International Journal of Biological Macromolecules 86 (2016) 646–655

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Mycobacterium tuberculosis PE9 protein has high activity binding peptides which inhibit target cell invasion

Diana P. Díaz[a,b], Marisol Ocampo[a,b,*], Laura Pabón[a,b], Chonny Herrera[a,b], Manuel A. Patarroyo[a,b], Marina Munoz[a,b], Manuel E. Patarroyo[a,c]

[a] *Fundación Instituto de Inmunología de Colombia (FIDIC), 111321 Bogotá, Colombia*
[b] *Universidad del Rosario, 111321 Bogotá, Colombia*
[c] *Universidad Nacional de Colombia, 11001 Bogotá, Colombia*

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 3 December 2015
Accepted 26 December 2015
Available online 2 February 2016

Keywords:
Tuberculosis
Synthetic peptide
PE9 protein

ABSTRACT

PE/PPE proteins are involved in several processes during *Mycobacterium tuberculosis* (*Mtb*) infection of target cells; studying them is extremely interesting as they are the only ones from the *Mycobacterium* genus, they abound in pathogenic species such as *Mtb* and their function remains yet unknown. The PE9 protein (Rv1088) was characterised, the *rv1088* gene was identified by PCR in *Mtb* complex strains and its expression and localisation on mycobacterial surface was confirmed by Western blot and immunoelectron microscopy. Bioinformatics tools were used for predicting PE9 protein structural aspects and experimental study involved the circular dichroism of synthetic peptides. The peptides were tested in binding assays involving U937 and A549 cells; two high activity binding peptides (HABPs) were found for both cell lines (39226-^{1}MSYMIATPAALTAAATDIDGI^{21} and 39232-^{125}YQRHFGTGGQPEFRQHSEHRR^{144}), one for U937 (39231-^{104}YAGAGRRQRRRRSGDGQWRLRQ^{124}) and one for A549 (39230-^{83}YGTGVFRRRRGRQTVTAAEHRA^{103}). HABP 39232 inhibited mycobacterial entry to A549 cells (~70%) and U937 cells (~50%), peptides 39226 and 39231 inhibited entry to U937 cells (~60% and 80%, respectively) and peptide 39230 inhibited entry to A549 cells (~60%). This emphasised HABPs' functional importance in recognition between *Mtb* H37Rv and target cell receptors. These peptide sequences could be involved in invasion and were recognised by the host's immune system, thereby highlighting their use when designing an efficient anti-tuberculosis multiantigenic vaccine.

1. Introduction

In spite of worldwide efforts at reducing tuberculosis (TB), this disease, caused by the bacillus *Mycobacterium tuberculosis* (*Mtb*), continues being the second cause of death by infection after HIV and millions of people die from it annually; the most recent statistics have reported 9.6 million new cases and 1.5 million deaths in 2014 [1].

Worldwide research has been focused on the search for alternatives aimed at eliminating this disease by 2050 as a worldwide public health problem and it has been recognised that the main strategy lies in developing an effective anti-TB vaccine [2], taking into account that the vaccine derived from the attenuated strain of *Mycobacterium bovis* Bacille Calmette-Guérin (BGC) (developed in 1921 and having very variable efficacy against pulmonary TB) is ineffective in adults [3]. It has been proposed that vaccines against TB should be made from attenuated mycobacterial strains, such as *Mycobacterium vaccae* [4], from recombinant proteins, such as Aeras 402/Ad35-85B-TB10.4 which uses a replication deficient adenovirus for expressing a fusion of *M. tuberculosis* antigens 85A, 85B and TB10.4 [5], or subunit-based ones, such as MVA85A which uses the modified vaccinia virus Ankara (MVA) expressing the 85A antigen which has been the most advanced one to date. However, it has been shown that the latter did not protect against the disease nor reduce contagion, according to an assay involving more than 2700 children in South Africa [6]. This means that none of those mentioned above have shown greater efficacy than that presented by the current vaccine to date.

