

Functional Dissection of the PE Domain Responsible for Translocation of PE_PGRS33 across the Mycobacterial Cell Wall

Alessandro Cascioferro^{1#a}, Maria H. Daleke^{2,3#}, Marcello Ventura^{1#}, Valentina Donà^{1#b}, Giovanni Delogu⁴, Giorgio Palù¹, Wilbert Bitter^{2,3}, Riccardo Manganeli^{1*}

1 Department of Histology, Microbiology and Medical Biotechnologies, University of Padua, Padua, Italy, **2** Department of Medical Microbiology and Infection Control, VU University Medical Centre, Amsterdam, The Netherlands, **3** Department of Molecular Microbiology, VU University, Amsterdam, The Netherlands, **4** Institute of Microbiology, Catholic University of the Sacred Heart, Rome, Italy

Abstract

PE are peculiar exported mycobacterial proteins over-represented in pathogenic mycobacterial species. They are characterized by an N-terminal domain of about 110 amino acids (PE domain) which has been demonstrated to be responsible for their export and localization. In this paper, we characterize the PE domain of PE_PGRS33 (PE_{Rv1818c}), one of the best characterized PE proteins. We constructed several mutated proteins in which portions of the PE domain were deleted or subjected to defined mutations. These proteins were expressed in different mycobacterial species and their localization was characterized. We confirmed that the PE domain is essential for PE_PGRS33 surface localization, and demonstrated that a PE domain lacking its first 30 amino acids loses its function. However, single amino acid substitutions in two regions extremely well conserved within the N-terminal domain of all PE proteins had some effect on the stability of PE_PGRS33, but not on its localization. Using *Mycobacterium marinum* we could show that the type VII secretion system ESX-5 is essential for PE_PGRS33 export. Moreover, in *M. marinum*, but not in *Mycobacterium bovis* BCG and in *Mycobacterium tuberculosis*, the PE domain of PE_PGRS33 is processed and secreted into the culture medium when expressed in the absence of the PGRS domain. Finally, using chimeric proteins in which different portions of the PE_{Rv1818c} domain were fused to the N-terminus of the green fluorescent protein, we could hypothesize that the first 30 amino acids of the PE domain contain a sequence that allows protein translocation.

Citation: Cascioferro A, Daleke MH, Ventura M, Donà V, Delogu G, et al. (2011) Functional Dissection of the PE Domain Responsible for Translocation of PE_PGRS33 across the Mycobacterial Cell Wall. PLoS ONE 6(11): e27713. doi:10.1371/journal.pone.0027713

Editor: Olivier Neyrolles, Institut de Pharmacologie et de Biologie Structurale, France

Received: August 10, 2011; **Accepted:** October 23, 2011; **Published:** November 16, 2011

Copyright: © 2011 Cascioferro et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°201762 and by Ministero dell'Università e della Ricerca Scientifica (MIUR), Progetti di Interesse Nazionale (PRIN) 2008 (code number 2008Y8RZTF). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: riccardo.manganeli@unipd.it

#a Current address: Integrated Mycobacterial Pathogenomics Unit, Institut Pasteur, Paris, France

#b Current address: Tuberculosis Research Section, Laboratory of Clinical Infectious Disease, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, United States of America

These authors contributed equally to this work.

Introduction

PE, together with PPE, are peculiar mycobacterial proteins over-represented in pathogenic mycobacterial species. Despite lacking typical secretion signals, both PE and PPE are secreted or located in the mycobacterial envelope [1,2,3,4,5]. PE proteins can be divided into three distinct subfamilies, of which the most abundant is represented by PE_PGRS. All PE proteins are characterized by an N-terminal highly conserved domain of about 110 amino acids, named PE after the conserved signature motif Pro-Glu (PE) present near the N-terminus. In PE_PGRS proteins, the PE domain is followed by a C-terminal domain with a highly variable Gly-Ala rich sequence [6,7], which has been suggested to be involved in antigenic variation [8,9]. In the other PE proteins the PE domain can be followed by an unrelated C-terminal domain, or the PE domain represents the entire protein [7]. In the latter case the PE-encoding gene is usually in tandem

with a PPE-encoding gene, and at least in one case the PE and PPE domains encoded by the coupled genes have been shown to interact [10,11,12]. Hardly any of the about 100 PE proteins encoded by the *Mycobacterium tuberculosis* genome have been associated with a physiological function, with the exceptions of LipY (Rv3097c), whose C-terminal domain shows lipase activity [13], PE_PGRS11, which was recently shown to encode a functional phosphoglycerate mutase [14] and PE_PGRS33, which might be involved in induction of macrophage necrosis and apoptosis through interaction with Toll-like receptor 2 [15,16].

We recently showed that PE_PGRS33 is surface exposed when expressed in *Mycobacterium smegmatis* and that its PE domain contains the information necessary for the surface exposure. Chimeric proteins based on this PE domain were expressed on the *M. smegmatis* and *Mycobacterium bovis* BCG cell surface [1,17,18], and this feature was used to develop a surface delivery system to

express heterologous antigen on *M. bovis* BCG envelope and increase its immunogenic potential [18].

As already mentioned, although most PE and PPE proteins lack classical secretion signals, many are exported to the mycobacterial surface, suggesting the involvement of a novel secretory pathway. Interestingly, in *Mycobacterium marinum* their secretion has recently been linked to ESX-5, a member of a novel family of secretion systems typical of mycobacteria (type VII secretion systems) [2,19,20].

In this paper, we show that PE_PGRS33 secretion in *M. marinum* is ESX-5 dependent, and by characterizing the cellular localization of several PE_PGRS33 mutants and PE-based chimeric proteins in *M. smegmatis*, *M. marinum*, *M. bovis* BCG and *M. tuberculosis* we identify portions of the PE domain that are required for protein translocation.

Materials and Methods

Bacterial strains, media and growth condition

M. smegmatis mc²155 [21], *M. bovis* BCG and *M. tuberculosis* were grown at 37°C. *M. marinum* wild-type strain E11 [22] and its ESX-5 mutant 7C1 [2] were grown at 30°C. All mycobacterial strains were grown in Middlebrook 7H9 broth or on 7H10 solid medium (Difco Becton-Dickinson), supplemented with 0.2% glycerol (Sigma-Aldrich), ADC 10% (Becton-Dickinson), and 0.05% v/v Tween 80 (Sigma-Aldrich). In *M. marinum* secretion experiments, cells grown to mid-logarithmic phase were washed to remove BSA (part of the ADC supplement). Washed cells were inoculated to an OD₆₀₀ of 0.3 in Middlebrook 7H9 supplemented with 0.2% dextrose and 0.05% Tween, and grown for an additional 24 hours. Strains processed for the proteinase K assay or cell fractionation were grown in Sauton's medium (Difco) for 14 hours from a starting OD₆₀₀ of 0.1. For cloning procedures *Escherichia coli* strains HB101 and DH5 α were grown in Luria-Bertani medium (LB) [23]. Hygromycin (Roche) was used at a final concentration of 100 $\mu\text{g ml}^{-1}$ (solid media) or 50 $\mu\text{g ml}^{-1}$ (liquid media) for *M. smegmatis*, *M. marinum*, *M. bovis* BCG, and *M. tuberculosis*, and at a final concentration of 200 $\mu\text{g ml}^{-1}$ for *E. coli*. Kanamycin was used at a concentration of 25 $\mu\text{g ml}^{-1}$ for *M. marinum*.

