Structural And Functional Characterization Of Cation-Transporting Atpase Of Mycobacterium Tuberculosis CDC 1551 By In Silico Methods

C Kumar, C Anuradha, K Venkateswara Swamy, N Khan

Citation

C Kumar, C Anuradha, K Venkateswara Swamy, N Khan. *Structural And Functional Characterization Of Cation-Transporting Atpase Of Mycobacterium Tuberculosis CDC 1551 By In Silico Methods*. The Internet Journal of Genomics and Proteomics. 2005 Volume 2 Number 1.

Abstract

The three dimensional (3-D) structure of cation-transporting ATPase (ctATPase) of Mycobacterium tuberculosis CDC 1551 has been generated from amino acid sequence using homology modeling with backbone structure of calcium ATPase of Oryctolagus cuniculus as template. The model was generated using the method of satisfaction of spatial restraints by mod6v2 version of MODELLER. The generated 3-D structure of ctATPase) of M. tuberculosis was evaluated at various web interfaced servers i.e., WHAT IF, PRODOM, Pfam and PROCHECK which checks the stereo chemical quality of the structure in terms of bonds, bond angles, dihedral angles and non-bonded atom-atom distances, structural as well as functional domains etc., The generated model was visualized using the RASMOL visualization tool.

INTRODUCTION

With it's re-emergence, tuberculosis (TB) has become a major health problem worldwide and now a days it became a global emergency. It causes nearly 3 million deaths annually and an estimated 8.8 million new cases are reported every year. M. tuberculosis strain CDC 1551 is a highly transmissible strain which exhibited unusually high levels of infectivity and virulence during a recent outbreak of tuberculosis in a rural community. The complete genome of CDC 1551 (EMBL Accession number AE000516) consists of 4,403,836 base pairs [9,13]. This strain develops lung and spleen infections that are nearly identical to those caused by Mycobacterium tuberculosis strains Erdman, H37RV and Indian. Iron acquisition is critical for M. tuberculosis survival and replication in the infected host. Iron is essential as a cofactor of enzymes involved in vital cellular functions ranging from respiration to DNA replication. Limitation of this metal triggers an extensive response aimed at enhancing iron acquisition while coping with iron deficiency. Pathogens have to compete for iron in the host in order to multiply and establish a successful infection [16]. Studies in M. tuberculosis using one- and two-dimensional gel electrophoresis combined with mass spectroscopy and sequence information have shown that the levels of several proteins change in response to iron availability [22]. Among these, a putative cation-transporting ATPase was induced in bacteria grown in a low-iron medium suggesting its importance in the maintenance of vital processes that depend on iron-containing cofactors including electron transport, energy metabolism and DNA synthesis [8]. The ctATPase plays a major role in the influx of ferrous ions during iron deficiency in the bacterium. They perform multiple vital functions in the cells which include ATPase activity coupled to trans-membrane movement of ions (ion-transport), phosphorylative mechanism, hydrolase activity and regulation of cation homeostasis [8]. Bacteria use ctATPase to pump out toxic heavy metal ions. These ATPases are important for intracellular bacteria like M. tuberculosis, which require transition metal cations for the synthesis of superoxide dismutases, and catalases which constitute an important line of defense against macrophage's killing mechanisms [10,16]. There is now compelling evidence that iron acquisition is critical for M. tuberculosis survival and replication in infected host. Hence present study has been conducted to understand and elucidate the 3-D structure of ctATPase of M. tuberculosis by homology modeling. Through the insights gained from its 3-D structure, the functionally important domains are well characterized. Knowledge of these structural features of ctATPase of M. tuberculosis is essential for establishing its catalytic action mechanism at molecular level as well to target the protein

for designing effective and selective drug.

