

# Insight to pyrazinamide resistance in *Mycobacterium tuberculosis* by molecular docking

Amirudeen Nusrath Unissa<sup>1\*</sup>, Nagamiah Selvakumar<sup>2</sup>, Sameer Hassan<sup>3</sup>

<sup>1</sup>Department of Mycobacteriology, Tuberculosis Research Centre (TRC), Indian Council of Medical Research (ICMR), Mayor V. R. Ramanathan Road, Chetput, Chennai 600 031, Tamil Nadu, India; <sup>2</sup>Department of Mycobacteriology, TRC, ICMR; <sup>3</sup>Department of Biomedical Informatics Centre, Indian Council of Medical Research (ICMR), Mayor V. R. Ramanathan Road, Chetput, Chennai 600 031, Tamil Nadu, India; A. Nusrath Unissa - E-mail: nusrathunissa@gmail.com

Received April 09, 2009; Revised June 10, 2009; Accepted June 13, 2009; Published August 18, 2009

## Abstract:

Pyrazinamide (PZA) - an important drug in the anti-tuberculosis therapy, activated by an enzyme Pyrazinamidase (PZase). The basis of PZA resistance in *Mycobacterium tuberculosis* was owing to mutation in *pncA* gene coding for PZase. Homology modeling of PZase was performed using software Discovery Studio (DS) 2.0 based on the crystal structure of the PZase from *Pyrococcus horikoshii* (PDB code 1im5), in this study. The model comprises of one sheet with six parallel strands and seven helices with the amino acids Asp8, Asp49, Trp68, Lys96, Ala134, Thr135 and Cys138 at the active site. Five mutants were generated with Gly at position 8, Thr at position 96, Arg at position 104, Tyr and Ser at position 138. The Wild-type (WT) and five mutant models were docked with PZA. The results indicate that the mutants Lys96Thr, Ser104Arg Asp8Gly and Cys138Tyr may contribute to higher level drug resistance than Cys138Ser. These models provide the first *in-silico* evidence for the binding interaction of PZA with PZase and form the basis for rationalization of PZA resistance in naturally occurring *pncA* mutant strains of *M. tuberculosis*.

**Key Words:** *Mycobacterium tuberculosis*, PZA resistance, PZase, mutants, docking.

## Background:

Despite being a controllable, preventable and curable disease, Tuberculosis (TB) still remains as a major public health problem in many parts of the world. The recent increase in multi-drug resistant (MDR-TB) and extensively drug resistant (XDR-TB) has further worsened the situation. Drug resistance in TB is essentially a potential threat to the TB control programmes. Genetic and molecular analysis of drug resistance in *M. tuberculosis* suggests that the bacilli usually acquire resistance either by alteration of the drug target through mutation or by titration of the drug through overproduction of the target. PZA is an important sterilizing drug and a principle component in the current six-month short course TB-chemotherapy. PZA plays a unique role in shortening the therapy from a period of 9 to 12 months down to 6 months, because PZA kills a population of semi dormant tubercle bacilli, residing in an acidic environment, which cannot be killed by other TB drugs [1].

In *M. tuberculosis*, the susceptibility to PZA correlates with the presence of a single enzyme with nicotinamidase and pyrazinamidase activities. Strains of *M. tuberculosis* that are resistant to PZA are often defective in PZase activity [2]. PZA-resistant (PZA<sup>r</sup>) *M. tuberculosis* clinical isolates are usually defective for PZase activity, and there is very good correlation between PZA resistance and loss of this enzyme. Scorpio and Zhang in 1996 had identified the PZase gene (*pncA*) from *M. tuberculosis* and had shown that *pncA* mutations are a major mechanism of PZA resistance [3]. The identified *pncA* mutations are largely missense mutations causing amino acid substitutions, and in some cases nucleotide insertions or deletions and nonsense mutations in the *pncA* structural gene or in the putative promoter region of *pncA* [4]. The uniqueness in the mutations of *pncA* gene is its diversity and scattering along the whole gene though there does appear to be some degree of clustering at three regions of PncA protein (3 to 17, 61 to 85, and 132 to 142). These regions are likely to contain catalytic sites for the PZase enzyme [5]. PZA as a prodrug needs to be activated by the bacterial nicotinamidase-PZase into pyrazonic acid (POA), the active moiety of the drug [3].

