J Mol Model (2010) 16:1357–1367
DOI 10.1007/s00894-010-0653-4

ORIGINAL PAPER

Exploring the molecular basis for selective binding of *Mycobacterium tuberculosis* Asp kinase toward its natural substrates and feedback inhibitors: A docking and molecular dynamics study

M. Chaitanya · B. Babajan · C. M. Anuradha · M. Naveen · C. Rajasekhar · P. Madhusudana · Chitta Suresh Kumar

Received: 28 October 2009 / Accepted: 5 January 2010 / Published online: 7 February 2010

Abstract Tuberculosis (TB) is still a major public health problem, compounded by the human immunodeficiency virus (HIV)-TB co-infection and recent emergence of multidrug-resistant (MDR) and extensively drug resistant (XDR)-TB. In this context, aspartokinase of *mycobacterium tuberculosis* has drawn attention for designing novel *anti*-TB drugs. Asp kinase is an enzyme responsible for the synthesis of 4-phospho-L-aspartate from L-aspartate and involved in the branched biosynthetic pathway leading to the synthesis of amino acids lysine, threonine, methionine and isoleucine. An intermediate of lysine biosynthetic branch, mesodiaminopimelate is also a component of the peptidoglycan which is a component of bacterial cell wall. To interfere with the production of all these amino acids and cell wall, it is possible to inhibit Asp kinase activity. This can be achieved using Asp kinase inhibitors. In order to design novel Asp kinase inhibitors as effective *anti*-TB drugs, it is necessary to have an understanding of the binding sites of Asp kinase. As no crystal structure of the enzyme has yet been published, we built a homology model of Asp kinase using the crystallized Asp kinase from *M. Jannaschii*, as template structures (2HMF and 3C1M). After the molecular dynamics refinement, the optimized homology model was assessed as a reliable structure by PROCHECK, ERRAT, WHAT-IF, PROSA2003 and VERIFY-3D. The results of molecular docking studies with natural substrates, products and feedback inhibitors are in agreement with the published data and showed that ACT domain plays an important role in binding to ligands. Based on the docking conformations, pharmacophore model can be developed by probing the common features of ligands. By analyzing the results, ACT domain architecture, certain key residues that are responsible for binding to feedback inhibitors and natural substrates were identified. This would be very helpful in understanding the blockade mechanism of Asp kinase and providing insights into rational design of novel Asp kinase inhibitors for *M.tuberculosis*.

Keywords AutoDock · Errat · Gromacs · Modeller · Pharmacophore · Procheck · Prosa2003 · What-if

M. Chaitanya · B. Babajan · C. Rajasekhar · C. S. Kumar (✉)
Department of Biochemistry, Sri Krishnadevaraya University,
Anantapur 515003 A.P, India
e-mail: chitt34c@gmail.com

C. M. Anuradha · P. Madhusudana
Department of Biotechnology,
University College of Engineering and Technology,
Sri Krishnadevaraya University,
Anantapur 515003 A.P, India

M. Naveen
Institute of Life Sciences,
University of Hyderabad Campus,
Gachibowli, Hyderabad

Introduction

Mycobacterium tuberculosis continues to be one of the world's deadliest pathogens, causing a prospective burden of one billion newly infected individuals and 36 million casualties within the next 20 years [1]. Despite the existence of effective chemotherapies, no new drugs have come onto the market during the past 40 years. In addition, the rise in drug resistance among *M.tuberculosis* strains is becoming a severe threat to public health, illustrated by the

recent emergence of extensively drug-resistant tuberculosis (XDR-TB) that has caused a mortality of >98% [2]. The XDR *M.tuberculosis* isolates were resistant to isoniazid and rifampin and to at least three of the six main classes of second-line drugs (aminoglycosides, polypeptides, fluoroquinolones, thioamides, cycloserines and *para*-aminosalicylic acid) [3]. They not only constitute a deadly threat to the affected patients with TB but also hamper the TB-control program. If these extensively resistant pathogens are allowed to develop and spread in the society, they will constitute a significant public health problem [4, 5]. This problematic situation illustrates an urgent need to find an effective medicine for treating such complicated cases.

In the present work we have focused on the end product regulation of the first enzyme in a branched biosynthetic pathway leading to the synthesis of multiple end products. In bacteria, the branched biosynthetic pathway leading to lysine, methionine and threonine is controlled by the end products and their biosynthetic intermediates in a process called feedback inhibition. In bacteria enzymes of this pathway have a high importance as they constitute targets for the development of new pharmaceutical compounds. In this scenario we aimed at selective prediction of interaction sites of aspartokinase (Asp kinase) (EC 2.7.2.4) shown in (Fig. 1), which catalyzes the first step *i.e.*, conversion of L-aspartate to 4-phospho-L-aspartate, of branched biosynthetic pathway leading to the synthesis of amino acids lysine, threonine, methionine and isoleucine [6, 7]. An intermediate of lysine-biosynthetic branch, mesodiaminopimelate is also a component of peptidoglycan, constituent of bacterial cell wall. The cell wall plays a vital role in bacterial growth and survival in hostile environment [8]. In this aspect we can conclude that Asp kinase also plays a role for the formation of the cell wall. Asp kinase control presents a unique situation, where in the presence of any deficiency in the production Asp kinase, results in the deficiency of amino acids lysine, threonine, methionine and isoleucine as well as peptidoglycan of the cell wall.

The present study is aimed at elucidating the 3D structural features of *M.tubercuosis* Asp kinase and selective prediction of interaction sites for substrates and feedback inhibitors. In this paper, we report that the 3D model of Asp kinase was derived using comparative modeling analysis [9] and that the generated 3D models would give insight into the influence of various interactive fields on the activity and thus, can help in designing and forecasting the translation inhibition activity of novel molecules. Further, refinement of the generated 3D model was done by subjecting it to molecular dynamics (MD) simulations. Molecular docking studies were also performed to analyze the interactions among Asp kinase and its ligands, which are found to be helpful in designing of novel *anti*-tuberculosis compounds.

Materials and methods

All the calculations were performed on a workstation AMD Opteron Duo-core 2.0 GHz and 4 GB RAM. Molecular modeling tasks were performed with Modeller9v3, MD simulations were analyzed with Gromacs 3.2.1; docking calculations were performed with AutoDock 4.0. If not otherwise stated, default settings were used during all calculations.

