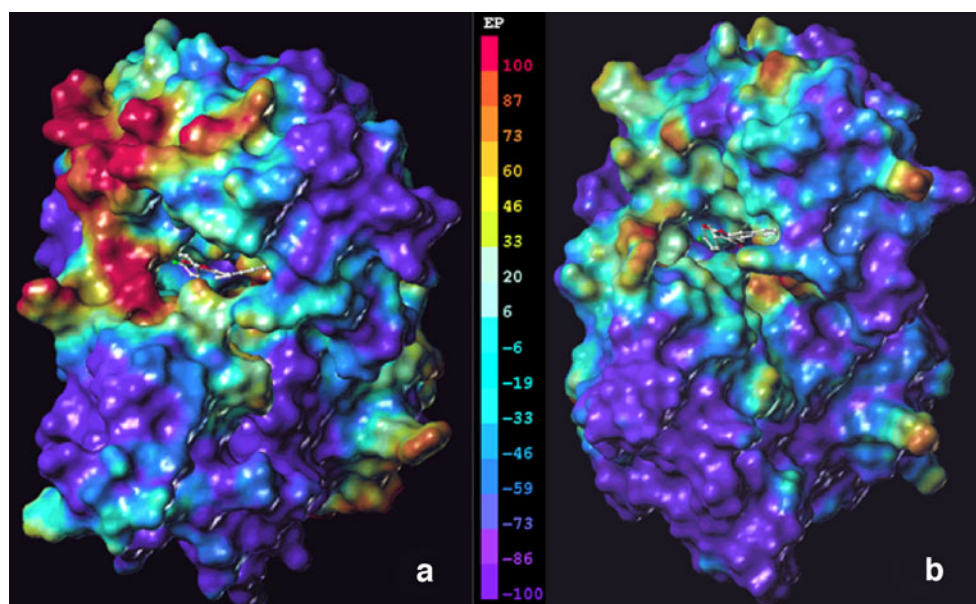
The MEPS maps of Mtb-MurB and (*E. coli*)-MurB are presented in Fig. 5a,b, respectively, with their MEPS color ramps. To facilitate comparison, MEPS of the two proteins were placed on the same scale (-100 to $+100$ kcal mol $^{-1}$). The MEPS map of (*E. coli*)-MurB protein (Fig. 5a) showed

blue color regions spanning most parts of its surface in comparison to that of the corresponding Mtb-MurB protein. It is clear from Fig. 5 that the MEPS around the ligand naphthyl tetronic acid showed variation in both (*E. coli*)-MurB and Mtb-MurB proteins. Striking differences were observed at the tetronic acid site (chlorobenzene), showing higher positive potential regions in Mtb-MurB vs neutral-to-negative MEPS regions in (*E. coli*)-MurB, respectively (Fig. 5). Thus, there is clear evidence of differences in charge complementarity observed near the chlorine atom of the ligand naphthyl tetronic acid, indicating tighter binding in Mtb-MurB than in (*E. coli*)-MurB (Fig. 5). In other words, ligand binding site in Mtb-MurB contained a wide range of charge distribution, while (*E. coli*)-MurB has a narrow range of charge distribution from electropositive to electronegative. The present analysis clearly indicates that (*E. coli*)-MurB has a very short range of MEPS values within the active site as compared to that of the Mtb-MurB protein. Thus, MEPS analysis revealed basic differences in the ligand binding pocket, which could be exploited in the future for the rational design of selective Mtb-MurB inhibitors.