DNA manipulation

All genes expressed in this work were amplified with Pfu DNA polymerase (Stratagene). Genes encoding HA-tagged proteins were amplified using upper primers designed to contain an NheI site immediately before the start codon, and lower primers designed to contain the HA-coding sequence in frame with the coding sequence of the gene of interest, followed by a stop codon and a BamHI site (Table S1). Mutations in the PE-encoding region were introduced by amplifying the gene with primers containing the mutations (Table S1). The resulting constructs were inserted in the mycobacterial expression vectors pMV10-25 and/or pMV4-36 [24] after NheI and BamHI digestion. For expression in *M. marinum*, fragments containing the P_{hsp60} or P_{Rv1818c} promoter sequence, the PE_PGRS33 encoding gene fragment and the in frame sequence coding for the HA tag were transferred to the pSMT3 vector [20,25].

To obtain translational fusions between the coding region of the N-terminal part of PE_PGRS33 and GFP, different portions of the sequence encoding the PE domain of PE_PGRS33 were amplified using an upper primer containing an XbaI site before the start codon of Rv1818c, and different lower primers containing an in frame XbaI site after the last codon of the PE-encoding region to be included in the construct. The resulting DNA fragments were subsequently cloned in frame with the

GFP-encoding gene present in pMV10-25, after digestion with XbaI. Subsequently, DNA fragments containing the promoter and the chimeric genes sequence were extracted from the original plasmids and subcloned in the integrative vector pMV306 [26].

Electroporation

Electroporation of mycobacteria was performed as previously described [17,27]. Briefly, cells grown to mid-exponential in phase were extensively washed in 10% glycerol and concentrated approximately 40-fold. 50 or 100 μl concentrated cells were mixed with 1 μg of DNA, and transferred to 0.2 cm gap cuvettes (Eppendorf). Samples were electroporated using an Electroporator 2510 (Eppendorf; capacitance 10 μF ; voltage 12.5 kV cm^{-1} ; resistance 600 Ω) (for *M. tuberculosis*), an Electroporator Gene Pulser Transfection Apparatus (Biorad; capacitance 25 μF ; voltage 12.5 kV cm^{-1} ; resistance 200 Ω) (for *M. smegmatis* and *M. bovis* BCG) or a BTX ECM600 Electroporation System (Harvard Apparatus; capacitance 25 μF ; voltage 2.5 kV cm^{-1} ; resistance 720 Ω) (for *M. marinum*). After the pulse, the cells were diluted in 900 μl of liquid medium, incubated for 3 h (*M. smegmatis*), 4 h (*M. marinum*) or 24 h (*M. tuberculosis* and *M. bovis* BCG) and then plated on selective solid medium.

Protein sample preparation

Mycobacterial cells grown to mid-logarithmic phase were separated from culture supernatants by centrifugation. The cells were washed with PBS and thereafter subjected to Proteinase K treatment, Genapol extraction or subcellular fractionation as described below, or sonicated to lyse bacteria. Secreted proteins were precipitated from culture supernatants with 10% TCA (w/v; Sigma-Aldrich). Proteins samples were boiled and separated by SDS-PAGE as described below.

Proteinase K sensitivity assay

Proteinase K sensitivity assay was performed as previously described [17]. Briefly, selected strains were grown in 20 ml of medium for 14 h starting from an OD₆₀₀ of 0.1. Cells were washed once in TBS buffer (Tris HCl pH 7.5, NaCl 150 mM, KCl 3 mM) and resuspended in 1 ml of the same buffer. Each sample was divided in two identical aliquots, of which one was treated with Proteinase K (Sigma-Aldrich) up to a concentration of 100 $\mu\text{g ml}^{-1}$, whereas the other was left untreated. Both aliquots were incubated for 30 min at 4°C. The reaction was stopped by the addition of 1X complete EDTA-free protease inhibitor (Roche). Subsequently, samples were washed once in TBS and resuspended in TBS plus loading buffer 5X (sucrose 50% w/v, SDS 10% w/v, 0.3 M Tris HCl pH 6.8, bromophenol blue 0.05% w/v, β -mercaptoethanol 5% v/v). Finally, samples were boiled for 10 min to allow bacterial lysis and loaded on a polyacrylamide gel in equal amounts. Treated and untreated samples were processed in parallel using the same procedure to allow their comparison. Each experiment was performed at least twice with different biological samples.

Genapol extraction of *M. marinum*

Surface-exposed proteins were extracted with the mild detergent Genapol X-080 (Sigma-Aldrich) as previously described [28]. Briefly, intact bacteria were incubated with 0.5% Genapol for 30 min at room temperature with head-over-head rotation. Extracted proteins were separated from the bacteria by centrifugation. Treated cells were lysed by sonication, and all samples were boiled and separated by SDS-PAGE as described below.

Cell fractionation

Cellular fractionation was performed as previously described [17]. Briefly, 20 ml bacterial cultures were harvested and washed once in 1X PBS at room temperature. The collected culture supernatants were filtered through low protein binding 0.22 μ m PVDF filters (Millipore) and concentrated to a final volume of 1 ml on Amicon filters (cut-off 3 kDa) to recover secreted proteins. The pellets were resuspended in 0.5% Genapol X-080 (150 μ l/10 mg of wet pellet) and incubated at 30°C for 30 minutes. Extracted proteins were separated from the bacteria by centrifugation. Bacterial pellets were resuspended in PBS 1X/phenyl methane sulphonyl fluoride (Sigma-Aldrich, PMSF) and subjected to sonication. The lysates were centrifuged at 1000 g^{-1} at 4°C to precipitate cellular debris and unlysed cells. Supernatants were transferred to fresh tubes and sedimented at 27.000 g for 30 min at 4°C in order to allow cell wall precipitation. Once again, the supernatant was precipitated at 100.000 g for 2 h to separate cytoplasmic membrane from cytosolic fraction. Cytosolic proteins were subsequently concentrated on Amicon centrifugal filters (cut-off 3 kDa) to a final volume of 1 ml. All sedimented samples were washed once after each step of centrifugation in PBS/PMSF 1 mM and finally resuspended in an appropriate volume of PBS (1 ml) plus Loading buffer 5X. Samples were boiled for 5 min before being separated on polyacrylamide gels and subjected to Western blotting as described below. Protein sorting/localization

was calculated by densitometric analysis using the Versadoc Imaging System (Bio-Rad) and Quantity One 4.2.3 software (Bio-Rad).

