UNIVERSITA` DEGLI STUDI DI SASSARI

UNIVERSITY OF SASSARI

Department of Biomedical Sciences

PhD Course in Life Sciences and Biotechnologies

XXIX Doctoral cycle

PE_PGRS3: a new player in *Mycobacterium tuberculosis* pathogenesis

PhD Student: Basem Battah

2016/2017

Tutor: Prof. Salvatore Rubino

Co-tutor: Prof. Giovanni Delogu

Reality cannot compete with imagination

Index

Chapter III: Material and methods

Abstract

The *M. tuberculosis* (*Mtb*) genome contains around 60 *pe_pgrs* genes, whose role and function remain elusive. In this study, two PE_PGRS proteins with high sequence homology were selected and investigated (PE_PGRS3 and PE_PGRS4), with PE_PGRS3 characterized by the presence of a C-terminal domain rich in arginine. Interestingly, full-length PE_PGRS3 protein is expressed by *Mtb* strains but not by other MTBC subspecies causing disease in animals. A gene reporter system was developed to investigate in *M. smegmatis* (*Msm*) the expression pattern of these genes. Fluorescence microscopy, FACS and transcriptional analysis indicated that the two genes are differentially regulated, with *pe_pgrs3* but not *pe_pgrs4* being expressed only when mycobacteria are cultivated in low inorganic phosphate (iPhos). Expression of *pe_pgrs3* in low iPhos correlated with the upregulation of *relA* in *Msm* recombinant strains and *Mtb*, suggesting that *pe_pgrs3* is involved in the stringent response. Overexpression of the PE_PGRS3, and of its functional deletion mutant (PE_PGRS3ΔCt), in *Msm* were obtained by expressing these genes under control of *hbh*A promoter. Interestingly, *Msm* strains overexpressing PE_PGRS3 showed enhanced ability to entry in macrophages and epithelial cells compared to *Msm* expressing PE_PGRS3ΔCt or *Msm* parental strain. No differences in the ability of these strains to survive intracellularly were measured. These results provide new insights on the role of PE_PGRS3 in TB pathogenesis.

Chapter I

Introduction

1. Tuberculosis

Tuberculosis (TB), one of the oldest recorded human catastrophes, is still one of the biggest killers among the infectious diseases (WHO, 2016). Despite the worldwide use of a live attenuated vaccine and effective antibiotics, new vaccines and drugs are needed to control the worldwide epidemic of TB that kills two million people each year. In order to develop new anti-tubercular agents, there is a need to study the genetics and physiology of tubercle bacillus and understand the host-pathogen interaction to learn how *M. tuberculosis* bacteria overcome host defences and cause disease.

In 2015, there were an estimated 10.4 million new (incident) TB cases worldwide, 480. 000 new cases of multidrug-resistant TB (MDR-TB), 1.4 million TB deaths and an additional 0.4 million deaths resulting from TB disease among people living with HIV (WHO, 2016). Therefore, new and effective tools against TB are urgently needed to control the disease at multiple levels.

1.1 The genus mycobacterium

Beside *M. tuberculosis*, more than 100 other mycobacterial species have been identified (Tortoli, 2006). These can be divided in two groups based on their growth rate: the rapid-growers, which produce visible colonies on solid medium within seven days and the slow – growers, which typically require 10-28 days for visible growth. Genetic analysis of the 16S rRNA genes indicates that the slow - growers have evolved from the fast growers (Fig 1.1) (Reva *et al.*, 2015). This division also reflects the virulence of mycobacterial species, as most fast grower are non-pathogenic, free – living saprophytes whereas the slowly growing group contain a number of important human and animal pathogens.

The latter include, apart from the species belonging to *M. tuberculosis* complex (*M. bovis*, *M. microti*, *M. canetti*, *M. tuberculosis*, *M. africanum*), *Mycobacterium leprae*, the causative agent of leprosy, *M. ulcerans*, the etiological agent of buruli ulcer, *M. avium*, which causes TB in birds and opportunistic infections in immunocompromised humans, and *M. marinum*, that causes chronic progressive disease in fish and amphibia, skin infections in humans.

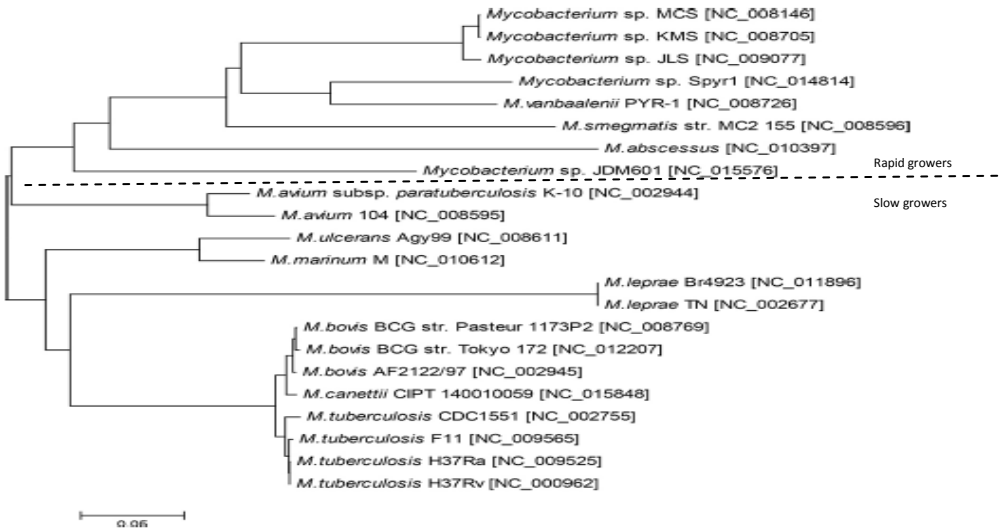


Figure 1.1: Species phylogenetic tree of the genus mycobacterium (Reva *et al.*, 2015).

1.2 The *M. tuberculosis* Genome

The *M. tuberculosis* H37Rv genome consists of 4.4 x10^6 bp and encodes approximately 4,000 genes (Fig 1.2). Analysis of the *M. tuberculosis* genome showed that this bacterium has some unique features. Over 200 genes, accounting to 6% of the total, have been annotated as encoding enzyme involved in the metabolism. Among these approximately 100 are predicted to be implicated in the oxidation of fatty acids, while in comparison *E. coli* only has 50 enzymes involved in fatty acid metabolism. The large number of enzymes that putatively use fatty acids may be related to the ability of the *M. tuberculosis* to grow in

infected host tissues, where fatty acids may provide the major carbon source (Smith, 2003). Interestingly, five gene locus encoding Type 7 secretion apparatus (ESX1-5) were also identified, highlighted an unexpected ability of *M. tuberculosis* to secrete proteins. Moreover, almost 8% of the genome coding capacity was devoted to genes belonging to protein with highly similar protein sequence, named PE and PPE.

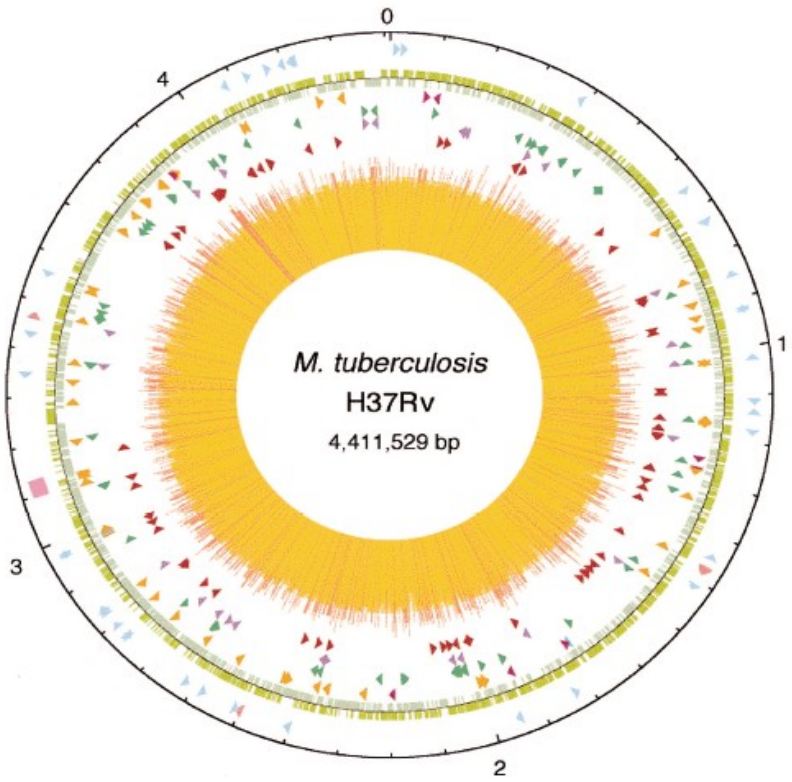


Figure 1.2: Circular map of the chromosome of *M. tuberculosis* H37Rv (Cole *et al.*, 1998).

2. *M. tuberculosis* virulence and the disease process

M. tuberculosis usually enters the alveolar pathways of exposed humans in an aerosol droplet, and its first contact is thought to be with macrophages, but it is also possible that bacteria can be initially ingested by alveolar epithelial type II pneumocytes which are found in greater numbers than macrophages in alveoli. In addition, dendritic cells play a very important role in the early stages of infection since they are much better antigen presenters than are macrophages and may play a key role in activating T cells with specific *M. tuberculosis* antigens. Dendritic cells are migratory and may play an important role in *M. tuberculosis* dissemination (Smith, 2003) (Fig 1.3).The recognition of *M. tuberculosis* is mediated by a set of surface receptor, which drive the uptake of bacteria and trigger the innate immune signalling pathways leading to the production of various chemokines and cytokines. Epithelial cells and neutrophils can also produce chemokines in response to the bacterial products, this promotes recruitment of other immune cells, more macrophages, dendritic cells and lymphocyte, to the infection site. They organize in spherical structure with infected macrophages in the middle surrounded by various categories of lymphocytes (mainly CD4+, CD8+). Macrophages can fuse to form multi nucleated giant cells or differentiate into lipid–rich foamy cells. B lymphocytes tend to aggregate in follicular-type structure adjacent to the granuloma. The bacteria can survive for decades inside granuloma with no symptoms of the disease in 90-95% of the cases (Mayra Silva Miranda, 2012). It is known that infected macrophages in the lung, through their production of chemokines, attract inactivated monocytes, lymphocytes, and neutrophils, which cannot kill the bacteria very efficiently. Mycobacteria can escape intracellular killing, multiply and further promote inflammation (van *et al.*, 2002). Then, granulomatous focal lesions composed of macrophage derived giant

cells and lymphocytes begin to form. These processes generally serve as effective means for controlling bacterial spread. As cellular immunity develops, macrophages loaded with bacilli are killed, and this results in the formation of the caseous center of the granuloma, surrounded by a cellular barrier of fibroblasts, lymphocytes, and blood-derived monocytes (Smith, 2003) (Fig 1.3).

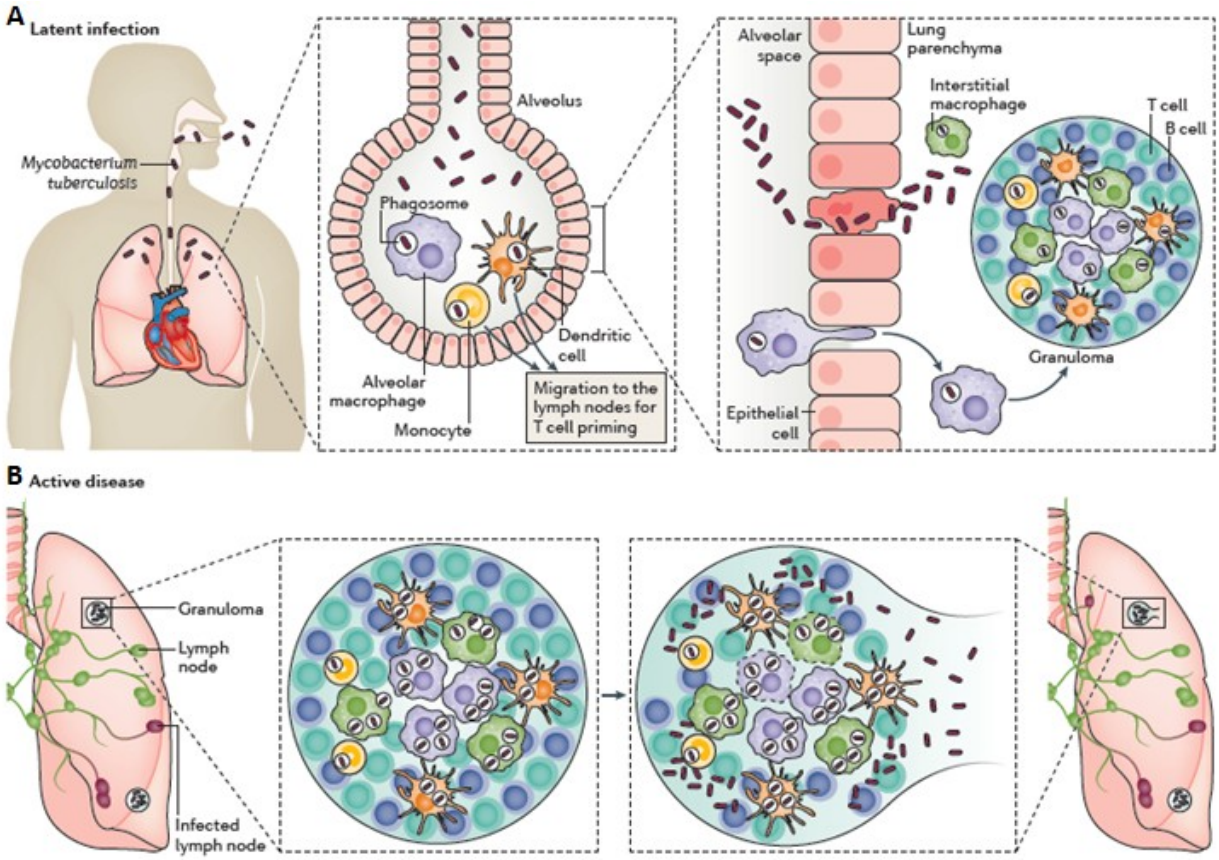


Figure 1.3: *Mycobacterium tuberculosis* infection. (A) The Infection begins when *M. tuberculosis* enters the lungs via inhalation, if this first line of defence fails to eliminate the bacteria, *M. tuberculosis* invades the lung interstitial tissue, either by the bacteria directly infecting the alveolar epithelium or the infected alveolar macrophages migrating to the lung parenchyma. Subsequently, either dendritic cells or inflammatory monocytes transport *M. tuberculosis*, migrates to the lymph nodes to recruit the immune cells T and B cells to form a granuloma. (B) If the bacteria replicated within the growing granuloma with a great load, the granuloma will fail to contain the infection and the bacteria will disseminate to other organs including the brain. At this phase, the bacteria can enter the bloodstream or re-enter the respiratory tract to be released, the infected host is now infectious, symptomatic and is said to have active TB (Pai *et al.*, 2016).