Sequence alignments

The sequence of the *M.tuberculosis* Asp kinase (accession number NP_218226) was obtained from NCBI database [10]. With the aim of finding an adequate template for homology modeling of *M.tuberculosis* Asp kinase, sequence alignments of its amino acid sequence against

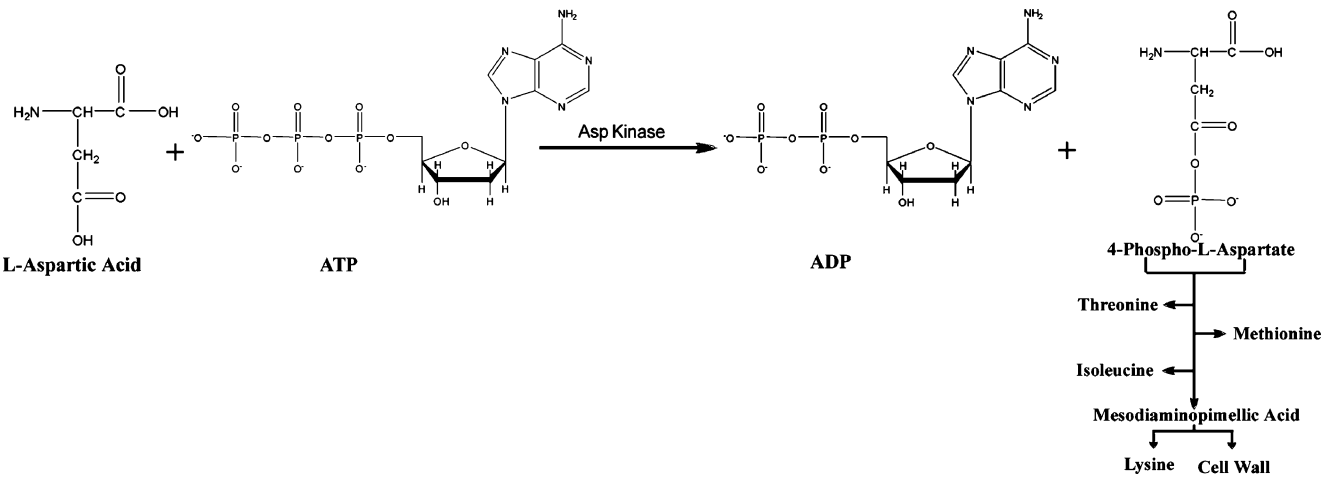

Fig. 1 Schematic representation of reaction catalyzed by Asp kinase from *M.tuberculosis*

Protein Data Bank (PDB) [11] were performed by means of the BLAST algorithm [12, 13]. Finally, an alignment between the selected templates and the *M.tuberculosis* Asp kinase was performed using the ClustalX 1.83 with default parameters [14].

Molecular model building

Among the BLASTp results for searching homology proteins in PDB, two structures were selected as templates: the threonine sensitive Asp kinase from Methanococcus jannaschii complexed with Mg-ADP and Aspartate (PDB ID: 2HMF) [15] and crystal structure of Threonine sensitive aspartokinase from *M. Jannaschii* wit Mg amp-PNP and Aspartate (PDB ID: 3C1M) [16]. The structure of *M. tuberculosis* Asp kinase was predicted by homology modeling on the basis of the structures of 2HMF and 3C1M using the program Modeller 9v3 [17]. This program is an automated approach to comparative modeling by satisfaction of spatial restraints [18–20]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). 200 models were generated and among them the one having lowest root mean square deviation (RMSD) value when superposed onto the templates 2HMF and 3C1M, was chosen for further analysis [21].

Molecular dynamics simulation

In order to validate the sequence-structure alignment, to remove bad contacts derived from the homology modeling and to achieve a good starting structure, the model was subjected to exhaustive molecular dynamics simulation of 8 ns with Gromacs 3.2.1 software using Gromacs 96.1 (43 A^2) force field [22]. The initial structure was placed in a box (7.552nmX6.339nmX8.751 nm) with a total volume of 418.93 nm^3. The box was filled with water (11,397 molecules), and 82 water molecules were replaced with 41 Na+ and 41 Cl^- ions in random order to neutralize the system and to represent physiological conditions (0.9% NaCl solution). The simulations were carried out using explicit solvent water molecules (described by the simple point charge, SPC/E) [23, 24] and periodic boundary conditions (cubic). In the MD protocol, all hydrogen atoms, ions, and water molecules were first subjected to 200 steps of energy minimization by steepest descent algorithm [25] to remove close van der Waals contacts. The system was then submitted to a short MD simulation with position restraints for a period of 1 ps and afterward it was subjected to a full MD without restraints. The temperature of the system was then increased from 50–310 K in five steps (50–100 K, 100–150 K, 150–200 K, 200–250 K, and 250–310 K) and the velocities at each step was reassigned according to the Maxwell-Boltzmann distribution at that temperature and equilibrated for 2 ps. Energy minimization and MD were carried out under periodic boundary conditions. The simulation was computed in the isobaric-isothermal (NPT) ensemble at 310 K with the Berendsen temperature coupling and constant pressure of 1 atm with isotropic molecules-based scaling [23]. The LINCS algorithm [26], with a 10^{-5} Å tolerance, was applied to fix all bonds containing a hydrogen atom, allowing the use of a time step of 2 fs in the integration of the equations of motion. No extra restraints were applied after the equilibration phase. The convergence of simulation was analyzed in terms of the potential energy, RMSD from the initial model structure and root mean- square fluctuation (RMSF). The analysis was calculated relative to the Cα back bone structures, and all coordinate frames from the trajectories were first superimposed on the initial conformation to remove any effect of overall translation and rotation.

Assessment of the homology model

To obtain an accurate homology model, it is very important that appropriate steps are built into the process to assess the quality of the model [27]. Therefore, in the modeling phase, the model quality was assessed by the geometric quality of the backbone conformation, the residue interaction, the residue contact and the energy profile of the structure using different methods, including PROCHECK [28] ERRAT [29], WHAT-IF [30], PROSA2003 [31] and Verify 3D [32, 33]. The organization of various domains in the modeled structure of *M.tuberculosis* Asp kinase was identified using Scansite [34] server. Scansite searches for motifs within the proteins by utilizing the entropy approach that assesses the probability of a site matching the motif using the selective values and sums the logs of the probability values for each amino acid in the candidate sequence.

Validation of the model by docking analysis

A docking study was conducted to evaluate the predictive ability of the *M.tuberculosis* Asp kinase homology model and its suitability for use in the structure-based drug design studies. As reported Asp kinase is inhibited by feedback inhibition, in the current study docking studies were performed to gain insights into the most probable binding conformations of *M.tuberculosis* Asp kinase with substrates, products and feedback inhibitors. The structures of Aspartic acid, ADP (Adenosine di- phosphate), substrates of Asp kinase and Lysine, Threonine feedback inhibitors of Asp kinase were built using Chemdraw (Cambridge soft Inc.) [35]. After a preliminary energy minimization to discard high-energy intramolecular interactions, the overall

geometry and the atomic charges were optimized using HyperChem release 7.5 [36]. In the validation phase, AutoDock 4.0 [37, 38] was used for performing docking. AutoDock combines a rapid energy evaluation through precalculated grids of affinity potentials with a variety of search algorithms to find suitable binding positions for a ligand on a given protein. When docking was performed, *M.tuberculosis* Asp kinase was kept rigid, but all the torsional bonds in ligands were set free to perform flexible docking. Polar hydrogens were added using the hydrogens module in AutoDockTools (ADT) for Asp kinase; after that Kollman united atom partial charges were assigned [39].

