Recent Research in Science and Technology 2012, 4(7): 23-26
ISSN: 2076-5061
Available Online: http://recent-science.com/

In-silico analysis of PtpA - an antigenic protein of *Mycobacterium tuberculosis*

Merly D. P., Anil Kumar S., K. Mohan Reddy, Tejaswini B., Darley S.P.[1] and M.B. Sulochana*

Department. of PG Studies and Research in Biotechnology, Gulbarga University, Gulbarga, Karnataka, India.
[1]Head of computer centre, Gulbarga University, Gulbarga.

Abstract
PtpA, a low-molecular weight tyrosine phosphatase, is a secreted protein of *Mycobacterium tuberculosis (Mtb)*. Many secretory proteins of *Mtb* are known to be the prominent targets of host immune response. It plays an important role in host-pathogen interaction and it interferes with the passing of the protein from one endosomal vesicle to the next endosomal vesicle in the infected macrophage. It inhibits host phagolysosomal fusion in the infected macrophages and thus allows the bacteria to survive within macrophages. Analysis of primary and secondary structure of the protein was done by ProtParam and GOR IV respectively. Since the number of negatively charged residues are higher than the positively charged residues, PtpA is an acidic protein. Immunity against *Mtb* is T-cell mediated Thus an important criterion in seeking protective antigens should be that they induce T-cell- mediated immunity. The characterization of PtpA inducing $CD4^+$ T-cell responses could critically contribute to the development of subunit vaccines for *Mtb.* Here we performed computational analysis by using Proped, T-cell epitope prediction program. *In-silico* antigenicity prediction of PtpA was done using VaxiJen. Owing to the resistance of this protein to the natural immune response, *in-silico* antigenicity and T-cell epitope prediction will be helpful to design better subunit vaccines to develop effective acquired immune response to *Mtb.*

Keywords: Secretory proteins, *in-silico*, PtpA, antigenicity, T cell epitope

INTRODUCTION

Tuberculosis is one of the most prevalent infection of human beings. The World Health Organisation has estimated that one third of the total world population is latently infected with *Mtb* and 5-10% of the infected individuals will develop active TB during their life time [1].

Mtb is an intracellular pathogen infecting primarily mononuclear phagocytes. The bacterium has developed intricate strategies to evade killing mechanisms of phagocytes. *Mtb* can also infect nonphagocytic cells in the alveolar space including M cells, alveolar endothelial, and type 1 and type 2 epithelial cells [2,3]. Dendritic cells play a very important role in the early stages of infection since they are much better antigen presenters than are macrophages [4] and presumably play a key role in activating T-cells with specific *Mtb* antigens [5,6]. While the majority of bacterial phagosomes undergo a process of acidification and maturation, *Mtb* inhibits this process [7, 8]. Since most macrophage killing of bacteria occurs in the phagolysosome [9], intracellular pathogens have evolved many ways to avoid this hostile vacuolar microenvironment. It is known that infected macrophages in the lung, through their production of chemokines, attract inactivated monocytes, lymphocytes, and neutrophils [10], none of which kill the bacteria very efficiently [9].

Proteins released by *Mtb* to the extracellular environment have been the focus of much of the research directed at identifying antigens that induce protective immunity or those that elicit immune responses of diagnostic value [11]. Secreted proteins of *Mtb* have been the focus of extensive research on the development of subunit vaccines because these proteins are considered as the important targets of recognition by the immune system [12,13]. The mycobacterial cell envelope has been a prime target for the identification and characterization of surface proteins with potential application in drug and vaccine development.

The secreted Mtb protein tyrosine phosphatase (PtpA) binds to subunit H of the macrophage vacuolar-H+-ATPase machinery, a multisubunit protein complex in the phagosome membrane that drives luminal acidification. Macrophages' engulfment of foreign bodies results in the formation of phagosome, which matures in a process that remodels its membrane and luminal contents through interaction and fusion with the endosomal network [14,15]. These membrane fusions allow the phagosome to acquire antimicrobial properties, including a profoundly acidic lumen, the hallmark of the macrophage maturation process [16]. More than a decade ago, Sturgill-Koszycki *et al*. showed that the macrophages fail to acidify phagosomes containing mycobacteria because these phagosomes do not accumulate the V-ATPase responsible for phagosomal acidification [17]. This finding suggested inhibition of the fusion between membrane vesicles harboring the V-ATPase complex and the mycobacterial phagosome [17].

