In silico Analysis of Novel Mutation Ala102Pro targeting \textit{pncA} Gene of \textit{M. tuberculosis}

Dhanurekha Lakshmipathy, Gayathri Ramasubban, Lily Therese*, Umashankar Vetrivel, Muthukumaran Sivashanmugam, Sunita Rajendiran, Sridhar R, Madhavan HN, and Meenakshi N

Vision Research Foundation, Chennai, Tamil Nadu, India

Abstract

This study reports on the structural and functional basis of pyrazinamide (PZA) resistance conferred by a novel mutation Ala102Pro in \textit{pncA} gene as sequenced from a PZA resistant 	extit{Mycobacterium tuberculosis} strain. Molecular modeling studies of Wild Type (WT) and Mutant Type (MT) of Pyrazinamidase (PZase) showed the mutation at Ala102Pro does not impact on the conformation of the protein. However, the docking studies infer that MT has a higher inhibitory constant (Ki=990.0m) compared to WT (Ki=822.42m), which is indicative of drug resistance in MT. Furthermore, molecular interaction studies also reveal that WT forms 4 hydrogen bonds involved in PZA-WT inhibitory interactions, whereas, in case of MT, it formed 5 hydrogen bonds with PZA. However, Ala102 in WT was found to be less fluctuating and more stable in molecular dynamics simulation when compared to Pro102 in MT which was highly fluctuating and unstable. This implies that Ala102 shall be a key residue involved in PZA inhibitory interactions. Moreover, MT does not show hydrogen bonding with PZA with Pro102 and also deviating in terms of PZA binding pose in comparison with WT. Hence, the observed deviations in terms of MT-PZA interactions shall be attributed to the drug resistance conferred.

Keywords: Pyrazinamide; M. Tuberculosis; \textit{pncA} gene

Introduction

Pyrazinamide (PZA) is the first-line drug used in the treatment of tuberculosis along with isoniazid and rifampicin and it inhibits semidormant mycobacteria only at low pH in vitro.

The \textit{pncA} gene encodes pyrazinamidase (PZase), and mutations in \textit{pncA} are associated with resistance to PZA \cite{1,2} or loss of PZase activity. PZA acts by targeting the fatty acid synthase/synthetase enzyme, and is responsible for the killing of persistent tubercle bacilli in the initial intensive phase of chemotherapy. It is a produg that is converted to its active form namely, pyrazinoic acid (POA) by the catalytic action of PZase enzyme, encoded by the \textit{pncA} gene in \textit{M. tuberculosis}. Interestingly, PZA is active only at low pH since acidic environment favours accumulation of POA in the cytoplasm due to an ineffective efflux pump, thereby leading to improper efflux out of the environment.

In \textit{M. tuberculosis} PZase could be key residues for hydrolysis of PZA \cite{8}. At the protein level, these regions were found to be well conserved among the amino acid sequences of \textit{pncA} proteins from different bacterial species.

The crystal structure of the \textit{M. tuberculosis} \textit{pncA} protein has been determined, showing significant differences in the substrate binding cavity when compared to the pyrazinamidases from \textit{Pyrococcus horikoshii} and \textit{Acinetobacter baumanii}. In \textit{M. tuberculosis}, this region was found to hold a Fe$^{2+}$ ion coordinated by one aspartate and three histidines, the most crucial structural element in this loop appears to be the specific positioning of residue His57 which is directly involved in the coordination of the Fe$^{2+}$ ion. The overall architecture of the pyrazinamidase of \textit{M. tuberculosis} is similar to that reported for the other pyrazinamidases of \textit{A. baumanii} and \textit{P. horikoshii} \cite{11}.