By the time the host immune response properly control bacterial replication, the tubercle bacilli is thought to be able to disseminate by the lymphatics and bloodstream to potentially any organ and tissue.

The strength of the host cellular immune response determines whether an infection is arrested here or progresses to the next stages, the resulted enclosed infection is referred to as latent or persistent TB and can persist throughout a person's life in an asymptomatic and non-transmissible state. In persons with efficient cell-mediated immunity, the infection may be arrested permanently at this point. The granulomas subsequently heal, leaving small fibrous and calcified lesions which is the hallmark of latent tuberculosis (Sandor *et al.*, 2003).

It is estimated that 5-10% of the *M. tuberculosis* infected subjects may develop disease during the lifetime, either because the infected person cannot control the initial infection or because a person with latent infection may lose the ability to control bacterial replication due to a weakened immune system (use of immunosuppressive drugs, HIV infection, malnutrition, aging, or other factors). Bacterial replication leads the granuloma centre to liquefy by unknown processes and then serves as a rich medium in which the now revived bacteria can replicate in an uncontrolled manner (Smith, 2003), (Russell *et al.*, 2010) (Fig 1.3, 1.4).

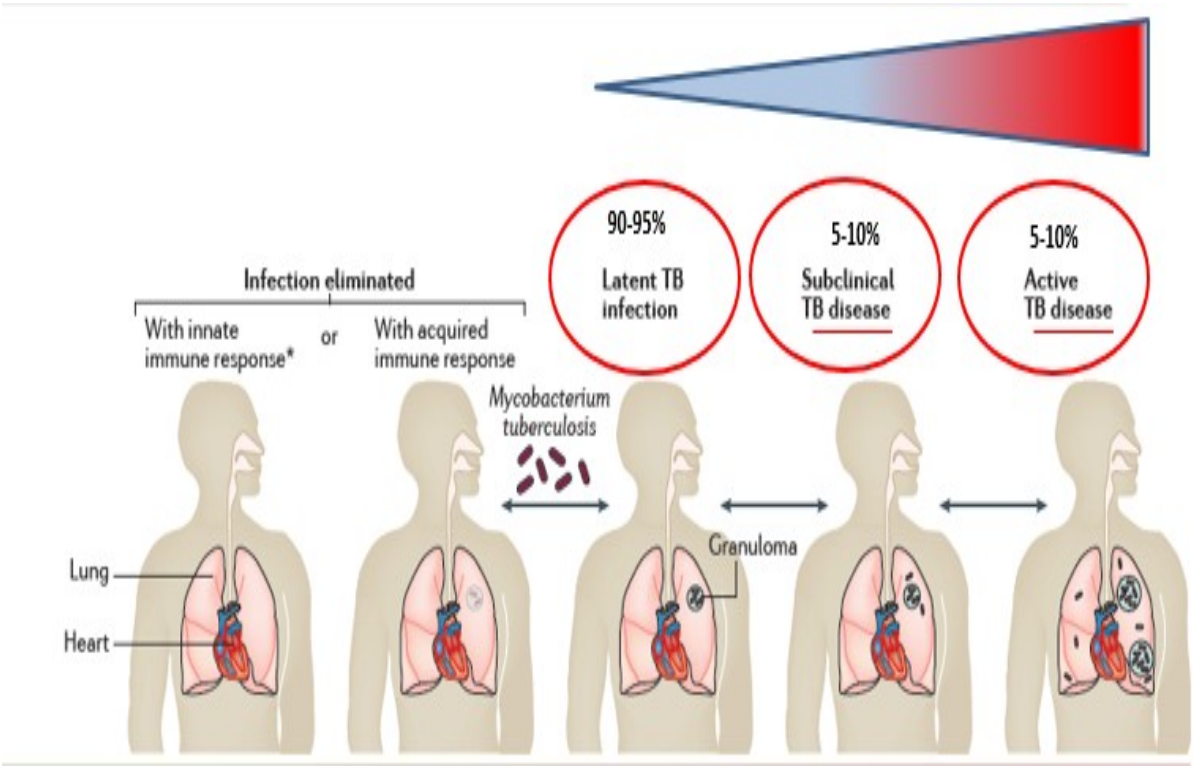


Figure 1.4: The spectrum of TB from *M. tuberculosis* infection to active disease (Pai *et al.*, 2016).

3. BCG vaccine

Today the only vaccine available against *M. tuberculosis* is Bacillus Calmette Guerin (BCG) obtained one hundred years ago by Calmette and Guerin at the Pasteur institute in France, after thirteen years of serial in vitro passage of a *M. bovis* strain isolated from cattles (Luca and Mihaescu, 2013). Recent genomic studies showed that the genetic determinants underlying the attenuation of *M. bovis* BCG is the deletion of long region of genome named Region of deletions (RD). A total of 14 RD regions have been identified, the most important of which is RD1 that is absent in all BCG strains and correspond to a 9.5 kbp region that in *M. tuberculosis* encodes 9 genes (Behr *et al.*, 1999), encoding among them two secreted low molecular weight proteins, CFP-10 and ESAT-6, which are transcribed together, that have been recognized as virulence factors and potential vaccine candidates (Brandt *et al.*, 2000). It has also been shown that deletion of RD-1 in *M. tuberculosis* strains leads to a strong attenuation of pathogenicity in mouse model of *M. tuberculosis* (Pym *et al.*, 2002). Protection against tuberculosis afforded by BCG is high in children, specifically against the most severe forms of the disease. However, protection is known to decrease over time resulting in variable outcomes of protection ranging from 20% to 80%. BCG protective effects tends indeed to wane in early adolescence (Colditz *et al.*, 1994), due to genetic and environmental factors. There are several TB vaccines in phase I or phase II trials. For example, Ad5-Ag85A vaccine in Phase I clinical trial, AERAS-402/crucel Ad35 vaccine, the recombinant BCG expressing lysteriolysin (Groschel *et al.*, 2014), (Cardona, 2006), and the M72/AS01E vaccine (Marisol Ocampo C, 2015). However, until now a vaccine effective in preventing TB in adults remains elusive (WHO, 2016).

4. *M. tuberculosis* treatment and emergence of MDR and XDR

Only very few antibiotics, rifampicin, isoniazid, pyrazinamide and ethambutol as first line and ethionamide, fluoroquinolones, streptomycin, aminoglicosides (amikacin and kanamycin) as second line are known to be active against active TB (Laurenzo and Mousa, 2011). Treatment of latent (asymptomatic) infection consists of INH taken for 6-9 months. Rifampin can be used in those exposed to INH resistant strains. The combination of rifampin-pyrazinamide should not be used because it caused a high rate of liver injury. Multiple drug resistant *M. tuberculosis* strains (MDR) have emerged primarily in AIDS patients and have resistance to both INH and rifampin, but some isolates are resistant to three or more drugs. The treatment of MDR organisms includes using of four or five drugs, ciprofloxacin, amikacin, ethionamide and cycloserine. *M. tuberculosis* strains resistant to INH, rifampin, a fluoroquinolone, and at least one additional drug are called XDR (extensively drug resistant) (levinsin W, 2008).

Drug	Gene	Mechanism Involved
Isoniazid	*katG*, *inhA*	Catalase/peroxidase; enoyl reductase
Rifampicin	*rpoB*	RNA polymerase
Pyrazinamide	*pncA*, *rpsA*	Pyrazinamidase; ribosomal protein 1
Ethambutol	*embB*	Arabinosyl transferase
Streptomycin	*rpsL*, *rrs*, *gidB*	S12 ribosomal protein, 16A rRNA, 7-methylguanosine methyltransferase
Quinolones	*gyrA*, *gyrB*	DNA gyrase
Capreomycin	*rrs*, *tlyA*	16S rRNA, rRNA methyltransferase
Kanamycin/Amikacin	*rrs*	16S rRNA
Ethionamide	*ethA*	Enoyl-ACP reductase
Para-aminosalicylic acid	*thyA*, *folC*	Thymidylate synthase A

Figure 1.5: First and second line anti-tuberculosis drugs currently in use and target of action (Palomino and Martin, 2014) .

Within the last 10 years, the mechanism of action of most of the anti-tuberculosis agents have been described, and many studies are beginning to elucidate some of the molecular mechanisms whereby *M. tuberculosis* becomes resistant.

The genetic basis of resistance for some anti-tuberculosis agents is not fully known. For example, streptomycin resistance emerges through mutations in *rrs* and *rpsL* that produce an alteration in the streptomycin binding site. Isoniazid-resistance is caused by modification of KatG, which is the enzyme that activate isoniazid to the active hydrazine derivative. Mutation in *KatG* lead to high-level resistance to isoniazid (Zhang *et al.*, 1993). A deficiency in enzyme activity produces high-level resistance and is found in more than 80% of isoniazid-resistant strains. Most pyrazinamide resistant organisms have mutations in the pyrazinamidase gene (*pncA*). Pyrazinamidase is essential in producing the active pyrazinoic acid derivative, and mutants are unable to produce an active drug (Gillespie, 2002), (Ramaswamy and Musser, 1998). Ethambutol resistance in approximately 60% of organisms is due to amino acids replacements at position 306 of an arabinosyltransferase encoded by *embB* gene (Ramaswamy and Musser, 1998) (Fig 1.5).

Bedaquiline has recently received conditional approval for the treatment of MDR-TB under the trade name Sirturo after the results of two phases, phase II clinical trials and phase III trials was scheduled to begin in 2013, The mechanism of action of bedaquiline is by inhibiting the ATP synthase of *M. tuberculosis*, which was a completely new target of action for an anti-mycobacterial drug. The only mutation found was in the *atpE* gene, which encodes the c part of the F0 subunit of the ATP synthase. Nevertheless, in a study to further assess the mechanisms of resistance to bedaquiline in *M. tuberculosis*, it was found that only 15 out of 53 resistant mutants had mutations in *atpE*. The other strains do not have

mutations in *atpE*, which suggests that other mechanisms of resistance are still possible (Huitric *et al.*, 2010).

Delamanid acts by inhibiting the synthesis of mycolic acid and is undergoing clinical evaluation in a phase III trial. Delamanid has more recently shown its safety and efficacy in a clinical evaluation for MDR-TB. It only inhibits methoxy- and keto-mycolic acid while isoniazid also inhibits α-mycolic acid requires reductive activation by *M. tuberculosis* to exert its activity. In experimentally generated delamanid-resistant mycobacteria, a mutation was found in the *Rv3547* gene, suggesting its role in the drug activation (Palomino and Martin, 2014). The increasing prevalence of drug-resistant strains of *M. tuberculosis* makes the development of novel drugs for tuberculosis and identify potential drug targets an urgent priority. For new drug targets there are several criteria that should be considered. First, the drug target must be essential to bacterial viability, virulence or the persistence of *M. tuberculosis* in granulomas. Second, targeting a novel pathway not inhibited by existing drugs may reduce the chance of cross-resistance with current drug-resistant strains. Third, targets not conserved in humans may reduce the likelihood of off-target effects. Fourth, the target must be accessible to inhibitors, which is particularly important for penetrating the unique and highly impermeable cell envelope of *M. tuberculosis*. in general, target that are positioned outside the cytoplasmic membrane will be more accessible (Feltcher *et al.*, 2010).

5. Cell structure

Mycobacterium tuberculosis is a long large non-motile aerobic acid- fast rod-shaped bacterium belonging to the order of actinomycetales (William A.Strohl, 2001). It grows slowly and has a doubling time of 18 hours (levinsin W, 2008). The rods are 2-4 µm in length and 0.2-0.5 µm in width (Fig 1.6A), *M. tuberculosis* colonies are small and buff colored when grown on solid medium Figure (Fig 1.6B) (Kenneth Todar, 2017).

