Understanding properties of the master effector of phage shock operon in *Mycobacterium tuberculosis* via bioinformatics approach

Zarrin Basharat*, Azra Yasmin

Microbiology & Biotechnology Research Lab, Department of Environmental Sciences, Fatima Jinnah Women University, 46000, Pakistan.

***Corresponding author email:** zarrin.iiui@gmail.com

Keywords: *Mycobacterium tuberculosis*, Phage shock protein, Structure modeling, Dynamics simulation.

NOTE: An expanded version of this manuscript is under-preparation. The current pre-print version (v.01) of this manuscript may contain grammatical & proofreading mistakes. Errors and omissions excepted.

Abstract

The phage-shock-protein (psp) is a part of the psp operon, which assists in safeguarding the survival of bacterium in stress and shields the cell against proton motif force challenge. It is strongly induced by bacterium allied phages, improperly localized mutant porins and various other stresses. Master effector of the operon, PspA has been modeled and simulated, illustrating how it undergoes significant conformational transition at the far end in *Mycobacterium tuberculosis*. Association of this key protein of the operon influences action of Psp system on the whole. We are further working on the impact of phosphorylation perturbation and the impact of structural fluctuations during complex formation of PspA with other moieties of interest.

We hereby report *ab initio* structure model of the *Mycobacterium tuberculosis* phage shock protein A (PspA). PspA is the central constituent of bacterial stress response machinery, encoded by phage shock operon (Huvet et al., 2010). PspA, regulates not only it's own transcription but that of the whole operon as well (Elderkin et al., 2005; Male et al., 2014). PspA is inferred to be a dual-function protein (Guegwen et al., 2009; Jovanovic et al., 2014) and localized amid cytoplasmic and inner membrane interface of the bacterium (Engl et al., 2009). It is responsible for maintaining the cell membrane integrity along with restoration of the proton motive force (Male et al., 2014; Engl et al., 2009; Wan et al., 2015).

Mycobacterium tuberculosis is a pathogenic bacterium responsible for causing disease in humans and veterinary species (Sakamoto, 2012). A lot of work has been carried out on *Mycobacterium* phages for therapeutic purpose. However, to the best of our knowledge, no report of the phage shock protein analysis for this pathogen exists at the moment. We retrieved PspA protein sequence of *Mycobacterium tuberculosis* from the Uniprot database with Accession number: R4M912 and analyzed the sequence and the structure using computational tools. Although, PspA belongs to the highy conserved PspA/IM30 family but *Mycobacterium tuberculosis* PspA shared a very low sequence homology with *Escherichia coli* PspA, revealed using Clustal Omega (Fig. 1) with seeded guide trees and HMM profile-profile technique for alignment generation at the backend (Sievers et al., 2011). The secondary structure analyzed by PROMOTIF tool (Hutchison and Thornton, 1996) revealed that the protein consisted of 9 helices, 6 helix-helix interacs, 14 β-turns and 3 □-turns (Fig. 2A).

Due to low homology with experimentally determined structures available in the RCSB Protein databank, 3D structure (Fig. 2B) was modeled by I-TASSER (Roy et al., 2010; yang et al., 2015) using *Escherichia coli* PspA as a template (PDB ID: 4WHE). The C-score based on the significance of template alignment threading and simulations of the structure assembly

convergence parameters of the model chosen for analysis was -1.29 (lies between ideal range of -5 to 2), indicating good quality model. Estimated root mean square deviation of the predicted model from the *Escherichia coli* model was 7.8±4.4Å. The constructed structure resembled a helical bundle with important binding site residues predicted to occur at position 66, 69, 70, 73, 92, 96, 99, 100, 102 and 103. Despite low sequence conservation, the structure is however, well conserved due to the underlying fact that the protein folds remain conserved in similar function proteins. PspA structure in *Escherichia coli* is known to self-assemble into ring (Standar et al., 2008) or striated and indented rod-shaped complexes (Male et al., 2014) based on electron microscopy and helical rods based on X-ray crystallography analysis (Osadnik et al., 2015). PspA homologue LiaH in *Bacillus subtilis* (Wolf et al., 2008) and holins of bacteriophage lambda (Savva et al., 2010) have also been reported to self-assemble to rod-like structures from ring shaped protein complexes. The *Mycobacterium tuberculosis* PspA is also rod shaped and it is implied that these rod-like structures could form a support framework and aid in the maintenance of membrane integrity during phage shock response (Male et al., 2014).