The exact mechanism by which PtpA is secreted across the phagosomal membrane remains unclear. However, using electron microscopy, neutralizing antibodies, and Western blot analysis, it has been shown that PtpA is present in the host cytosol milieu [18]. There is evidence suggesting that bacterial proteins having molecular size less than 70 kDa can cross the phagosomal membrane. It has also been shown that *Mtb* PtpA (MPtpA) dephosphorylates the host macrophage protein VPS33B, a key

Received: May 08, 2012; Revised: June 19, 2012; Accepted: July 25, 2012.

*Corresponding Author
Dr. M .B. Sulochana
Associate Professor, Dept.of Biotechnology, Gulbarga University,Gulbarga, Karnataka, India.

Tel: +91-9449618676
Email: mbsulochana@rediffmail.com

regulator of membrane fusion, leading to inhibition of phagosome–lysosome fusion [18]. The secreted phosphatases MptpA and MptpB are key virulence factors that play important roles in survival of *Mtb* during macrophage infection. These enzymes are therefore attractive alternative targets for chemotherapy. The analysis of Mtb genome indicated the presence of genes coding for protein tyrosine phosphatases MPtpA and MPtpB.

The physicochemical and the structural properties of the proteins are well understood with the use of computational tools. The statistics about a protein sequence such as number of amino acids, sequence length, and the physico-chemical properties of a protein such as molecular weight, atomic composition, extinction coefficient, GRAVY, aliphatic index, instability index, etc. can be computed by computational tools for the prediction and characterization of protein structure. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function and physical and chemical properties.

Mtb survives phagocytosis and replicates within macrophages. The secretory and cell surface proteins of *Mtb* which stimulates specific host immune response in different cell types like T-cells, B-cells and macrophages control the pathogenesis of the disease. Therefore, the identification of T-cells epitopes from proteins may serve to define candidate antigens with vaccine potential. Following infection with *Mtb*, both healthy subjects and patients with active tuberculosis develop T-cell responses against mycobacterial antigens. The identification of antigens capable of eliciting $CD4^+$ T-cell responses and the characterization of immunodominant T-cell epitopes are of primary importance for the development of subunit vaccines for tuberculosis. The role of $CD8^+$ T-cells in immunity to *Mtb* is less defined [19].

The human leukocyte antigen (HLA) system is the major histocompatibility complex in humans. The primary function of the HLA system is to regulate the immune response [20]. HLA studies have revealed that the allele DRB1*15 is associated with tuberculosis in Indians [21].

MATERIALS AND METHODS

Primary sequence analysis

The sequence of *Mtb* PtpA of accession number P65716 was retrieved from UniProt (http://www.uniprot.org/), a protein database, in FASTA format. The primary structure, the basic physico-chemical properties, of PtpA was analysed using ProtParam (http://web.expasy.org/protparam/ (Gasteiger E). ProtParam is available through the ExPASy server.

Identification of domains

PtpA sequence was analysed using SMART (Simple Modular Architecture Research Tool) and Pfam. SMART is an online resource (http://smart.embl.de/) for the identification and annotation of protein domains and analysis of protein domain architecture. Pfam is a widely used database of protein families and domains. Pfam is available via (http://pfam.sanger.ac.uk/) and is used to organise sequences, to find the origin and evolution of proteins and for identifying interesting new targets for the structure determination.

Prediction of function

The biological function of PtpA was studied using ESG (http://kiharalab.org/web/esg.php), a sequence similarity based protein function prediction server.

Secondary structure prediction

GOR IV (http://npsapbil.ibcp.fr/cgibin/npsam) was used to predict the secondary structure of PtpA. The present version, GOR IV uses all possible pair frequencies within a window of 17 amino acid residues.