Molecular dynamics simulation

Molecular dynamics simulations were performed to determine the stability of the predicted 3D structure of Mtb-MurB enzyme. The trajectories were stable during the whole production phase of the 5,000 ps (or 5 ns) MD simulation run. Trajectory stability was monitored and was confirmed by analysis of backbone RMSD and potential energy (Fig. 6) as a function of time. RMSD values for Mtb-MurB showed a rise in the first 2,000 ps,

Fig. 5 Molecular electrostatic potential surface (MEPS) map of Mtb-MurB (a) and (*E. coli*)-MurB (b) along with the MEPS color ramp from ($+100$ to -100) kcal mol $^{-1}$. Most positive potential regions are shown by red color while most negative potential regions with blue color, on the same potential scale for comparison. The ligand naphthyl tetronic acid is shown in ball and stick model



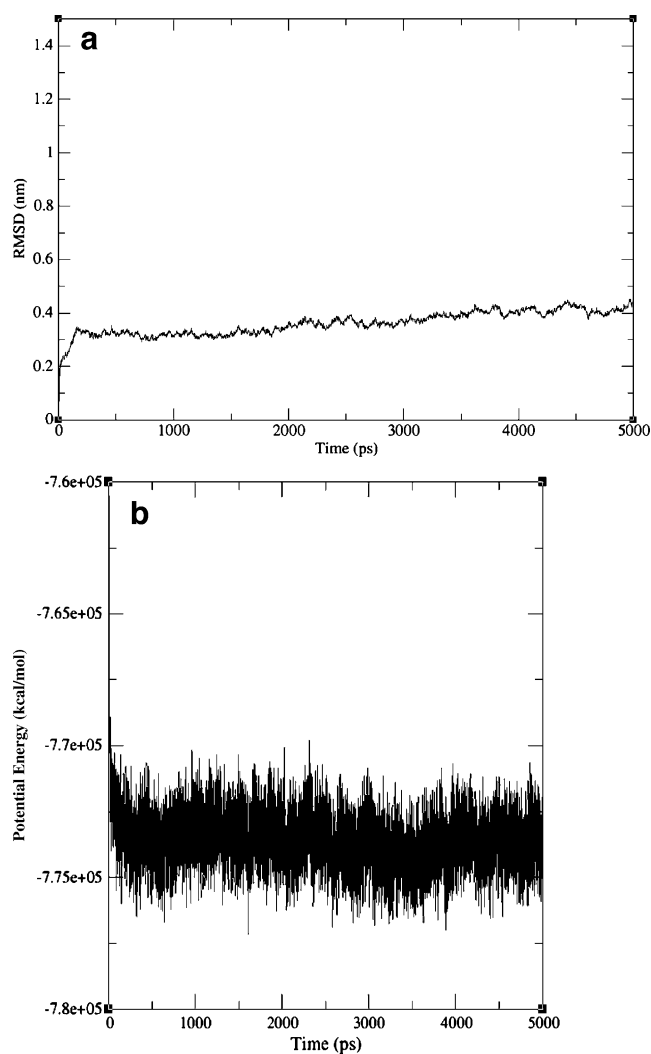


Fig. 6 Graphical representation of RMSD of backbone carbons vs time (a) and potential energy vs simulation time (b) for the Mtb-MurB model

and then remained stable for the remaining simulation time. The average RMSD for the Mtb-MurB model when measured from 5,000 ps was found to be 0.36 nm, and the potential energy (kJ mol^{-1}) was found to be stable throughout the simulation time. Root mean square fluctuation (RMSF) of C_{α} carbon atoms (Fig. S7), RMSF of backbone (N, C_{α} , C) atoms (Fig. S8) and RMSF of side chain atoms (Fig. S9) were plotted for the developed model in order to understand the dynamic stability of the protein system. It was observed that, throughout the dynamic simulations, very few fluctuations exceeded 0.5 nm, and even fewer fluctuations exceeded 0.6 nm for the total protein. The graph showed that residues at the N-terminus, 184–208 and 257–272, are very flexible and have fluctuations close to 0.5 nm, while active site residues showed less fluctuation and remained stable.

Molecular docking analysis

Docking analysis was performed on 28 Mtb-MurB inhibitors to identify key amino acid residues involved in making interactions with the Mtb-MurB model structure. All inhibitors were docked successfully in the Mtb-MurB homology model, showing good docking scores. The docking program Autodock computes binding energy (BE) and inhibition constant (K_i) with respect to the docked inhibitors. The structure of 28 Mtb-MurB inhibitors, their experimental activity (IC_{50}) and Autodock-computed BE and K_i values along with H-bond interacting residues are presented in Table 3.