METHODOLOGY

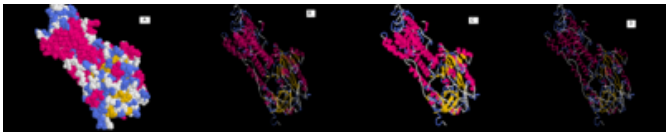
The ctATPase of M. tuberculosis CDC 1551 (gil15839483l) having 752 amino acids was modeled by using ATPase Oryctolagus cuniculus as a reference template (PDB ID :1VFP) whose crystal structure is known at 2.6A resolution (http://www.rcsb.org/pdb). Structural homologue to target sequence was find by BLAST-p against PDB database at http://www.ncbi.nlm.nih.gov/BLAST [1]. The 3-D structure of ctATPase of M. tuberculosis CDC 1551 was generated by homology modeling routine of MODELLER release 6v2. This routine includes the satisfaction of spatial and geometric restraints and a very fast molecular dynamic annealing: no other refinements were applied. MODELLER accepts the alignment of the sequence to be modelled with related known structures and automatically calculates a model with all non-hydrogen atoms that satisfies the restraints as well as possible and 3-D model was obtained by optimization of a molecular probability density function [2]. Thus generated 3-D structure was visualized by RASMOL [19]. The stereo chemical quality of the structure was analyzed using PROCHECK [15] which uses Ramchandran map [17] to check peptide bond planarity, bond lengths, bond angles, hydrogen bond geometry, and side chain conformations of protein structures as a function of atomic resolution. The functional domains of ctATPase of M. tuberculosis were obtained by submitting 3-D structure to Pfam at http://www.sanger.ac.uk/Software/Pfam [4, 5]. Prediction of secondary structure was done at WHAT IF (http://swift.cmbi.kun.nl/whatif/) [21]. The sequence segment domains for ctATPase were found in Swiss Prot database by BLASTp and the same were aligned by multiple sequence alignment using PRODOM at http://prodes.toulouse.inra.fr/prodom.html [7].

RESULTS

Most functional restraints on evolutionary divergence operate at the level of tertiary structure hence 3-D structures are more conserved in evolution than sequence [3]. The 3D structure of protein is an important source of information to better understand the function of a protein, its interactions with other compounds (ligands, proteins, DNA) and to understand phenotypical effects of mutations [20]. A large number of techniques have been developed to predict 3-D structures of proteins among which homology modeling is one of the best [14]. Homology modeling is historically the first [6] and MODELLER has been considered as most accurate method [18]. Hence in the present study also we used homology modeling using MODELLER to generate the 3-D structure of ctATPase. The 3-D structure of ctATPase in M. tuberculosis CDC 1551 generated from MODELLER contains large bifurcated beta-sheets (mainly anti parallel) covered on both sides with helices and loops (Fig1 A-D). The ctATPase structure showed trans-membrane helices (Figure 1-C), some of which are well conserved throughout the super family thus they all operate via a similar mechanism. The beta sheet covered on both sides with helices and loops is probably the region where the ferric ion or other cations bind before being channeled down through the trans-membrane helices.

Figure 1

Figure 1: Cation transporting ATPase (ctATPase) of CDC 1551. Model generated MODELLER mod.6v2 version. The 3-D structure of ctATPase was visualized with RASMOL and represented in the form of spacefill (A), cartoons (B), ribbons (C) and backbone (D).



The predicted secondary structure of ctATPase obtained from WHAT IF server (Table 1) clearly indicated the sequence positions of different turns (79), helices (30) and sheets (28). Tentatively ctATPase structure was verified at PROCECK (Table 2) and correspondingly Ramachandran plot (Figure 2) was also constructed to validate the structure [17].

