I J C B
International Journal of Computational Biology
ISSN: 2229–6700, E-ISSN: 2229–6719, Vol. 2, Issue 2, 2011, pp-38-40
Available online at http://www.bioinfo.in/contents.php?id=96

STRUCTURAL ANALYSES OF HEPARIN-BINDING HEMAGGLUTININ (HBHA) OF *MYCOBACTERIUM TUBERCULOSIS* – A POTENTIAL VACCINE

SAYAK GANGULI[1], DEBDOOT GUPTA[2]* AND ABHIJIT DATTA[3]
[1]DBT Centre for Bioinformatics, Presidency University, Kolkata
[2]Dept. of Microbiology, University of Burdwan
[3]Department of Botany, Presidency University, Kolkata
*Corresponding Author: Email- Debdoot Gupta/ dg3004@gmail.com

Received: October 08, 2011; Accepted: November 30, 2011

Abstract- Heparin-binding hemagglitinin (HBHA) has been identified as a surface-exposed adhesin that is responsible for mediating interactions with non-phagocytic cells and subsequent dissemination of the pathogen from the primary infection site. There has been a suggestion that HBHA is a good marker for the immunodiagnosis of latent tuberculosis, and HBHA-specific Th1 responses may contribute to protective immunity against active tuberculosis. This can be established from the fact that latently infected human individuals mount a strong T cell response to HBHA, whereas patients with active disease do not. However the structural analysis of the protein is still a virgin area which is the main focus of this work. Several important structural features were identified through the various methods used and data suggests that the protein is an orphan protein due to its low similarity rate (<10%) with known proteins and there are multiple epitopes comprising the largest binding pockets in the protein structure. This data would enable researchers to successfully design their in vivo studies with the said protein.
Key Words- HBHA, Binding pockets, Orphan protein, Epitope

Introduction

Tuberculosis which is caused by Mycobacterium tuberculosis, still to date accounts for over million deaths every year (1, 2). The developing world is also plagued by Leprosy, a result of infection with Mycobacterium leprae, which remains a significant unsolved health problem (3). Patients having acquired immunodeficiency syndrome are susceptible to the opportunistic infections that are caused by members of the Mycobacterium avium- intracellulare (4, 5). The recent dramatic increase in the number of tuberculosis affected patients in developed countries along with the striking emergence diversity of drug- resistant M. tuberculosis strains (6) poses some of the most challenging areas of disease informatics. Immunoelectron microscopy has been able to establish the site of expression of HBHA to be on the mycobacterial cell surface, and this suggests that the protein is structurally positioned for cell-cell interactions. BLAST searches have not been able to detect suitable homologues of the protein in most cases. Various workers strongly suggest that interactions of sulfated carbohydrates with pathogenic mycobacteria on the surface of cells other than macrophages may contribute to virulence. Although alveolar macrophages have been reported to play a significant role in mycobacterial infection through the lungs, it can be stated that dissemination of the tubercule bacilli to other sites via the lymphatic or circulatory system requires epithelial or endothelial cell attachment. Studies suggest that direct adherence and penetration of alveolar epithelial cells is the route through which M. tuberculosis may indeed gain access to the lymphatic and circulatory systems (7) Methylation pattern of the methylated protein HBHA regulates the antigenicity of the protein in subjects with latent infection and it also controls its protective immunogenicity. Locht et.al (2006) has shown the production of HBHA-specific IFN-γ by both the CD4+ and the CD8+ T cells in both mice and man. Furthermore, the CD8+ T cells which were HBHA-specific were also shown to express bactericidal and cytotoxic activities to mycobacteria- infected macrophages (8). As a whole these observations strongly establishes the potential of methylated HBHA as an important protein vaccine component, against tuberculosis in future.

Discussion

Initial sequence analysis and comprehensive secondary structure analyses revealed that the protein had <10% similarity (14) with other known protein classes; this leads us to conclude that the protein is an orphan protein as no annotated domains were identified in the existing secondary structure databases. The evolution of orphan proteins is a debated topic still and many workers feel that there evolution either occurred far away from the nearest neighbor and thus the problem with their assignment to a domain family or they originated by some de novo mechanism.