The docked pose of the most potent molecule, **10a** (IC_{50} = 5.8 μM), and naphthyl tetronic acid obtained using Autodock4 software is presented in Fig. 7. Naphthyl tetronic acid showed strong hydrogen bonding interactions between the oxygen atom of the carbonyl group of its five-membered ring and Asn241 (Fig. 7a); while molecule **10a** forms hydrogen bonds with residues Asn241 and Ser237 in Mtb-MurB model (Fig. 7b and Table 3, respectively). In addition, Tyr122, Gly123, Arg156, Arg218, Gly262 and Lys274 are among other residues involved in making hydrogen bonding interactions with the other docked inhibitors in the Mtb-MurB model (Table 3).

Multiple sequence alignment of Mtb, *E. coli* and *S. aureus* MurB proteins revealed that Ser237 is conserved in all three species. Residue Asn241 is conserved in Mtb and *E. coli* MurB proteins and the corresponding residue in *S. aureus* MurB is Arg230. Residues Tyr122, Gly123, Arg156 and Arg218 are conserved in all three species; residue Lys274 is conserved in Mtb and *E. coli* MurB protein; and residue His304 is conserved in Mtb and *S. aureus* MurB. Thus, docking analysis provided direct evidence of the key inhibitor binding residues in Mtb-MurB, which are in agreement with those of the *E. coli* and *S. aureus* MurB proteins.

BE and K_i values of the most potent molecule **10a** is $-8.73 \text{ kcal mol}^{-1}$ and 0.39 μM , respectively, while those of the least potent molecule **3c** are $-7.16 \text{ kcal mol}^{-1}$ and 5.66 μM respectively (Table 3). Using 28 Mtb-MurB inhibitors, we obtained a good correlation of 0.83 between BE and activity value (IC_{50}), and 0.81 between K_i and activity value (IC_{50}), respectively (Fig. S10).

Docking analysis was also performed on the mutant models of the Mtb-MurB enzyme. Tyr155 was mutated to phenylalanine and Ser237 to alanine. Site-directed mutagenesis experimental analysis of *S. aureus* species for these mutants (Tyr175 mutated to Phe175; Ser226 mutated to Ala226) was determined and the corresponding K_d (ligand dissociation constant) values are presented in Table 4 [34]. Experimental kinetic values and our molecular docking analysis for the mutants and wild type

Table 3 Structure of 28 molecules along with their experimental activity and Autodock-computed binding energy (BE) and inhibition constant (K_i) along with the H-bond interacting residues in the Mtb-MurB model

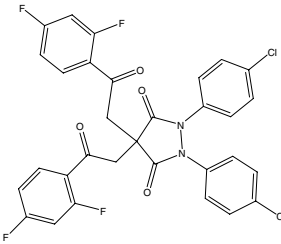
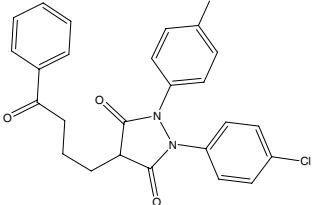
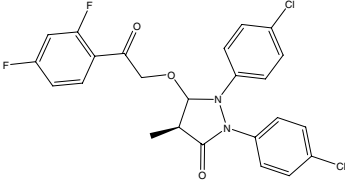
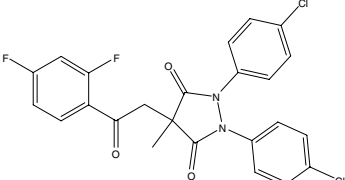
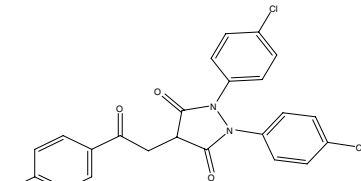
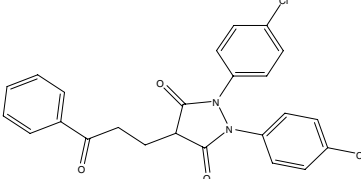
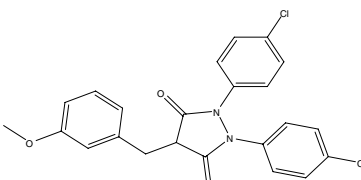
^a Molecule ID	Structure of compound	^a IC ₅₀ (μM)	Autodock		H-bonding interactions
			Binding Energy (kcal/mol)	Inhibition constant (K_i) μm	
10a		5.8	-8.73	0.39	Ser237, Asn241
7m		10	-8.41	0.68	Asn241, Lys274
15a		50	-7.54	2.98	Ser237, Tyr122
14a		49	-7.54	2.98	Ser237
9a		5.4	-8.79	0.36	Asn241, Tyr122, Gly123
7n		13	-8.73	0.39	Asn241, Arg156
3e		50	-7.29	4.51	His304, Gly262