Antigenicity prediction

The server VaxiJen (http://www.jenner.ac.uk/VaxiJen) was used to analyse the antigenicity of the query protein. VaxiJen was developed to allow antigen classification based on the physicochemical properties of proteins. Protein sequences can be submitted as single protein or as a multiple sequence file in fasta format.

T-cell epitope prediction

Prediction of human HLA-DR-restricted determinants within this antigen was performed using a virtual matrix-based prediction program (ProPred). The ProPred analysis [22] of this protein was performed using the server (http:www. imtech. res.in/ raghava/ ProPed/).

RESULTS AND DISCUSSION

The low-molecular-weight protein tyrosine phosphatase belongs to a distinctive class of phosphotyrosine phosphatases (PTP). PTPs have long been considered to be confined to eukaryotes and only recently the genes encoding PTPs been found in bacteria. A low molecular weight PTP, YwlE, was identified from *Bacillus subtilis.* YwlE shares a low sequence identity (31%) with LMW PTP of *Mtb* [23]. MptpA shows 37% sequence identity and high structural similarity to the human LMW PTP.

Primary sequence analysis

The sequence of MPtpA, as given below, contains 163 AA with molecular weight of 17,892 Da.

>sp|P65716|PTPA_MYCTU Probable low molecular weight protein-tyrosine-phosphatase OS=*Mtb* GN=ptpA PE=1 SV=1

MSDPLHVTFVCTGNICRSPMAEKMFAQQLRHRGLGDAVRVTSAG
TGNWHVSCADERAAGVLRAHGYPTDHRAAQVGTEHLAADLLVAL
DRNHARLLRQLGVEAARVRMLRSFDPRSGTHALDVEDPYYGDHS
DFEEVFAVIESALPGLHDWVDERLARNGPS

The parameters shown by ProtParam in Table 1 reveal the physicochemical properties of the protein. Some amino acids like alanine, arginine, aspartic acid, glycine, leucine are present in high percentage. Lysine is present in lowest percentage and the number of negatively charged residues are slightly higher than positively charged residues. Earlier studies showed that the calculated molecular masses of two putative tyrosine phosphatases, designated as MPtpA and MPtpB ,in *Mtb* were 17.5 and 30 kDa respectively.

Table 1. Physicochemical properties of PtpA shown by ProtParam

No. of amino acids	163
Molecular weight	17892 Da
Total number of negatively charged residues (Asp + Glu)	22
Total number of positively charged residues (Arg + Lys)	17
Total no. of atoms	2470
Theoretical pI	6.03
Estimated half-life (mammalian reticulocytes, in vitro).	30 hours
Instability index (unstable)	42.47
Aliphatic index	80.25
Grand average of hydropathicity (GRAVY)	-0.330

Identification of domains

SMART and PFAM identifies the specific region that encodes the domain as shown in Table 2.1 and 2.2 respectively. SMART revealed that the domain LMWPc is present between amino acids 5 and 152. Pfam classified the domain in PfamA family group with id PF01451.

Table 2.1. Region identified by SMART that encodes the domain

Name	Begin	End	E-value
LMWPc	5	152	3.80e-62

Table 2.2. Region identified by PFAM that encodes the domain

Source	Domain	Start	End
PfamA	LMWPc	7	150

In human, there are 38 known Protein Tyrosine Phosphatases, which are subdivided into receptor and non-receptor. The non-receptor phosphatases contain a single catalytic Protein Tyrosine Phosphatase domain.The receptor tyrosine phosphatases are localized to the plasma membrane and have one or two Protein Tyrosine Phosphatase domains and a characteristic transmembrane domain. Low molecular weight phosphatases lack regulatory domains, unlike other tyrosine phosphatases that contain both catalytic and regulatory domains [24].

Prediction of function

ESG reveal the molecular function, biological function and the percentage distribution of the Protein in different cellular components as shown in Table 3.1,3.2 and 3.3 respectively. It predicted that 99 percent molecular function is on both phosphatase and hydrolase activity. The main biological function of the protein is amino acid dephosphorylation and it is abundantly present in the cytoplasm than other cellular components.