Docking of ligands to *M.tuberculosis* Asp kinase was carried out using the empirical free energy function and the Lamarckian genetic algorithm, applying a standard protocol with an initial population of 300 randomly placed individuals, a maximum number of $2.5X10^7$ energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1, where the average of the worst energy was calculated over a window of the previous 10 generations. For the local search, the so-called Solis and wets algorithm was applied [37], using a maximum of 300 iterations. The probability of performing a local search on an individual in the population was 0.06, and the maximum number of consecutive successes or failures before doubling or halving the local search step size was 4; 100 independent docking runs were carried out for each ligand. Results were clustered according to the 1.0 Å root-mean-square deviation (RMSD) criterion. All torsion angles for each compound was considered flexible. The grid maps representing the proteins in the actual docking process were calculated with AutoGrid. The grids (one for each atom type in the ligand plus one for electrostatic interactions) were chosen to be sufficiently large to include not only active site but also significant portions of the surrounding surface. The dimensions of the grids were thus 60 Å× 60 Å×60 Å, with a spacing of 0.375 Å between the grid points. After docking the ligand-receptor complexes were analyzed by Pymol program [40].

Results and discussion

Sequence alignments

The search using the BLASTp alignment algorithm within the PDB database showed various potential templates for molecular modeling purposes. More than 60 crystallographic structures were found showing high identity score with respect to *M.tuberculosis* Asp kinase. No crystallographic data was observed when BLASTp analysis against humans was performed for *M.tuberculosis* Asp kinase. This allowed us to use this enzyme for the development of potential *anti*-tuberculosis drugs. Among the BLASTp results, two structures were selected as templates: the threonine sensitive Asp kinase from *M.jannaschii* complexed with Mg-ADP and Aspartate (PDB ID: 2HMF) [15], and threonine sensitive Asp kinase from *M.jannaschii* with MgAMP-PNP and L-aspartate (PDB ID: 3C1M) [16]. The sequence identities between *M.tuberculosis* Asp kinase and templates 2HMF and 3C1M were 35% and 33% respectively. The most significant step in homology modeling process is to obtain the correct sequence alignment of the target sequence with the homologues. The sequence alignment performed using the ClustalW 1.83 for homology modeling is shown in Fig. 2a, and it reveals that the residues involved in binding of various feedback inhibitors in templates (Ser7, Gly9, Gly10, Ser41, Thr47, Glu74, Phe135, Arg151, Gly152, Gly153, Ser154, Thr174, Asp175, Val176, Gly178, Asp183, Pro184, Arg185, Ala190, and Val211) were conserved in *M.tuberculosis* Asp kinase.

Molecular modeling

The coordinates of the crystal structures of *M.jannaschii* Asp kinase, (PDB ID: 2HMF, 2.70 Å resolution) [15] and crystal structure of *M.Jannaschii* Asp kinase (PDB ID: 3C1M, 2.30 Å resolution) [16] were used as templates to build the structure of *M.tuberculosis* Asp kinase. The 3D models of the *M.tuberculosis* Asp kinase was built by Modeller 9v3. 200 models were generated and the model showing the least RMSD with respect to trace (Cα atoms) of the crystal structure of the template was saved for further refinement and validation (Fig. 2b). Further, refinement was performed in order to obtain the best conformation of the developed model of *M.tuberculosis* Asp kinase.

Analysis of the MD simulation

MD simulation of the *M.tuberculosis* Asp kinase was carried in order to check the stability of the structural model proposed for Asp kinase. Analysis of 8 ns dynamics shows that the *M.tuberculosis* Asp kinase structure is stable, after a rapid increase during the first 500 ps the protein backbone RMSD average and standard deviation over the last 3 ns of the Asp kinase trajectory was 2.7±0.2 Å. A plateau of RMSD for the system was achieved within 3 ns of unrestrained simulation, suggesting that 8 ns unrestrained simulation was sufficient for stabilizing fully relaxed models. Figure 3a shows the evolution of the RMSD during the dynamics. The superposition of the average structure of the *M.tuberculosis* asp kinase with the initial model Fig. 3b does not show major structure conformational changes in comparison to the initial model, which is consistent with the relatively low RMSD values. Proteins unbound as previously demonstrated in

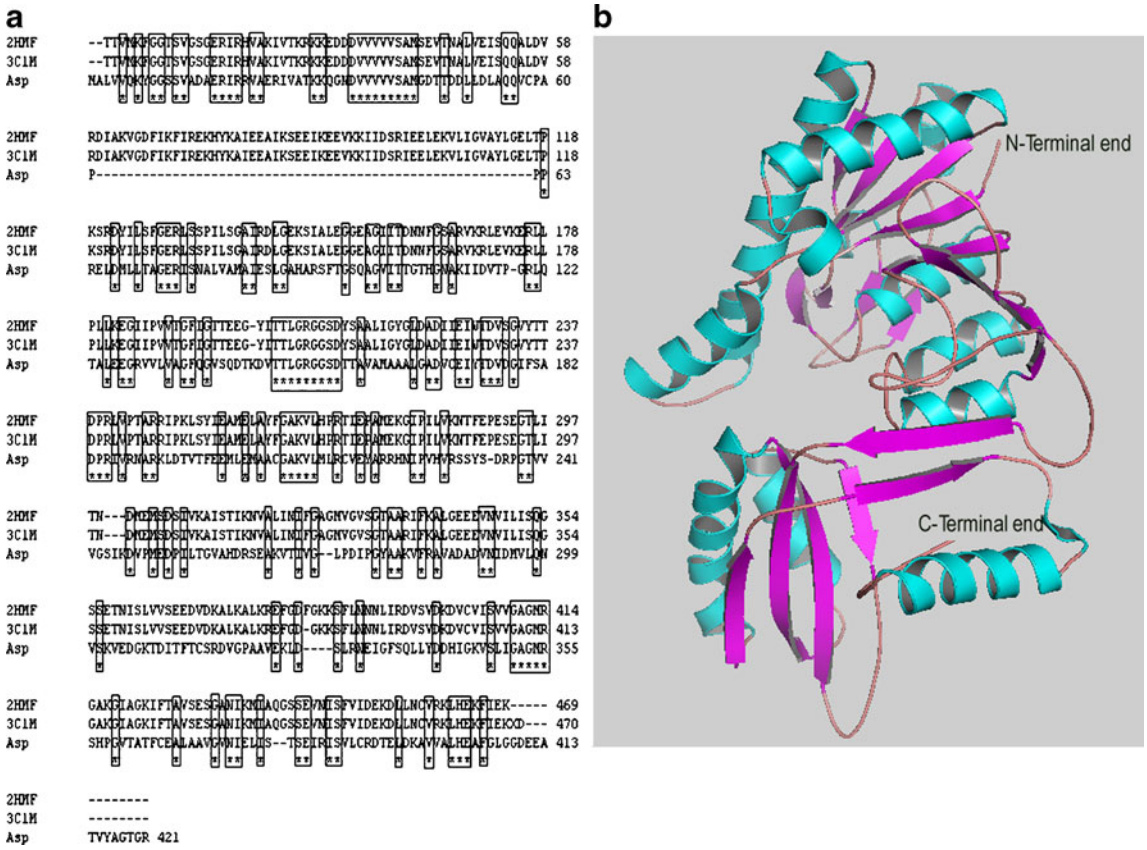

Fig. 2 (**a**) Sequence alignment of *M.tuberculosis* Asp Kinase with the crystallized 2HMF and 3C1M from M. Jannaschii. Highly conserved residues are represented in rectangular boxes. (**b**) The final 3D-structure of *M.tuberculosis* Asp kinase. The α-helix is represented by cyan, the β-sheet is represented by pink and turns are represented by wheat