Figure 2

Table 1: Secondary structure of ctATPase of CDC1551 obtained from WHAT IF (). The primary sequence of protein along with predicted secondary structures displayed simultaneously (H – helx, T – turns and S – sheets).

```
          10    20    30    40    50    60
           |     |     |     |     |     |
 1 - 60  MAAPVVGDADLQSVRRIRLDVSGMSCAACASRVETKLNKIPGVRASVNFATRVATIDAVG
 1 - 60  TTTT TTT  T TTTT TTTTT T          HHHHHHHHHHHHHHHH TTTH
          70    80    90   100   110   120
           |     |     |     |     |     |
61 - 120 MAADELCGVVEKAGYHAAPHTETTVLDKRTKDPDGAHARRLLRRLLVAAVLFVPLADLST
61 - 120 HHHHHHHT   HHHHHHHHHHHHHHHHHHHT SSSSTTTTSSSSTTT TTT SSSSST
         130   140   150   160   170   180
           |     |     |     |     |     |
121 - 180 LFAIVPSARVPGWGYILTALAAPVVTWAAWPFHSVALRNARHRTTSMETLISVGIVAATA
121 - 180 T SS T SSSSSS TT SSS TTTT  T SS  TT T  TTSSSSS SSSSSSTT
         190   200   210   220   230   240
           |     |     |     |     |     |
181 - 240 WSLSSVFGDQPPREGSGIWRAILNSDSIYLEVAAGVTVFVLAGRYFEARAKSKAGSALRA
181 - 240 33THHHHHTTTT  T T THHHHHHHHHHHHHHHHHHHHHHHHHHHTTT  TTTT HHH
         250   260   270   280   290   300
           |     |     |     |     |     |
241 - 300 LAELGAKNVAVLLPDGAELVIPASELKKRQRFVTRPGETIAADGVVVDGSAAIDMSAMTG
241 - 300 HHHHHHHHHHHHT TTHHHHHHHHHHHHHHHHHHT SSSTTHHHHHHHHH TSSSSST T
         310   320   330   340   350   360
           |     |     |     |     |     |
301 - 360 EAKPVRAYPAASVVGGTVVMDGRLVIEATAVGADTQFAAMVRLVEQAQTQKARAQRLADH
301 - 360 TT     TT   TTTTTT TTTT   TT HHHHTTTT HHHHHH
         370   380   390   400   410   420
           |     |     |     |     |     |
361 - 420 IAGVFVPVVFVIAGLAGAAWLVSGAGADRAFSVTLGVLVIACPCALGLATPTAMMVASGR
361 - 420 TTTT   TTT    HHHH   TT HHHHHHHHHHHHHHHHHTTT SSSSS T T
         430   440   450   460   470   480
           |     |     |     |     |     |
421 - 480 GAQLGIFIKGYRALETIRSIDTVVFDKTGTLTVGQLAVSTVTMAGSGTSERDREEVLGLA
421 - 480 TT  SSSSS   TTTHHHHHHHHHHTT SSSSS TT HHHHHHHHHHHT SS T
         490   500   510   520   530   540
           |     |     |     |     |     |
481 - 540 AAVESASEHAMAAAIVAASPDPGPVNGFVAVAGCGVSGEVGGHHVEVGKPSWITRTTPCH
481 - 540 HHHHHHHHH SS T  TTT T  TTT SSSSS TTTTHHHHHHTTSSSS TTT HHHHH
         550   560   570   580   590   600
           |     |     |     |     |     |
541 - 600 DAALVSARLDGESRGETVVFVSVDGVVRAALTIADTLKDSAAAAVAALRSRGLRTILLTG
541 - 600 T TSSST THHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHTTTT   HHHHHH
         610   620
           |     |
601 - 624 DNRAAADAVAAQVGIDSAVADMLP
601 - 624 HHHHHHTHHHH333  TTT


625 - 625 E
625 - 625 ?
       630   640   650   660   670   680
         |     |     |     |     |     |
626 - 685 GKVDVIQRLREEGHTVAMVGDGINDGPALVGADLGLAIGRGTDVALGAADIILVRDDLNT
626 - 685  HHHHHHHHHHH TTTT TTT    TTTT TTTTT  TTTT  THHHHHHHHHHH
       690   700   710   720   730   740
         |     |     |     |     |     |
686 - 745 VPQALDLARATMRTIRMNMIWAFGYNVAAIPIAAAGLLNPLIAGAAMAFSSFFVVSNSLR
686 - 745 3333TTTTT TTTT 333 HHHHHTTTTHHHHHHHTTTTTT TTTTTTTT    TTTT
       750
         |
746 - 752 LRNFGAQ
746 - 752 HHHHHT
```