SDS-PAGE and Immunoblot

SDS-PAGE was performed according to standard protocols. Briefly, proteins were separated on 10%, 12% or 18% polyacrylamide gels [23], and subsequently transferred to polyvinylidene fluoride membranes (PVDF; Bio-Rad) or nitrocellulose membranes (Amersham Biosciences) by Western blotting. Proteins were visualized by immunoblotting using monoclonal antibodies directed against the HA epitope (Anti-HA.11; Covance, dilution 1:2000), GFP (Chemicon; dilution 1:2500), GroEL2 (Rv0440, BEI Resources dilution 1:200) or ESAT-6 (Hyb 76-8; Statens Serum Institut, Copenhagen, Denmark, dilution 1:200), and polyclonal antibody directed against a peptide of Rv1698 (Houben et al., in preparation) or against Mpt64 (Delogu et al. in preparation). Secondary goat anti-mouse (Santa Cruz Biotechnology; dilution 1:2000) or goat anti-rabbit (Santa Cruz Biotechnology; dilution 1:1000) horseradish peroxidase conjugates were used to detect proteins. The West Dura Signal Kit (Pierce) was used to develop the chemiluminescent signal. Image acquisitions and quantifications were performed using a Versadoc Imaging System (Bio-Rad) and Quantity One 4.2.3 software (Bio-Rad).

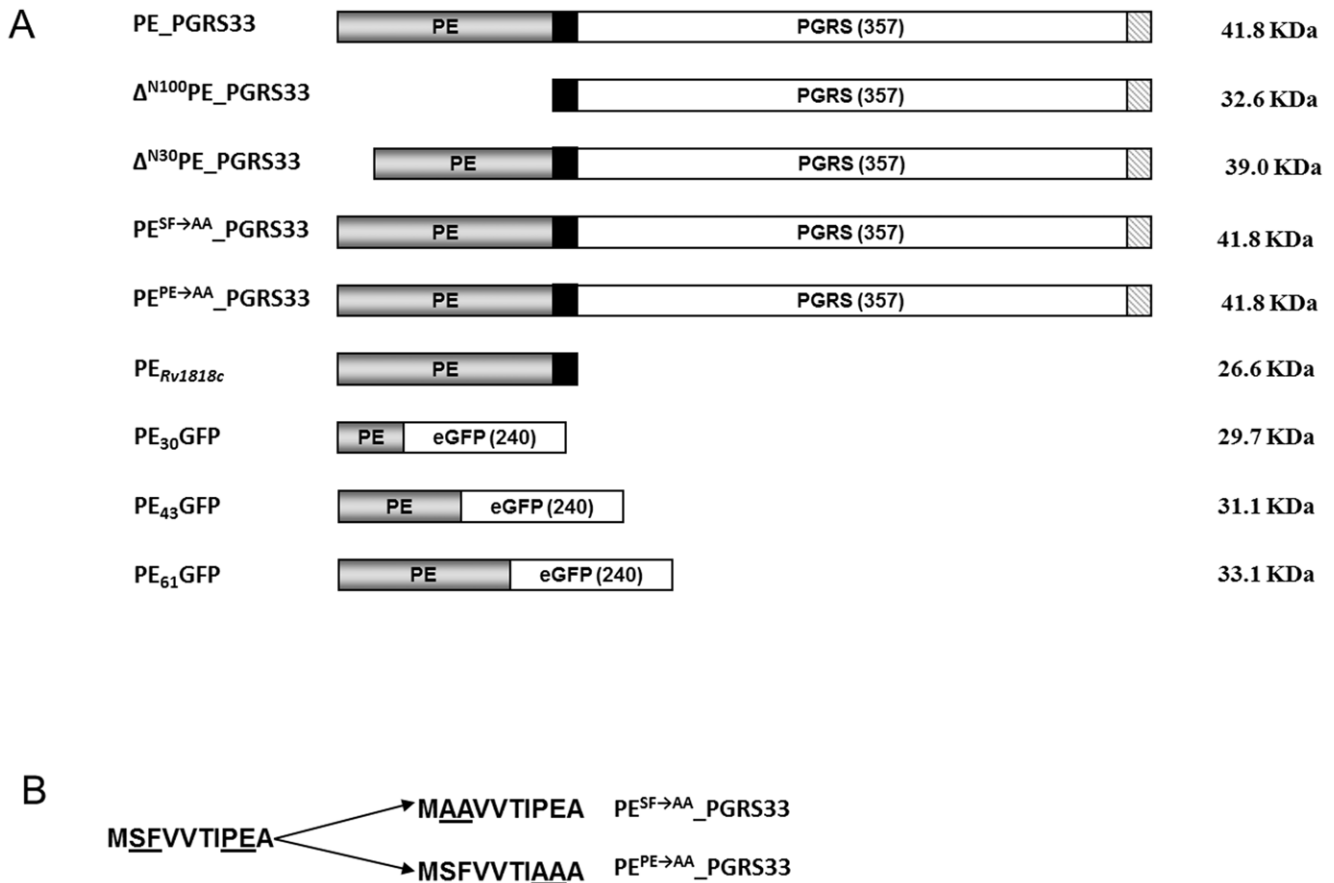


Figure 1. Maps of PE_PGRS33 and its derivatives used for this study. A) The PE domain is shown in gray, the 41 bp transition domain in black, the PGRS domain and GFP sequence in white, while the 9 amino acid HA epitope is striped; B) sequence of the first 10 amino acids of the wild type PE domain and of the PE domains of the two mutant proteins, in which the conserved SF and PE residues were replaced by alanine residues are shown. Altered amino acids are underlined.