CABS-flex procedure based on the well-established coarse-grained CABS protein model (Fraga et al., 2014) was employed for the fast simulation of near-native dynamics of PspA. CABS is a computationally efficient alternative to classical all-atom molecular dynamics (Jamroz et al., 2013). The 3D modeled structure was input and used as a starting point for the all-atom, explicit water, 10-nanosecond dynamic simulation. Analysis was carried out at the backend and automatically analyzed trajectory (Fig. 3) was obtained to study the dynamic behaviour of protein. A set of eight (all-atom) protein model sets were obtained with global distance test score ranging from 0.6-0.7. Most dominant structural fluctuations appeared at the last beta turn region including 8^{th} and 9^{th} helix. Relative propensity of protein residues to deviate from an average dynamics structure increased substantially at the ending helices with

fluctuation increasing from 100 Å at residue 160 to to >600 Å at residue 172. Understanding of flexibility of PspA can be of aid in research areas as molecular evolution (Boehr et al., 2009).

Phage infection has also been demonstrated to induce substantial fluctuations in host protein phosphorylation (Rieul et al., 1987; Russel and Model, 2006) and this was suggestive of PspA potential for phosphorylation as well. NetPhos Bac 1.0 (Miller et al., 2009) was used for prediction of possible phosphorylation residues. Seven serine residues (S144, S149, S156, S157, S158, S164, S166) were predicted to have phosphorylation potential based on neural network approach. However, none of these exhibited a knack to occur on predicted binding residues and hence, their exact role in PspA function of *Mycobacterium tuberculosis* yet remains to be elucidated.

Our findings pave way for further experimental studies and are of aid in understanding the *Mycobacterium tuberculosis* PspA response to the extracytoplasmic stresses that may damage the cytoplasmic membrane. We have used computational approach for the prediction of 3D structure of this protein but to furthur understand the function of the rod-like structure of *Mycobacterium tuberculosis* PspA, additional studies are required which can confirm and enhance the reported information along with elucidation of in depth biological function and interactions of *Mycobacterium tuberculosis* PspA.

Figure Legends

Figure 1. Multiple sequence alignment of the *Escherichia coli* and *Mycobacterium tuberculosis* PspA. Conserved residues are shown in red color. Helices are denoted by squiggles at the top of the alignment. Solvent accessibility is depicted by a bar below the sequence (blue = accessible, cyan = intermediate, white = buried).

Figure 2. (A) Secondary structure of *Mycobacterium tuberculosis* PspA (helices labelled H1, H2…H9). β depicts beta turn and γ is for gamma turn. (B) 3D structure of *Mycobacterium tuberculosis* PspA.

Figure 3. (A) Structural flexibility profile of simulated *Mycobacterium tuberculosis* PspA with fluctuations for individual protein residues shown via red line. The output is based on all-atom model via trajectory clustering. (B) Refinement of the model and superpositioning (Provided 3D model as base) is centred on maximum likelihood superposition method of THESEUS (Theobald and Wuttke, 2006).

Competing interests

The authors declare that no competing interests exist.

Acknowledgements

The authors are thankful to Dr. Alexandra Elbakyan for help with literature.