other MD simulations [41, 42] present a higher RMSD value if compared with the same protein complexed. The potential energy curve shows a smooth decrease until 4 ns, after which it was found fluctuating around a constant value of −5,51,500 kJ mol^{-1} (Fig. 4a). Berendsen thermostat have limitations of maintaining velocity rescaling approaches to ensure that the average kinetic energy of the system corresponding to the expected value at the desired temperature. In these MD simulations use of Berendsen thermostat have not been adversely affected by these limitations.

The flexibilities of the proteins were assessed by the RMSF values from MD of the trajectory which reflected the flexibility of each atom residue in a molecule Fig. 4b. The major backbone fluctuation was seen to occur in the loop regions, whereas regions with low RMSF correspond exclusively to the rigid *beta-alpha-beta* fold. In a typical RMSF pattern, a low RMSF value indicates the well-structured regions while the high values indicate the loosely structured loop regions or domains terminal [43]. In addition, analysis of the structure during the dynamics simulation indicates that the regions (Fig. 4b) L1 (turn composed by Gly9-Ser12), L2 (loop Gly150-Gly153), L3 (turn Gly166-Asp168), L4 (loop Asn224-Pro226), L5 (turn Glu251-Ile254), L6 (turn Ile275-Gly277), L7 (loop Val303-Gly306), L8 (big turn His357-Gly359) and L9 (big loop composed by Ala417-Gly420) present higher RMSF values, which strongly shows that these regions are the most flexible in the predicted structure of *M.tuberculosis* Asp kinase.

Validation of homology model

The first validation was carried out using Ramachandran plot calculations computed with PROCHECK program by checking the detailed residue-by-residue stereo-chemical quality of a protein structure [28]. The φ and ψ distributions of the Ramachandran's plot of non-glycine, non-proline residues are summarized in Table 1 and Fig. 5a. Altogether, 100% of the residues in homology model were in favored and allowed regions. In comparison with the templates, the homology model had a similar Ramachandran plot with 0.0 % residues in disallowed regions. ERRAT is a so-called "overall quality factor" for non-bonded atomic interactions, and higher scores mean higher quality [29]. The normally accepted range is >50 for a high

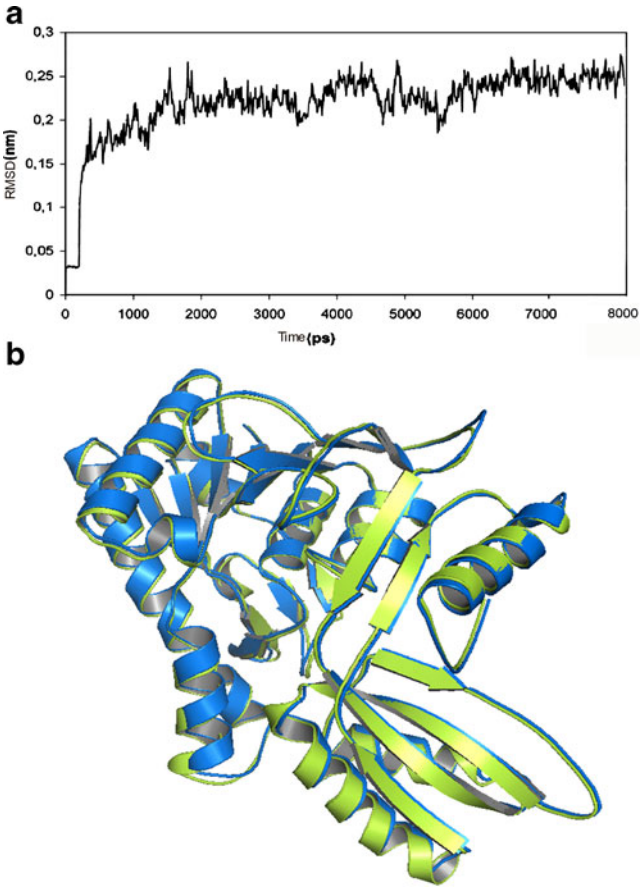

Fig. 3 (**a**) The graphical representation of root-mean-square deviation (RMSD) of all Cα from root structure of *M.tuberculosis* Asp kinase as a function of time. (**b**) Superposition of the average structure during the MD simulation with the initial minimized structure of *M. tuberculosis* Asp kinase. The structures are presented as cartoon diagram. The average structure is colored in limon; the initial structure is colored in marine

quality model [29]. In the current case, the ERRAT score for the Asp kinase model is 88.704, well within the range of a high quality model, in the mean time the ERRAT score for the templates 2HMF is 96.388 and 3C1M is 96.703 (Table 1). Thus, the above analysis suggests that the backbone conformation and non-bonded interactions of Asp kinase homology model is all reasonable within a normal range.

WHAT-IF is used to check the normality of the local environment of amino acids [30]. For the WHAT-IF evaluation, the quality of the distribution of atom types is determined around amino fragments. For a reliable structure, the WHAT-IF packing scores should be above −5.0 [30]. In this case, none of the scores for each residue in the homology model is lower than −5.0 as depicted in Fig. 5b. Therefore, the WHAT-IF evaluation also shows that the homology mode structure is very reasonable. The interaction energy per residue was also calculated by the PROSA2003 program [31]. In this analysis, the interaction energy of each residue with the remainder of a protein is computed to judge whether or not it fulfills certain energy criteria. The PROSA Z-Score indicates overall model quality. For the *M.tuberculosis* Asp kinase homology model, it gains Z-Score of −10.3, when compared with Z-Scores of −12.17 for 2HMF and −11.8 for 3C1M templates. Figure 5c displays the PROSA2003 energy profiles calculated for the Asp kinase homology model along with the templates. The energy profile of the *M.tuberculosis* Asp kinase homology model is consistent with a reliable conformation based on its similarity to that of the templates 2HMF and 3C1M. The final evaluation of the built *M. tuberculosis* Asp kinase structure was checked by Verify 3D [32, 33]. Figure 5d represents the Verify 3D graph of the predicted *M.tuberculosis* Asp kinase model. It is to be noted that compatibility scores above zero correspond to acceptable side chain environment. From Fig. 5d, we can see that almost all residues are reasonable, but only a few regions are variable (Pro59-pro61, Leu66 and Gly73-Ser77) and are built poorly.