doi:10.1371/journal.pone.0027713.g001

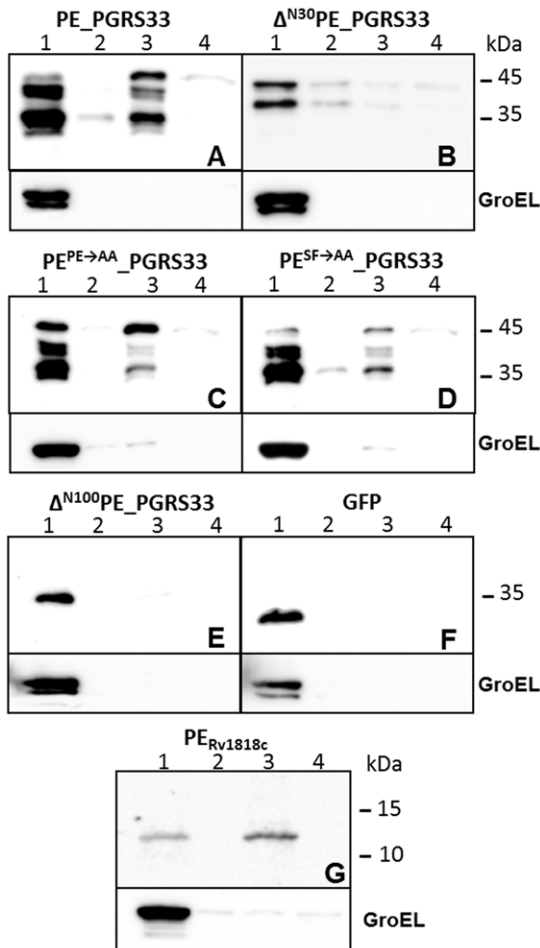


Figure 2. Subcellular fractionation analysis of *M. smegmatis* expressing different proteins. A) PE_PGRS33; B) Δ^{N30} PE_PGRS33; C) PE^{PE→AA}_PGRS33; D) PE^{SF→AA}_PGRS33; E) Δ^{N100} PE_PGRS33; F) GFP; G) PE_{Rv1818c}. All constructs were expressed from P_{Rv1818c} with the exception of that shown in panel B, which was expressed from P_{hsp60}. Lane 1: cytoplasmic fraction; lane 2: membrane fraction; lane 3: cell wall fraction; lane 4: Genapol supernatant. Proteins were detected by Western blot using monoclonal antibodies against GFP, HA, or GroEL. doi:10.1371/journal.pone.0027713.g002

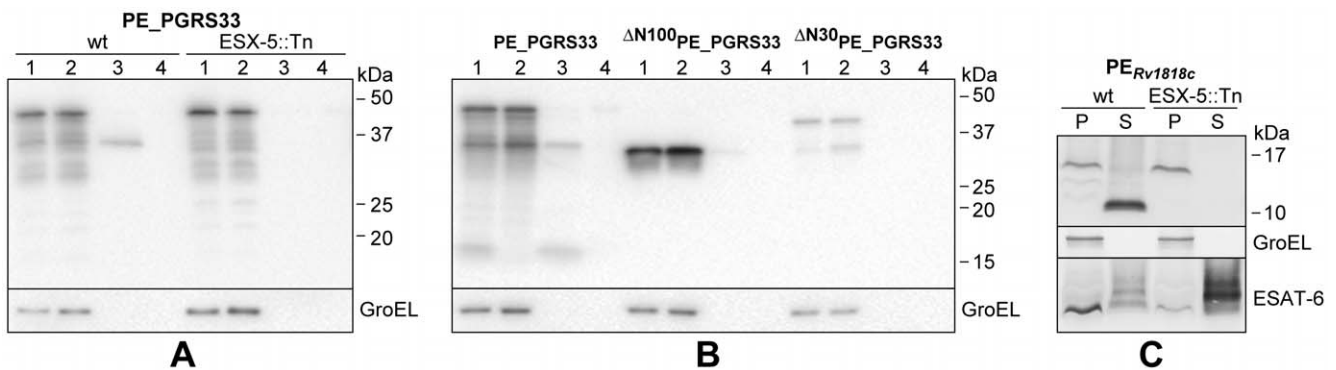


Figure 3. ESX-5-dependent secretion of PE_PGRS33 in *M. marinum*. Immunoblot analysis of wt and truncated forms of PE_PGRS33, expressed from the physiological promoter P_{Rv1818c} (A) or P_{hsp60} (B-C), in *M. marinum* wt strain E11 and its ESX-5 mutant 7C1 (A and C), or in E11 alone (B). In A-B) fractions containing equal OD units of cell pellets (1), cells treated with Genapol (2), Genapol supernatants (3) and culture supernatants (4) are shown. In C) cell pellet (P) and two-fold more culture supernatant (S) were included. PE_PGRS33 proteins were detected with the HA antibody, GroEL was included as control for bacterial lysis, and in panel C ESAT-6 was detected as control for protein secretion. doi:10.1371/journal.pone.0027713.g003

Results

Construction and expression of different PE_PGRS33 mutants

We recently demonstrated that PE_PGRS33 localizes in the external part of the mycobacterial envelope, and that chimeric proteins in which the PE domain is fused at the N-terminus of an heterologous protein are translocated across the mycobacterial envelope and are surface exposed [17,18]. Based on these data, which suggest that the PE domain contains a translocation signal, we generated a panel of four mutant PE_PGRS33 proteins to further characterize the functional role of this domain. As shown in Fig. 1, the first mutant protein (Δ^{N100} PE_PGRS33) lacked the entire PE domain (100 aa), while the second lacked only the first 30 amino acids (Δ^{N30} PE_PGRS33). In addition, two mutant proteins that had the same length as wt PE_PGRS33 were generated, in which either the SF residues at position 2-3 ($\Delta^{SF\rightarrow AA}$ PE_PGRS33) or the PE residues at position 8-9 ($\Delta^{PE\rightarrow AA}$ PE_PGRS33), were replaced by two alanine residues. These SF and PE residues were chosen since they are well conserved across PE domains, suggesting that they may have an important function. All of the mutant proteins were labeled with the HA epitope at their C-terminus to facilitate their detection on Western blots. The genes encoding these proteins were cloned in a replicative plasmid in which they were placed under transcriptional control of the physiological promoter of PE_PGRS33 (P_{Rv1818c}). Constructs encoding HA-tagged wt PE_PGRS33 and an HA-labeled version of its PE domain (PE_{Rv1818c}) were available from previous work [17].

Subcellular fractionation of *M. smegmatis*

The plasmids expressing wt, truncated and mutated forms of PE_PGRS33 from P_{Rv1818c} were introduced in *M. smegmatis*, and the subcellular localization of the recombinant proteins was investigated in the resulting strains. Even if *M. smegmatis* genome does not encode PE_PGRS proteins, it encodes a small number of PE proteins and was previously shown to be able to export PE_PGRS33 [17]. We chose to use a plasmid expressing PE_PGRS33 from its own promoter, since we previously demonstrated that when this protein was expressed from P_{hsp60} in *M. smegmatis* it could be found in all fractions due to overexpression [17]. Intact cells were subjected to low temperature extraction with Genapol, a detergent previously shown to extract

loosely associated proteins from the mycobacterial surface in these conditions [17,28,29], and thereafter a subcellular fractionation procedure was carried out. Cellular fractions were then analysed by Western blot, using GroEL as a control for the cytoplasmic fraction. The results, shown in Fig. 2, confirm that without its PE domain, PE_PGRS33 remains in the cytoplasm (Fig. 2E). However, the localization of the two proteins carrying the PE and SF mutations did not differ from that of the wt protein (Fig. 2C and D), suggesting that these amino acids, despite their high conservation, do not have a major role in protein localization. Interestingly, the amount of protein extracted with Genapol was extremely low for all proteins, suggesting that previous data showing high level of extraction of this protein in *M. smegmatis* was probably due to overexpression [17]. No PE_PGRS33 was detected in culture supernatants (not shown). GFP and PE_{Rv1818c} were used as control for cytoplasmic and cell wall proteins, respectively [17], while GroEL was used as a control for cellular integrity. Due to construct instability, it was not possible to obtain a plasmid in which the mutant protein without the first 30 amino acids was placed under P_{Rv1818c} transcriptional control (not shown). However, we could obtain and express this protein from P_{hsp60}. As shown in the Fig. 2B, this protein clearly localized in the cytoplasm, although low levels of protein were detectable also in the other fractions (probably due to overexpression).