In the \textit{pncA} model, the putative catalytic centre would be located in a pocket formed by one α-helix (αE), four β-strands (β1, β2, β3 and β4) consisting of β1 (Asp-8 and Phe-13), β2 (Asp-49), β3 (Lys-96), β4 (αA-134 and Thr-135) and αE (Cys-138). In this pocket, the conserved active cysteine residue Cys-138 is located close to the conserved residues: Asp-8, Trp-68, Lys-96, Ser-104, Ala-134 and Thr-135. In the \textit{pncA} model, the side chains of the two residues Asp-8 and Lys-96 are found to point towards Cys-138 of the active-site. The modification of the amino acid residues Asp-8, Lys-96 and Ser-104 in the mutants DSG, K96T and S104R resulted in enzymes showing specific activities drastically impaired (%0.004 unit mg), thus suggesting that these...
residues are essential for the pncA activity. The amino acids found at positions 8, 13, 61, 69, 96, 103, 104 and 146 are functionally and structurally important in pncA [9]. Hence, in this study we attempt to utilize this empirical structural data to analyze the impact of Ala102Pro mutation on PZA resistance.

Materials and Methods

Phenotypic drug susceptibility testing

Phenotypic PZA (100 µg/mL) susceptibility testing was performed by BACTEC MicroMGIT culture system following manufacturer’s instruction (Becton Dickinson, USA) [12].

DNA extraction

DNA was extracted from the M. tuberculosis isolate by boiling at 80°C for 10 minutes followed by centrifugation at 3,000 rpm. 5 µl of supernatant was used as the template DNA.

PCR targeting pncA gene

The amplification reactions contained 200 µM of each dNTPs [dATP, dGTP, dCTP (Bangalore Genei)], 1 µM of outer and inner set of primers, 1x buffer [10 mM Tris – HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2], 3 units of Taq DNA polymerase enzyme and 5 µl of template DNA. PCR targeting pncA gene was performed with the following forward and reverse primers pncA1: 5’GGCGTACATGGACCCTATATC3’ and pncA2: 5’CAAACGTTTATCCCCGTTTC3’ [13] with the thermal profile for 30 cycles of denaturation at 95°C for 1 min, annealing at 60°C for 30 sec and extension at 72°C for 1 min which yielded a 670 bp product. For each PCR run, negative control and positive control were included in the study. The PCR results were considered valid only when the negative control was negative without amplicon and the positive control yielded specific band.

Detection of amplified product

10 µl of the amplified product was subjected to electrophoresis on 2% agarose gel incorporated with 0.5 µg/ml ethidium bromide for visualization by UV transilluminator (Vilber Lourmet, France).

Standard Precautions for PCR

Strict and adequate precautions were taken to prevent amplicon getting contaminated. Separate rooms were used for the preparation of DNA, its amplification and analysis of amplified products. PCR cocktail was prepared in a laminar flow workbench; the microfuge tubes and milliQ water used for PCR were double sterilized.

DNA sequencing and Data analysis

DNA Sequencing of amplicons was carried out using an ABI prism 3110 automated DNA sequencer. Cycle sequencing of the amplified products was carried out using ABI Prism BigDye terminator kit (Applied Biosystems, USA) following manufacturer’s instruction. The sequences were analyzed by BioEdit sequence alignment software [14]. The sequences generated were compared with the wild type sequence (Genbank accession nos. for X68081 for katG) by using Multalin software [15] to identify the presence of mutation or polymorphism.

Bioinformatics

Data sets

The protein sequence for Pyrazinamidase (Pzase) pncA (UniProtKB id: Q50575) was obtained from UniProtKB to perform in silico sequence analysis. The 3D atomic coordinates (PDB ID:3PL1) were obtained from Protein Data Bank (http://www.pdb.org/pdb/home/home.do) for structure analysis, mutation modeling (Ala102Pro), molecular docking and molecular dynamics simulation studies.

Predicting stability change on single amino acid polymorphism based on support vector machine (I-Mutant 2.0)

Protein structural stability of the mutants were assessed using I-Mutant 2.0 server which is a support vector machine (SVM) – based tool for automatic prediction of protein stability changes upon single-point mutations; predictions are performed for both sequence and structure of proteins using 1-Mutant 2.0 server. The output of this program displays the predicted free energy change value (ΔΔG) which is calculated from the unfolding Gibbs free energy value of the native type (kcal/mol). Positive ΔΔG values infer that the mutated protein possesses high stability and vice versa (http://gpcr.biocomp.unibo.it/cgi/predictors/ I-Mutant 2.0/1-Mutant 2.0.cgi).