In order to investigate the organization of various domains in the developed model of *M.tuberculosis* Asp kinase, it was subjected to Scansite server. It was reported that Asp kinase has regained the ACT domains. ACT domain is a structural motif in proteins of 70–80 amino

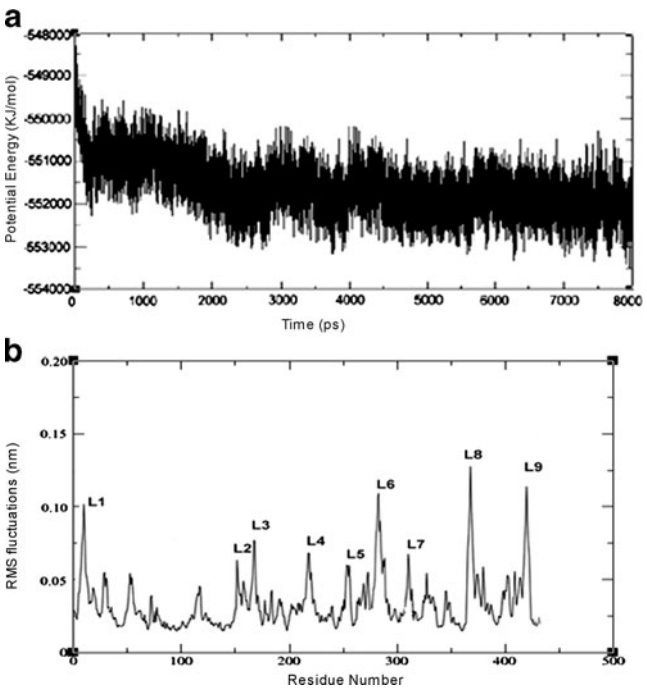

Fig. 4 (**a**) The potential energy curve of the system during the MDS. Every 20th observation point has been plotted on the graph. (**b**) Graphical representation of root-mean-square fluctuations (RMSF) of all the Cα from starting structure of models as a function of time. The graphic shows the average RMSF of the last 3 ns of simulation of *M. tuberculosis* Asp kinase

Table 1 Quality of structures checked by PROCHECK and ERRAT

Procheck								ERRAT score
Ramachandran plot quality (%)					Goodness factor			
	Most favored	Additional allowed	Generously allowed	Dis-allowed	Dihedrals	Covalent	Overall	
2HMF	91.2	8.3	0.0	0.5	−0.16	0.54	0.11	96.338
3C1M	87.9	12.1	0.0	0.0	−0.11	0.42	0.10	96.703
Ask kinase	93.5	6.5	0.0	0.0	−0.18	0.25	0.08	88.704

Ramachandran plot qualities show the percentage (%) of residues belonging with the most favored, additionally allowed, generously allowed and disallowed regions of the plot; goodness factors show the quality of covalent and overall bond/angle distances; these scores should be above −0.5 for a reliable model.

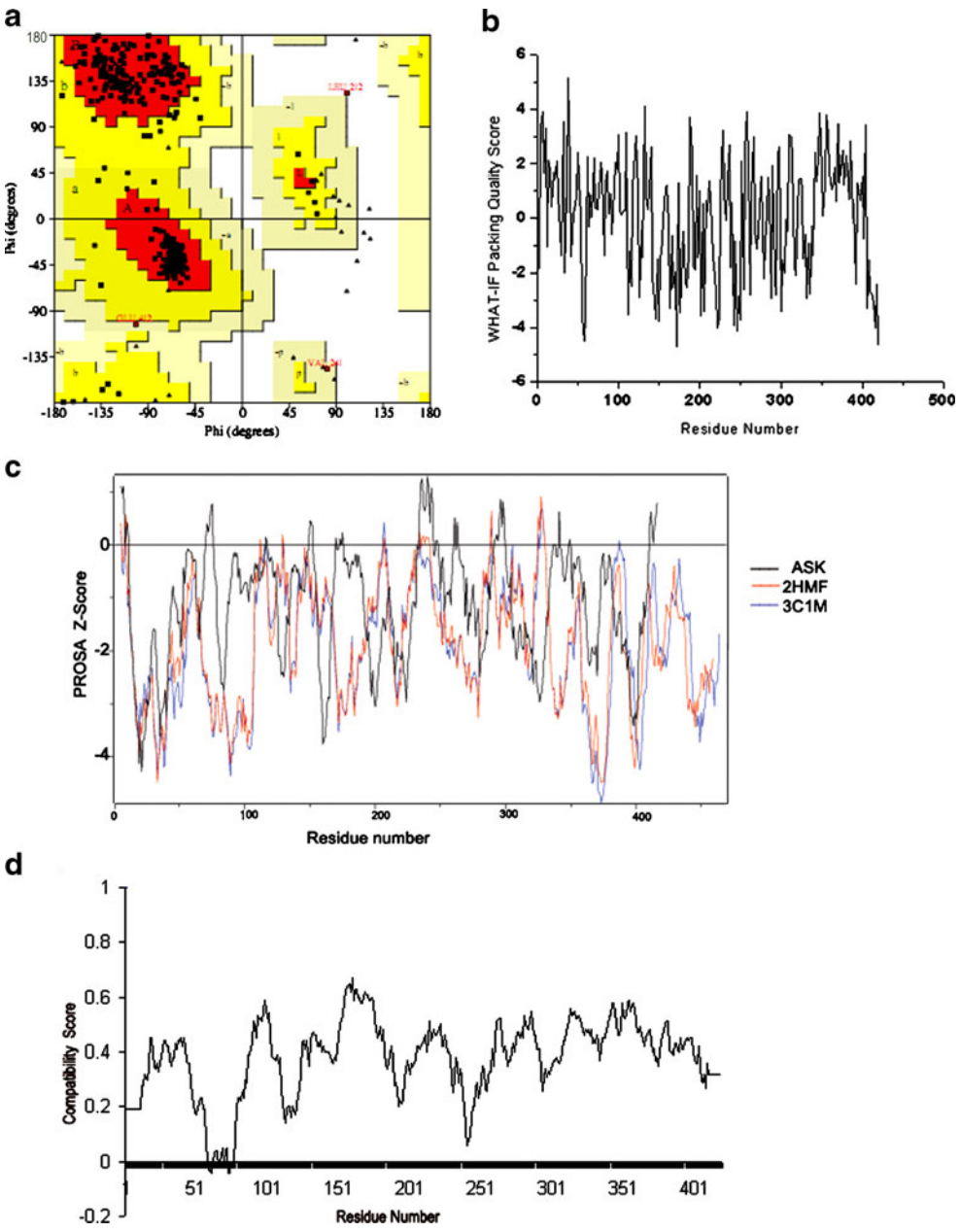

Fig. 5 (**a**) Ramachandran plot for predicted *M.tuberculosis* Asp kinase. (**b**) The WHAT-IF packing quality scores calculated for the homology model of *M.tuberculosis* Asp kinase. The score should be above −5 for a reliable model. (**c**) Prosa 2003 energy profiles calculated for the templates 2HMF (Red), 3C1M (blue) and built Asp kinase from *M.tuberculosis* (black). (**d**) The 3D profiles verified results of predicted *M.tuberculosis* Asp kinase, residues with positive compatibility score are reasonably folded