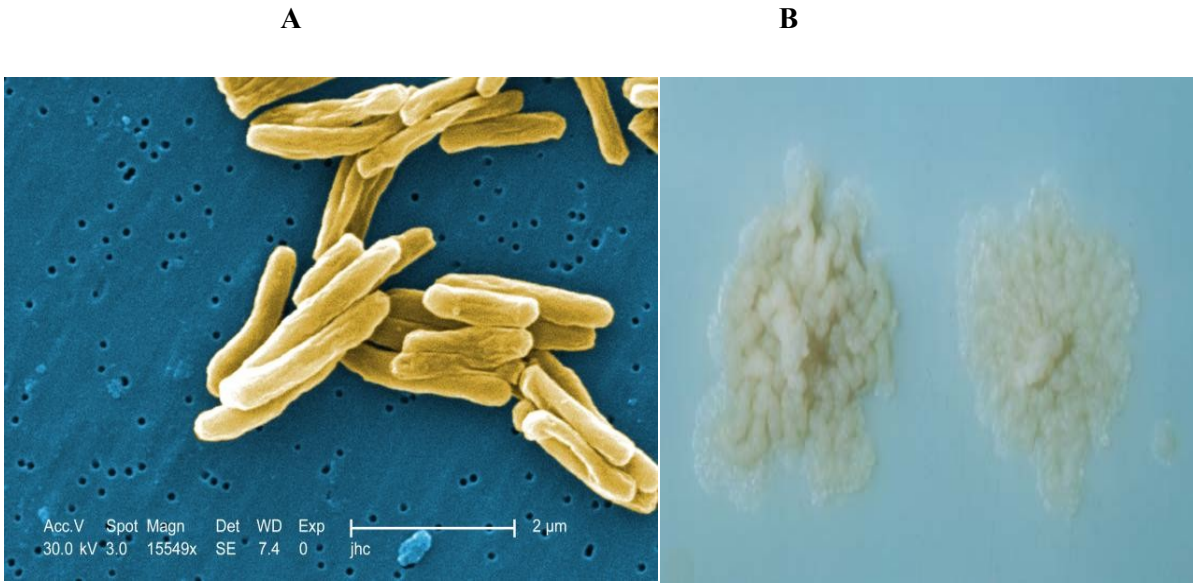


Figure 1.6 : *M. tuberculosis* scanning electron micrograph. Mag 15549X (a), Colonies of *M. tuberculosis* on Lowenstein-Jensen medium (B) (Kenneth Todar, 2017).

5.1 The mycobacterial cell wall

Mycobacterium are surrounded by a unique cell wall with unusual low permeability. A peptidoglycan layer surrounds the plasma membrane and long chain polysaccharides such as lipoarabinomannan (LAM), lipomannan and arabinogalactan covalently link the peptidoglycan layer with the outer membrane.

The mycobacterial outer membrane, also named mycomembrane, is functionally similar the membrane of gram-negative bacteria, though its composition is peculiar. The inner leaflet of the mycomembrane is composed of mycolic acids and the outer leaflet of small glycolipids, sulfolipids and other lipids. The mycomembrane provides strength and impermeable barrier to mycobacteria, due to its high hydrophobicity (Zuber *et al.*, 2008) (Brennan and Nikaido, 1995), (Brennan and Besra, 1997). Many of the drugs used to combat mycobacteria are effective because they specifically target the biosynthesis of the mycobacterial cell wall components.

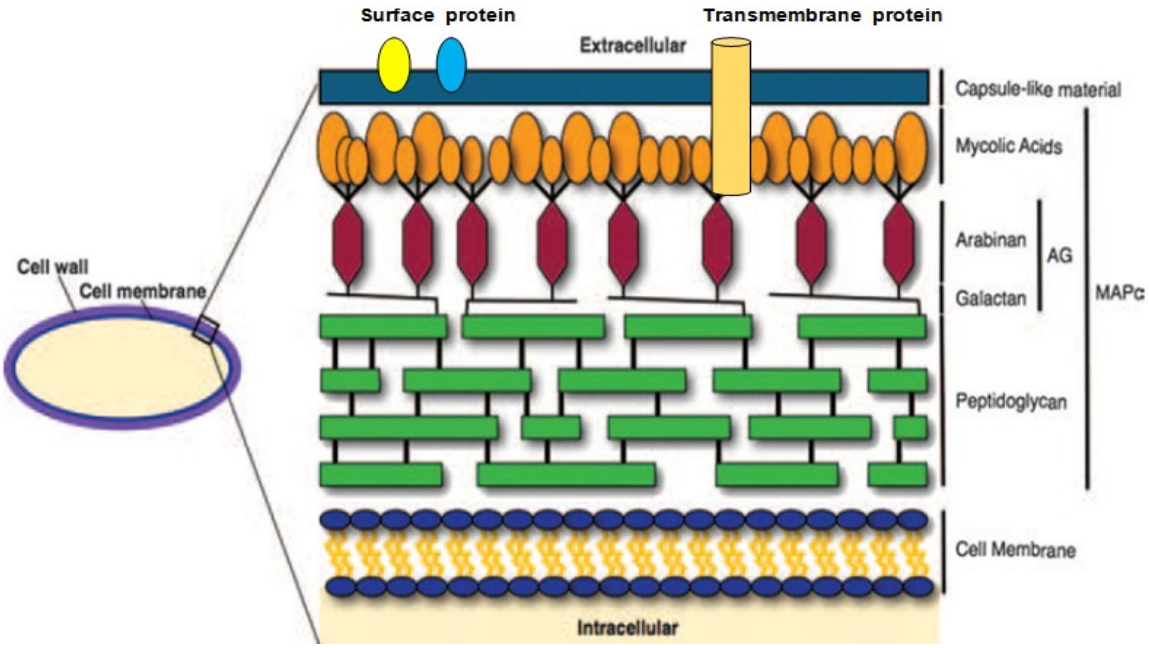


Figure 1.7: Basic components of the mycobacterial cell wall. MAPc, MA-AG-PG complex (Hett and Rubin, 2008).

6. The type VII protein secretion pathways

The virulence of bacterial pathogens highly depends on the ability to secret proteins and molecules to the bacterial surface, external milieu or directly into host cells (Abdallah *et al.*, 2007). Because the biological membranes in the bacterial cell wall hinder export of proteins, translocation across these barriers is mediated by dedicated proteins secretion systems. Similar to other bacteria, mycobacteria secret proteins across inner membrane via ubiquitous general secretory (sec) pathway or twin-arginine translocation (tat) system (Champion and Cox, 2007). In gram positive bacteria, which have only one lipid bilayer, the Sec/Tat pathways are generally sufficient for protein export. Gram negative bacteria, have evolved a number of specialized secretion systems for transport of protein across the outer membrane. Six pathways, generally known as the type I to type VI secretion systems, either secrete proteins that are delivered into periplasm by sec or Tat system (type II and V) in two step process, or via a signal peptide-independent one-step mechanism across the entire cell envelope (type I, III, IV and VI). Although mycobacterium also contains a diderm cell envelope, they lack type I to type VI pathways and evolved a unique specialized secretion system which known as the ESX or type VII secretion pathway. In *M. tuberculosis* there are five type VII secretion systems encoded by gene clusters and called ESX1 to ESX5 (Fig 1.8). ESX1 and ESX5 secrete different proteins involved in the virulence of *M. tuberculosis*, while ESX1 is missing in the attenuated *M. bovis* vaccine strain Bacille Calmette and Guerin (Delogu *et al.*, 2013). ESX1 is required for the full virulence of *M. tuberculosis*, which uses this secretion system to escape from the phagosome into the host cell cytosol of infected macrophages where it may persist in a protected environment (Romagnoli *et al.*, 2012). ESX1 mediates secretion many antigens, ESAT-6 and CFP-10, both small highly immunogenic proteins that

form the basis of the immunological diagnosis of *M. tuberculosis* infection in the interferon-gamma release assays (IGRAs), and appears that the ESX-1 secreted proteins have the ability to disrupt the biological membrane (Brennan and Nikaido, 1995), (Gao *et al.*, 2004). It has also been demonstrated that ESX3 secretion system is responsible for the secretion of some soluble factors required for growth that are probably involved in optimal iron and zinc uptake (Serafini *et al.*, 2009). ESX5 is restricted to the slow growing species, while it is absent from the genome of fast growing bacteria such as *M. smegmatis.* ESX-5 is found in *M. tuberculosis* complex (MTBC), *M. marinum*, *M. ulcerans*, *M. leprae*, and *M. tuberculosis* (Gey van Pittius *et al.*, 2001), and it may represents a secretion systems specifically evolved to interact with a complex immune system such as that of mammals. Indeed the ESX-5 was shown to play a role in immunomodulation (Abdallah *et al.*, 2008), and induce cell death which facilitate cell to cell spread and it was hypothesized that ESX-5 effectors will interact and manipulate the host cell after ESX-1 mediated escape from phagosome into cytosol during the infection (Abdallah *et al.*, 2011). ESX5 appears to be a major export pathway for PE/PPEs *M. marinum* proteins, especially for the most recently evolved members, the PE_PGRS and PPE_MPRT proteins (Abdallah *et al.*, 2009), while the role and function of ESX2 and ESX4 are still unknown, the study of the role of ESX systems in TB pathogenesis is one of the major advancements of the last decade in the TB field. The ESX export system represent other potential targets for new anti TB drugs, ESX-1 and ESX-3 are known to be essential for virulence and growth of *M. tuberculosis*, respectively. An inhibitor that targets a conserved core component of the ESX pathways (EccB, EccD, EccE and MycP), has the potential to disrupt all ESX systems simultaneously, which could reduce evolution of drug resistence. There are also secreted proteins of the ESX system

that may function in the ESX secretion process (EspA and ESAT-6/CFP-10) that could be accessible to inhibition (Feltcher *et al.*, 2010).

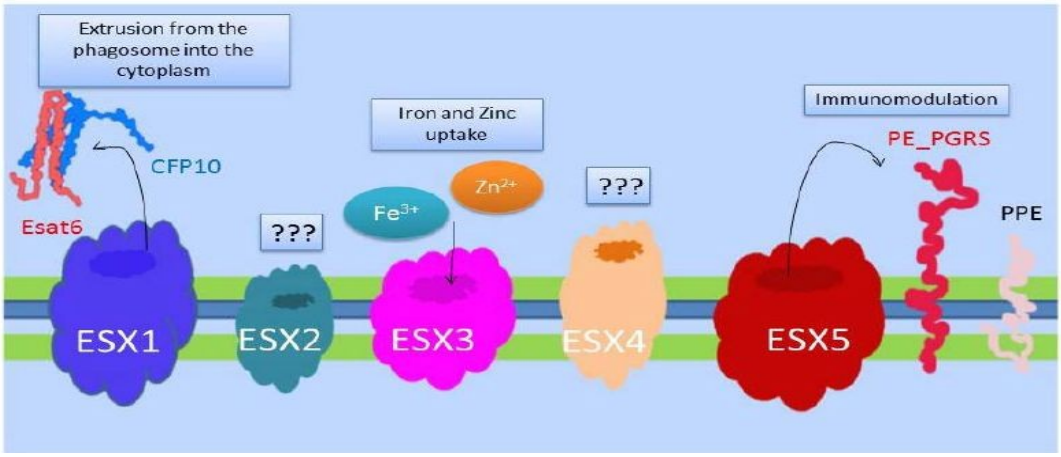


Figure 1.8: Protein secretion systems. ESX1 secretes antigens that interfere with the integrity of the phagosomal membrane, leading to phagosomal rupture and bacterial emission into the cytosol. ESX5 is present only in slow growing mycobacteria (such as *M. tuberculosis* and *M. marinum*) and it is thought to be involved in the secretion of proteins (PPE and PE-PGRS) with immunomodulatory properties. ESX3 is involved in Zinc and Iron uptake and homeostasis and as such is essential for growth. The role of ESX2 and ESX4 remain still unknown (Delogu *et al.*, 2013) .

6.1 Genetic organization of ESX systems and their secreted proteins

ESAT-6 (early secreted antigentic target of 6 KDa) and CFP-10 (culture filtrate protein of 10KDa) which are encoded by an operon are both secreted in the culture medium of *M. tuberculosis* (Berthet *et al.*, 1998), (Sorensen *et al.*, 1995) and belong to WXG 100 family of proteins, characterized by Trp-x-Gly motif. Comparative genomic methods revealed that *esat-6* and *cfp-10* genes are found in a region known as region of difference 1 (RD1) which is present in virulent strains of *M. tuberculosis* and *M. bovis* but absent from the genome of *M. bovis* BCG and *M. microti* because of independent deletion events (Daniel, 2006).

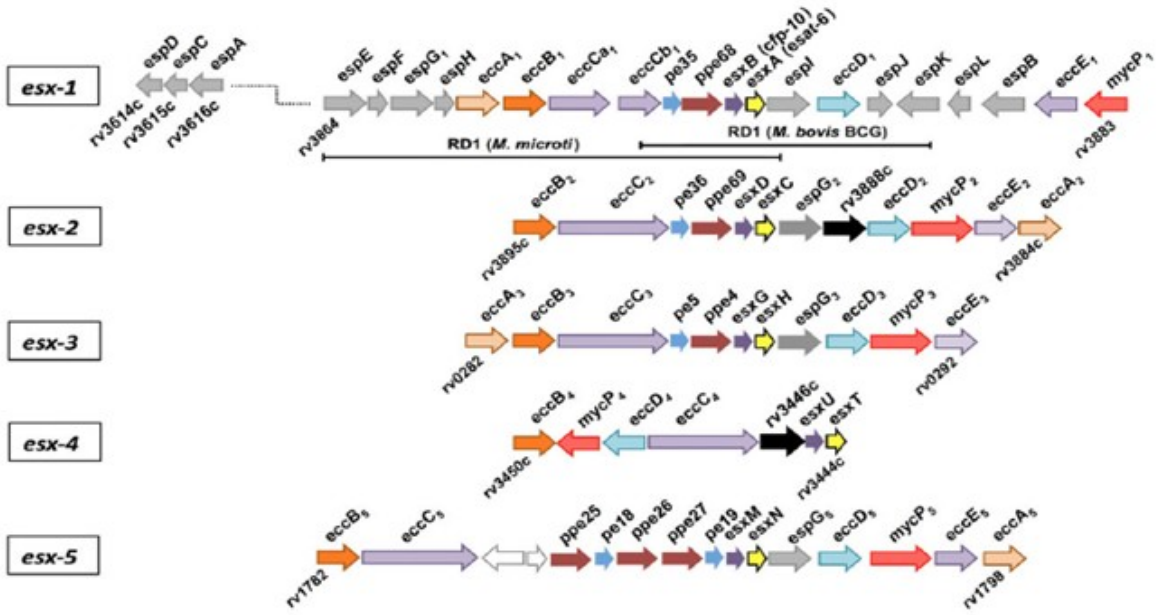


Figure 1.9: Genetic organization and gene names of the five ESX loci and espA operon in *M. tuberculosis*. The RD1 deletions of *M. Bovis* BCG and *M. Microti* are marked in ESX-1 cluster. Ecc stands for ESX conserved component and esp for ESX-1 secretion system (Majlessi *et al.*, 2015).