Analysis of secondary structure using Protein Structure Prediction Server (PSIPRED v3.0)

The PSIPRED is a reliable sequence based Protein Secondary Structure Prediction Server and was used in this study to predict the secondary structure of the mutant sequence so as to assess the impact of the novel mutation at the secondary structure elements. The output gives the details of presence of helix (H), sheets (E) and coils(C) in the protein sequence with graphical representation (http://bioinf.cs.ucl.ac.uk/psipred/).

Detection of conserved amino acids in the human pathogenic organisms containing pyrazinamidase/ nicotinamidase

Consurf analysis can be utilized to reveal the conservation of residues among the group of organisms with common metabolic functions. Here, in this study, the bacteria which exhibit PZase activity were grouped together in terms of sequence identity and were further analyzed for conserved residues through consurf analysis.

Homology modeling of mutant Pzase

Homology modeling helps in predicting the 3D structure of a protein with unknown structure by comparing it with a known structure sharing high sequence similarity. In this study, MODELLER9v7 [16] was used for modeling the novel mutant protein. Since the crystal structure of the pyrazinamidase M. tuberculosis (PDB id: 3PL1) was available [12], the same was used as template to predict the mutant form. Finally, the stereochemical property of WT and MT of the protein structure was predicted using SAVS server (http://nihserver.mbi.ucla.edu/SAVES/).

Structure optimization and validation

The WT and the modeled MT PZase structures were subjected to energy minimization by steepest descent algorithm using GROMOS96 force field [17]. The quality of the protein structure modeled was again checked using Q-Mean server [18] and ProQ (http://www.sbc.su.se/~bjornw/ProQ/ProQ.cgi).

Active site analysis of Pzase enzyme

The active site of Pzase from M. tuberculosis was analyzed using Q-SiteFinder, CASTp (http://sts-fw.bioengr.uic.edu/castp/calculation.php) and Pocket Finder server (http://www.modelling.leeds.ac.uk/).
pocketfinder/help.html) to identify the binding site for docking studies. The results obtained were correlated with the active site residues from published literature [7,8,19].

Ligand optimization

The initial structure of PZA was obtained from PubChem in 3D SDF format. It was converted to PDB format using Open Babel (http://www.molecularnetworks.com/online_demos/corina_demo.html). Hydrogen bonds were added to the structure and was geometry optimized (Broyden- Fletcher-Goldfarb-Shanno line search method set to 1000 steps) using ArgusLab [20], wherein, the energy minimization was carried out using Universal Force Field (UFF) [21].

Docking of PZA-Pzase

PZA was docked into the structures of WT and MT using AutoDock 4.0 [22]. In this docking simulation, semiflexible docking protocols [23] were used, in which the protein structures were kept rigid and the PZA being docked was kept flexible. Blind docking was performed using grid point value (X, Y and Z) of 100Å and spacing between the grid points was 0.375Å. Lamarckian genetic algorithm was selected for ligand conformational searching and default docking parameters were used. The best docked complexes based on the lowest binding energy were further analyzed for hydrogen bonding interactions and the binding energy of WT and MT was compared. Finally, the Protein-ligand complexes were analyzed using Discovery Studio Visualizer [24] and Pymol visualization tool (http://www.pymol.org/).

Docking complex simulation

To study the stability of PZA ligand with wild and mutant forms, molecular dynamics simulation was carried out for the docked Protein-ligand complexes. All simulations were performed using GROMACS [19], with GROMOS96 43al force field. The topology file for the ligand was generated using Dundee PRODRG Server [http://davapc1.bioch.dundee.ac.uk/prodrg/].