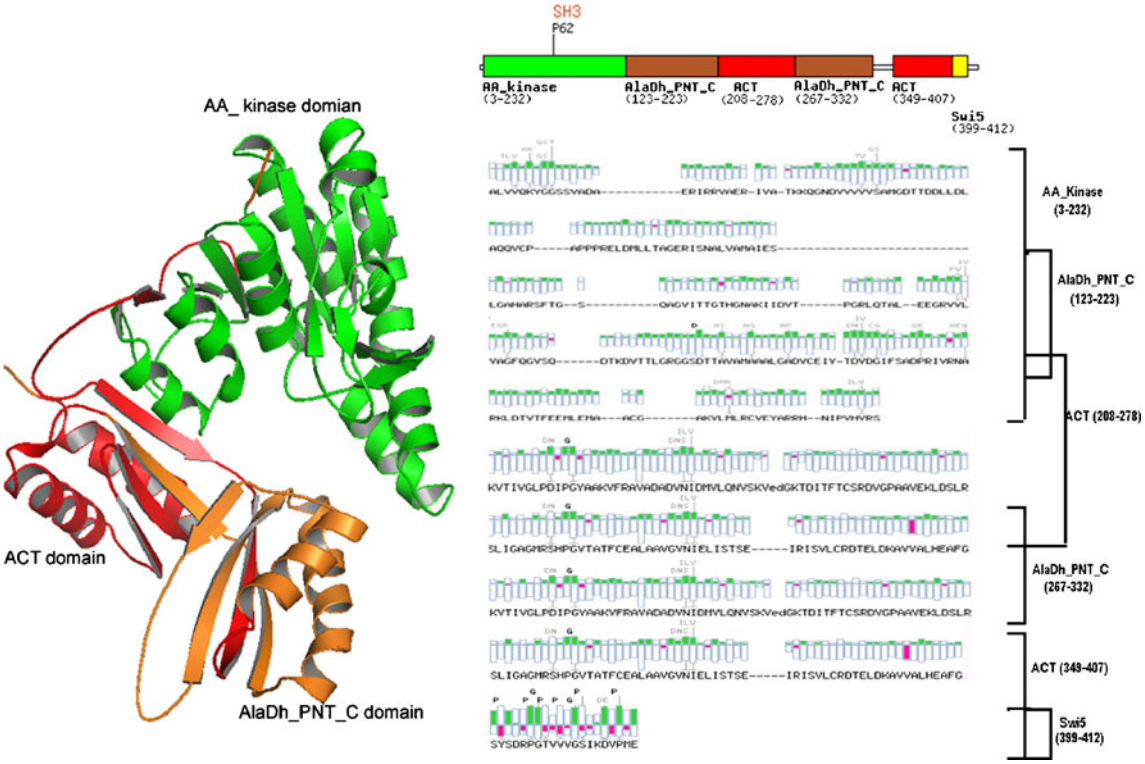

Fig. 6 Organization of various domains in *M.tuberculosis* Asp kinase

acids that are one of a growing number of intracellular small molecule binding domains that function in the control of metabolism, solute transport and signal transduction [44–48]. ACT domain is composed of four β strands and two α helices arranged in a βαββαβ fold. In *M.tuberculosis* Asp kinase two ACT domains are organized; one is from 267–332 and another is from 349–407 towards the N-terminal end (Fig. 6).

In brief, the geometric quality of the backbone conformation, the residue interaction, the residue contact and the energy profile of the structure is all well within the limits established for reliable structures. All evaluations suggest that a reasonable homology model for *M.tuberculosis* Asp kinase has been obtained that can be exposed for examination of protein-substrate and protein inhibitor interactions.

Molecular docking of Aspartic acid into Asp kinase homology model

To further verify our constructed homology structure, we used the molecular docking approach described in Sect. 2.5 to dock the natural substrate molecule (*i.e.*, Aspartic acid) to the modeled *M.tuberculosis* Asp kinase. We obtained almost the same binding mode of Aspartic acid as that observed in the crystal structure of the threonine sensitive Asp kinase from M.jannaschii complexed with Mg-ADP and Aspartate (2HMF) and threonine sensitive Asp kinase from M.jannaschii with MgAMP-PNP and L-aspartate (PDB ID: 3C1M). The docking conformation and corresponding complex analysis of aspartic acid is depicted in Fig. 7a. Aspartic acid is trapped in a cavity, in which both the carboxylic groups and amino group of aspartic acid are involved in hydrogen bonding interactions with Arg151, Lys210 and Ser263. The binding energy, the intermolecular energy, cluster rank and cluster number for docking of Aspartic acid are represented in Table 2. The computational result is consistent with the result reported by Grant [49] which showed that the proteins containing ACT domain as a structural motif reside in a growing number of different intracellular small molecule binding domains that function in the control of metabolism. Docking of aspartic acid into the homology model of Asp kinase showed that the binding site of aspartic acid is situated in ACT domain.

Molecular docking of ADP into Asp kinase homology model

Figure 7b represents the docking conformation and corresponding complex analysis of ADP natural product of Asp kinase. We obtained almost the same binding mode of ADP as that observed in the crystal structure of the threonine sensitive Asp kinase from M.jannaschii com-