A study [6] depicting the structure-function relationships of PncA protein (PZase) has clearly shown that the decrease in PZase activity observed in the mutant proteins correlates well with the structural

modifications. However, the report also emphasizes the need for further structural studies to validate the positioning of the catalytically important amino acid residues (Asp8, Asp49, Lys96, Trp68, Ala134, Thr135 and Cys138). Although the report has provided both kinetic and structural data related to nine different mutants of PZase in connection to Cys138 a putative active site residue, but it did not provide any information regarding the mutants of Cys138. In another recent report [7] comprehensive enzymatic characterization of PZase was carried out with the generation of nine different mutants. It suggested that the Asp8, Lys96 and Cys138 were key residues for catalysis and Asp49, His51, His57 and His71 were essential for metal ion binding, this report too demanded the need to promote structural studies. In the light of which, we therefore have made an attempt to develop a three dimensional (3D) model of PZase along with the generation of five mutant models with Cys138Ser, Cys138Tyr, Asp8Gly, Lys96Thr and Ser104Arg and in the process, to explore its interactions with PZA. Out of the five, two (Cys138Ser and Cys138Tyr) were created in our study whereas the other three mutants Asp8Gly, Lys96Thr and Ser104Arg were created on the basis of prior kinetic studies. In addition, a basic analysis was performed to determine the correlation between kinetic and the docking data based upon the above-mentioned studies.

## Methodology:

### Homology modeling of PZase:

Homology modeling helps in predicting the 3D structure of a macromolecule with unknown structure (target) by comparing it with a known template from another, structurally highly similar, macromolecule. The sequence of PZase from *M. tuberculosis* was searched against PDB (Protein databank) using PSI-BLAST (Basic Local Alignment Search Tool) [8] program to identify the related protein structure to be used as a template. The X-ray structure of PZase (PDB code, 1im5) from *Pyrococcus horikoshii* (*P. horikoshii*) was chosen as the template for the modeling PZase of *M. tuberculosis* (SWISSPROT accession no. Q5UFH2). Sequence and structural alignment of 1im5 and 2hor with PZase sequence was performed using ClustalW [9] and MSDFOLD. Aligning was carried out carefully without insertion of any gaps in the conserved secondary structural regions. From the alignment, spatial restraints were derived and used in the 3D model construction with MODELLER (Discovery Studio;

Accelrys). The stereochemical quality of the model structure was analyzed with RAMPAGE program using Ramachandran plot [10]. This model was further used for the identification of active site and for docking with PZA.

Modeling of mutant PZase was similar to the WT showing 99% similarity owing to change in single amino acid. Sequence alignment of WT with the five mutant sequences was performed using ClustalW (see supplementary materials available with authors).