ESX-5-dependent PE_PGRS33 localization in *M. marinum*

The ESX-5 system has previously been shown to be a major secretion pathway for PE and PPE proteins in *M. marinum* [2,19,20]. To investigate whether PE_PGRS33 is also a substrate of ESX-5, we introduced the plasmid encoding wt PE_PGRS33 under transcriptional control of its natural promoter in *M. marinum* strain E11 (wild-type) and its ESX-5 mutant 7C1 [2], and analyzed the localization of the expressed proteins. GroEL and ESAT-6 were used as controls for the cytoplasmic and secreted fractions, respectively. As shown in Fig. 3A we could not detect any secretion in the supernatant, but a processed form of PE_PGRS33 was found in Genapol extracts in the wild-type strain. In the ESX-5 mutant strain this processed form was not present in the Genapol extract, indicating that PE_PGRS33 is secreted to the bacterial surface by the ESX-5 system of *M. marinum*.

To elaborate on this, we also analyzed a strain expressing only PE_{Rv1818c}. In a previous study [17] we showed that in *M. smegmatis* this protein is associated to the cell envelope and that it is not released in the supernatant, even when overexpressed from P_{hsp60} [17]. However, as shown in Fig. 3C, in *M. marinum* it was substantially released in the supernatant in an ESX-5-dependant manner, confirming ESX-5 dependency and the fact that the PE domain triggers secretion. Interestingly, the secreted protein showed a lower molecular weight suggesting the presence of a maturation process.

Next we tested the effect of the different PE mutations on cell wall localization and secretion. The chimeras carrying the substitution of the PE or the SF residues could not be detected when introduced in *M. marinum* (not shown), suggesting that the PE and SF motifs near the N-terminus of PE_PGRS33 are important for the stability of the protein in this species. As expected, the proteins lacking either the entire PE domain or its first 30 amino acids were not exported (see Fig. 3B for results with wt *M. marinum*, results for the ESX-5 mutant strain not shown), indicating that this domain is indeed important for secretion.

Subcellular fractionation of *M. bovis* BCG

Finally, we decided to study the localization of our chimeric proteins in their natural hosts. For this purpose, the genes

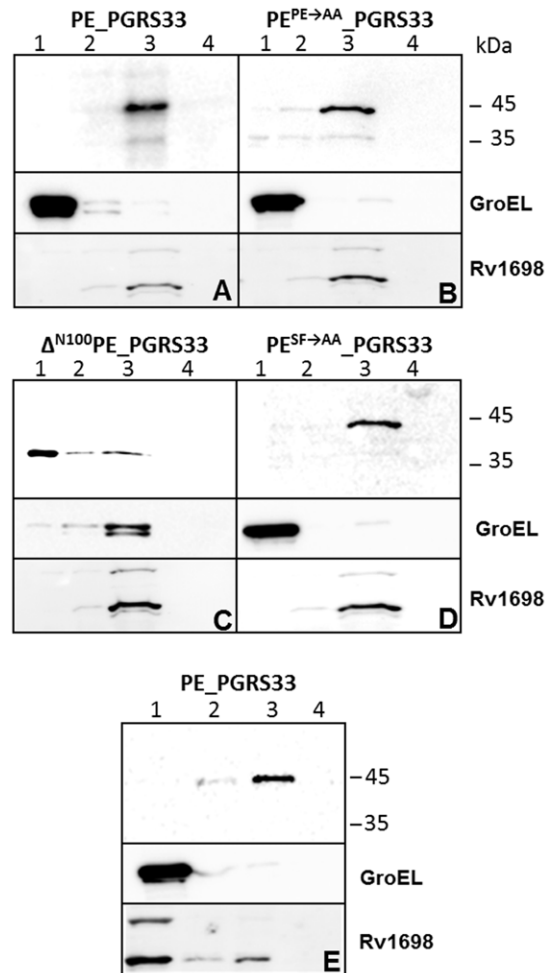


Figure 4. PE_PGRS33 localization in *M. bovis* BCG and *M. tuberculosis*. Subcellular fractionation analysis was carried out on different cellular fractions of *M. bovis* BCG (A-D) or *M. tuberculosis* (E) expressing PE_PGRS33 (A and E) or different PE_PGRS33 mutants (B-D). Lane 1: cytoplasmic fraction; lane 2: membrane fraction; lane 3: cell wall fraction; lane 4: Genapol supernatant. Proteins were detected by Western blot using monoclonal antibodies against, HA, or GroEL, and a rabbit serum against Rv1698

doi:10.1371/journal.pone.0027713.g004

encoding the different chimeric proteins under transcriptional control of the PE_PGRS33 natural promoter (P_{Rv1818c}) were subcloned in an integrative plasmid, and introduced in *M. bovis* BCG. The resulting strains were subjected to Genapol extraction and subcellular fractionation. Samples were analyzed by Western blot using GroEL and the outer membrane protein Rv1698 as controls for the cytoplasmic and cell wall fractions, respectively [30]. As shown in Fig. 4, the wt protein (Fig. 4A) and its PE and SF mutants localized in the cell wall fraction (Fig. 4B and D), while the Δ^{N100} PE_PGRS33 protein mainly localized in the cytoplasm (Fig. 4C), confirming the results obtained in *M. smegmatis*, with the difference that cell wall localization was more efficient in *M. bovis* BCG. The presence of small amounts of the Δ^{N100} PE_PGRS33 protein in the insoluble fractions, together with that of GroEL in these samples, suggests that the PGRS domain might produce insoluble aggregates that associate with the chaperon GroEL. The construct expressing the protein missing the first 30 amino acids could not be electroporated in *M. bovis* BCG suggesting that this protein is toxic in this species.

The construct expressing the wt protein was also introduced in *M. tuberculosis* and shown to localize entirely in the cell wall fraction (Fig. 4E).