Developing an efficient vaccine requires detailed knowledge of the pathogen's biology, recognising the specific antigens used in infection and those which could induce a protection-inducing immune response. Consequently, our approach has been based on the search for potential antigens which are significant from a

* Corresponding author at: Fundación Instituto de Inmunología de Colombia (FIDIC), 111321 Bogotá, Colombia. Fax: +57 1 3158920X120.
E-mail address: marisol.ocampo@urosario.edu.co (M. Ocampo).

http://dx.doi.org/10.1016/j.ijbiomac.2015.12.081
0141-8130/

functional point of view, mainly *Mtb* cell surface proteins participating in host–pathogen interaction and which could also be involved in intracellular pathogenicity and multiplication [7,8]. These amino acid (aa) sequences must be recognised by the immune system to counteract infection by the mycobacteria; experience gained using sequences from *Plasmodium falciparum* parasite-derived proteins has shown that binding sequences are not immunogenic and require modifications to their sequences to reach the desired levels of protection and antigenicity when designing an efficient vaccine [9,10]. This approach has been adopted for studying different *Mtb* H37Rv proteins, identifying peptide sequences specifically interacting with target cells for infection and which inhibit mycobacterial entry in *in vitro* assays [11].

The *rv1088* gene was selected for this study; it encodes the PE9 (Rv1088) surface protein and forms part of the two sets of highly polymorphic genes from mycobacteria (known as the PE and PPE families) which are particularly abundant in pathogenic mycobacterial strains such as *Mtb*. These families account for around 10% of the mycobacterial genome, involving 99 PE genes and 69 PPE genes in *Mtb* H37Rv [12]. The PE protein's name is due to the N-terminal proline-glutamic acid (PE) and proline-proline-glutamic acid (PPE) conserved domains; these are of great interest due to their immunogenic potential [13,14]. It has been suggested that they contribute towards the bacillus' virulence [15,16]; this is related to this type of protein being associated with the mycobacterial membrane or them being secreted [17–21], meaning that they are functionally important in host-pathogen interaction. Some PE/PPE proteins are selectively induced in macrophages for modulating their function, thereby favouring intracellular bacillus survival [15].

The PE9 protein's function is unknown; it has a 15.36 kDa molecular weight and its sequence consists of 144 aa. It has been presumed that it is a mycobacterial membrane protein by prediction and two-dimensional electrophoresis and mass spectrometry (2DLC–MS/MS) since three PE9 protein peptides were found in this cell fraction in an evaluation of *Mtb* subcellular profile [22]; the SGDGQWRLR sequence was found using liquid chromatography coupled to mass spectrometry (LC–MS/Ms), covering this protein's residues 115–123 in culture supernatant [23]. On the other hand, the protein has been identified by LC–MS in guinea pig lungs 90 days after they had been infected by *Mtb* [24]. The PE9 RLFNANAEEYHALSA peptide has been identified as covering residues 53–67 as a $CD8^+$ T-cell epitope; this has led to proposing their inclusion in TB diagnostic tests based on T-cell response and as a vaccine candidate [13]. A recent study found that this protein forms an atypical heterodimer with the PE10 protein which interacts with TLR4 receptor, induces THP-1 macrophage apoptosis and is likely to play a crucial role in *Mtb* immune evasion strategies [25].

The foregoing suggests that the PE9 protein could play an important role in intracellular bacillus infection and persistence, meaning that it is a promising antigen for being included in the design of a synthetic, multiantigenic anti-TB vaccine. Such approach would involve following the proposed methodology for developing chemically-synthesised vaccines which could be applicable regarding different infectious diseases [10,26] and has been based on using synthetic peptides covering the protein's complete sequence for identifying peptide sequences being capable of high specific binding to target cells of infection (HABPs) and which also inhibit mycobacterial entry in *in vitro* assays. It has been found that functionally important sequences could not be recognised by the host's immune system; however, when precise modifications are made to these sequences' three-dimensional structure, their immunogenic nature has become enhanced.