Phylogenetic analysis and genomic comparison suggest that the five ESX systems ESX-1 to ESX-5 have evolved by duplication events, where the ESX-4 cluster contains the lowest number of genes, is thought to be the most ancestral cluster which duplicated to give arise ESX-1, ESX-2, ESX-3 and finally ESX-5 (Fig 1.9). Interestingly, ESX-5 is restricted to slow growing pathogenic mycobacteria such as *M. tuberculosis, M. leprae and M. marinum* and absent from the genome of fast-growing *M. smegmatis*, and seems to be a major secretion pathway of PE and PPE proteins like PPE41 (Abdallah *et al.*, 2009) (Abdallah *et al.*, 2006). Conversely, ESX-1 to ESX-4 clusters are distributed in the genome of mycobacteria and ESX-1 cluster required for secretion of ESAT-6 and CFP-10, encodes a functionally secretion system in mycobacteria such as *M. marinum* and *M. smegmatis* (Converse and Cox, 2005).

7. PE-PPE protein family

During the analysis of complete genome sequence of *Mycobacterium tuberculosis* H37Rv, two gene families were identified, encoding proteins with conserved N-terminal domains characterized by motifs Pro-Glu (PE) or Pro-Pro-Glu (PPE) near their respective N-termini, which occupy around 8% of the coding capacity of the genome. These two families consist of 100 and 67 members, respectively and each family has been divided into subfamilies. PE/PPE proteins can consist of only these so called PE and PPE domains (Fig 1.10A), or they may have extended C-termini (Fig 1.10B). In the latter case, the C-terminal domains are composed of unique sequences or of sequences with characteristic glycine-rich repeat, such as those encoded by polymorphic CG-rich-repetitive sequence (PGRS) subfamily of PE proteins and the major polymorphic tandem repeat (MPTR) PPE subfamily mainly encode glycine and aspargine.

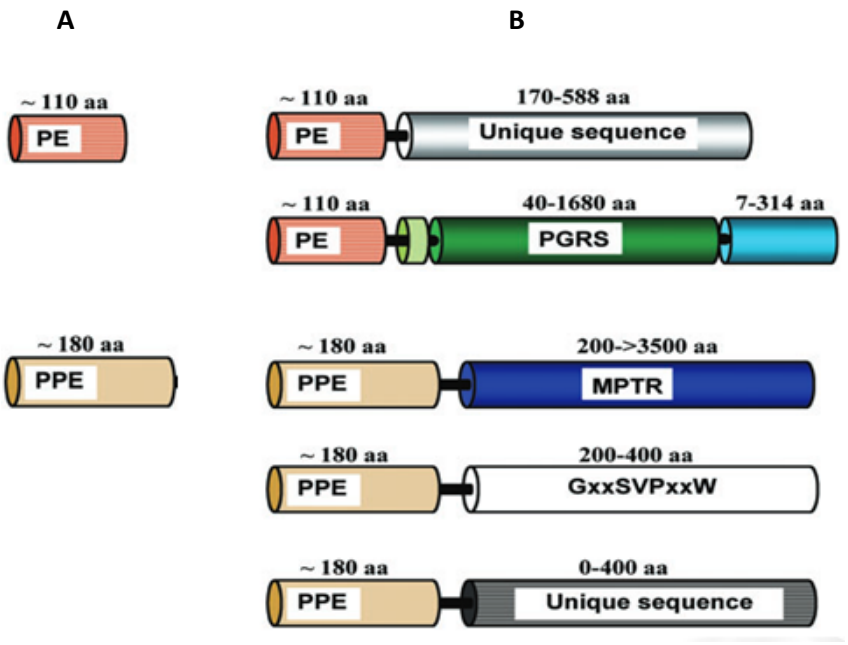


Figure 1.10: Domain organization of the PE and PPE proteins. (A) the most ancestral members of these family consist of only PE/PPE domain. (B) During evolution, the PE/PPE proteins appear to have acquired extended C-terminal domains (Bottai and Brosch, 2009).

Comparative genomic studies of mycobaterial species has shown that *pe/ppe* genes are largely unique to mycobacteria. Moreover, fast – growing mycobacteria carry only few of these genes and lack genes of the PE_PGRS and PPE-MPTR subfamilies, which are enriched in the genome of slow-growing, pathogenic species (Gey van Pittius *et al.*, 2006). Many members of PE and PPE proteins were shown to localize on the mycobacterial surface, where they are available to interact with host components (Banu *et al.*, 2002), (Malen *et al.*, 2010). The highly polymorphic nature of the C-terminal of these proteins (PE_PGRS and PPE_MPTR) which exhibit the most sequence variation suggests their involvement in the anti-genetic variation (Sampson, 2011). This hypothesis is supported by a study showing that a DNA vaccine expressing the N-terminal PE region of the PE_PGRS33 protein is able to elicit a cellular immune response, while mice immunized with 1818PE_PGRS develop primarily a humoral response (Delogu and Brennan, 2001). It also proved that the PE domain of the PE_PGRS33 is necessary for the subcellular localization while the PGRS domain but not PE domain affect the bacterial shape and colony morphology (Delogu *et al.*, 2004). However, the accumulation of *pe/ppe* genes into the genomes of mycobacterial pathogens highlights on the important role of these proteins in mycobacterial pathogenesis. Many of *pe/ppe* genes are located within ESX clusters and both loci are evolutionary correlated. The expansion of PE/PPE family is associated with the duplication of ESX gene cluster and the emergence of repeat proteins PGRS and MPTR is a recent evolutionary event (Gey van Pittius *et al.*, 2006) (Fig. 9). The PE/PPE families and ESX systems also appear to be functionally linked, as several PE and PPE proteins have shown to be substrate to the ESX system. For example, PE35 which is encoded directly upstream of *ppe68*, *esxB* (CFP-10) and *esxA* (ESAT-6), influence the expression of these three genes.

Transposon insertion in the PE35 gene in two independently derived *M. tuberculosis* strains failed to express PPE68, ESAT-6 or CFP10 and were both attenuated, while in contrast, transposon mutagenesis or partial deletion of *ppe86* did not have attenuated effect in mouse infection model (Delogu G and Cole S.T and borsch R, 2008). Another example is the secretion of PPE42, a hydrophilic protein expressed within the ESX-5 system, which has been shown to induce a strong B cell response in human. PPE41 forms a heterodimeric protein with the neighbouring PE25 and expressed as an operon (Strong *et al.*, 2006). However, while the export of a single PE/PPE couplets has been attributed to the ESX-1 and to all the other ESX systems, (Sani *et al.*, 2010), the ESX-5 seems to be the major export pathway for the most recently evolved proteins such as PE_PGRS and PPE_MPRT, which were shown to play an essential role for the full virulence of *M. tuberculosis* (Abdallah *et al.*, 2009), (Abdallah *et al.*, 2006), (Ahmed *et al.*, 2015). Moreover, several transcriptional factors have been reported to be involved in the modulation of *pe/ppe* genes expression (Ahmed *et al.*, 2015), (Mohareer *et al.*, 2011). For instance, the iron dependent regulator ideR, responsible for the induction of genes involved in iron uptake, regulates around 11 *pe/ppe* genes (Rodriguez *et al.*, 2002). Furthermore, PhoPR, which is a two component system involved in the regulation of genes encoding the type VII secretion system ESX1 and genes involved in synthesis of several cell wall components, regulates *pe/ppe* genes (Solans *et al.*, 2014).

7.1 PE_PGRS subfamily

All PE_PGRS proteins have a common molecular architecture. The PGRS domain is linked to a highly conserved N-terminal domain that is around 90-100 amino acids along and contains the PE motif (Fig 1.10). The PGRS domain is linked to the PE domain with a linker region 35-40 amino acids long and contains a conserved GRPLI motif (AC domain). This domain may form a putative helix and has been suggested to play a role in the localization of PE_PGRS proteins (Delogu G and Cole S.T and borsch R, 2008) or may serve to anchor the PE_PGRS to a specific, yet unidentified portion of the cell wall. The PE_PGRS is characterized by the presence of multiple tandem repetitions (Gly-Gly-Ala and Gly-Gly-Asn motifs) that vary in number from several tens to hundreds amino acids. These repeats are intercalated by short regions of diverse sequence composition and size, and these differences may be used to define subgroups within PE_PGRS proteins (Delogu G and Cole S.T and borsch R, 2008).

Group A1: which has intercalating sequences of 4-10 amino acids long and present unique C-terminal domain that may extend up to 15-27 amino acids.

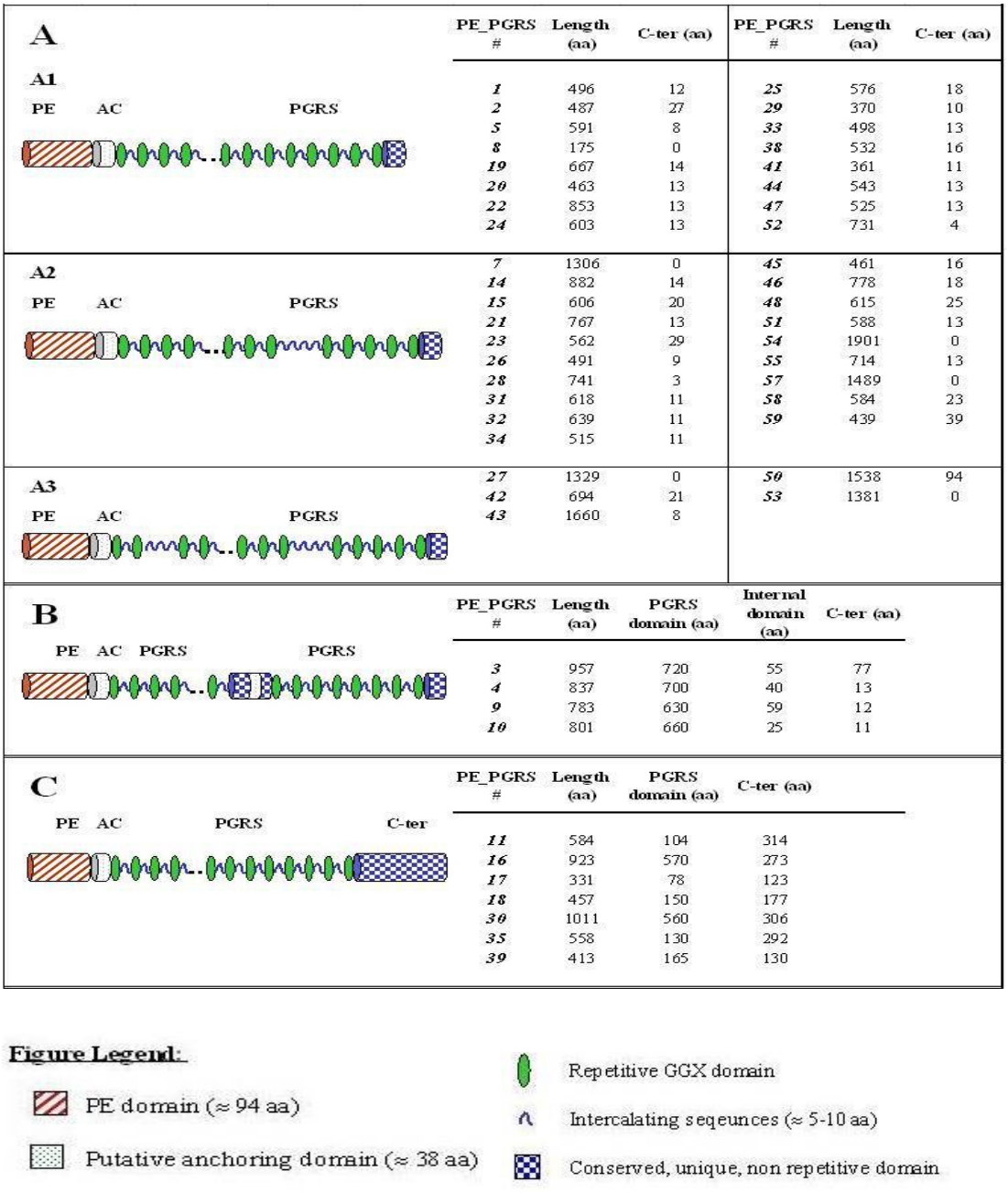
Group A2: contains intercalating sequences that may extend up to 20 amino acids.

Group A3: contain high number of these intercalating sequence.

Group B: consists of four PE_PGRS which have an intercalating domain that varies in size from 25 to 59 amino acids and containing inside PGRS domain a second GRPLI motif.

Group C: contains a large C-terminal domain that can be as large as 300 amino acids (Fig. 1.11).

However, the PGRS domain cannot simply be considered a repetitive or redundant domain but rather contains a specific sequence interspaced among GGX-GGX regions that make this domains unique for each PE_PGRS protein. The lack of experimental data about PE_PGRS domains make it necessary to study.



A	PE_PGRS #	Length (aa)	C-ter (aa)	PE_PGRS #	Length (aa)	C-ter (aa)
A1	*1*	496	12	*25*	576	18
	2	487	27	*29*	370	10
	5	591	8	*33*	498	13
	8	175	0	*38*	532	16
	19	667	14	*41*	361	11
	20	463	13	*44*	543	13
	22	853	13	*47*	525	13
	24	603	13	*52*	731	4
A2	*7*	1306	0	*45*	461	16
	14	882	14	*46*	778	18
	15	606	20	*48*	615	25
	21	767	13	*51*	588	13
	23	562	29	*54*	1901	0
	26	491	9	*55*	714	13
	28	741	3	*57*	1489	0
	31	618	11	*58*	584	23
	32	639	11	*59*	439	39
	34	515	11			
A3	*27*	1329	0	*50*	1538	94
	42	694	21	*53*	1381	0
	43	1660	8			

B	PE_PGRS #	Length (aa)	PGRS domain (aa)	Internal domain (aa)	C-ter (aa)
	3	957	720	55	77
	4	837	700	40	13
	9	783	630	59	12
	10	801	660	25	11

C	PE_PGRS #	Length (aa)	PGRS domain (aa)	C-ter (aa)
	11	584	104	314
	16	923	570	273
	17	331	78	123
	18	457	150	177
	30	1011	560	306
	35	558	130	292
	39	413	165	130

Figure 1.11: Division of PE_PGRS proteins into subfamilies (Delogu G and Cole S.T and borsch R, 2008).