The PE domain of PE_PGRS33 is not secreted in *M. bovis* BCG and *M. tuberculosis*

The finding that PE_{Rv1818c} is secreted in *M. marinum* (Fig. 3C) prompted us to study its localization in *M. bovis* BCG and in *M. tuberculosis*, the natural hosts of this protein. However, as clearly shown in Fig. 5, this protein prevalently localized in the cell wall fraction and was not detected in the culture supernatant in these two species, even when overexpressed from P_{hsp60}. The secreted protein Mpt64 was used as a positive control.

Identification of a translocation signal in PE_PGRS33 PE domain

In order to identify the cell wall translocation signal of the PE domain, we constructed three expression plasmids containing translational fusions between the first 30, 43 or 61 aa of the PE_PGRS33 PE domain and the green fluorescent protein (GFP) under transcriptional control of P_{Rv1818c} (Fig. 1). Subsequently, the three plasmids were introduced in both *M. smegmatis* and *M. bovis* BCG. The resulting strains were subjected to cell fractionation after Genapol extraction. Even if with different efficiencies, all the constructs were translocated to the cell wall of both species (Fig. S1) suggesting that the first 30 amino acids of the PE domain contain a sequence that allows protein translocation. The *M. smegmatis* strain expressing the PE₆₁-GFP chimera was subjected to

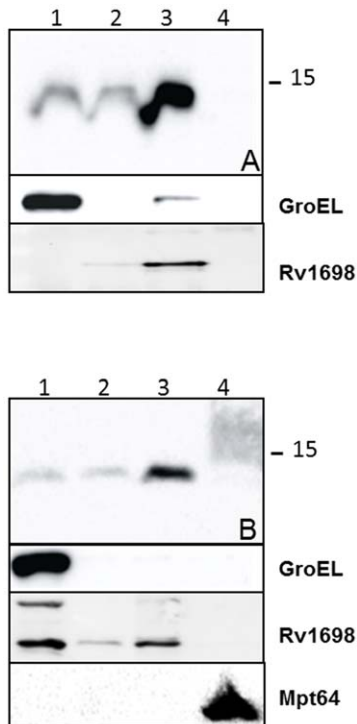


Figure 5. PE_{Rv1818c} localization in *M. bovis* BCG and *M. tuberculosis*. Subcellular fractionation analysis was carried out on different cellular fractions of *M. bovis* BCG (A) or *M. tuberculosis* (B) expressing the PE domain of PE_PGRS33: lane 1: cytoplasmic fraction; lane 2: membrane fraction; lane 3: cell wall fraction; lane 4: culture supernatant. Proteins were detected by Western blot using antibodies recognizing HA, Rv1698, GroEL, or Mpt64. doi:10.1371/journal.pone.0027713.g005

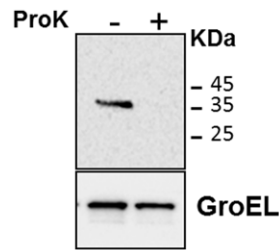


Figure 6. Proteinase K sensitivity assay. The assay was performed on the *M. smegmatis* strain expressing PE₆₁-GFP. GroEL was used as a control of cellular integrity. Proteins were detected by Western blot using monoclonal antibodies against GFP or GroEL. doi:10.1371/journal.pone.0027713.g006

Proteinase K degradation assay to further prove that this chimera was exposed on the bacterial surface: as shown in Fig. 6, the protein was totally degraded in the presence of the protease, while GroEL (a cytoplasmic protein used as a negative control) was not, demonstrating that the PE₆₁-GFP chimera was indeed available on the bacterial surface of *M. smegmatis*.

Finally, we analyzed the culture supernatant of these recombinant bacteria. While no recombinant protein was found in the supernatant of the *M. smegmatis* strains (Fig. 7A), surprisingly we found that in *M. bovis* BCG the PE₃₀-GFP chimera was partially secreted, (Fig. 7B).

Discussion

PE are exported mycobacterial proteins characterized by a well conserved N-terminal domain of about 100 amino acids that we recently showed to be required for their export [17,20]. The lack in the PE domains of clear secretion signals led to the hypothesis that these proteins might be secreted through a new type of mycobacterial-specific secretion system. Indeed, it was recently showed that in *M. marinum* PE protein secretion is abrogated in mutants lacking the type VII secretion system ESX-5 [2,19,20]. The aim of this paper is to better characterize the functional portions of the PE domain and their role in translocation through the mycobacterial cell wall. Different PE_PGRS mutant proteins were expressed in different model organisms to study their localization: *i*) *M. smegmatis*, whose chromosome encodes neither ESX-5 nor PE_PGRS proteins, but was previously shown to be able to export PE_PGRS33 and chimeric proteins based on its PE domain [17]; *ii*) *M. marinum*, whose chromosome encodes ESX-5 and many PE_PGRS proteins, but not an apparent PE_PGRS33 orthologue, and *iii*) *M. bovis* BCG and *M. tuberculosis*, whose chromosomes encodes both ESX-5 and PE_PGRS33. The rationale to use different mycobacterial species was to compare the behaviour of a single PE_PGRS protein in different genetic backgrounds to investigate the conservation of the PE-exporting pathway(s).

As expected, the protein lacking the entire PE domain was not able to translocate and was exclusively found in the bacterial cytoplasm. The same expression profile was found in the mutant protein missing the first 30 amino acids of the PE domain, suggesting that this portion of the protein is essential for the translocation. Surprisingly, mutation of the SF or of the PE conserved residues did not result in any clear phenotype with the exception of a strong instability of the resulting proteins in *M. marinum*.

It is worth noting that PE_PGRS33 was found in Genapol extracts of *M. marinum*, but not in those of the other tested mycobacterial species suggesting a weaker association of this

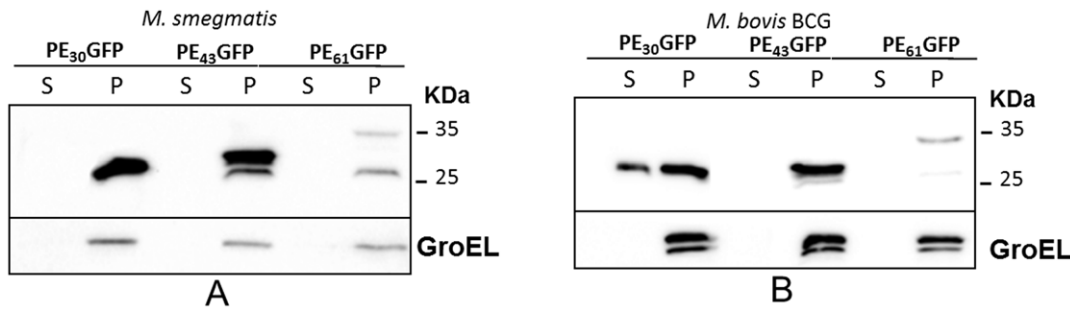


Figure 7. Secretion and surface exposure of PE-based GFP chimeric proteins in *M. smegmatis* and *M. bovis* BCG. Pellet (P) and culture supernatant (S) analysis was carried out on *M. smegmatis* (A) or *M. bovis* BCG (B) strains expressing different fusions of the PE domain of PE_PGRS33 with GFP. Proteins were detected by Western blot using monoclonal antibodies against GFP or GroEL.
doi:10.1371/journal.pone.0027713.g007

protein to the cell wall in this species. Of course it is still possible that using different experimental procedures (as increased temperature during extraction) PE_PGRS33 and/or some of the mutant protein might be extracted with this detergent even in *M. smegmatis*, *M. tuberculosis* or *M. bovis* BCG.