References

1. Huvet, M., Toni, T., Sheng, X., Thorne, T., Jovanovic, G., Engl, C., Buck, M., Pinney, J.W. and Stumpf. M.P.H. (2010). The evolution of the Phage shock protein response system: Interplay between protein function, Genomic organization, and system function. Mol. Biol. Evol. 28, 1141-1155.
2. Elderkin, S., Bordes, P., Jones, S., Rappas, M. and Buck, M. (2005). Molecular determinants for PspA-mediated repression of the AAA transcriptional activator PspF. J. Bacteriol. 187, 3238–3248.
3. Male, A.L., Oyston, P.C.F. and Tavassoli, A. (2014). Self-Assembly of *Escherichia coli* Phage Shock Protein A. Adv. Microbiol. 4, 353-359.
4. Guegwen, E., Savitzky, D.C., Darwin, A.J. (2009). Analysis of the *Yersinia enterocolitica* PspBC proteins defines functional domains, essential amino acids and new roles within the phage-shock-protein response. Mol. Microbiol. 74, 619-633.
5. Jovanovic, G., Mehta, P. and Buck, M. (2014). Anionic lipids and the cytoskeletal proteins MreB and RodZ define the spatio-temporal distribution and function of membrane stress controller PspA in *Escherichia coli*. Microbiology. 160, 2374-86.
6. Engl, C., Jovanovic, G., Lloyd, L.J., Murray, H., Spitaler, M., Ying, L., Errington, J. and Buck, M. (2009). *In vivo* localizations of membrane stress controllers PspA and PspG in *Escherichia coli.* Mol. Microbiol. 73, 382-396.
7. Wan, F., Mao, Y., Dong, Y., Ju, L., Wu, G. and Gao, H. (2015). Impaired cell envelope resulting from arcA mutation largely accounts for enhanced sensitivity to hydrogen peroxide in *Shewanella oneidensis.* Sci. Rep. 5, 10228.
8. Sakamoto, K. (2012).The Pathology of *Mycobacterium tuberculosis* Infection. Vet. Pathol. 49, 423-439.

9. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Sys. Biol. 7, 539.
10. Hutchison, E.G. and Thornton, J.M. (1996). PROMOTIF-a program to identify and analyze structural motifs in proteins. Protein Sci. 5, 212–220.
11. Roy, A., Kucukural, A. and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725-738.
12. Yang, J., Yang, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nat. Methods. 12, 7-8.
13. Standar, K., Mehner, D., Osadnik, H., Berthelmann, F., Hause, G., Lunsdorf, H. and Bruser, T. (2008). PspA Can Form Large Scaffolds in *Escherichia coli*. FEBS Lett. 58, 3585-3589.
14. Osadnik, H., Schöpfel, M., Heidrich, E., Mehner, D., Lilie, H., Parthier, C., Risselada, H.J., Grubmüller, H., Stubbs, M.T. and Brüser, T. (2015). The PspF-binding domain PspA1-144 and the PspA·F complex-New insights into the coiled-coil dependent regulation of AAA+ proteins. Mol. Microbiol. In press. doi: 10.1111/mmi.13154.
15. Wolf, D., Kalamorz, F., Wecke, T., Juszczak, A., Mader, U., Homuth, G., Jordan, S., Kirstein, J., Hoppert, M., Voigt, B., Hecker, M. and Mascher, T. In-depth profiling of the LiaR response of *Bacillus subtilis*. J. Bacteriol. 192, 4680-4693.
16. Savva, C.G., Dewey, J.S., Deaton, J., White, R.L., Struck, D.K., Holzeburg, A. and Young, R. (2008). The holin of bacteriophage Lambda forms rings with large diameter. Mol. Microbiol. 69, 784-793.

17. Rieul, C., Cortay, J.C., Bleicher, F. and Cozzone, A.J. (1987). Effect of bacteriophage M13 infection on phosphorylation of dnaK protein and other *Escherichia coli* proteins. Eur. J. Biochem. 168, 621-627.

18. Russel, M. and Model, P. (2006). Filamentous phage. In The Bacteriophages, R.L. Calendar, ed. (NY, USA: Oxford University Press), pp. 152-160.

19. Miller, M.L., Soufi, B., Jers, C., Blom, N., Macek, B. and Mijakovic, I. (2009). NetPhosBac-a predictor for Ser/Thr phosphorylation sites in bacterial proteins. Proteomics, 9, 116-125.

20. Theobald, D.L. and Wuttke, D.S. (2006). THESEUS: maximum likelihood superpositioning and analysis of macromolecular structures. Bioinformatics. 22, 2171-2172.

21. Jamroz, M., Kolinski, A. and Kmiecik, S. (2013). CABS-flex: server for fast simulation of protein structure fluctuations. Nuc. Acids Res. 41, W427-W431.