The system was placed at the center of a cubic periodic box with an area of 37×37×37 Å and was solvated with a simple point charge (SPC213) water model. The distance between solute and edge of the box was 0.9 Å for the system. It included water molecules for all the complexes. In order to neutralize the systems charge to zero, adequate number of Na+ or Cl- ions were added. Each protein-ligand complex was subjected 2000 steps of energy minimization using steepest descent algorithm. Equilibration of the ensembles was performed by position restraint of the system for 100 ps. A constant temperature of 300 K and 1 atm pressure was maintained in the system. For long range electrostatics calculation, Partial Mesh Ewald coulomb type was used with a cut off of 1.0 nm. All bonds including H-bonds were constrained using LINCS algorithm. The MD simulation was carried out for 1ns for each complex. After simulation, the trajectories were analyzed by ngmx software package. All the molecular modeling and simulation studies were carried on OPEN Discovery 2 Linux Platform [25].

Results and Discussion

Phenotypic drug susceptibility testing

Phenotypic drug susceptibility testing by BACTEC method showed the M. tuberculosis isolate to be resistant to PZA.

PCR targeting the pncA gene and DNA sequencing

PCR targeting the pncA gene resulted in a 670-bp amplified product. DNA sequencing of the amplified product with forward and reverse primers followed by MultAlin analysis of the M. tuberculosis strain revealed the presence of novel substitution mutation Ala102Pro (GCC→CCG).

SNP Analysis of I-Mutant Server

In this study, wild type sequence of pncA at 102nd position Ala was replaced by Pro to predict protein stability changes through I-mutant server. The results infer loss of stability by the mutant protein with negative Gibbs free energy value of -1.10 at pH 5.5 and 37°C.

Secondary structure analysis for Wild and Mutant Type of PZA (Pdbsum SERVER)

PDBpred results for the MT shows no significant change when compared to WT and MT (Figure1A and figure 1B).

ConSurf analysis

Consurf analysis revealed the following residues to be conserved across all the selected Pzase activity exhibiting pathogenic organisms: Asp8, Phe13, Val44, Asp49, His51, His71, Gly78, Lys96, Tyr103, Thr114, Leu120, Gly132, Ala134, Asp136, Cys138, Val139, Ala146, Val155, Ala171, and Met175 and shall contribute to structural and functional aspects. As the novel mutation Ala102Pro was predicted to occur near the Tyr103 conserved region, it shall also contribute in the modulation of activity of the protein. Moreover, proline insertion shall also confer structural deviation as it forms partial peptide bond.

![Figure 1a](image1a.png) ![Figure 1b](image1b.png)

**Figure 1:** Secondary structure prediction of wild (1A) and mutation (1B) of PZA using pdbsum server.
Homology modeling and loop refinement of MT Pbase

The structure of MT Pbase (Ala102Pro) was modeled using Modeledr9v7 with WT structure as template. The modeled structure was found to be highly plausible as it had 99% sequence identity with that of the template. Moreover, the Ramachandran plot also showed 83.5% of residues in most favored regions with no residues in disallowed region (Figure 2).

Structure optimization and Validation

The structures of WT and MT were energy minimized using GROMOS96 force field by Gromacs software package. The potential energy of WT and MT were found to be -2.6391059e+04 Kcal/mol and -2.3403371e+04 Kcal/mol, respectively. The optimized structure energy of WT and MT were found to be -2.6391059e+04 Kcal/mol and -2.3403371e+04 Kcal/mol, respectively. The optimized structure was validated using Q-mean and ProQ Server (Table 1 and figure 3).