These results confirm previous proteomics data on Genapol extracts from *M. marinum* as compared to *M. tuberculosis* or *M. smegmatis* [29]. Moreover, the size of PE_PGRS33 in *M. marinum* Genapol extracts showed a molecular weight lower than that predicted for the entire protein, suggesting a maturation process, as we recently reported for LipY, another PE protein not belonging to the PE_PGRS family [20]. In that case the protein was also processed in *M. tuberculosis*, but only when bacteria were grown inside macrophages, suggesting that still unknown factors only expressed during infection are required for full export and maturation of PE proteins in *M. tuberculosis*. For this reason, PE_PGRS33 might also be expected to be processed in *M. tuberculosis* during growth inside macrophages. In *M. smegmatis* some of the recombinant proteins showed multiple bands. However, since bands of the same size were present in both cytoplasmic and cell wall fractions, they were likely not due to processing during translocation but rather the result of degradation during cell lysis. The lack of processing observed in *M. smegmatis* might be due to the absence of ESX-5 in this species. Finally, the release of PE_PGRS33 was totally abrogated in the *M. marinum* ESX-5 mutant, confirming that PE protein translocation in *M. marinum* is dependent on this secretion system.

From previous work [17], we knew that the PE domain of PE_PGRS33, when expressed in *M. smegmatis* in the absence of the PGRS domain, localizes in the cell wall without being exposed to the external environment, suggesting that it remains embedded in the mycobacterial outer membrane. Surprisingly, when the same construct was expressed in *M. marinum* it was efficiently secreted in the culture supernatant in an ESX-5-dependent manner, while it remained associated with the cell wall when expressed in *M. tuberculosis* or *M. bovis* BCG. Moreover, in *M. marinum* the secreted form of the PE domain had a lower apparent molecular weight than predicted, indicative of a maturation process, as was previously demonstrated for PE_PGRS33 wt and for LipY [20].

In the cell fractionation experiments involving *M. tuberculosis* extracts, Rv1698 was not only found in the cell wall fraction (as in those involving *M. bovis* BCG extracts), but also in the cytoplasmic fraction. This imperfect fractionation could be due to the fact that, for biosafety reasons, *M. tuberculosis* was lysed by bead beating instead of sonication.

Since our data suggested that the first portion of the PE domain contains functions required for protein export, we constructed

three chimeric proteins in which the first 30, 43 or 61 residues of the PE domain of PE_PGRS33 were fused to the coding sequence of GFP. All these proteins were able to localize in the cell wall, although in *M. smegmatis* only the construct including the first 61 residues of the PE domain was translocated with a fair efficiency. These result clearly indicated that PE_PGRS33-mediated translocation is more efficient in *M. bovis* BCG than in *M. smegmatis*. It should be noted that, while *M. bovis* BCG (as *M. tuberculosis* and *M. marinum*) encodes an ESX-5 secretion system, *M. smegmatis* does not. Moreover, this species encodes only few PE and no PE_PGRS proteins. However, *M. smegmatis* chromosome encodes other type VII secretion systems (as ESX-1 and ESX-3) that might, with low efficiency, complement the absence of ESX-5 and have a role in the secretion of PE_PGRS33 and its derivatives in this species. Finally, the chimera including only the first 30 amino acids of the PE domain fused to GFP was partially released in *M. bovis* BCG culture supernatant, but not in that of *M. smegmatis*. This interesting finding suggests that the first 30 amino acids of the PE domain contain sufficient information to allow protein translocation. The only structural information available for a PE protein derives from PE25, a protein including only the PE domain, whose structural gene is followed by the gene encoding a protein of the PPE family (PPE41). These two proteins were shown to interact to form a heterodimer, for which the crystal structure was solved [10]. The PE structure included two antiparallel α -helices (the first between residues 8 and 37, the second between residues 45 and 84) connected by a loop (residues 38-44) [10]. If the different PE domains have a similar folding, the first chimera (PE₃₀-GFP) would include most of the first α -helix, while the second (PE₄₃-GFP) would include both the first α -helix and the loop, and the last chimera (PE₆₁-GFP) would include the first α -helix, the loop and part of the second α -helix (Fig. S2). Our data suggest an involvement of the first α -helix in directing the protein to the secretion system. The fact that the other two chimeric proteins were not secreted and were detected in the cell wall fraction, suggests that the loop between the two α -helices might be involved in the association of the PE domain with the cell wall, even though no similarity with cell wall anchoring domains of other bacteria was identified. However, it is also possible that the PE₄₃-GFP and the PE₆₁-GFP chimeric proteins form insoluble aggregates that co-localize with the cell wall, or that they are misfolded and bind the membrane via exposed hydrophobic patches.

The identification of an N-terminal portion of a PE domain responsible for translocation and association to the bacterial surface opens new avenues to study the interaction of type VII substrates and the secretion machinery and could be used to

identify novel substrates. The fact that at least one of the ESX-1 substrates (Cfp10) was shown to have a C-terminal signal peptide [31] opens the interesting possibility that multiple secretion signals might be recognized by type VII secretion system and be present in their targets.

Supporting Information

Figure S1 Localization of PE-based GFP chimeric proteins in *M. smegmatis* and *M. bovis* BCG. Subcellular fractionation analysis was carried out on different cellular fractions of *M. smegmatis* (A-D) or *M. bovis* BCG (E-H) expressing GFP or different fusions of the PE domain of PE_PGRS33 with GFP: lane 1: cytoplasmic fraction; lane 2: membrane fraction; lane 3: cell wall fraction; lane 4: Genapol supernatant. Proteins were detected by Western blot using monoclonal antibodies against GFP. (TIF)

Figure S2 Putative structure of the PE domain of PE_PGRS33 and of the chimeric proteins in which small

fragments of the PE domain were fused to GFP. The figure was drawn assuming that the PE domain of PE_PGRS33 has a structure similar to the PE domain of PE25 [10].

(TIF)

Table S1 Primers used in this study. (DOC)

Acknowledgments

The authors wish to thank Ida Rosenkrands for providing anti-ESAT-6 antibodies, and Aniek van der Woude and Edith Houben for providing the anti-Rv1698 antibody.