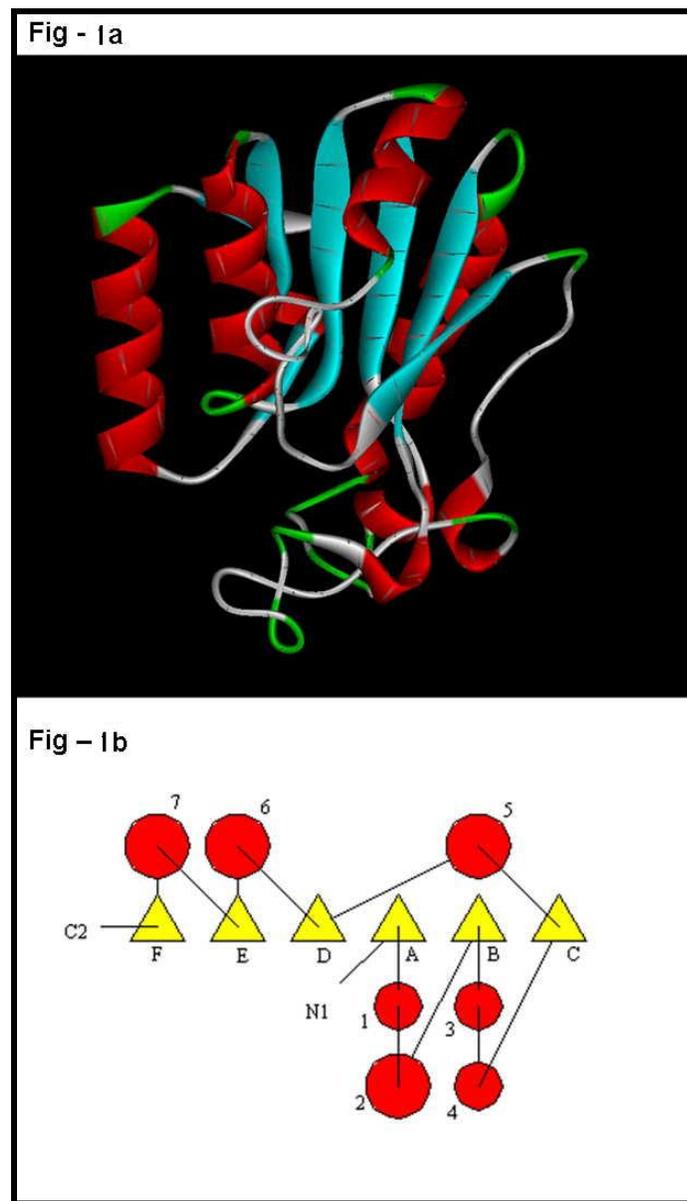
**Active site analysis of PZase enzyme:**

The active site of PZase from *M. tuberculosis* was analyzed using DS package. To identify a binding site, the receptor is first mapped to a grid. Grid points within a given distance of a receptor atom are marked as occupied by the receptor, and thus designated as undesirable locations for ligand atoms. The remaining unoccupied grid points are

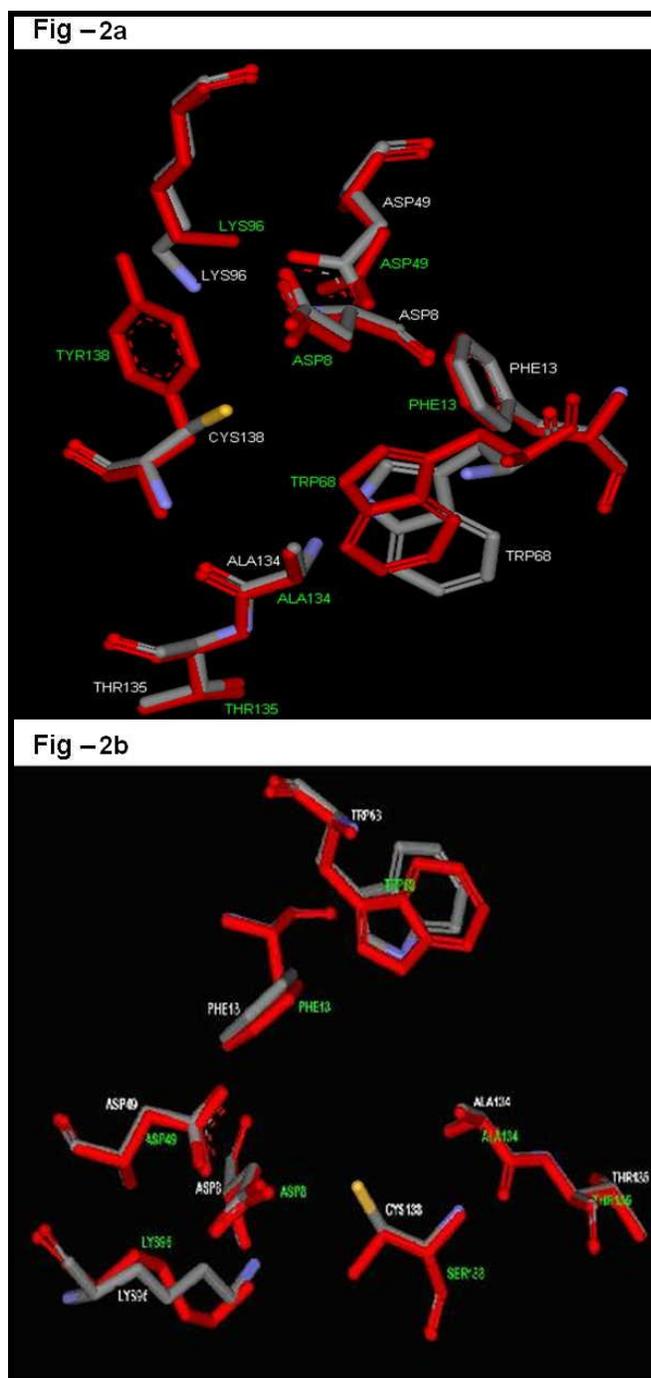
used to define the binding site (See supplementary material available with the authors for procedure to analyze active site) [11].