22. Fraga, H., Grana-Montes, R., Illa, R., Covaleda, G. and Ventura, S. (2014). Association Between Foldability and Aggregation Propensity in Small Disulfide-Rich Proteins. Antioxid. Redox Signal. 21, 368-383.

23. Boehr, D.D., Nussinov, R. and Wright, P.E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. Nat. Chem. Biol. 5, 789-796.

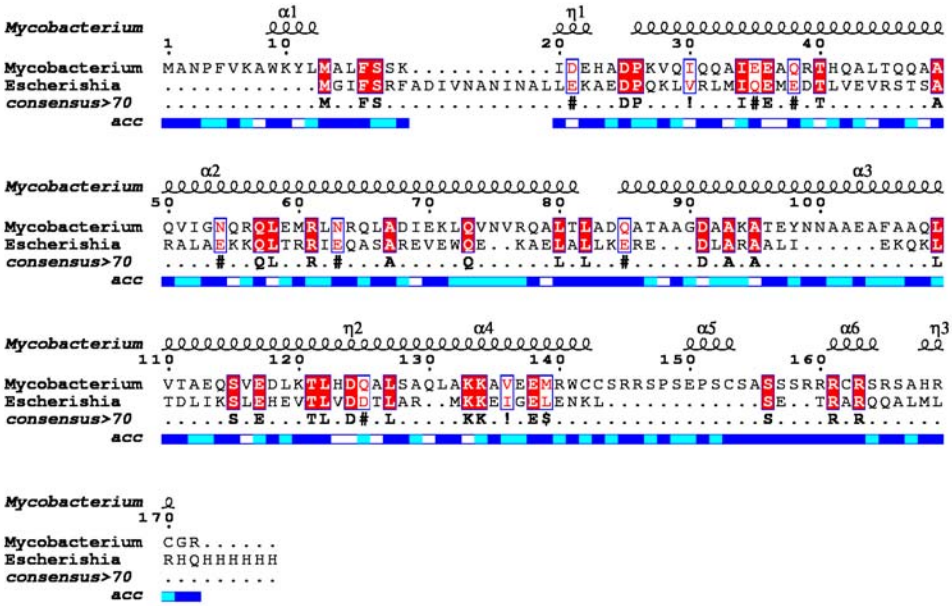

Fig 1.

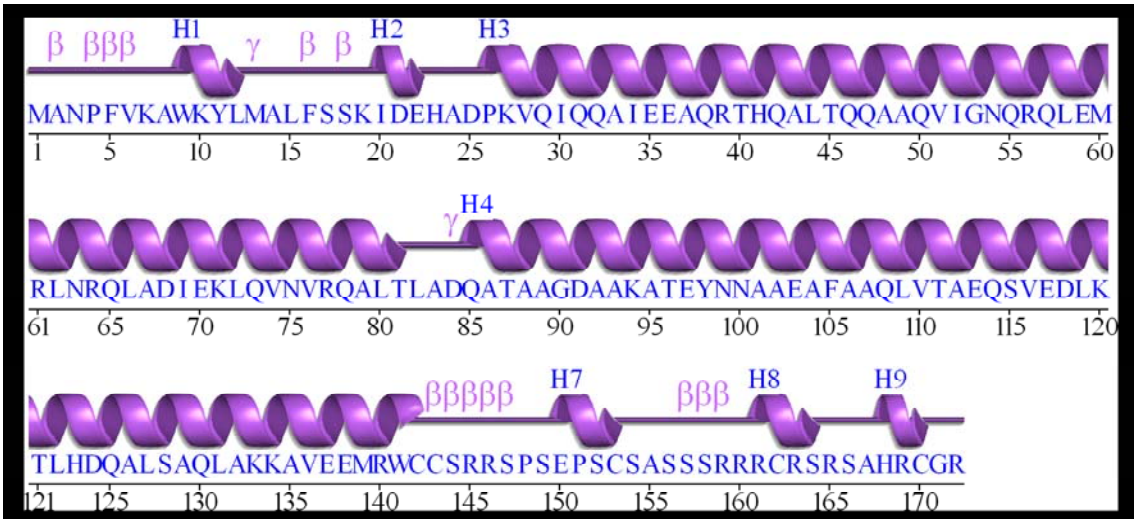

Fig 2A.