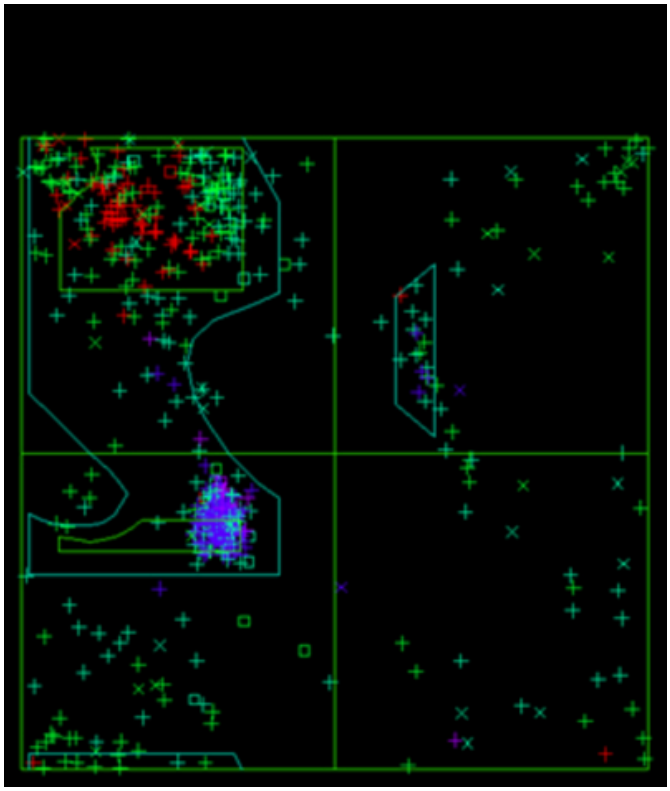
Figure 3

Table 2: Procheck summery for ctATPase of CDC1551 obtained at .

PROCHECK SUMMARY				
Test 2.0		752 residues		
Ramachandran plot	76.0% core	16.0% allow	5.1% gener	2.9% disall
Gly & Pro Ramach:	7 labeled residues (out of 99)			
Chi1-chi2 plots:	9 labeled residues (out of 332)			
Main-chain params:	5 better	0 inside	1 worse	
Side-chain params:	5 better	0 inside	0 worse	
Residue properties:	Bad contacts: 89			
Bond len/angle:	16.9		Morris et al class: 1 1 3	
G-factors	Dihedrals: - .39	Covalent: -1.53	Overall: -.81	
M/c bond lengths:	58.6% within limits	41.4% highlighted	21 off graph	
M/c bond angles:	46.5% within limits	53.5% highlighted	41 off graph	
Planar groups:	100.0% within limits	0% highlighted		

Figure 4

Figure 2: Ramachandran Plot for generated 3-D structure of ctATPase of CDC 1551 obtained from WHAT IF web interface server. Color codes for secondary structures – blue -helix, red - strand and green - turns and random coil.



The function of a protein is generally determined by shape, dynamics and physiochemical properties of its solvent exposed molecular surface. Likewise functional differences

among the members of the same protein family are usually a consequence of the structural differences on the protein surface [11, 12]. Hence the functional domains of ctATPase were established through Pfam server and PRODOM and the results clearly revealed three major domains namely HMA (18-79 amino acids), E1-E2-ATPase (215-436 amino acids), and hydrolase (440-667 amino acids) which are responsible for heavy metal binding, phosphorylation activity and hydrolase activity of the protein respectively (Figure 3, Figure 4 and Table 3).