**Docking of PZA-PZase:**

Docking was performed using CDOCKER. The CDOCKER protocol is an implementation of the CDOCKER algorithm [12] in the DS environment. It allows running a refinement docking of any number of ligands with a single protein receptor. CDOCKER is a grid-based molecular docking method that employs CHARMM. The ligand was added with H (Hydrogen) atoms using DS. The models were energy minimized with CHARMM force field and metal ion Zn was added before performing docking. The receptor is held rigid while the ligands are allowed to flex during the refinement. Docking was performed using the default settings. The default speed selection was used to avoid a potential reduction in docking accuracy.



**Figure 1:** (a) Three-dimensional model of PZase. (b) A TOPS cartoon of PZase model. Strands are shown in yellow triangles and helices in red circles. The direction of elements can be deduced from the connecting lines. 'Up' strands are indicated by upward pointing triangles.



**Figure 2:** (a) Superimposition of binding residues of WT (Cys138) and mutant (Tyr138) model. (b) Superimposition of binding residues of WT (Cys138) and mutant (Ser138) model.

**Discussion:**

**Homology modeling of PZase:**

In the results of BLAST search against PDB, 1im5 and 2hor were identified as template and reference proteins respectively. 3D-PSSM algorithm was used for fold recognition analyses by threading and provided similar results to that of BLAST. PZase of *P. horikoshii* (1im5), a member of  $\alpha/\beta$  isochorismatase-like hydrolases family, having high level sequence identity of 37%, was chosen as a template

for modeling PZase of *M. tuberculosis*. Previously, a study [6] provided structural details about the of PZase on the basis of model obtained from *Arthrobacter CSHase* as template which had sequence homology of about 26%, whereas in the present study the chosen template showed 37% identity; indicating a greater degree of reliability. In general, 30% sequence homology is required for generating useful models.

Secondary structure of PZase was predicted using JPRED. Amino acid sequence of the target was aligned to the template sequence based on the secondary structure information for building an accurate model. Structurally conserved regions (SCRs) for the target sequence and the templates were determined by superimposition of two structures and multiple sequence alignment. Coordinates from the reference protein (1im5) to the SCRs, structurally variable regions (SVRs), N-termini and C-termini were assigned to the target sequence based on the satisfaction of spatial restraints. The initial model was thus generated with the above procedure. On the basis of PFAM analysis it was found that the model contains isochromatase domain (residues 2-185) consisting of seven  $\alpha$ -helices and six strands of a parallel  $\beta$ -sheet, which is similar to those in alpha and beta (a/b) proteins. The 3D structure of PZase is as shown in Figure 1a. The connectivity between the secondary structures is represented in TOPS diagram (Figure 1b).