Table 3.1. Molecular function terms revealed by ESG

Sl No.	Probability	Term	Description
1	99%	GO:0004721	Phosphoprotein phosphatase activity
2	99%	GO:0004725	Protein tyrosine phosphatase activity
3	99%	GO:0016787	Hydrolase activity
4	83%	GO:0003993	Acid phosphatase activity
5	82%	GO:0004726	Non-membrane spanning protein tyrosine phosphatase activity
6	10%	GO:0042802	Identical protein binding
7	9%	GO:0016791	Phosphoric monoester hydrolase activity

Table 3.2. Biological Process Terms revealed by ESG

Sl No.	Probability	Term	Description
1	99%	GO:0006470	Protein amino acid dephosphorylation
2	9%	GO:0007165	Signal transduction
3	9%	GO:0051726	Regulation of cell cycle

Table 3.3. Cellular component terms revealed by ESG

Sl No.	Probability	Term	Description
1	83%	GO:0005737	Cytoplasm
2	10%	GO:0005625	Soluble fraction
3	9%	GO:0005634	Nucleus
4	9%	GO:0005829	Cytosol
5	83%	GO:0005737	Cytoplasm

Secondary structure prediction

The secondary structure prediction defined each residue into either alpha helix, beta sheet or random coil secondary structures. GOR IV analysis revealed that alpha helix is more than beta strand and random coil, as shown in Table 4.

Table 4. GOR IV analysis results

Secondary structure	No. of residues	Percentage
Alpha helix	81	49.69%
Beta sheet	18	11.04%
Random coil	64	39.2%

Antigenicity prediction

VaxiJen classified PtpA as protective antigen because it has a value of 0.5014 which is above the normal threshold value of 0.5.

T cell epitope prediction

Table 5. The predicted epitopes using the ProPred

Predicted HTL epitopes sequences
LVALDRNHA
VRM
LRSFDPRSG

ProPred, as shown in Table 5, predicted three T cell epitopes with threshold setting as three and number of alleles as one. *Mtb* contain around four thousand proteins. ProPred could identify three T-cell epitopes from PtpA that can induce immune response. The predicted epitopes may be used for safe subunit vaccine development against tuberculosis. Therefore, there is a clear need for programs which can effectively reduce the number of candidates inducing an immune response. These results demonstrate that the ProPred analysis can facilitate the selection of promiscuous peptides recognized by Th1 cells, and thus it can be useful in the identification of peptide-based vaccine candidates against TB. T-cell epitope prediction on some of the secretory proteins has been done earlier by [25].

CONCLUSION

The studies on the secreted proteins in *Mtb* help in the immunological characterization of these proteins to define their potential for immunological diagnosis of tuberculosis or vaccine design. Identifying molecules present on the surface of the organism

that can act as ligand and help in the entry of the pathogen into the host cell. Immune response developed against these proteins can stop receptor-ligand interactions and prevent invasion into host cells. Prediction of antigenic epitopes on the protein surfaces is important for vaccine design and only a few structure-based epitope prediction algorithms are available and they have not yet shown satisfying performance.

REFERENCES

[1] World Health Organization, 2009. "Global tuberculosis control: surveillance, planning and financing," WHO/HTM/TB/2009.411, WHO, Geneva, Switzerland.

[2] Bermudez L. E., F. J. Sangari, P. Kolonoski, M. Petrofsky, and J. Goodman. 2002.The efficiency of the translocation of *Mycobacterium tuberculosis* across a bilayer of epithelial and endothelial cells as a model of the alveolar wall is a consequence of transport within mononuclear phagocytes and invasion of alveolar epithelial cells, *Infection and Immunity*, vol. 70, no. 1, pp. 140–146

[3] Kinhikar A. G., I. Verma, D. Chandra *et al*., 2010. Potential role for ESAT6 in dissemination of *Mtb* via human lung epithelial cells, *Molecular Microbiology*, vol. 75, no. 1, pp. 92–106.

[4] Tascon, R. E., C. S. Soares, S. Ragno, E. Stavropoulos, E. M. Hirst, and M. J. Colston 2000. *Mycobacterium tuberculosis*-activated dendritic cells induce protective immunity in mice. *Immunology* 99:473-480.