In the genome of *M. tuberculosis* 63 open reading frames have annotated as *pe_pgrs* genes, but a number of them have a frameshift mutations that prevent the synthesis of some functional proteins and two proteins (PE_PGRS62 and 63) lack the typical PGRS domain and the linker region, 51 potentially functional *pe_pgrs* genes scattered throughout the *M. tuberculosis* H37Rv genome. *pe_pgrs* genes were found only in the MTBC or other pathogenic mycobacteria, like *M. ulcerans* and *M. marinum,* highlighting their potential involvement in the pathogenesis of *M. tuberculosis*. Many studies raised the attention on these proteins which are directly involved in the pathogenesis of *M. tuberculosis* infection and in the evasion from the host immune response (Brennan and Delogu, 2002), (Forrellad *et al.*, 2013), (Lalita Ramakrishnan, 2000). The differences between the genetic organization of *pe_pgrs* genes and the *pe/ppe* genes families suggests an autonomous and independent regulation of gene expression for most of PE_PGRS proteins. *M. tuberculosis* can differently regulate the expression of PE_PGRS. For example *pe_pgrs30* gene expression increased following intracellular growth in bone marrow-derived macrophages but not in type-II human pneumocytes, while *pe_pgrs9* was induced in both in vitro systems (Iantomasi *et al.*, 2012). Another study showed that PE_PGRS16 and PE_PGRS26 are inversely regulated in macrophages and in mice infected with *M. tuberculosis* (Dheenadhayalan *et al.*, 2006). Interestingly, *M. smegmatis* strains expressing PE_PGRS33 and PE_PGRS26 were able to persist at higher level in spleen and liver tissues compared with *M. smegmatis* expressing PE_PGRS16, suggesting a differential role of these proteins in mycobacterial pathogenesis (Singh *et al.*, 2008). Upregulation of *pe_pgrs9*, -16 and -30 genes were observed in *M. tuberculosis* infected macrophages and in host tissues during the chronic step of infection and the upregulation was higher in the spleen compared to the lung infected mice (Delogu G and Cole S.T and borsch R, 2008). Detailed analysis identified

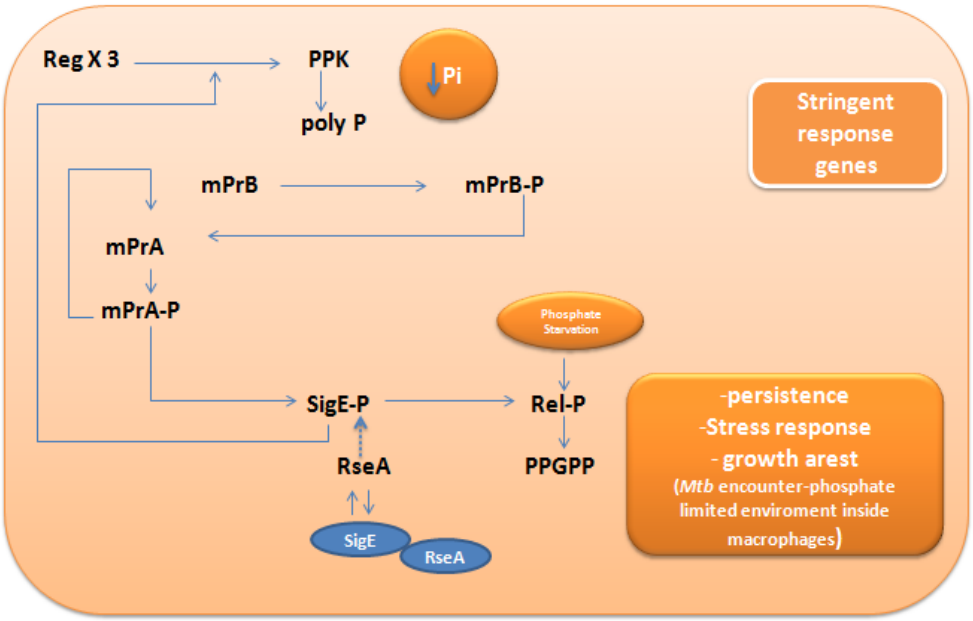
PE_PGRS47 as an inhibitor of autophagy and a factor contributing to evasion of both innate and adaptive immunity by *M. tuberculosis* (Saini *et al.*, 2016). Another data have been demonstrated that some sigma factors are involved in the expression of some PE_PGRS like SigA which mediates in-vitro transcription of PE_PGRS33 and their expression is repressed in stress condition (Vallecillo and Espitia, 2009). Therefore, we can summarize that *pe_pgrs* genes are differently expressed and regulated by *M. tuberculosis* in host tissue depending on the different environmental conditions that mycobacteria faced during infection. Both innate and adaptive immune responses play an important role in *M. tuberculosis* infection, and some studies demonstrated that the PE and PGRS domains have a role in cellular and humoral immune responses (Cohen *et al.*, 2014). It has been suggested that the PGRS domain could be the target of the host immune response because of its extensive variability, although the link between the genetic variability and antigenic variation is still hypothetical (Delogu G and Cole S.T and borsch R, 2008). Interestingly, the Epstein–Barr Virus nuclear antigen 1 (EBNA1) shows significant similarity with the PGRS domain of PE_PGRS proteins, containing numerous Gly-Ala repeats that are known to inhibit antigen processing and presentation through the major histocompatibility complex I (MHCI pathway) (Cole *et al.*, 1998), (Brennan and Delogu, 2002). PE_PGRS proteins have domains that confer resistance to ubiquitin/proteosome dependent protein degradation and may use this mechanism to evade immune detection and killing of mycobacterium infected cells (Koh *et al.*, 2009). Another study has demonstrated that two PE_PGRS proteins, PE_PGRS 17 and PE_PGRS11, recognize TLR2 and induce the maturation and activation of human dendritic cells, enhancing the ability of dendritic cells to stimulate CD4+ T cells. In this way PE_PGRS proteins could contribute in the initiation of innate immune response during *M. tuberculosis* infection (Bansal *et al.*, 2010).

Moreover, studies carried out on another PE_PGRS proteins showed that the unique C-terminal domain of the PE_PGRS30, which is around 300 AA, is not required for the full virulence, further implicating the PGRS domain in TB pathogenesis (Iantomasi *et al.*, 2012). In line with these findings, a *M. marinum* mutant for a *pe_pgrs* gene (*MMAR_0242*), encoding for a protein containing an extended and unique C-terminal domain, was shown to be impaired in its ability to survive intracellulary. Attenuation of the mutant was the result of lack of inhibition of phagosomal/lysosomal fusion (Singh *et al.*, 2016). Recently, a study has demonstrated that even a small PGRS region of PE_PGRS33, containing few repeats, can activate the TLR2 depending entry in macrophages (Palucci *et al.*, 2016). These experimental evidences provide support to the role of PE_PGRS proteins in TB pathogenesis.

8. Stringent responses and phosphate depletion in *M. tuberculosis*

The mycobacterial pathogen *M. tuberculosis*, has a remarkable adaptation against various physiological and environmental stresses including that induced by drugs. The granuloma formation during the tubercular infection, with the encircling and enclosure of bacilli and infected cells, is a classical example of the physical, chemical and biological changes encountered by *M. tuberculosis* during infection (Ghosh *et al.*, 2011). Interestingly, *M. tuberculosis* can survive over years in a latent state and under the pressure of the host immune responses. To resist this harsh environment, *M. tuberculosis* is able to modulate a number of metabolic processes which are regulated by the so called stringent response. The stringent response has been characterized by a number of studies, and expression of *relA* has been shown to initiate the expression of a number of genes that may lead to the dormant state. The importance of RelA arises from the fact that it synthesizes the stringent response regulator ppGpp (Guanosine Tetraphosphate) which is essential for the long term survival of *M. tuberculosis* and persistent infection in mice by altering the expression of antigenic and enzymatic factors that may contribute to successful latent infection (Dahl *et al.*, 2003), (Sureka *et al.*, 2008). *M. tuberculosis* and *M. smegmatis* both have the ability to survive for a long time under stress condition and partly share the elements of stringent response pathway (Ojha *et al.*, 2000). Recent studies have provided information about the stress signalling pathways in mycobacteria starting from ppK1 and poly phosphate (poly p) which serve as a phosphate donor in the conversion of MprB to MprB-P, facilitate transcription of *sigE* which regulate the transcription of *relA* and play a key role in activation of the stringent response in mycobacteria (Sureka *et al.*, 2007) (Fig 1.12.A).

A



B

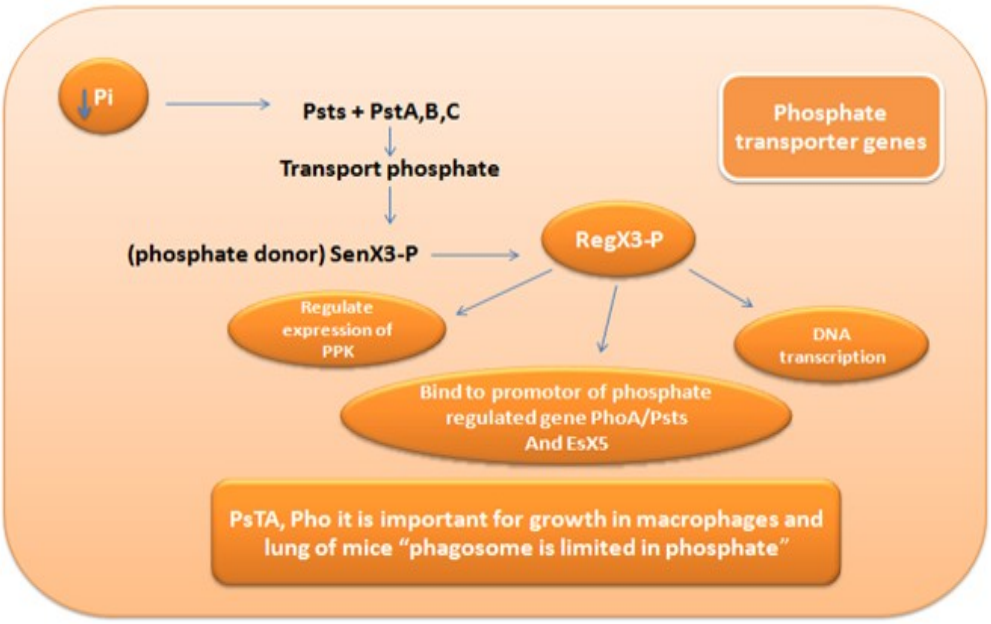


Figure 1.12: The important components of stringent response pathway (A) and phosphate transporter pathway (B) in *M. tuberculosis* and *M. smegmatis* (Ghosh *et al.*, 2011), (Rifat *et al.*, 2009), (Sureka *et al.*, 2007).

The two genes *mprAB* encode the histidine kinase sensor MprB and its cytoplasmic partner response regulator MprA, which responds to the environmental changes sensed by MprB by regulating adaptive transcriptional programs. Poly phosphate kinase 1 (PPK1) catalyses the poly-p providing phosphate to MprAB, possibly to face the phosphate limited environment inside the macrophages (Rifat *et al.*, 2009). MprA-P activates the transcription of *sigE*, which activates transcription of $relA_{Msm}$ (Sureka *et al.*, 2007). Thus, *sigE* and $relA_{Msm}$ are indirectly responsive to poly phosphate levels which increase under stress condition. The *relA* expression in mycobacteria is controlled by a complex signalling cascade that depends on the amount of polyphosphate present in the cell. High levels of polyphosphate lead to elevated expression of *sigE* and $relA_{Msm}$ which is correlated with slowed growth and increased isoniazide tolerance (Thayil *et al.*, 2011). This transcriptional network has a positive feedback where MprA-P activates its own transcription. This leads to a high and low level expression of $relA_{Msm}$ which maintain two cell populations. One with high level expression of $relA_{Msm}$ more likely to exhibit a persister cell phenotype, slow growth and greater resistance to antibiotic and other stresses generated by the host (Dahl *et al.*, 2005). The other cell population with low level expression which has the bias to grow (Boutte and Crosson, 2013). Also in *M. tuberculosis* the *ppk1* is significantly up regulated due to phosphate starvation resulting in the synthesis of inorganic poly-p, the two component system SenX3-RegX3 is known to be activated in phosphate starvation in both *M. tuberculosis* and *M. smegmatis* and required for the virulence of *M. tuberculosis* (Glover *et al.*, 2007), (Parish *et al.*, 2003). RegX3 has been shown to regulate the expression of *ppk1* when the phosphorylated RegX3 binds to *ppK1* promoter of *M. tuberculosis* and both SigE and RegX3 were found to regulate the transcription of *ppK1* promoter (Sanyal *et al.*, 2013).