#### Validation of PZase:

The models were validated using Ramachandran plot. For the WT, a total of 94.5 % of the residues were in favored region. The number of favorable region residues for the mutants viz., Asp8Gly, Lys96Thr, Ser104Arg, Cys138Tyr and Cys138Ser are 95.1, 96.7, 96.2, 93.4, and 94.5% respectively (see supplementary material for Figures).

#### Superimposition of 1IM5 with PZase enzyme:

The 3D structures of PZase (final model) was superimposed with 1im5 and 2hor using MSDFOLD (refer supplementary material for figures). Based on the above analysis, it was evident that the secondary structures were highly conserved among the three proteins but 2hor showed 25 residues (forming two antiparallel strands) insertion between  $\beta$ B and  $\alpha$ 3. The weighted root mean square deviation (RMSD) for 2hor was 1.4 Å with PZase. The RMSD of C $\alpha$  trace between 1im5 and PZase was 0.5 Å with a Z score of 7.0 supporting that 1im5 is the right template for modeling PZase protein.

The mutant models obtained was also validated using the Combinatorial Extension (CE) method. The structural alignment was performed for the WT and mutant models and was indicative of good structural alignment (99.5%) due to the change in single amino acid. The 3D structures of WT and mutants were also superimposed. The RMSD score was calculated using the sequence identity and gaps in the alignment. The RMSD value of 0.1 Å and Z-score of 7.3 was shown for the mutants Asp8Gly, Lys96Thr and Ser104Arg, and for the mutants Cys138Tyr and Cys138Ser the values were of 0.6 Å and 1.0 Å respectively (see supplementary materials available with the authors).

#### Active site analysis of PZase:

The final model was analyzed for identifying the possible binding sites of PZase using binding site tool in DS and Qsite finder tool [13]. Six binding sites were identified using DS. Out of 6 sites, site one was considered as active site consisting of Asp8, Phe13, Asp49, Trp68, Lys96, Ser104, Ala134, Thr135 and Cys138, which was found to be in concordance with the earlier reports [6, 7]. Of these, Asp8 occur within the loop region connecting  $\beta$ A and  $\alpha$ 1, Phe13 within  $\alpha$ 1, Asp49 within  $\beta$ B, Trp68 within 20 residues long loop connecting  $\alpha$ 3 and  $\alpha$ 4, Lys96 emerge from  $\beta$ C, Ala134 and Thr135 within loop region connecting  $\beta$ B and  $\alpha$ 6. Ser104 emerge from the loop connecting  $\beta$ C and  $\alpha$ 5. The putative active site residue Cys138, which is highly conserved, is located within  $\alpha$ 6 region of the model. Hence, site one was chosen as the most favorable binding site for exploring the interaction with drug molecule.

#### Mutated models:

In this study, PZase was modeled using the DS 2.0 and the models were constructed based on the sequence of *pncA* gene for WT and five mutants Asp8Gly, Lys96Thr, Ser104Arg, Cys138Tyr and Cys138Ser (refer supplementary materials available with the authors). No significant conformational change was observed in the structure of the mutant models Lys96Thr and Ser104Arg whereas in the mutant