2. Materials and methods

2.1. Bioinformatics analysis of the PE9 protein

The *M. tuberculosis* H37Rv PE9 protein sequence was obtained from the NCBI database at http://www.ncbi.nlm.nih.gov/protein/CCP43840.1. The sequence was obtained in FASTA format for subsequent bioinformatics analysis.

Given the importance of studying the proteins present in different pathogenic strains of *Mycobacterium*, an *in silico* analysis was made for determining homology between the PE9 protein sequence reported for *Mtb* H37Rv and that found in different mycobacterial species and strains using the Basic Local Alignment Search Tool (BLAST) http://blast.ncbi.nlm.nih.gov/Blast.cgi [27].

Bearing in mind that our interest lay in studying *Mtb* H37Rv surface proteins, the PE9 protein's subcellular localisation was predicted with TBpred http://www.imtech.res.in/raghava/tbpred/ [28], PSORTb v 3.0 http://www.psort.org/psortb/ [29] and Gpos-PLoc tools http://www.csbio.sjtu.edu.cn/bioinf/Gpos-multi/ [30].

The presence of a signal sequence characteristic for proteins which are secreted or exported to the surface by classical route were analysed using SignalP 3.0 http://www.cbs.dtu.dk/services/SignalP-3.0/, [31] and Phobius tools http://phobius.sbc.su.se/index.html [32]. Topological characterisation included using TMHHMM version 2.0 http://www.cbs.dtu.dk/services/TMHMM/ [33], TMPred http://www.ch.embnet.org/software/TMPRED_form.html [34], SPLIT 4.0 http://split4.pmfst.hr [35] and PRED-TMR tools http://athina.biol.uoa.gr/PRED-TMR/input.html [36] for identifying the presence of transmembrane domains in the protein's sequence.

The protein's secondary structure was predicted with PSIPRED http://bioinf.cs.ucl.ac.uk [37] and RaptorX tools http://raptorx.uchicago.edu [38] and the tertiary structure was obtained using the Robetta protein structure prediction server [39] http://robetta.bakerlab.org; 3D models were validated using the RAMPAGE server, http://mordred.bioc.cam.ac.uk/~rapper/rampage.php whose evaluations were based on torsion angles and the protein's geometry [40], and the QMEAN server, http://swissmodel.expasy.org/qmean/cgi/index.cgi, a composite scoring function which can derive both global and local error estimates on the basis of a single model [41]. The 3D model was designed by using visual molecular dynamics (VMD) structural modelling and visualisation software http://www.ks.uiuc.edu/Research/vmd/ [42].

2.2. rv1088 *gene presence and transcription*

The *rv1088* gene's presence and transcription were determined in the following mycobacterial species and strains: *M. tuberculosis* H37Rv (ATCC 27294), *M. tuberculosis* H37Ra (ATCC 25177), *M. bovis* (ATCC 19210) and *M. bovis* BCG (ATCC 27291, Pasteur substrain). All mycobacteria were grown for 5–15 days in Middlebrook 7H9 (Difco Laboratories, Detroit, MI, USA) with oleic albumin dextrose catalase (OADC) growth supplement and then incubated at their optimum temperature until cultures reached 0.5–1.0 OD_{600}. Mycobacteria were harvested at mid- to late-log phase culture by spinning at 12,500 × g for 20 min at 4 °C, suspended in PBS and stored at −20 °C.

An Ultra Clean Microbial DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) was used for isolating genomic DNA (gDNA) from mycobacterial species and strains, following the manufacturer's instructions. Extracted DNA quality was assessed by PCR amplification with primers targeting the *hsp65* gene, (Tb11: 5′-ACCAACGATGGTGTGTCCAT-3′ and Tb12: 5′-CTTGTCGAACCGCATACCCT-3′). The rv1088 gene presence and transcription in *in vitro* culture conditions were assessed using the following primers: rv1088-sense: 5′-ATACATGATTGCCACACCAG-3′ and rv1088-antisense: 5′- TGCTCGCTGTGTTGTCGG -3′.