Table 1: Energy minimization and Protein structure Quality assessment of WT and MT of Pbase.

| SAVS (Structure Analysis and Validation Server) | Before Energy Minimization | After Energy Minimization |
|-----------------------------------------------|---------------------------|--------------------------|
| *WILD*                                       | *WILD*                    | *MUTANT*                 |
| Most favoured regions                        | 84.3%                     | 90.5%                    | 87.3 %                    | 91.9% |
| Additional allowed regions                   | 12.6%                     | 7.0%                     | 12.7 %                    | 8.1%  |
| Generously allowed regions                   | 1.3%                      | 1.9%                     | 0%                        | 0%    |
| Disallowed regions                           | 1.9%                      | 0%                       | 0%                        | 0%    |
| ProQ LG score                                | 5.167                     | 5.340                    | 6.064                     | 5.006 |
| Max sub                                      | 0.187                     | 0.228                    | 0.594                     | 0.511 |
| Q mean score                                 | -                         | -                        | 0.82                      | 0.79  |

Active site analysis of PbZA

Three different methods namely, CASTp, Q-site and pocket finder were implemented to predict the active site residues of Pbase. All the predictions were found to be synonymous and representing the same amino acid residues: Asp8, Phe13, Leu19, Asp49, His57, Trp68, His71, Lys96, Ala102, Tyr103, Thr135, Asp136, and Cys138 (Figure 4a). Moreover, the predicted residues also correlated well with the documented studies. [8,9,19].

Docking studies of WT &MT type of Pbase with PbZA

The docking studies were performed using the Autodock software. The aminoacid residues Asp8, Phe13, Thr61, Pro69, Lys96, Tyr103, and Cys138 were assigned as catalytic region as these residues were proven to be potentially active site [8,9,19].

The WT had binding energy of -4.21 Kcal/mol and theoretical inhibitory constant (Ki) of 822.42 µM (Table 2). The residues Asp8, Lys96, and Ala102 interacted through 4 hydrogen bonds with PbZA. The interaction was found to span within the chosen active site [8,9,19].

The MT had binding energy of -4.1 Kcal/mol and theoretical inhibitory (Ki) constant of 990.0 µM (Table 2). The results of docking study infer that both WT and MT interact with PbZA at the same binding pocket region. However, MT showed deviation in terms of interaction with PbZA by excluding hydrogen bond formation with Lys96 and Ala102 and including Ala134 and Ile133 (Figures 5A and 5B).

Dynamics Simulation analysis

To further validate the molecular docking studies, high performance molecular dynamics simulation protein-ligand complex were performed and from the resultant trajectory, root mean square fluctuation (RMSF) and gyration analysis were performed. The RMSF results showed that the MT has high fluctuation when compared to WT due to the acquired mutation (Figure 6A). Further, the Radius of gyration analysis also showed MT to be slightly compact than WT which shall interfere with the flexibility of the protein (Figure 6B).
Conclusion

The novel mutation Ala102Pro was found to span in the putative active site (Lys96-Tyr103) region of PZase. The SNP analysis by I-mutant server showed that the stability of the protein was affected with negative Gibbs free energy value at pH 5.5 and 37°C. Molecular modeling and docking analysis showed that MT has a higher inhibitory constant than WT in terms of PZA binding, which indicates that the drug affinity is highly affected in the MT due to structural changes. The molecular dynamics simulation and gyration analysis together infer the change in flexibility at the active site cavity of MT. Hence, this study gives insight on the impact of novel mutation on the activity of this protein which can be attributed to the drug resistance observed. This study also presents a need for the discovery of potential newer drug molecules against both mutant and wild types.

Acknowledgements

The authors gratefully acknowledge the financial support by Chennai Willingdon Corporate Foundation (Chennai, India) for funding the research project.