Table 3 (continued)

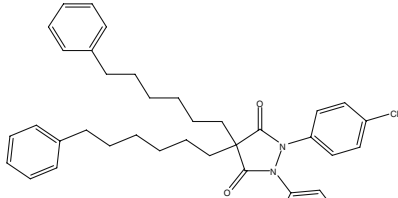
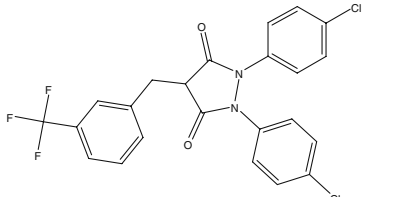
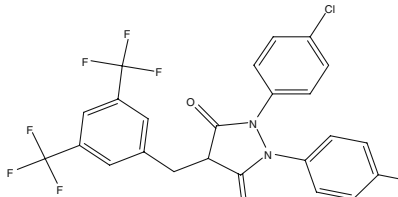
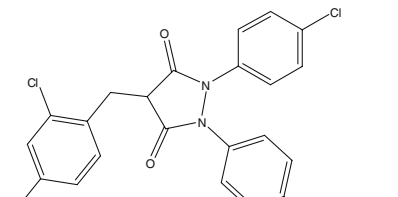
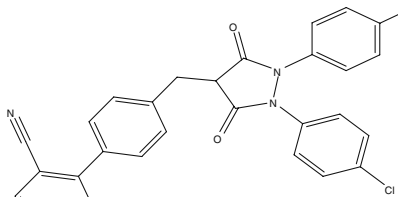
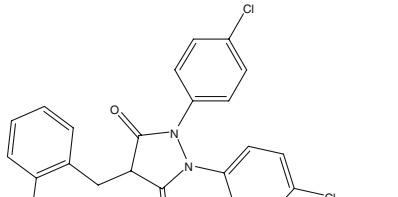
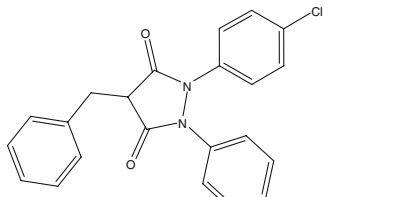
8a		7.2	-8.61	0.49	Tyr122, Asn241
7i		18	-7.26	4.78	Asn241
7k		12	-7.81	1.90	Arg218
7j		11	-8.00	1.36	Arg218
7l		11	-9.04	0.23	Tyr122, Asn241
7h		50	-7.38	3.87	Ser237
3c		50	-7.16	5.66	Arg218

Table 3 (continued)