In both mycobacteria, it seems that poly-p regulates the stringent response by the MprA-SigE-Rel pathway. The extraordinary ability of *M. tuberculosis* to survive for a long period in oxygen and nutrient-limited granuloma is facilitated by the stringent response (Boutte and Crosson, 2013). In this phase the so called persister cells are characterized by growth stasis and antibiotic tolerance. The stringent response in *M. tuberculosis* is controlled by $RelA_{Mtb}$. Phosphate starvation, hypoxia and activation of the alternative sigma factor *sigE*, increase $relA_{Mtb}$ transcription, which leads to up regulation of (p)ppGpp (Boutte and Crosson, 2013). The $RelA_{Mtb}$ enzyme transfers pyrophosphate from ATP to GDP or GTP to synthesize ppGpp and pppGp, respectively. (p)ppGp then influences numerous metabolic processes. $relA_{Mtb}$ also encodes a second catalytic domain that hydrolyzes (p)ppGp into pyrophosphate and GDP or GTP. It is known that Rel_{Mtb} is required for chronic *M. tuberculosis* infection in mice and demonstrated that the $RelA_{Mtb}$ (p)ppGp synthetase activity is required for maintaining the bacteria during chronic infection where the mutants didn't persist in mice, while hydrolase mutant $RelA_{Mtb}$ during acute or chronic infection in mice was lethal to the infecting bacteria and this also confirms the distinct role of $RelA_{Mtb}$ mediated (p)ppGp hydrolysis in *M. tuberculosis* pathogenesis (Weiss and Stallings, 2013). On the other hand over expression of *rseA* (anti-SigE) attenuated *ppK1* expression under phosphate starvation supporting the role of *sigE* in *ppK1* transcription which regulates the *sigE* expression via MprAB two component systems, so there are multiple feedback loops in this signalling circuit which could be linked to the bistability in the system which could play a key role in *M. tuburculosis* persistence (Sanyal *et al.*, 2013), (Sureka *et al.*, 2008) (Fig 1.12A).

On the other hand, *M. tuberculosis* once encounters the phosphate limited environment inside macrophages is able to regulate other genes encoding proteins involved in phosphate transportation. In fact, phosphate limitation is known to restrict *M. tuberculosis* growth in a dose-dependent manner. *M. tuberculosis* genes *ppK1* and *relA* were shown to be significantly upregulated after phosphate starvation, followed by inorganic polyphosphate accumulation and *M. tuberculosis* stringent response stimulation. The phosphate specific transporter operon *pstS3-pstC2-pstA1* was induced in phosphate starvation and its expression was dependent on the two-component regulatory system SenX3-RegX3 (Fig 1.12B). RegX3 appears to regulate the *M. tuberculosis* phosphate starvation response and it is essential for bacillus survival during phosphate depletion in the mammalian lung tissue, where the *regX3* mutated strains showed a reduced persistence in mouse and guinea pig lungs 56 days after infection. On the other hand, *phoY1*, *pstS1, pstS2, pstC1* and *pKnD* were not required for *M. tuberculosis* survival in animal lungs (Rifat *et al.*, 2009). *M. tuberculosis pstA1* is essential for virulence in mice and persistence in front of IFN-γ dependent host immunity (Tischler *et al.*, 2013). Another study showed that the Pst/SenX3-RegX3 system directly regulates ESX-5 secretion at the transcriptional level in response to phosphate availability and defines phosphate limitation as an environmental signal that activates ESX-5 secretion (Elliott and Tischler, 2016). The expression of phosphate starvation response is important for *M. tuberculosis* persistence to encounter the phosphate limited condition in mammalian lung infection (Rifat *et al.*, 2009). Poly phosphate deficiency is associated with increased susceptibility to certain drugs by *M. tuberculosis* and certain polyphosphate levels are required for *M. tuberculosis* survival in guinea pigs (Singh *et al.*, 2013).

9. M. tuberculosis between dormancy and reactivation

In the human host, *M. tuberculosis* is equipped for persistence in a dormant stage that cause latent tuberculosis without clinical disease and the term persister is used for *M. tuberculosis* organisms that are phenotypically resistant to drugs although they are genetically susceptible to these drugs. During latent infection, in addition to the dormant non replicating bacteria, some actively replicating bacteria are present, and the equilibrium balance between dormant/replicating *M. tuberculosis* determines the development an active or latent TB. In the latent infection, it has been suggested that most bacilli are in a dormant state and few have the ability, depending on the environmental conditions, to "awake", start replicating. Most of these scouts, under normal condition, are readily killed by the host immune response (Fig 1.13) (Gengenbacher and Kaufmann, 2012). When, for a number of reasons, the host cannot effectively and rapidly kill these scouts, bacilli start replicating and active TB disease may ensue (Chao and Rubin, 2010).

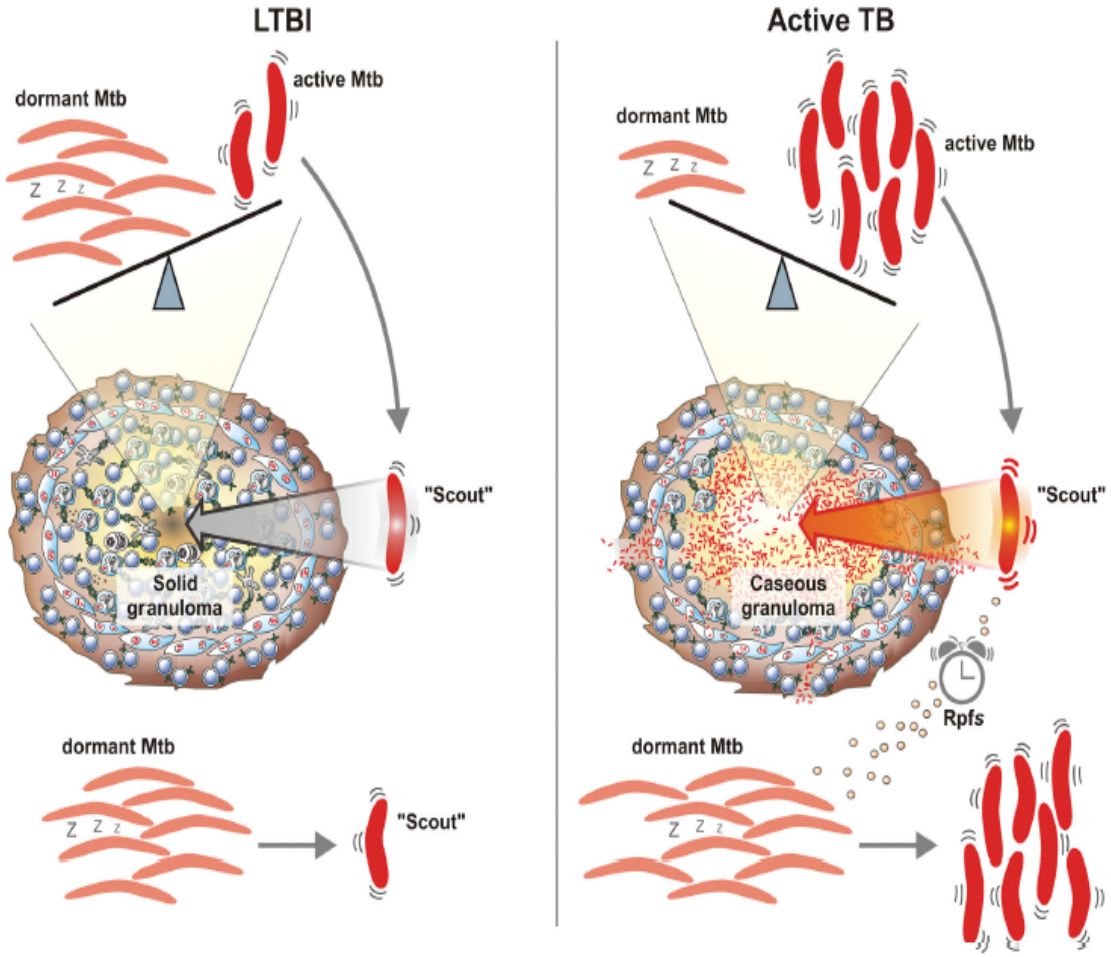


Figure 1.13: Dynamic models for latent tuberculosis infection (Gengenbacher and Kaufmann, 2012).

Identification of target genes and characterization of their respective antigens involved in primary infection, dormancy and reactivation and more in general bacterial factors known to play a key role in this complex interplay with the host, may help us to better understand *M. tuberculosis* pathogenesis and identify new and more effective tools to control disease at multiple levels.

The aim of the study

The few PE_PGRS proteins that have been so far characterized were found to be involved in key events during TB pathogenesis. These proteins are differentially expressed and were found to be involved in different steps of *M. tuberculosis* infection and host-pathogen interaction (Brennan and Delogu, 2002), (Sampson, 2011). The aim of the present study is to investigate the role of other PE_PGRSs in *M. tuberculosis* pathogenesis and gain new insights on the role in the biology of the tubercle bacillus. In this work, PE_PGRS3 and PE_PGRS4 were studied for the first time and have been chosen because they are in the same genome region and close to the ESX3 secretion system gene locus, present a 71.8% similarity and 67% identity, and present unique features. In fact, the presence of a second GRPLI motif, which has been suggested to have a role in anchoring the PE_PGRS domain to the *M. tuberculosis* outer membrane, has been detected in both proteins (Delogu G and Cole S.T and borsch R, 2008). Moreover, PE_PGRS3 have a unique C-terminal domain rich in arginine contains (≈ 80 aa in length contains 30aa arginine), that may be involved in the interaction with the host extracellular matrix components. In light of these features, we investigated the gene expression profile of these protein and, using recombinant strains, shed light on the role of the different domains in the interaction with a mammal host.

Chapter II

Results

1. Main features of PE_PGRS3 and PE_PGRS4

After a careful analysis of the whole protein family by specific bioinformatics tools, two PE_PGRS proteins were selected: PE_PGRS3 and PE_PGRS4. These proteins have 67% identity and 71.8% similarity, are found in the same gene locus though they seem to be transcriptionally regulated by two different promoters and are located immediately downstream of the ESX3 secretion system gene locus, which is important in *M. tuberculosis* pathogenesis (Serafini *et al.*, 2009) (Tufariello *et al.*, 2016) (Figure 2.1A and 2.1B). Each protein shows the presence of two GRPLI motifs, which have been suggested to represent the transmembrane domain that anchor the PE_PGRSs to the *M. tuberculosis* outer membrane (Delogu G and Cole S.T and borsch R, 2008). Moreover, the C-terminal domain of PE_PGRS3 is highly hydrophilic and contains numerous arginine amino acids that are typically found in proteins involved in the interaction with the host extracellular matrix components (Karsdal *et al.*, 2013) (Figure 2.1C).

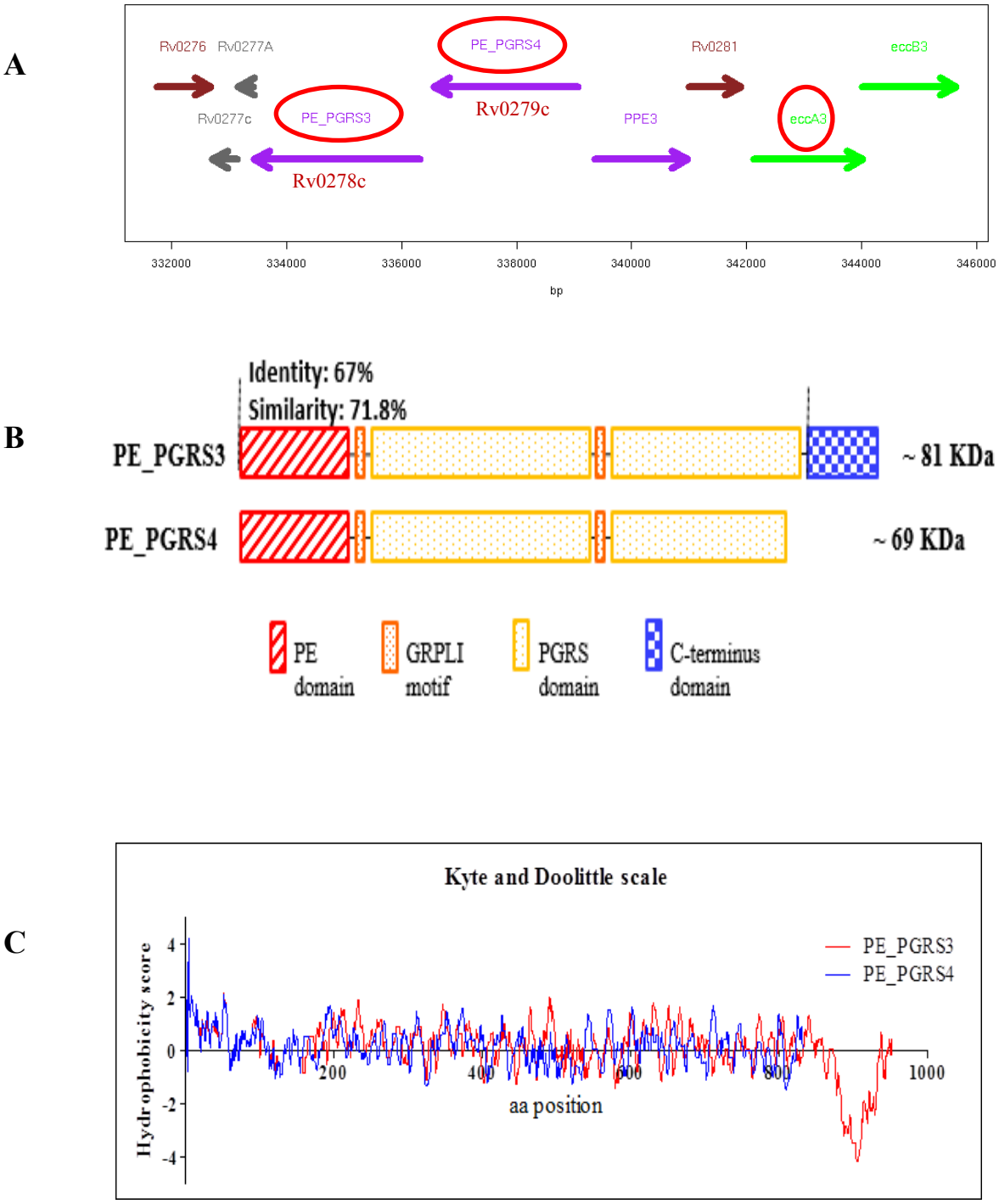


Figure 2.1: Shows in **(**A) Schematic representation of *pe_pgrs3* and *pe_pgrs4* localization on *M. tuberculosis* genome. (B) PE_PGRS3 and PE_PGRS4 proteins and their structural characteristic domains. (c) PE_PGRS3 and PE_PGRS4 hydrophobicity analysis using ExPASy tools (ProtScale – Kyte & Doolittle).