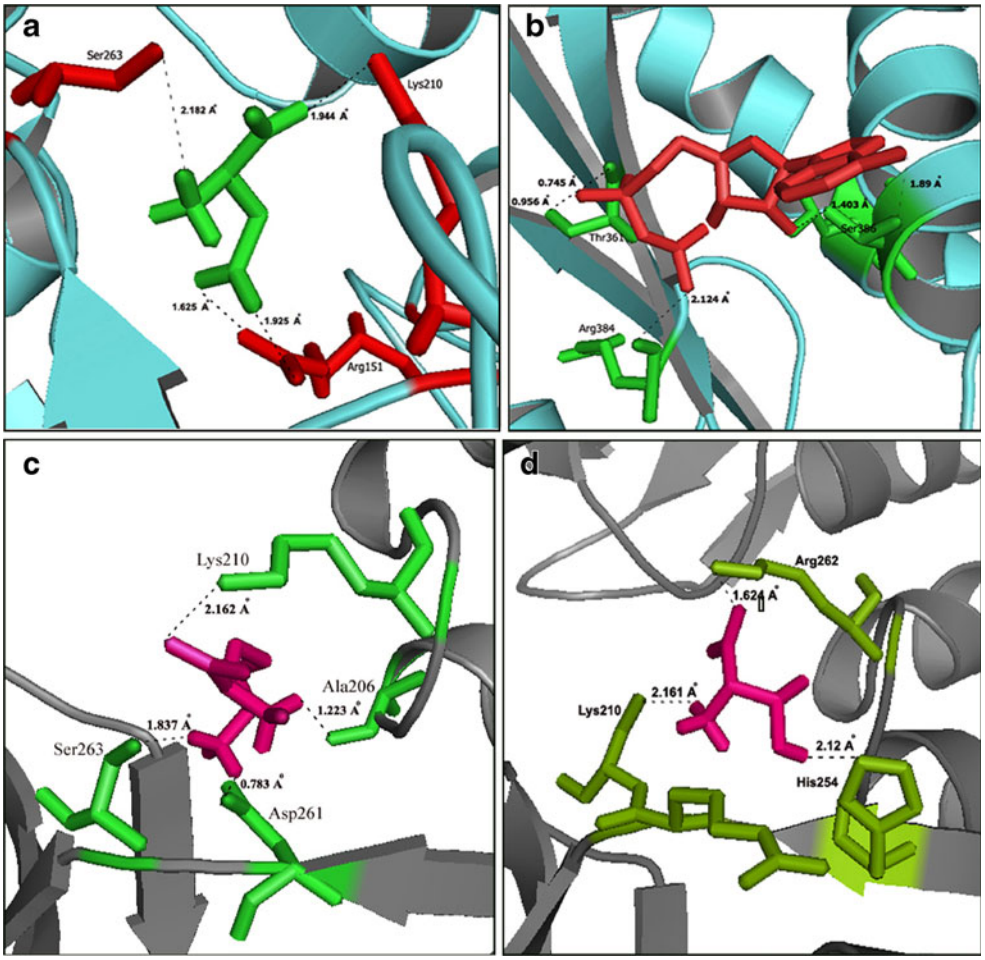

Fig. 7 (**a**) The docking conformation of Aspartic acid in the ACT domain of *M.tuberculosis* Asp kinase homology model. (**b**) The docking conformation of ADP (Adenosine diphosphate) in ACT domain of *M.tuberculosis* Asp kinase homology model. (**c**) The docking conformation of lysine in the ACT domain of *M.tuberculosis* Asp kinase homology model. (**d**) The docking conformation of Threonine in the ACT domain of *M.tuberculosis* Asp kinase homology model. Built model of Asp kinase is represented in cartoon. Ligands and the residues interacting with ligands are represented by sticks

plexed with Mg-ADP and Aspartate (2HMF). The phosphate groups of ADP are involved in hydrogen bonding interactions with Thr361 and Arg384. The nitrogen atom in the purine ring is engaged in the hydrogen bond interaction with Ser386. The docking results of ADP are represented in Table 2. The docking analysis of ADP into predicted model of *M.tuberculosis* Asp kinase is consistent with the result reported by Grant [49], which showed that the proteins containing ACT domain as a structural motif resides a growing number of different intracellular small molecule binding domains that function in the control of metabolism. Docking of ADP into the homology model of Asp kinase showed that the binding site of ADP is situated in ACT domain.

Table 2 Summary of docking results of ligands onto *M.tuberculosis* Asp kinase

Docked molecule	Cluster rank	Cluster number	Binding energy (kcal/mol)	Final intermolecular energy (kcal/mol)
Aspartic acid	1	49	−8.59	−14.59
ADP	1	21	−10.86	−12.21
Threonine	2	17	−6.58	−8.52
Lysine	1	39	−8.12	−9.33

Molecular docking of threonine into the homology of Asp kinase

Investigating feedback inhibitor binding sites may help to explain the action mechanism of feedback inhibitors and other potentiators and thus contributes to drug design for tuberculosis treatment. With this in mind we performed a docking calculation of threonine and lysine by running the AutoDock program on modeled *M.tuberculosis* Asp kinase, as described in the methods section. In case of threonine, the interactions with the *M.tuberculosis* Asp kinase showed (Table 2) very similar binding profile of binding to ACT domain. Threonine is involved in hydrogen bonding interactions with Lys210, His254 and Arg262. Figure 7c represents the predicted conformation and binding mode of threonine with *M.tuberculosis* asp kinase homology model, which was also consistent with the binding mode reported by Grant [49].

Molecular docking of Lysine into the homology of Asp kinase

In case of lysine, another feedback inhibitor of Asp kinase, the docking result showed very similar binding profile of binding to ACT domain where lysine is involved in hydrogen bonding interactions with Ala206, Lys210, Asp261 and Ser263. Table 2 shows the docking results of Lysine onto the *M.tuberculosis* Asp kinase 3D model. Figure 7d represents the predicted conformation and binding mode of lysine with *M.tuberculosis* Asp kinase homology model, which was also consistent with the binding mode reported by Grant [49].

In summary, the docking results of Asp kinase with natural substrates, products and feedback inhibitors around the ACT domains provide strong confidence about the homology model. Therefore, it is obvious that this docked model would provide more detailed information and accuracy in its description of ligand binding with *M. tuberculosis* Asp kinase. We presume that such an application of feedback inhibitor binding can be of great use in the detection of new ligands.

Conclusions

In this paper, we have developed high accurate homology model of *M.tuberculosis* Asp kinase homology model using 2HMF and 3C1M crystal structures. The quality of the homology model depends on the level of sequence identity between the templates of known 3D structures and the protein to be modeled. In this case, the level of sequence identity between the target and the templates (35% &33%) is observed to reach the threshold value of 30% sequence identity for obtaining a homology model. In addition, the molecular dynamics simulation improved the general structure of the generated model. This model has been qualified using several validation methods, including PROCHECK, ERRAT, WHAT-IF, PROSA2003 and VERIFY-3D. All evidences suggest that the geometric quality of the backbone conformation, the residue interaction, the residue contact and the energy profile of the structure is well within the limits established for reliable structures. The docking studies of this homology model with natural substrates, products and feedback inhibitors were found to be consistent with the published data of the key role played by ACT domain to bind to ligands. Furthermore, based on the docking conformations, a pharmacophore model can be built for *M.tuberculosis* Asp kinase. Different from classical homology modeling, the combination of homology modeling, substrate and feedback inhibitor docking will be very useful in helping us understand the binding modes of Asp kinase and its ligands, and avoid obvious pitfalls in our further design. In conclusion, we believe that these results may help in the understanding of the mechanism of action of *M.tuberculosis* Asp kinase. Further it will provide information for those who want to pursue their work in the design of selective Asp kinase blockers for the treatment of tuberculosis.

Acknowledgments This work is supported by the DBT-BIF facility (F.No. BT/BI/25/2001/2006). The authors gratefully acknowledge University Grant Commission (UGC), New Delhi (F.No.33-222/2007 (SR) 13-3-08) for financial support. We are grateful to Dr. Kusuma Kumari, Vice-chancellor, Sri Krishnadevaraya University for encouragement to our research work. We also thank Prof. P. Gautam, Department of Biotechnology, Anna University, Chennai, for rendering critical valuable suggestions on the present work.