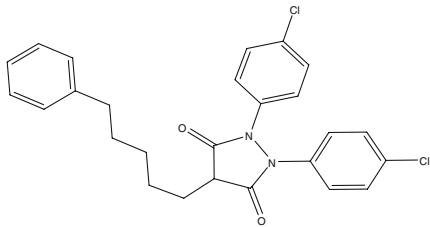
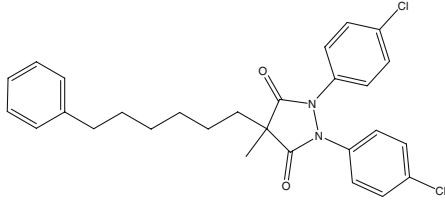
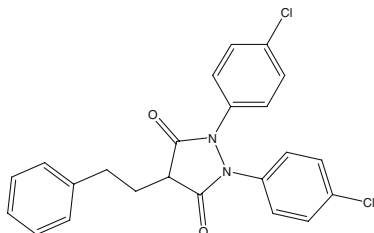
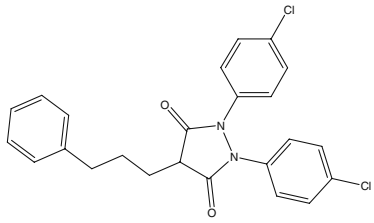
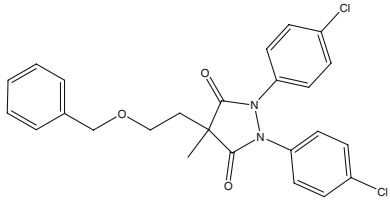
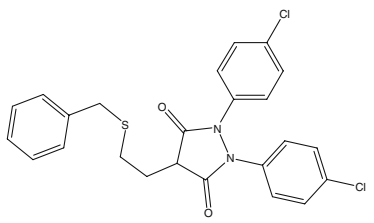
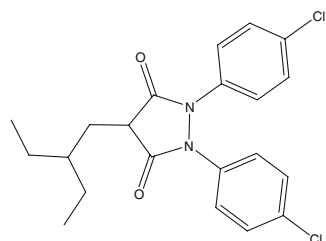
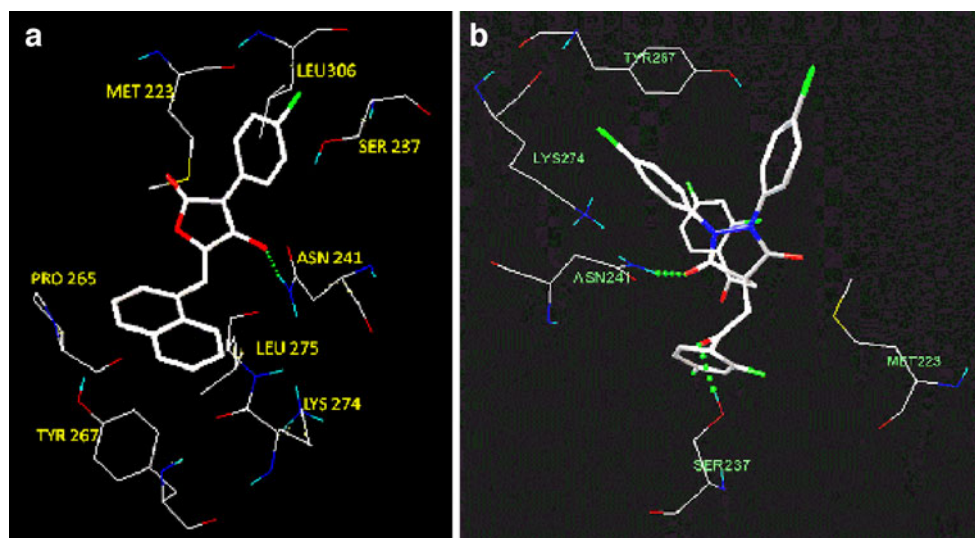
7c		10	-8.27	0.87	Tyr122, Asn241
13a		50	-7.40	3.78	Ser237
3d		35	-7.66	2.41	Ser237, Asn241
7a		10	-8.16	1.04	Tyr122, Asn241
13b		20	-8.17	1.02	Arg156
7f		5.2	-9.07	0.23	Arg218
3a		50	-7.49	3.23	Arg218, Gly123

Table 3 (continued)