Asp8Gly changes were detected at Lys96 (a key active site residue). Mutated model with Cys138Tyr substitution imparted certain conformational changes at the active site with deviation in the side chain of residues viz., Asp49, Trp68 and Lys96. Structural evaluation with RAMPAGE showed a decrease in number of residue in the most favored region (93.4%). Model structure with Cys138Ser substitution also produced some degree of conformational changes at the active site. Structural evaluation showed 94.5% residues in the most favored region using RAMPAGE using (see supplementary material). Superimposition of WT with mutant (Asp8Gly, Lys96Thr and Ser104Arg) models using CE implies that there was a little degree of deviation from the WT (on the basis of RMSD score of 0.1 Å); hence the change in single amino acid did not produce any significant change in the overall conformation of the protein. In case of mutants such as Cys138Tyr and Cys138Ser (0.6 Å and 1.0 Å) the deviation was more pronounced affecting the conformation drastically, this was quite apparent in Cys138Ser model (refer supplementary material available with authors). Superimposition of binding residues of WT (Cys138) with mutant (Tyr138) and (Ser138) models are shown in Figure 2a and Figure 2b respectively.

#### Docking of PZA-PZase:

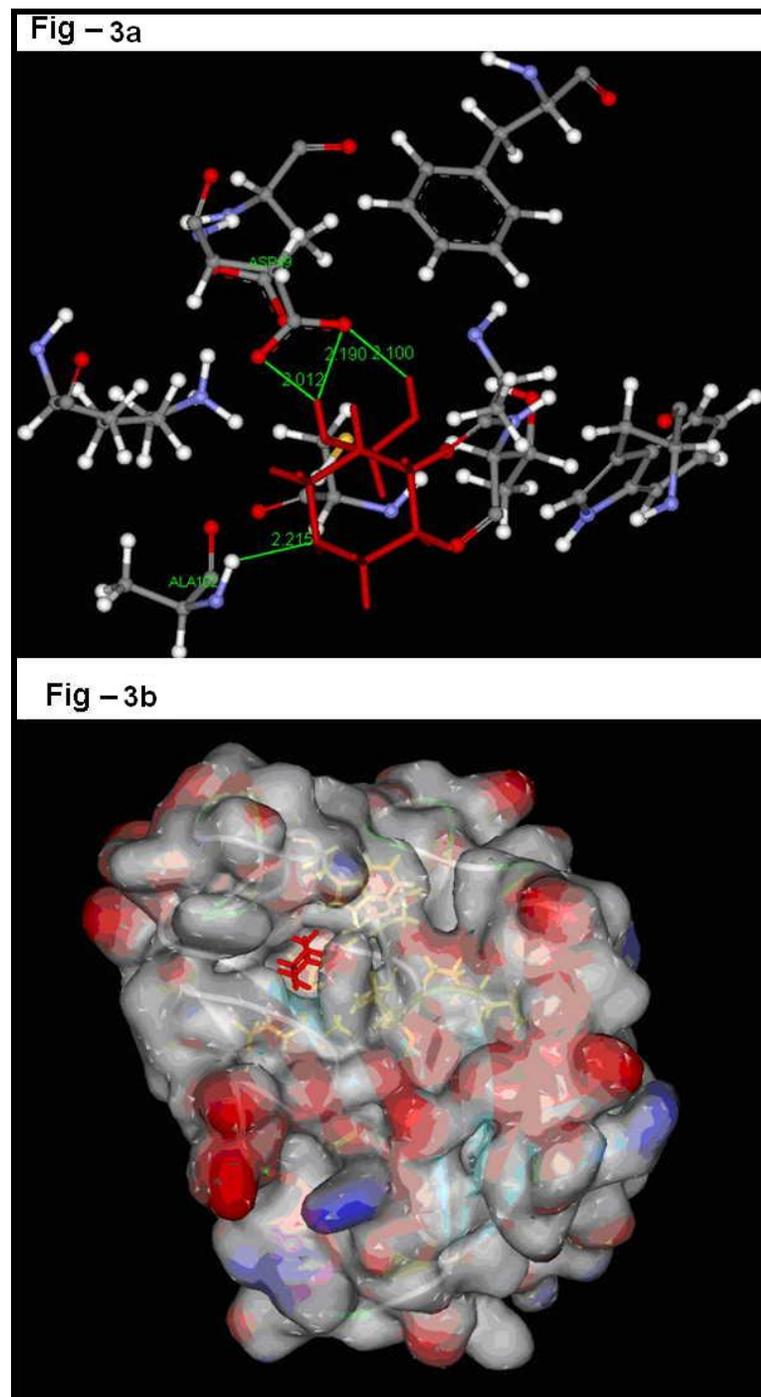
All the six models were flexibly docked with PZA using CDOCKER. Of the ten poses produced, the best ligand pose was selected based on CDOCKER top score and the target structure was selected for further analysis. The ligand poses were analyzed and a heat map was produced to count H bonds made by the poses to the receptor molecule and count close contacts (van der Waals clashes) between the poses and receptor molecule. No van der Waals clashes were seen between the PZA and model structure (data not shown). Docking results indicate that conserved amino acid residue (Cys138) in PZase play an important role in maintaining a functional conformation and directly involved in drug donor binding.