References

1. Dalmier O, Walter E (1936) Verfahren zur Herstellung von Derivaten der Morlock. Nature 92: 662-667.
2. Sreevatsan S, Pan X, Zhang Y, Kreiswirth BN, Musser JM (1997) Mutations associated with pyrazinamide resistance in pncA of Mycobacterium tuberculosis complex organisms. Antimicrob Agents Chemother 41: 636-640.
3. Lemaître N, Sougakoff W, Truffot-Pernot C, Jarlier V (1999) Characterization of New Mutations in Pyrazinamide-Resistant Strains of Mycobacterium tuberculosis and Identification of Conserved Regions Important for the Catalytic Activity of the Pyrazinamidase pncA. Antimicrob Agents Chemother 43: 1761-1763.
4. Lakshmipathy D, Ramasubban G, Therese L, Vetrivel U, Sivashanmugam M, et al. (2013) In silico Analysis of Novel Mutation Ala102Pro targeting pncA Gene of M. tuberculosis. J Comput Sci Syst Biol 6: 083-087. doi:10.4172/jcssb.1000103
5. Petrella S, Gelus-Ziental N, Maudry A, Laurans C, Boudjelloul R, et al. (2011) Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci. Wiley-Blackwell: 95-98.
6. Scorpio A, Zhang Y (1996) Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antimycobacterial drug pyrazinamide in Mycobacterium tuberculosis. Nature Med 2: 662-667.
7. Sreevatsan S, Pan X, Zhang Y, Kreiswirth BN, Musser JM (1997) Mutations associated with pyrazinamide resistance in pncA of Mycobacterium tuberculosis complex organisms. Antimicrob Agents Chemother 41: 636-640.
8. Lemaître N, Callebaud I, Frenois F, Jarlier V, Sougakoff W (2001) Study of the structure-activity relationships for the pyrazinamidase (PncA) from Mycobacterium tuberculosis. Biochem J 353: 453-458.
9. Ramaswamy S, Musser JM (1998) Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. Tuberc Lung Dis 79: 3-29.
10. Petrella S, Gellus-Ziental N, Maudry A, Laurans C, Boudjelloul R, et al. (2011) Crystal structure of the pyrazinamidase of Mycobacterium tuberculosis: insights into natural and acquired resistance to pyrazinamide. PLoS One 6: e15765.
11. Hoffman H (2003) Siddiqui BACTEC 460TB System; Procedure and Product Manual MA0029, Linton Dickinson and Company, Maryland, USA.
12. Lakshmipathy D, Ramasubban G, Therese L, Vetrivel U, Sivashanmugam M, et al. (2013) In silico Analysis of Novel Mutation Ala102Pro targeting pncA Gene of M. tuberculosis. J Comput Sci Syst Biol 6: 083-087. doi:10.4172/jcssb.1000103
13. Sreevatsan S, Pan X, Zhang Y, Kreiswirth BN, Musser JM (1997) Mutations associated with pyrazinamide resistance in pncA of Mycobacterium tuberculosis complex organisms. Antimicrob Agents Chemother 41: 636-640.
14. Lemaître N, Callebaud I, Frenois F, Jarlier V, Sougakoff W (2001) Study of the structure-activity relationships for the pyrazinamidase (PncA) from Mycobacterium tuberculosis. Biochem J 353: 453-458.
15. Petrella S, Gellus-Ziental N, Maudry A, Laurans C, Boudjelloul R, et al. (2011) Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci. Wiley-Blackwell: 95-98.
16. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comp 4: 435-447.
17. Benkert P, Künnmi M, Schwede T (2009) QMEAN server for protein model quality estimation. Nucleic Acids Res 37: W510-514.
18. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comp 4: 435-447.
19. Du X, Wang W, Kim R, Yakota H, Nguyen H, et al. (2001) Crystal structure and mechanism of catalysis of a pyrazinamidase from Pyrococcus horikoshii. J Mol Biol 310: 1416-14172.
20. Stoermer MJ (2006) Current status of virtual screening as analysed by target class. Med Chem 2: 89-112.
21. da Silva CH, Campo VL, Carvalho I, Taft CA (2006) Molecular modeling, docking and ADMET studies applied to the design of a novel hybrid for treatment of Alzheimer’s disease. J Mol Graph Model 25: 169-175.
22. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comp 4: 435-447.
23. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19:1639-1652.
24. Discovery Studio Modeling Environment (2011) Release 3.1
25. Vetrivel U, Pillai K (2008) Open discovery: An integrated live Linux platform of Bioinformatics tools. Bioinformation 3: 144-146.