References

1. www.tballiance.org
2. Masjedi MR, Farnia P, Sorooch S, Pooramiri MV, Mansoori SD, Zarifi AZ, Akbarvelayati A, Hoffner S (2006) Extensively drug-resistance tuberculosis: 2 years of surveillance. Iran Clin Infect Dis 43:841–847
3. Centres for Diseases Control and Prevention (CDC). Emergence of Mycobacterium tuberculosis with extensive resistance to second-line drugs worldwide, 2000–2004. Atlanta, GA: US Department of Health and Human Services, CDC, 2006. Available at: http://www.cdc.gov/nchstp/tb/worldtbday/2006/activities.htm. Accessed April 2006
4. Robert J, Trystram D, Truffot pernort C, Jarlier V (2003) Multidrug-resistant tuberculosis:eight years of surveillance in France. Eur Respir J 22:833–837
5. Cox HS, Orozco JD, Male R, Ruesch-Gerdes S, Falzon D, Small L, Doschetov D, Kebede Y, Aziz M (2004) Multidrug-resistant tuberculosis in central Asia. Emerg Infect Dis 10:865–872
6. Cohen GC (1983) The common pathway to lysine, methionine, and threonine. In: Davies JE, Hermann KM, Somerville RL (eds) Amino acids:biosynthesis and genetic regulation. Addision-Wesly, Reading, pp 147–171
7. Stadtman ER (1966) Allosteric regulation of enzyme activity. Adv Enzymol 28:1041–1054
8. Anuradha CM, Mulakayala C, Babajan B, Naveen M, Kumar CS (2009) Probing ligand binding modes of Mycobacterium tuberculosis MurC ligase by molecular modeling, dynamic stimulation and docking. J Mol Model. doi:10.1007/s00894-009-0521-2
9. Bohm M, Sturzebecher J, Klebe G (1999) Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. J Med Chem 42:458–477
10. NCBI, http://www.ncbi.nlm.nih.gov/
11. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
13. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 50:3389–3402
14. Chenna R, Sugawana H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequene alignment

with the Clustal series of programs. Nucleic Acids Res 31:3497–3500
15. Faehnle CR, Liu X, Pavlovsky A, Viola RE (2006) The initial step in the archaeal aspartate biosynthetic pathway catalyzed by a monofunctional aspartokinase. Acta Crystallogr Sect F Struct Biol Cryst Commun 62:962–966
16. Liu X, Pavlovsky AG, Viola RE (2008) The structural basis for allosteric inhibition of a threonine-sensitive aspartokinase. J Biol Chem 283:16216–16225
17. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815
18. Sali A, Overington JP (1994) Derivation of rules for comparative protein modeling from a database of protein structure alignments. Protein Sci 3:1582–1596
19. Sali A (1995) Modeling mutations and homologous proteins. Curr Opin Biotechnol 6:437–451
20. Sali AL, Potterton F, Yuan H, van Vlijmen M, Karplus (1995) Evaluation of comparative protein modeling by MODELLER. Proteins 23:318–326
21. Maiti R, van Domselaar GH, Zhang H, Wishart DS (2004) SuperPose: A simple server for sophisticated structural superposition. Nucleic Acids Res 32:W 590–W 594
22. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) GROMACS: fast, flexible and free. J Comput Chem 26:1701–1718
23. Berendsen HJC, Grigera JR, Straatsma TP (1987) The missing term in effective pair potentials. J Phys Chem 91:6269–6271
24. Chowdhuri S, Tan ML, Ichiye TJ (2006) Dynamical properties of the soft sticky dipole–quadrupole–octupole water model:a molecular dynamics study. J Chem Phys 125:14451–14453
25. Arfken G (1985) The Method of Steepest Descents.n§7.4. In: Mathematical Methods for Physicists, 3rd ed. Orlando, FL Academic Press, pp 428–436
26. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18:1463–1472
27. Kleywegt GJ (2000) Validation of protein crystal structures. Acta Crystallogr D Biol Crystallogr 56:249–265
28. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486
29. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of non-bonded atomic interactions. Protein Sci 2:1511–1519
30. Vriend G, Sander C (1993) Quality-control of protein models-directional atomic contact analysis. J Appl Cryst 26:47–60
31. Tomii K, Hirokawa T, Motono C (2005) Protein structure prediction using a variety of profile libraries and 3D verification. Proteins 61:114–121
32. Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253:164–170
33. Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356:83–85
34. http:// scansite.mit.edu
35. Buntrock RE (2002) ChemOffice Ultra 7.0. J Chem Inf Comput Sci 42:1505–1506
36. HyperChem, Release 7.5 for Windows, Molecular Modeling System. Hypercube Inc and Autodesk Inc. [http://www.hyper.com/]
37. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19:1639–1662
38. Huey R, Morris GM, Olson AJ, Goodsell DS (2007) A semiempirical free energy force field with charge-based desolvation. J Comput Chem 28:1145–1152
39. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins J Am Chem Soc 106:765–784
40. DeLano WL (2006) The PyMOL Molecular Graphics System. DeLano Scientific, SanCarlos, CA, USA. http://www.pymol.org
41. Caceres RA, Timmers LFS, Vivan AL, Schneider CZ, Basso LA, de Azevedo WF Jr (2008) Molecular modeling and dynamics studies of /ytidylate kinase from Mycobacterium tuberculosis H37Rv. J Mol Model 14:427–434
42. Caceres RA, Timmers LFS, Dias R, Basso LA, Santos DS, De Azevedo WF Jr (2008) Molecular modeling and dynamics simulations of PNP from Streptococcus agalactiae. Bioorg Med Chem 16:4984–4993
43. Alexander CS, Yan X, Pei T (2005) Homology modeling and molecular dynamics simulations of transmembrane domain structure of human neuronal nicotinic acetylcholine receptor. J Biophys 88:1009–1017
44. Anantharaman V, Konnin EV, Aravind L (2001) Regulatory potential, phylogenic distribution and evolution of ancient, intracellular small-molecule-binding domains. J Mol Biol 307:1271–1292
45. Aravind L, Koonin EV (1999) Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. J Mol Biol 287:1023–1040
46. Ettema TJG, Brinkman AB, Tani TH, Rafferty JB, van der Oost J (2002) A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. J Biol Chem 37:37464–37468
47. Chipman DM, Shaanan B (2001) The ACT domain family. Curr Opin Struct Biol 11:694–700
48. Liberles JS, Thórólfsson M, Martínez A (2005) Allosteric mechanisms in ACT domain containing enzymes involved in amino acid metabolism. Amino Acids 28:1–12
49. Grant GA (2006) The ACT domain: a small molecule binding domain and its role as a common regulatory element. J Biol Chem 281:33825–33829