Author Contributions

Conceived and designed the experiments: RM WB GD AC MHD. Performed the experiments: AC MHD MV. Analyzed the data: RM WB GD MHD AC. Contributed reagents/materials/analysis tools: VD GP. Wrote the paper: RM WB MHD AC.

References

- Delogu G, Pusceddu C, Bua A, Fadda G, Brennan MJ, et al. (2004) Rv1818c-encoded PE_PGRS protein of *Mycobacterium tuberculosis* is surface exposed and influences bacterial cell structure. *Mol Microbiol* 52: 725–733.
- Abdallah AM, Verboom T, Weerdenburg EM, van Pittius NC, Mahasha PW, et al. (2009) PPE and PE_PGRS proteins of *Mycobacterium marinum* are transported via the type VII secretion system ESX-5. *Mol Microbiol* 73: 329–340.
- Sampson SL, Lukey P, Warren RM, van Helden PD, Richardson M, et al. (2001) Expression, characterization and subcellular localization of the *Mycobacterium tuberculosis* PPE gene Rv1917c. *Tuberculosis (Edinb)* 81: 305–317.
- Song H, Sandie R, Wang Y, Andrade-Navarro MA, Niederweis M (2008) Identification of outer membrane proteins of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)*.
- Brennan MJ, Delogu G, Chen Y, Bardarov S, Kriakov J, et al. (2001) Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. *Infect Immun* 69: 7326–7333.
- Brennan MJ, Delogu G (2002) The PE multigene family: a ‘molecular mantra’ for mycobacteria. *Trends Microbiol* 10: 246–249.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393: 537–544.
- Banu S, Honore N, Saint-Joanis B, Philpott D, Prevost MC, et al. (2002) Are the PE-PGRS proteins of *Mycobacterium tuberculosis* variable surface antigens? *Mol Microbiol* 44: 9–19.
- Talarico S, Cave MD, Marrs CF, Foxman B, Zhang L, et al. (2005) Variation of the *Mycobacterium tuberculosis* PE_PGRS 33 gene among clinical isolates. *J Clin Microbiol* 43: 4954–4960.
- Strong M, Sawaya MR, Wang S, Phillips M, Cascio D, et al. (2006) Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 103: 8060–8065.
- Riley R, Pellegrini M, Eisenberg D (2008) Identifying cognate binding pairs among a large set of paralogs: the case of PE/PPE proteins of *Mycobacterium tuberculosis*. *PLoS Comput Biol* 4: e1000174.
- Gey van Pittius NC, Sampson SL, Lee H, Kim Y, van Helden PD, et al. (2006) Evolution and expansion of the *Mycobacterium tuberculosis* PE and PPE multigene families and their association with the duplication of the ESAT-6 (*esx*) gene cluster regions. *BMC Evol Biol* 6: 95.
- Mishra KC, de Chastellier C, Narayana Y, Bifani P, Brown AK, et al. (2008) Functional role of the PE domain and immunogenicity of the *Mycobacterium tuberculosis* triacylglycerol hydrolase LipY. *Infect Immun* 76: 127–140.
- Chaturvedi R, Bansal K, Narayana Y, Kapoor N, Sukumar N, et al. (2010) The multifunctional PE_PGRS11 protein from *Mycobacterium tuberculosis* plays a role in regulating resistance to oxidative stress. *J Biol Chem* 285: 30389–30403.
- Dheenadhayalan V, Delogu G, Brennan MJ (2006) Expression of the PE_PGRS 33 protein in *Mycobacterium smegmatis* triggers necrosis in macrophages and enhanced mycobacterial survival. *Microbes Infect* 8: 262–272.
- Basu S, Pathak SK, Banerjee A, Pathak S, Bhattacharyya A, et al. (2007) Execution of macrophage apoptosis by PE_PGRS33 of *Mycobacterium tuberculosis* is mediated by Toll-like receptor 2-dependent release of tumor necrosis factor- α . *J Biol Chem* 282: 1039–1050.
- Cascioferro A, Delogu G, Colone M, Sali M, Stringaro A, et al. (2007) PE is a functional domain responsible for protein translocation and localization on mycobacterial cell wall. *Mol Microbiol* 66: 1536–1547.
- Sali M, Di Sante G, Cascioferro A, Zumbo A, Nicolo C, et al. (2010) Surface Expression of MPT64 as a fusion with the PE Domain of PE_PGRS33 enhances *Mycobacterium bovis* BCG protective activity against *Mycobacterium tuberculosis* in mice. *Infect Immun* 78: 5202–5213.
- Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, et al. (2007) Type VII secretion—mycobacteria show the way. *Nat Rev Microbiol* 5: 883–891.
- Daleke MH, Cascioferro A, de Punder K, Ummels R, Abdallah AM, et al. (2011) Conserved Pro-Glu (PE) and Pro-Pro-Glu (PPE) protein domains target LipY lipases of pathogenic mycobacteria to the cell surface via the ESX-5 pathway. *J Biol Chem* 286: 19024–19034.
- Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR Jr. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* 4: 1911–1919.
- Puttinaowarat S, Thompson KD, Adams A (2000) Mycobacteriosis: detection and identification of aquatic *Mycobacterium* species. *Fish Veterinary Journal* 5: 6–21.
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor N.Y.: Cold Spring Harbor Laboratory.
- Delogu G, Bua A, Pusceddu C, Parra M, Fadda G, et al. (2004) Expression and purification of recombinant methylated HBHA in *Mycobacterium smegmatis*. *FEMS Microbiol Lett* 239: 33–39.
- Hayward CM, O’Gaora P, Young DB, Griffin GE, Thole J, et al. (1999) Construction and murine immunogenicity of recombinant Bacille Calmette Guerin vaccines expressing the B subunit of *Escherichia coli* heat labile enterotoxin. *Vaccine* 17: 1272–1281.
- Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, et al. (1991) New use of BCG for recombinant vaccines. *Nature* 351: 456–460.
- Maciag A, Dainese E, Rodriguez GM, Milano A, Provvedi R, et al. (2007) Global Analysis of the *Mycobacterium tuberculosis* Zur (FurB) Regulon. *J Bacteriol* 189: 730–740.
- Heinz C, Niederweis M (2000) Selective extraction and purification of a mycobacterial outer membrane protein. *Anal Biochem* 285: 113–120.
- Sani M, Houben EN, Geurtsen J, Pierson J, de Punder K, et al. (2010) Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog* 6: e1000794.
- Siroy A, Mailaender C, Harder D, Koerber S, Wolschendorf F, et al. (2008) Rv1698 of *Mycobacterium tuberculosis* Represents a New Class of Channel-forming Outer Membrane Proteins. *J Biol Chem* 283: 17827–17837.
- Champion PA, Stanley SA, Champion MM, Brown EJ, Cox JS (2006) C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*. *Science* 313: 1632–1636.