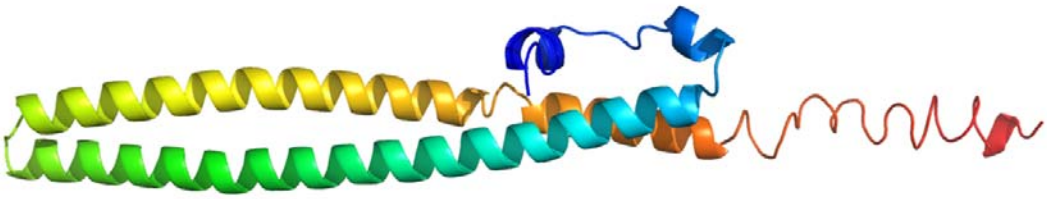

Fig 2B.

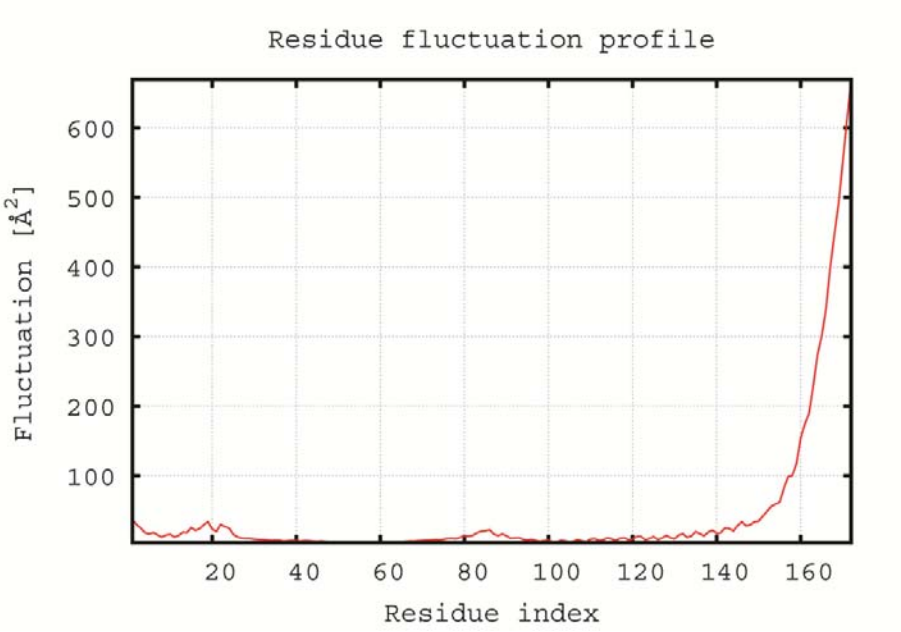

Fig 3A

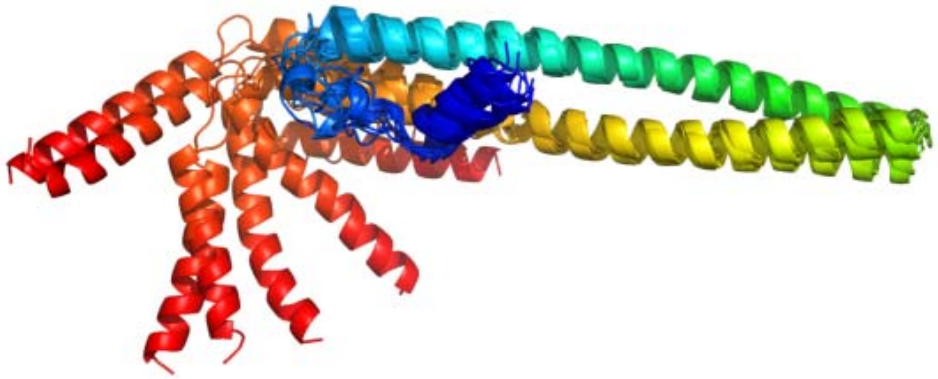

Fig 3B.