Figure 5

Figure 3: The ctATPase of showing three major domains. The HMA domain -green (18-79), E1-E2 ATPase domain - red (215-436) and Hydrolase domain - yellow (440-667). Results obtained from Pfam server and visualized with RASMOL(Ribbons)

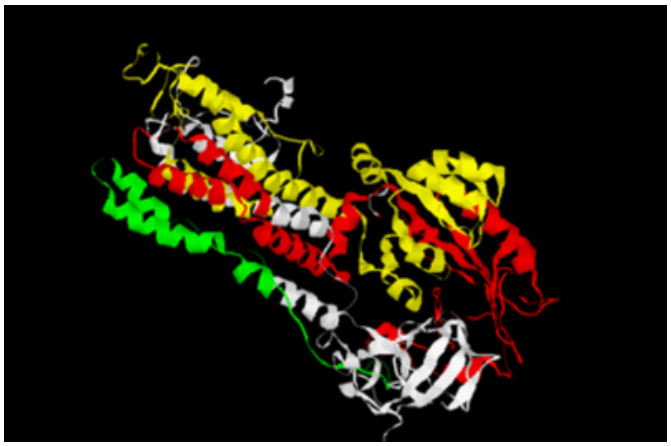


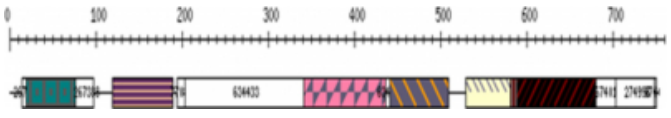
Figure 6

Table 3: Major functional domains of ctATPase of CDC1551. Domains, related bits score and e-values obtained from Pfam server which gives relationship of match sequences from server.

Domain	Start	End	Bits	E-value
HMA	18	79	59.90	7.2e-15
E_1-E_2 ATPas	215	436	298.10	1.4e-86
Hydrolase	440	667	118	2e-32

Figure 7

Figure 4: Sequence segments from PRODOM for ctATPase of CDC1551 displaying with different color codes for different domain families. Obtained from PRODOM server at



DISCUSSION

Iron transport by ctATPase in M. tuberculosis is of great significance as iron is an essential nutrient whose concentration is critical in determining the outcome of infection with pathogenic CDC 1551 strain of M. tuberculosis and hence competition for iron plays a pivotal role in the interaction between host and parasite. Such clues can help in boosting up medical research to combat tuberculosis. Iron is an obligate cofactor for at least 40 different enzymes encoded in the Mycobacterium tuberculosis genome [8]. Cytochromes involved in electron transport, ribonucleotide reducates for DNA synthesis and superoxide dismutase for defense against macrophages are the three important enzymes that require iron [2]. Restricting access of this organism to iron can be adopted as an effective strategy for the treatment of the disease caused by it. M. tuberculosis CDC 1551 has evolved a complex web of biosynthetic and metabolic pathways to obtain, store and regulate the use of iron. Unraveling of the intricacies of this system may help in the development of new drugs for effective regulation of tuberculosis. Further progress towards new therapies to counter mycobacterial infection requires an even greater understanding of the complexities of proteins that interact with iron within the bacterium. Hence it is predicted that characterization of proteins involved in cation transport specifically ctATPase will provide complete understanding of M. tuberculosis iron acquisition machinery. This knowledge should provide new targets for therapeutic intervention against tuberculosis. The 3-D structure generated for ctATPase of M. tuberculosis CDC-1551 is further under investigation in our lab to establish its interaction with small inhibitor molecules (SIM) from ligand database using AUTODOCK and TRIPHOS software, which will results in emerging with new drugs to combat tuberculosis more effectively.

ACKNOWLEDGEMENTS

Authors are grateful to Prof. A. Sali, California Institute for Quantitative Biomedical Research, University of California, San Francisco, for permission to use MODELLER6V2. SKC thanks to UGC-New Delhi for financial support to establish computer networking Lab.

References

1. Altschul S.F., Grish W., Miller W., Myers E. W. and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol.Biol. 215, 403-410.
2. Andrews, S. C.(1998). Iron storage in bacteria. Adv. Microb. Physiol. 40, 281-351.