Lomino et.al (2011) has investigated the structure of HBHA exists as a dimer in solution using circular dichroism spectroscopy and analytical ultracentrifugation techniques. They have demonstrated that the heparan sulfate (HS) binding region of HBHA does not play a role in dimerization in solution, however, the structure of the linker region that is situated between the predicted N-terminal coiled-coil and the HS binding region at C-terminal affect dimer stability. They also establish that most of the residues responsible for various folding processes such as dimerization, along with the stability of the entire structure lie within the coiled-coil region. The N-terminal helix which precedes the coiled-coil region appears to trigger the folding and dimerization of HBHA. The structure generated reveals the presence of five alpha helical segments connected by coiled coil domains (fig 2). Six tri amino acid repeats were identified bearing the KAA signature. Ohno and Ohno (1986) have discussed the evidence that variations of two small primordial sequences -- the decamer AAGGCTGCTG (=the peptide KAA) and a smaller derivative AAGCTG (=KL) are repeated continuously as primary themes in the sequences of genes, where they are found to alternate with secondary themes composed of different sequences (fig 1). Thus the identification of the repeat can be matched with the general prokaryotic pattern of the organism. Pockets were identified using the Delaunay Triangulation score (12) and five major pockets were observed in the final structure; with the largest pocket having 33 amino acids (table 1)(fig 3). All the hydrophilic amino acids were found to be clustered around the exposed surface of the protein as observed in the spiral plot for the accessible surface area(fig 4) (11).

A large number of B – Cell Epitopes were identified using an SVM based algorithm using various established toxin sequences (endo and exotoxins) and were validated further by comparing the predictions of Bcepred(13) (table 2). The detection of such large number of epitopes signifies the antigenicity of the molecule and further implies the potential use of it as a future protein vaccine.

References

[1] Smith P.G. and Moss A.R. (1994) *Tuberculosis, Pathogenesis, Protection, and Control. B.R. Bloom, editor. ASM Press, Washington*, DC. 47-59.

[2] Raviglione M.C., Snider D.E. and Kochi A. (1995) *Am. Med. Assoc*. 273:220-226.

[3] Noordeen S.K., Bravo U.L. and Sundaresan T.K. (1992) *WHO.* 70:7-10.

[4] Horsburgh C.R. (1991) *N. Engl. J. Med.* 324:1332-1338.

[5] Barnes P.F., Bloch A.B., Davidson P.T. and Snider D.E. (1991) *N. En, J. Med*. 324:1644-1650.

[6] Bloch A.B., Cauthen G.M., Onorato G.M., Dansbury K.G., Kelly G.D., Driver C.R., and Snider D.E. (1994) *Am. Med. Assoc*. 271:665-671.

[7] McDonough K.A. and Kress Y. (1995) *Immun.* 63:4802-481I.

[8] Locht C., Hougard M.J, Rouanet C, Place S and Mascart F. (2006) *Tuberculosis* 86, (3-4), 303-309.

[9] Lomino J.V., Tripathy A., Redinbo M.R. (2011) *J Bacteriol;* 193(9):2089-96.

[10] Ohno S and Ohno M. (1986) *Immunogenetics* 24: 71-78

[11] Ahmad S., Gromiha M., Fawareh H., Sarai A. (2004) *BMC Bioinformatics*. May 1; 5:51.

[12] Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y and Liang J. (2006) *Nucl. Acids Res*., 34:W116-W118.

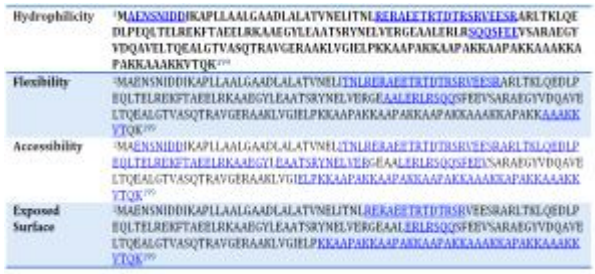
[13] Saha S and Raghava G.P.S. (2004) In G. Nicosia, Cutello V., Bentley P.J. and Timis J. *ICARIS LNCS* 3239, 197-204, Springer, 2004.

[14] Altschul Stephen F, Warren Gish, Webb Miller, Eugene W, Myers and David J Lipman (1990) *J. Mol. Biol.* 215:403-10

Table 1- The 5 Largest Pockets with Their Area & Volumes

ID	AREA	VOL
35	822.8	1016
34	603.1	748.7
33	387.5	429.9
32	132.3	111.1
31	157.5	124.9

Table 2- Predicted B-Cell Epitopes Shown In Blue and Underlined.