7d		8.9	-8.22	0.94	Ser237
7b		5.6	-8.89	0.30	Ser237
3b		42	-7.56	2.87	Gly123, Asn241
8b		7.2	-8.91	0.29	Asn241, Ser237
7e		50	-7.48	3.31	Ser237
3f		50	-6.98	7.62	Asn241, Tyr122
7g		11	-7.82	1.86	Ser237

^aMolecular numbering convention and Mtb-MurB inhibitory activity value from Ref. [31]

Fig. 7 Docked poses of naphthyl tetronic acid (**a**) and potent molecule **10a** (**b**) in the Mtb-MurB homology model



Mtb-MurB, (*E. coli*)-MurB and (*S. aureus*)-MurB enzymes are compared in Table 4. The most active molecule **10a** was docked in the above mentioned mutant models of Mtb-MurB using the Autodock4 program. BE and K_i values obtained for the mutants were less than values obtained with the wild type Mtb-MurB enzyme (BE = $-8.73 \text{ kcal mol}^{-1}$ and $K_i = 0.39 \text{ }\mu\text{M}$) since the contribution from these individual key amino acid residues at the binding site is abolished, affecting the function of the enzyme (Table 4). This analysis provided clear evidence of the importance of these amino acid residues in the active site of Mtb-MurB enzyme, and also the importance of our developed homology model. The smaller the dissociation constant, the more tightly the ligand is bound to the protein, or the higher the affinity between ligand and protein. The values presented in Table 4 clearly depict the good correlation between our computed docking results and the experimentally obtained ligand dissociation constant for different Mtb-MurB mutant models [34].

In summary, the 28 molecules belonging to 3,5-dioxopyrrolidine derivatives studied are potent inhibitors targeting Mtb-MurB. We observed that the homology model of Mtb-MurB assisted molecular docking analysis of this class of

inhibitors, establishing a strong relationship between enzyme inhibition, decreased soluble peptidoglycan synthesis, and antibacterial activity. Thus, the present structure-based inhibitor design study is encouraging, and indicates the possibility of developing novel anti-bacterial inhibitors targeting an early step in peptidoglycan biosynthesis.

Conclusions

We report the development of a homology model of Mtb-MurB enzyme based on the high-resolution structure of the (*E. coli*)-MurB enzyme. The Mtb-MurB homology model provided for the first time a 3D structural model that can be used for screening different molecules for Mtb-specific MurB inhibitory activity. The developed model showed good overall structural quality and was confirmed using several different validation tools. The overall stability of the Mtb-MurB model structure was finally investigated using an unconstrained MD simulation. The results showed that RMSD and potential energy remained stable during the 5 ns MD simulation.

Table 4 Comparison of the experimental kinetic values and molecular docking analysis for the mutants and wild type Mtb-MurB, (*E. coli*)-MurB and (*S. aureus*)-MurB enzymes. BE Binding energy

Mtb-MurB	Autodock ^a		(<i>S. aureus</i>)-MurB ^b	(<i>E. coli</i>)-MurB ^c
	BE (kcal/mol)	Inhibition constant (μM)		
Y155F (Mutant)	-8.48	0.61	173 (Y175F)	
S237A (Mutant)	-7.45	3.44	180 (S226A)	7.3 (S229A)
Wild Type	-8.73	0.39	41	4.1

^a Autodock analysis was carried out using the most potent inhibitor (**10a**) from the series

^b K_d values were obtained for the physiological ligand UDP-GlcNAcEP [15]

^c K_d values were obtained for the physiological ligand UDP-GlcNAcEP [35]

Twenty-eight Mtb-MurB inhibitors were docked in the homology model, and reasonably good correlation was found between *in vitro* activity (IC₅₀) and the BE obtained using the program Autodock. The docking study on mutant models also provided clear evidence of the importance of Tyr 155 and Ser237 as ligand-binding amino acid residues. There is a good correlation between our computed BE and K_i values and the experimental kinetic values of different Mtb-MurB mutant models. Thus, the present work forms the basis for further molecular modeling and biochemical studies on targeting the Mtb-MurB enzyme for TB therapy.

Acknowledgments V.K. is thankful to Department of Biotechnology, Government of India for providing a Junior Research Fellowship. The authors are grateful to the Department of Pharmaceuticals, Government of India for financial assistance.