[5] Gonzalez-Juarrero, M., and I. M. Orme. 2001. Characterization of murine lung dendritic cells infected with *Mycobacterium tuberculosis*. Inf*ect. Immun*. 69:1127-1133.

[6] Bodnar, K. A., N. V. Serbina, and J. L. Flynn. 2001. Fate of *Mycobacterium tuberculosis* within murine dendritic cells. *Infect. Immun*. 69:800-809.

[7] Vergne, I., J. Chua, S. B. Singh, and V. Deretic. 2004. Cell biology of *Mycobacterium tuberculosis* phagosome. *Annu. Rev. Cell Dev. Biol.* 20:367–39.

[8] Jutras, I. D. M. 2005. Phagocytosis: at the crossroads of innate and adaptive immunity. *Annu. Rev. Cell Dev. Biol.* 21:511–527.

[9] Fenton, M. J., and M. W. Vermeulen. 1996. Immunopathology of tuberculosis: roles of macrophages and monocytes. *Infect. Immun.* 64:683-690.

[10] Van Crevel, R., T. H. Ottenhoff, and J. W. van der Meer. 2002. Innate immunity to *Mycobacterium tuberculosis*. *Clin. Microbiol. Rev.* 15:294-309.

[11] Andersen A B, Brennan P. 1994. Proteins and antigens of *Mycobacterium tuberculosis*. In: Bloom B R, editor. Tuberculosis: pathogenesis, protection, and control. Washington, D.C.: American Society for Microbiology; pp. 307–327

[12] Andersen P. 1997. Host responses and antigens involved in protective immunity *Mycobacterium tuberculosis*. *Scand J Immunol* 45: 115–131.

[13] Mustafa A S. 2001. Biotechnology in the development of new vaccines and diagnostic reagents against tuberculosis. Curr *Pharm Bio Technol* 2: 157-173.

[14] Bright NA, Gratian MJ, Luzio JP. 2005. Endocytic delivery to lysosomes mediated by concurrent fusion and kissing events in living cells. *Curr Biol* 15:360–365.

[15] Sun-Wada GH, Tabata H, Kawamura N, Aoyama M, Wada Y. 2009. Direct recruitmentof H+-ATPase from lysosomes for phagosomal acidification. *J Cell Sci* 122:2504–2513.

[16] Tjelle TE, Lovdal T, Berg T. 2000. Phagosome dynamics and function. *Bioessays* 22:255–263.

[17] Sturgill-Koszycki S, *et al*., 1994. Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase. *Science* 263:678–681

[18] Bach H, Papavinasasundaram KG, Wong D, Hmama Z, Av-Gay Y. 2008. *Mycobacterium tuberculosis* virulence is mediated by PtpA dephosphorylation of human vacuolar protein sorting 33B. *Cell Host Microbe* 3:316–322

[19] Flynn, J. L., and J. Chan. 2001. Immunology of tuberculosis. *Annu. Rev.Immunol.* 19:93–129.

[20] Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, *et al*., 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329: 512-518

[21] Sriram U, Selvaraj P, Kurian SM, Reetha AM, *et al*., 2001. HLA-DR2 subtypes & immune responses in pulmonary tuberculosis. *Indian J. Med. Res.* 113: 117-124.

[22] Singh, H., and G. P. Raghava. 2001. ProPred: prediction of HLA-DR binding sites.*Bioinformatics* 17**:**1236-1237. [PubMed]

[23] Huimin Xu, Bin Xia and Changwen Jin *J. Bacteriol.* 2006, 188(4):1509. DOI:10.1128/JB.188.4.1509-1517.2006

[24] Fauman E B, Saper M A. 1996. Structure and function of the protein tyrosine phosphatases. *Trends Biochem Sci.*21:413–417

[25] Pallavi S, Vijai S, PK Seth. 2008. *In Silico* Prediction of Epitopes in Virulence Proteins of *Mycobacterium Tuberculosis* H37Rv for Diagnostic and Subunit Vaccine Design. *J Proteomics Bioinform* 1: 143-153. doi:10.4172/jpb.1000020.