Received: October 08, 2011; Accepted: November 30, 2011

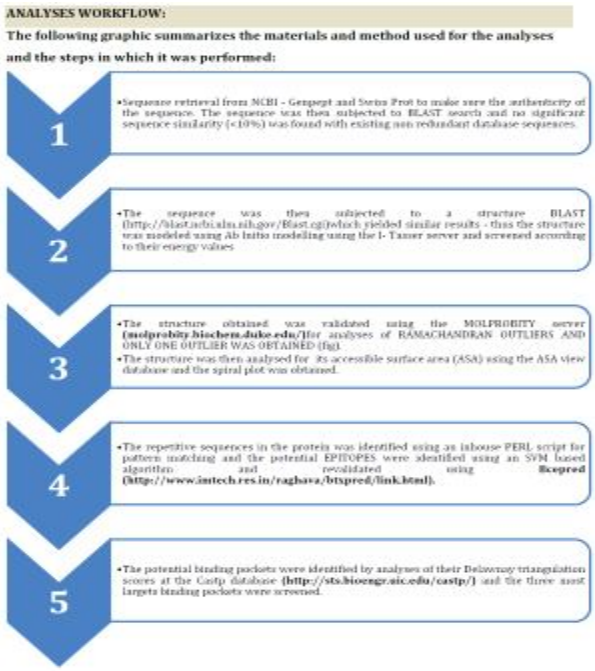


Fig. 1- Flowchart of the Work

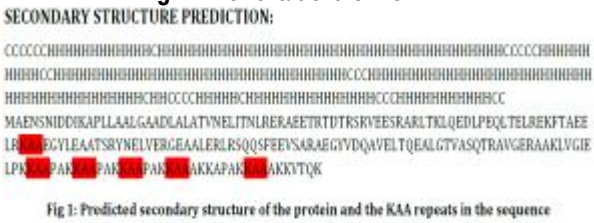


Fig 1: Predicted secondary structure of the protein and the KAA repeats in the sequence

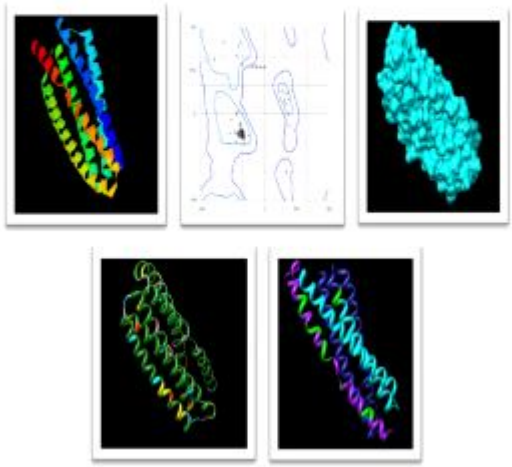


Fig. 2- Tertiary Structure of the Protein: Ribbon Diagram; Ramachandran Plot Showing 1 Outlier (Ala 178) – Surface View and Potential Phosphorylation Sites in the Structure (Serine – Red; Tyrosine – Pink; reonine – Cyan; Arginine (For RNA Interacting Sites) – Yellow); Sites of Kaa Repeats in the Structure (Marked In Green)

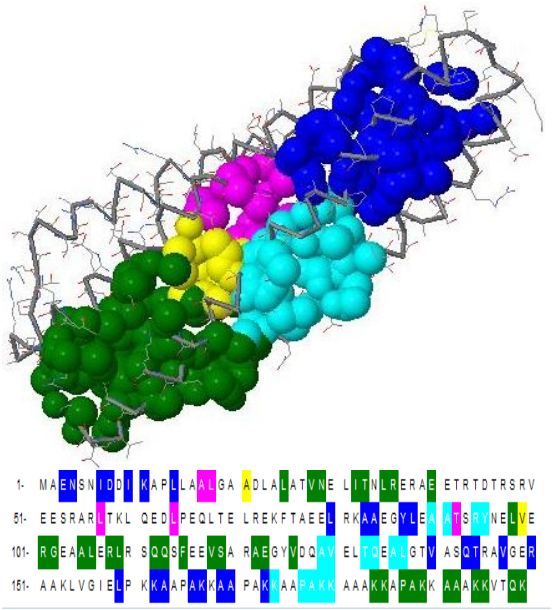


Fig. 3- Binding pockets in the 3D structure of the protein (Calculated using Delaunay TriangulationScore at the CASTp server)

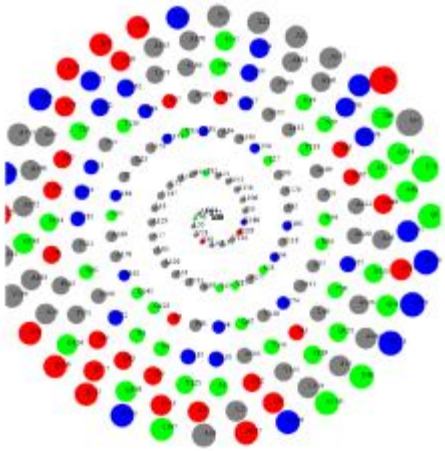


Fig. 4- Accessible Surface area of the residues of the protein (spiral plot – most buried in the middle to most expose on the outside)