Interaction of the drug molecule with mutated models was also studied on the basis of H bonding. As it is well known, H bonds play an important role for the structure and function of biological molecules, especially for the enzyme catalysis. The H bonds present in the enzyme-drug complex along with their distances are shown in Figure 3a. In this study, it was found that mutated model with Cys138Tyr forms five H bonds: 3 with Asp49, 1 with His71 and 1 with Ala102. The mutated model with Cys138Ser forms three H bonds with Asp49. The mutated model with Asp8Gly forms three H bonds. There occurred no H bond formation during the docking process for the mutants Lys96Thr and Ser104Arg. WT model with Cys138 forms four H bonds: Three H bonds were identified to form with Asp49 and one H bond with Ala102 these are important for strong interactions with the drug. The PZA-PZase complex was visualized using DS in order to get insight into the interaction between the drug and target. Surface representation of model structure with PZA docked into the cavity is shown in Figure 3b. It is evident from the figure that PZA is located in the center of the active site, and is stabilized by H bonding interactions. The drug is bound between  $\alpha$ 1-  $\alpha$ 4 alpha helices and  $\beta$ A –  $\beta$ D beta strands. The binding energy between the enzyme and drug was found to be -27.9397 kcal/mol. The electrostatic and Van der Waals energies were -5.48845 and -3.35432 kcal/mol respectively. It can be speculated that the conformational changes which occurs due to mutations at the active site of target protein, causes alteration in the drug binding ability of the protein. At instances when the mutations are severe, the binding ability gets diminished resulting in survival of the organism in the presence of drug, leading to drug resistance.

According to the inverse relationship between the binding affinity of drug and drug resistance (higher the affinity-lower will be the resistance). It can be pointed out that docking, although an important parameter to judge the affinity of protein (high score leads to high affinity) towards the ligand, cannot be considered as a sole factor for the determination of binding affinity. In the present context it has

provided the results contrary to the prior kinetic studies. The mutants Lys96Thr and Ser104Arg displayed a similar and low docking value (-7.21 kcal/mol) indicating lower affinity hence may contribute to high-level drug resistance, is in agreement with the reported [6, 7] kinetic data showing no activity and conferring high degree of resistance. On the contrary, the mutant with Asp8Gly conversion showed relatively high docking value (-16.51 kcal/mol) suggesting its higher affinity

towards the drug molecule leading to low-level drug resistance in concordance with the report [7] showing higher activity, but in sharp opposition with the report [8] with virtually no activity. The mutated model with substitution of Cys to Ser at position 138 showed higher value (-19.18 kcal/mol) compared to Cys to Tyr model (-17.75 kcal/mol). Therefore, Cys to Tyr substitution may lead to higher degree of PZA resistance than Cys to Ser.



**Figure 3:** (a) PZA docked with PZase. The dotted line represents H bonds with Asp49. (b) Surface representation of model structure with PZA (shown in red stick form) docked into the cavity. The residues surrounding the active site are colored yellow. Surface color red and blue represents electro negative and electro-positive.