To gain insights on the genetic pressure exerted during MTBC evolution on the *pe_pgrs3* gene, multiple sequence alignment was carried out on the *pe_pgrs3* gene found in different MTBC strains belonging to different phylogeographic lineages. Sequences were obtained from publicly available databases for the following strains: *M. tuberculosis*, *M. bovis*, *M. bovis BCG*, *M. africanum, M. canetti, M. tuberculosis* of the EAI lineage. Nucleotides sequence alignments (Fig 2.2A) and amino acids sequence alignments (Fig 2.2B) indicate that not all *M. tuberculosis* strains have full length PE_PGRS3. Instead, a full length *pe_pgrs3* single gene is observed only in *M. tuberculosis* (new and ancient strains) that cause disease in human but not in those that cause disease in animal (*M. bovis*), where a duplication events appears to have led to the presence of an extra copy of the *pe_pgrs3* gene (putatively expressing a protein of 957 aa). A frameshift due to single base deletion splits this copy into two parts PE_PGRS3a similar to 5` end of *Rv0278c* and *PE_PGRS3b* equivalent to the 3` end of the *Rv0278c*. Hence, it appears that of the two copies of *pe_pgrs3* found in *M. bovis*, none is able to express the full length protein containing the arginine-rich domain. Hence, fully functional *pe_pgrs3*, including expression of the arginine-rich C-terminal domain, exists only in MTBC strains that cause disease in human but not that cause disease in animal (Fig 2.2).

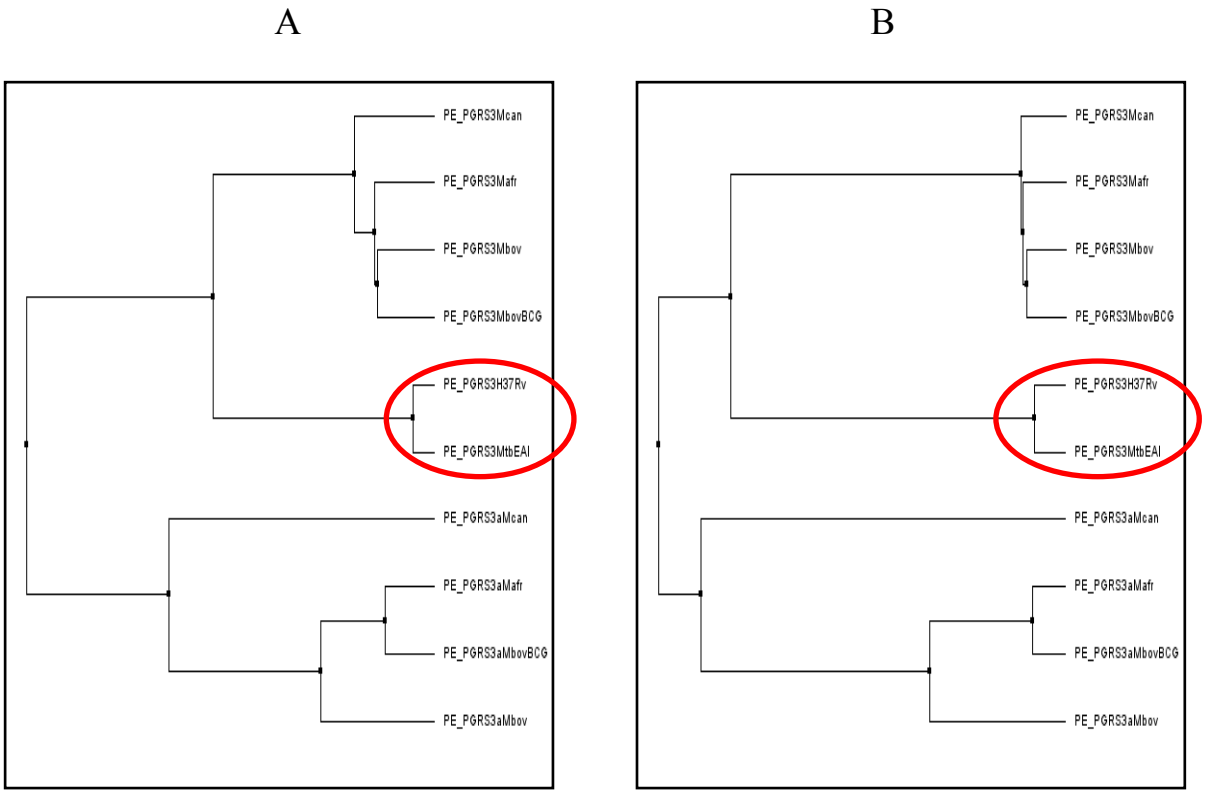


Figure 2.2: Phylogenetic tree obtained from multiple sequence alignments of PE_PGRS3 between different MTBC strains (*M. tuberculosis*, *M. bovis*, *M. bovis* BCG, *M. africanum*, *M. canetti*, *M. tuberculosis* of the EAI lineage) using multi-alignments tools (Clustel Omega and Jail view software). Figure A represents nucleotides sequence alignments, whereas figure B shows amino acids sequence alignments.

2. PE_PGRS3 and PE_PGRS4 are differentially expressed

To start investigating the role of these proteins, we decided to generate a gene reporter system, where the putative promoter and coding sequences of the two selected genes were cloned in mycobacterial shuttle plasmids to be expressed in *M. smegmatis*. Cloning was started by amplifying *pe_pgrs3* and *pe_pgrs4* genes with their own promoters from *M. tuberculosis* H37Rv genome. The primers used are showed in table. 1. The PCR 2.1-T/A cloning vector was used for sub cloning of the two genes and then these genes were inserted in several expression vectors (pMV-based plasmids).

In pMV206, the *pe_pgrs* genes were fused at 3` with the gene encoding GFP, while in pMV306 the genes were fused to the sequence coding the HA epitope. *M. smegmatis* mc^2155 were electroporated with the pMV306-based vectors and pMV-206 based vectors, alone or in combination, as indicated in figure 2.3, also to assess whether these neighboring genes are co-expressed.

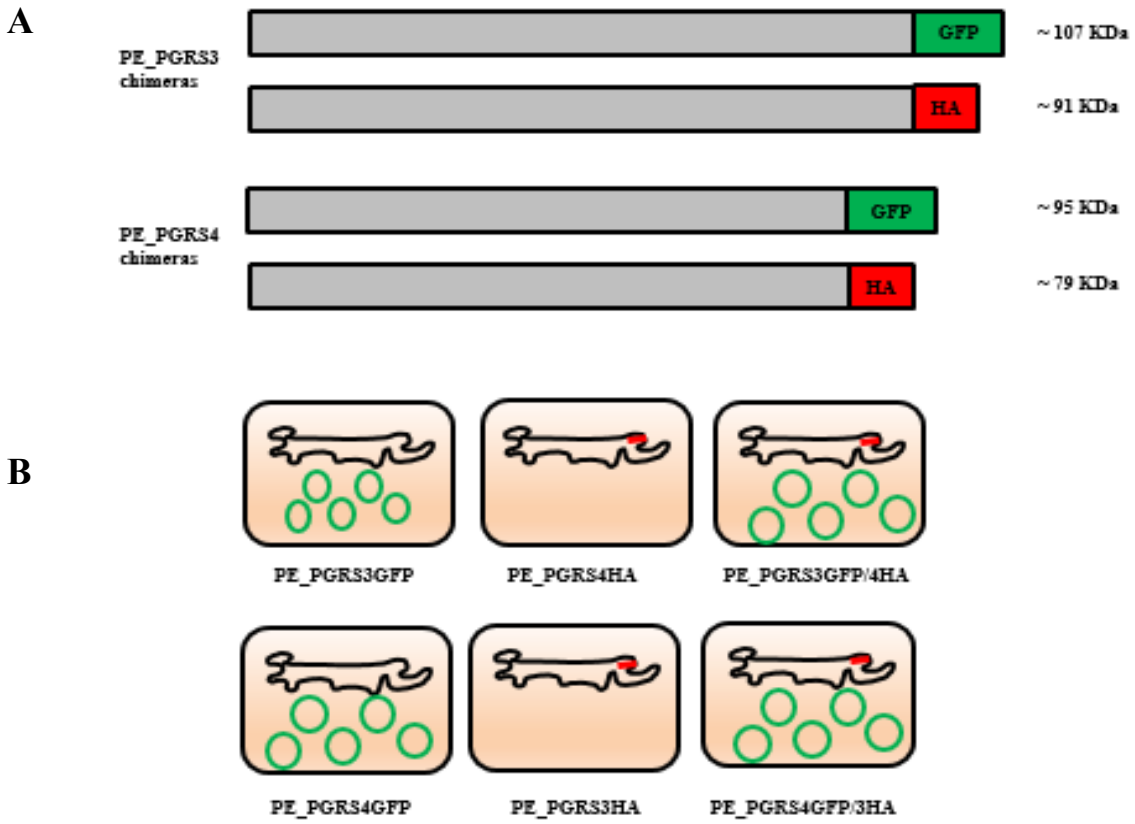


Figure 2.3: (A) Schematic representation of PE_PGRS3 and PE_PGRS4 protein chimeras, in pMV206 vector (PE_PGRS3GFP, PE_PGRS4GFP) and pMV306 vector (PE_PGRS3HA, PE_PGRS4HA). (B) *M. smegmatis* recombinant strains expressing PE_PGRS3, PE_PGRS4 protein chimeras (*M. smegmatis* PE_PGRS3GFP, *M. smegmatis* PE_PGRS4HA, *M. smegmatis* PE_PGRS3GFP/4HA, *M. smegmatis* PE_PGRS4GFP, *M. smegmatis* PE_PGRS3HA, *M. smegmatis* PE_PGRS4GFP/3HA).

The recombinant *M. smegmatis* strains were grown in 7H9/ADC/Tween liquid medium at 37c° and then analyzed at the fluorescence microscope. The *M. smegmatis* expressing PE_PGRS4-GFP showed an intense fluorescence while the

M. smegmatis expressing PE_PGRS3-GFP did not show any fluorescence (data not shown). Whole cell lysates from the four recombinant *M. smegmatis* strains expressing PE_PGRS3-GFP, PE_PGRS4-GFP, PE_PGRS3-GFP/PE_PGRS4-HA, PE_PGRS4-GFP/PE_PGRS3-HA chimeras were prepared and then analyzed by SDS-PAGE and immunoblot probed with anti-GFP and anti-HA specific antibodies. As shown in figure 2.4, we observed a signal corresponding to PE_PGRS4 with anti-GFP and anti-HA (~ 95 KDa for PE_PGRS4-GFP and ~ 79KDa for PE_PGRS4-HA), no signal was observed for PE_PGRS3 with both antibodies (anti-GFP and anti-HA), confirming proper expression of PE_PGRS4 and absence of expression of PE_PGRS3.

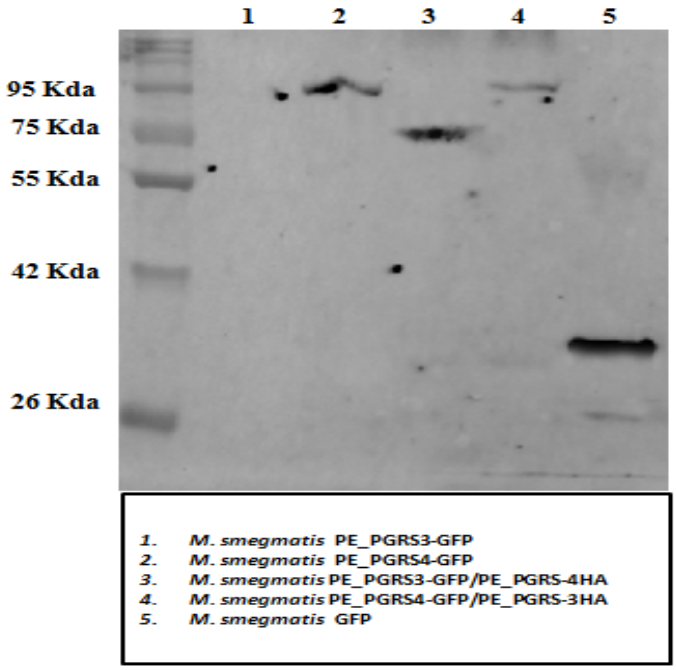


Figure 2.4: Shows SDS-PAGE and immunoblotting of the *M. smegmatis* expressing (PE_PGRS3GFP, PE_PGRS4GFP, PE_PGRS3GFP/4HA, PE_PGRS4GFP/3HA) whole cell lysates by using anti GFP and anti HA as primary antibody. The blott represents PE_PGRS4-GFP in lane 2,4 and PE_PGRS4-HA, GFP in lane 3,5 respectively.

The lack of expression of the two PE_PGRS3 chimeras, prompted us to verify for a second time by Sanger sequencing the pMV-PE_PGRS3-GFP and-HA plasmids, which however confirmed the presence of a correct sequence. To investigate any

specific condition that could be required for the PE_PGRS3 protein expression, the above mentioned recombinant strains were grown under different stress conditions that are known to be relevant during *M. tuberculosis* pathogenesis such low pH (pH = 5), low iPhos (~51 µm), low oxygen, low Mg^{2+}, low Fe^{3+}, PBS (Gengenbacher and Kaufmann, 2012). Interestingly, while PE_PGRS4 was always found expressed in the standard 7H9 medium, in Sauton minimal and low iPhos Sauton media, PE_PGRS3 seems not to be expressed in all the conditions tested but when the *M. smegmatis* PE_PGRS3-GFP strain was grown under low iPhos condition (Fig 2.5). These results indicate that PE_PGRS3 appears to be expressed only under low iPhos and repressed under common growth conditions or other conditions that have been associated with survival of *M. tuberculosis* in host tissues.