References

- McKinney JD, Höner Zu Bentrup K, Muñoz-Elías EJ et al (2000) *Nature* 406:735–738
- Ducati RG, Ruffino-Netto A, Basso LA et al (2006) *Mem Inst Oswaldo Cruz* 101:697–714
- Corbett EL, Watt CJ, Walker N et al (2003) *Arch Int Med* 163:1009–1021
- Nacheha JB, Chaisson RE (2003) *Clin Infect Dis* 36:24–30
- Blumberg HM, Leonard MK, Jasmer RM (2005) *Am Med Assoc* 293:2776–2784
- Zhang Y, Post-Martens KS (2006) *Drug Discov Today* 11:21–27
- Zhang Y (2004) *Front Biosci* 9:1136–1156
- Zhang Y (2005) *Annu Rev Pharmacol Toxicol* 45:529–564
- Anishetty S, Pulimi M, Pennathur G (2005) *Comput Biol Chem* 29:368–378
- Green DW (2002) *Exp Opin Ther Targ* 6:1–20
- Yang Y, Severin A, Chopra R et al (2006) *Antimicrob Agents Chemother* 50:556–564
- El Zoeiby A, Sanschagrin F, Levesque RC (2003) *Mol Microbiol* 47:1–12
- Kimura K, Bugg TD (2003) *Nat Prod Rep* 20:252–273
- Mansour TS, Caufield CE, Rasmussen B (2007) *Chem Med Chem* 2:1414–1417
- Garg A, Tewari R, Raghava G (2010) *BMC Bioinformatics* 11 (Suppl 1):S53. doi:10.1186/1471-2105-11-S1-S53
- Khedkar SA, Malde AK, Coutinho EC (2005) *J Mol Graph Model* 23:355–366
- Anuradha CM, Chaitanya M, Babajan B et al (2010) *J Mol Model* 16:77–85
- da Cunha EE, Barbosa EF, Oliveira AA et al (2010) *J Biomol Struct Dyn* 27:619–625
- Tatusova TA, Madden TL (1999) *FEMS Microbiol Lett* 174:247–250
- Henikoff S, Henikoff JG (1992) *Proc Natl Acad Sci USA* 89:10915–10919
- Marti-Renom MA, Stuart AC, Fiser A et al (2000) *Annu Rev Biophys Biomol Struct* 29:291–325
- Laskowski RA, MacArthur MW, Moss DS et al (1993) *J Appl Crystallogr* 26:283–291
- Colovos C, Yeates TO (1993) *Protein Sci* 2:1511–1519
- Sippl MJ (1993) *Proteins Struct Funct Genet* 17:355–362
- Baker BR, Garrell RL (2004) *Faraday Discuss* 126:209–222
- Bowie JU, Luthy R, Eisenberg D (1991) *Science* 253:164–170
- Wiederstein M, Sippl MJ (2007) *Nucleic Acids Res* 35:407–410
- Lindahl E, Hess B, van der Spoel D (2001) *J Mol Model* 7:306–317
- Berendsen HJC, Postma JPM, van Gunsteren WF (1981) In: Pullman B (ed) *Intermolecular Forces*. Reidel, Dordrecht, p 331
- Morris GM, Goodsell DS, Halliday RS et al (1998) *J Comput Chem* 19:1639–1662
- Kutterer KMK, Davis JM, Singh G et al (2005) *Bioorg Med Chem Lett* 15:2527–2531
- Benson TE, Walsh CT, Hogle JM (1996) *Structure* 4:47–54
- Benson TE, Filman DJ, Walsh CT et al (1995) *Nat Struct Biol* 2:644–653
- Nishida S, Kurokawa K, Matsuo M et al (2006) *J Biol Chem* 281:1714–1724
- Timothy EB, Christopher TW, Vincent M (1997) *Biochemistry* 36:796–805
- Awale M, Kumar V, Saravanan P et al (2010) *J Mol Model* 16:475–488