The mutant models Lys96Thr and Ser104Arg showed significantly lower score (60%) compared to WT value (-17.95 kcal/mol). The models Asp8Gly and Cys138Tyr displayed a low score of 9% and 2% respectively. However the mutant Cys138Ser presented higher score (7%) compared to WT and others. Overall the interaction between the WT and mutant models with the drug indicates differences in the binding ability, with the WT values being higher than the mutants such as Lys96Thr, Ser104Arg, Asp8Gly and Cys138Tyr and lower than Cys138Ser. Thus it can be hypothesized that the above mutants may contribute to high-level drug resistance in the following descending pattern: Lys96Thr/Ser104Arg > Asp8Gly > Cys138Tyr > Cys138Ser.

#### Conclusion:

With the help of predicted models, the conformational changes due to mutations at the active site residues were determined. On the basis of docking score, it can be interpreted that the steric constraints resulting from such mutations could significantly alter the folding of the mutant protein and or the positioning of the side chain of the catalytic residues. This may readily explain the loss of PZase activity thereby inhibiting the conversion of PZA to POA. Also, how point mutations can dramatically alter the ability of drug binding which eventually may lead to differences among the WT and mutants. Thus, we suggest that the mutant Cys138Ser represented with high docking score may contribute to low-level drug resistance than other mutants. The findings suggest that redesign of the PZA molecule to improve drug binding may be a viable approach to overcome resistance in PZase especially for mutant with Lys96Thr, Ser104Arg and Asp8Gly substitutions. Such studies will be helpful in better understanding about the mechanism of drug resistant. Detection of mutations at the active site of the target protein enables to gain novel insight into the drug-target interactions, leading to the rational design of more

efficacious wonder drugs that not only shorten TB-therapy but also can prevent the emergence of drug resistance.

#### Acknowledgement:

A. Nusrath Unissa received financial support from the Indian Council of Medical Research as Senior Research Fellow. We are grateful for the computational support provided by Biomedical Informatics Centre of TRC.

#### References:

- [1] A Somoskovi *et al.*, *Respir Res.* **2**:164 (2001) [PMID: 11686881]
- [2] S Ramaswamy, JM Musser, *Tuber. Lung Dis.* **79**:329 (1998) [PMID: 10645439]
- [3] A Scorpio, Y Zhang, *Nat. Med.* **2**:662 (1996) [PMID: 8640557]
- [4] S Sreevatsan *et al.*, *Antimicrob. Agents. Chemother.* **41**: 636 (1997) [PMID: 9056006]
- [5] N Lemaitre *et al.*, *Antimicrob. Agents. Chemother.* **43**:1761-1763 (1999) [PMID: 10390238]
- [6] N Lemaitre *et al.*, *Biochem. J.* **353**: 453 (2001) [PMID: 11171040]
- [7] H Zhang *et al.*, *FEBS J.* **275**: 753 (2008) [PMID: 18201201]
- [8] SF Altschul *et al.*, *Nucleic Acids Res.* **25**: 3389 (1997) [PMID: 9254694]
- [9] JD Thompson *et al.*, *Nucleic Acids Res.* **22**:4673 4680 (1994) [PMID: 7984417]
- [10] SC Lovell *et al.*, *Proteins.* **50**:437 (2003) [PMID: 12557186]
- [11] CM Venkatachalam *et al.*, *J. Mol. Graph. Model.* **21**:289 (2003) [PMID: 12479928]
- [12] G Wu *et al.*, *J. Comp. Chem.* **24**:1549 (2003) [PMID: 12925999]
- [13] AT Laurie, RM Jackson, *Bioinformatics*, **21**:1908 (2005) [PMID:15701681]

Edited by P. Kanguane

Citation: Unissa *et al.*, Bioinformation 4(1): 24-29 (2009)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.