3. Bajaj, M and Blundell, T. (1984). Evolution and the tertiary structure of proteins. Ann. Rev. Biophys. Bioeng. 13, 453-492.
4. Bateman A., Birney, E., Durbin, R., Eddy, S. R., Howe K L and Sonnhammer, E.L. (2000). The Pfam protein families database. Nucleic Acids Res. 28, 263-266.
5. Bateman A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffith Jones, S., Khanna, A., Marshal, M., Moxin, S.,and Sonnhammer, E.L. (2004). The Pfam protein families database. Nucleic Acids Res. 32, D138-D141.
6. Browne, W.J., North, A.C. and Philips, D.C. (1996). A possible three dimensional structure of bovine alpha-lactalbumin based on that of hen's egg-white lysozyme. J.Mol.Biool.42, 65-86.
7. Bru , C., Coudrcelle, E., Carrere, S., Beause, Y., Dalmar, S and Kahn Daniel. (2005). The ProDom database of protein domain families: more emphasis on 3D. Nucleic Acids Res. 33, D212-D215.
8. Calder, K.M., and Horwitz, M.A. (1998). Identification of iron-regulated proteins of Mycobacterium tuberculosis and cloning of tandem genes encoding a low iron-induced protein and metal transporting ATPase with similarities to two-component metal transport systems. Microb. Pathog 24, 133-143.
9. Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E 3rd., Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver K., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton J., Squares R., Squares S., Sulston J.E., Taylor K., Whitehead S. and Barrell B.G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature (London) 393, 537-544.
10. Dubnau, E., Fontan, P., Manganelli, R., Soares-Appel, S., and Smith, I. (2002) Mycobacterium tuberculosis genes induced during infection of human macrophages. Infect Immun. 70,2787-2795.
11. Fiser. A. and Sali. A. (2003). MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. Methods Enzymol. 374,461-491.
12. Fiser. A., Do, R.K. and Sali. A. (2000). Modeling of loops in protein structure. Protein Sci., 9, 1753-1773.
13. Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn. M., Haft. D, Hickey E, Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg .S.L., Delcher. A, Utterback. T., Weidman. J., Khouri. H., Gill. J., Mikula. A., Bishai. W., Jacobs. Jr. W.R Jr, Venter. J.C. and Fraser C.M. (2002). Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains. J Bacteriol. 184(19), 5479-90.
14. Lambert. C., Leonard.N., De Bolle. X. and Depiereux. E. (2002). ESyPred3D: Prediction of proteins 3D structures. Bioinformatics, 18, 1250-1256.
15. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M., (1993). PROCHECK: a program to check the stereo chemical quality of protein structures. J. Appl. Crystallogr. 26, 283-291.
16. Marcela Rodriguez G. and Issar Smith,(2003) Mechanism of iron regulation in Mycobacteria: role in physiology and virulence.J. Mol.Micro. 47(6),1485-1494.
17. Ramachandran G.N. and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. Adv Protein Chem. 23, 283-438.
18. Sanchez, R and Sali, A. (1997) Advances in comparative protein structure modeling. Curr. Opin. Struct. Biol.,7, 206-214.
19. Sayle. R.A. and Milner White. E.J. (1995). RASMOL: bimolecular graphics for all. Trends Biochem. Sci. 20, 374.
20. Tramontano, A. (1998). Homology modeling with low sequence identity. Methods:14, 293-300.
21. Vriend G. (1999). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. 8, 52-56.
22. Wong. D., Lee, B.Y., Horwitz, M.A. and Gibson, B. (1999). Identification of Fur, aconitase and other proteins expressed by Mycobacterium tuberculosis under conditions of low and high concentrations of iron by combined two-dimensional gel electrophoresis and mass spectrometry. Infect Immun. 67, 327-336.

Author Information

Chitta Suresh Kumar, Ph.D.
Bioinformatics Center, Department of Biochemistry, Sri Krishnadevaraya University

C. M. Anuradha, M.Sc., M.Phil.
Bioinformatics Center, Department of Biochemistry, Sri Krishnadevaraya University

K. Venkateswara Swamy, M.Sc.
Bioinformatics Center, Department of Biochemistry, Sri Krishnadevaraya University

Niamat Ali Khan, M.Sc.
Bioinformatics Center, Department of Biochemistry, Sri Krishnadevaraya University