Figure 2.5: Shows the different stress conditions such low pH (pH = 5), low iPhos (~51 µm), low oxygen, low Mg^{2+}, low Fe^{3+}, PBS that were applied on the recombinant *M. smegmatis* strains expressing PE_PGRS3-GFP, PE_PGRS-4GFP, PE_PGRS3-GFP/4HA, PE_PGRS4-GFP/3HA and GFP and were grown in 48 well plate at 37Cº in order to investigate a possible stress source can trigger the protein expression then the fluorescence of the mycobacteria which correspond to the protein expression was observed by phase-contrast fluorescence microscopy.

3. PE_PGRS3 has a specific expression under low inorganic phosphate condition

To further investigate the observed PE_PGRS3-specific expression profile, as emerged under growth in low iPhos condition, confocal microscopy analysis was used as a more sensitive system. The recombinant *M. smegmatis* strains expressing PE_PGRS3-GFP, PE_PGRS4-GFP and, as controls, the recombinant *M. smegmatis* strain expressing another well-characterized protein of the family (PE_PGRS33-GFP) (Delogu *et al.*, 2004) and *M. smegmatis* expressing cytosolic GFP, were grown until mid-log phase and then sub inoculations were made for all strains in low iPhos concentration and normal iPhos concentration (which for the sake of simplicity we here arbitrary define as high iPhos) in Sauton medium. After 15 days of incubation at 37C°, plated in chamber slides and then observed at confocal microscopy, the fluorescence microscopy images were analyzed by Image J software. The results obtained confirmed a strong fluorescence for *M. smegmatis* expressing PE_PGRS3 in low iPhos condition, while no fluorescence was observed when the same strain was grown in high iPhos condition. Conversely, no significant differences in fluorescence intensity were observed when the other recombinant strains were grown in low or high iPhos medium (Fig 2.6). Quantification of the fluorescence intensity as shown in figure 2.7 clearly indicates that PE_PGRS3-GFP is specifically expressed under low iPhos concentration.

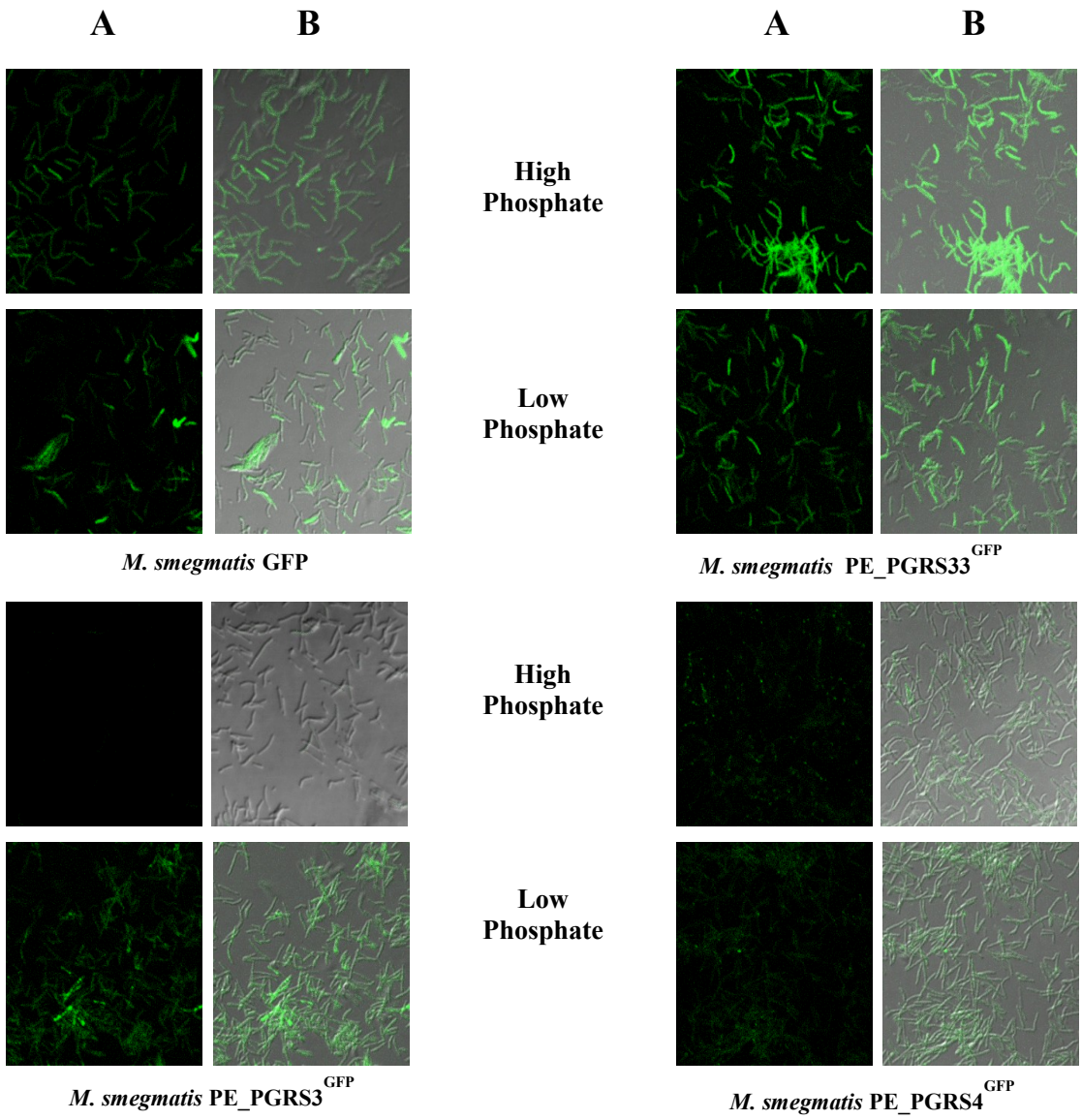


Figure 2.6: (A) Confocal microscopy images of *M. smegmatis* expressing PE_ PGRS3, PE_PGRS4 PE_PGRS33 (another well-characterized protein of the PE_PGRSs family) GFP tagged and *M. smegmatis* GFP grown in high and low iPhos Sauton medium, obtained with x60 objective. (B) Overlapping green channel and transmission microscopy images.

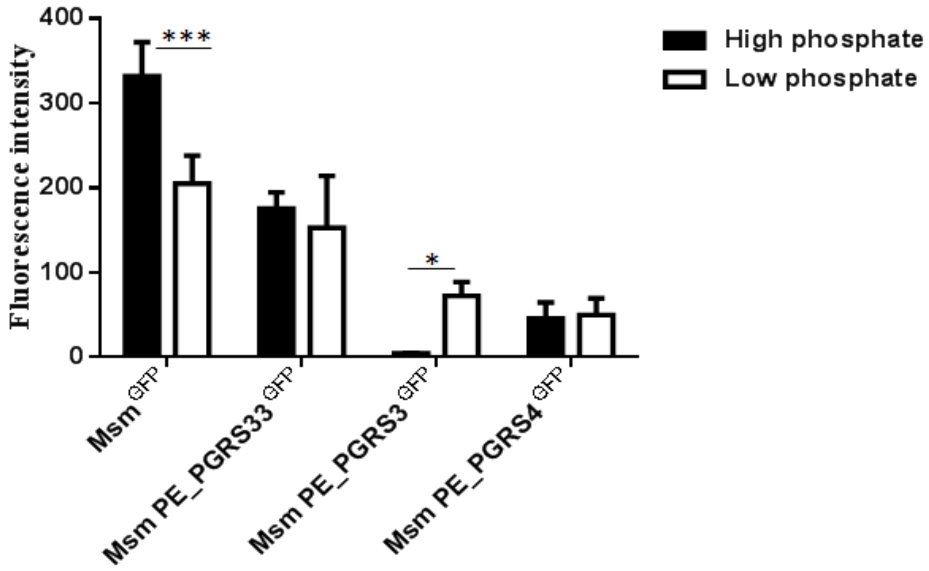


Figure 2.7: Shows confocal microscopy images fluorescence intensity analysis for *M. smegmatis* expressing PE_ PGRS3, PE_PGRS4, PE_PGRS33 (another well-characterized protein of the PE_PGRSs family) GFP tagged and *M. smegmatis* GFP grown in High and low iPhos Sauton medium. A strong fluorescence for *M. smegmatis* (*Msm*) expressing PE_ PGRS3-GFP was observed after 15 days of incubation in low iPhos Sauton medium and no fluorescence was observed in high iPhos Sauton medium *P<0.01. No significant difference in fluorescence intensity was observed for *M. smegmatis* expressing, PE_PGRS33-GFP, PE_PGRS4-GFP in both high and low iPhos Sauton medium. A significant difference was observed for *M. smegmatis* expressing GFP in low and high iPhos condition***P<0.001. The fluorescence microscopy images were analyzed using image J program.

3.1 Quantification of protein expression by measuring the fluorescence of single mycobacteria by FACS-Canto flow cytometer

To investigate the level of GFP expression at single cell level, we employed the flow cytometry (FACS-canto) for measuring the fluorescence of *M. smegmatis* expressing PE_PGRS3-GFP, PE_PGRS4-GFP, PE_PGRS33-GFP, GFP and *M. smegmatis* mc^2155wt as a negative control. All strains were grown in a low iPhos Sauton medium until we a observed a maximum and plateau fluorescence at the day 15 for the *M. smegmatis* PE_PGRS3-GFP. The % of fluorescent cells increased from 0% to arrive 2.5% at the 15^{th} day of incubation. To see if the phosphate is a specific regulator for the expression of *M. smegmatis* PE_PGRS3-GFP, bacterial cells were washed and suspended in a standard Sauton medium

(high iPhos) and the bacterial fluorescence measured until the 20th day of incubation. Interestingly, we observed a significant decrease in the % of *M. smegmatis* PE_PGRS3-GFP showing fluorescence, while we did not observe significant changes in the fluorescence expressed by the other recombinant strains (Fig 2.8).

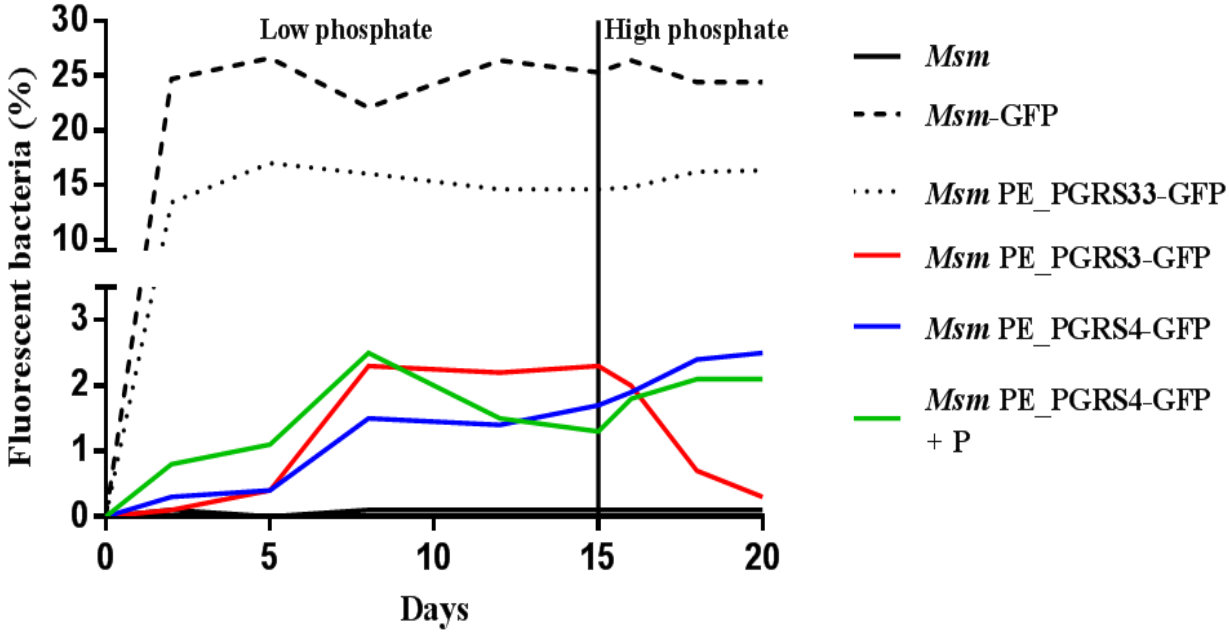


Figure 2.8: Shows the FACS results, the fluorescence was measured for *M. smegmatis* (*Msm*) expressing PE_PGRS3-GFP, PE_PGRS4-GFP, PE_PGRS33-GFP, GFP and *M. smegmatis* mc^2155wt grown in low iPhos Sauton medium. The fluorescence for *Msm* expressing PE_PGRS4-GFP grown in complete Sauton standard medium - high iPhos (PE_PGRS4+P) was also analyzed to confirm confocal microscopy results. Bacterial cells were washed at the 15th day and resuspended in Sauton standard medium and the fluorescence was measured until the 20th day of incubation at 37 C°.

The results obtained confirmed that PE_PGRS3, but not the other PE_PGRS proteins analyzed, has a specific expression in low iPhos condition and the phosphate is a specific regulator for the expression of the PE_PGRS3, since addition of iPhos to the culture repressed PE_PGRS3 expression.

3.2 PE_PGRS3 expression increased in low inorganic phosphate condition and correlated with RelA in *M. smegmatis* and *M. tuberculosis*

The inorganic phosphate regulation is a critical element in *M. tuberculosis* pathogenesis and many studies focused on the importance of the role of iPhos depletion in *M. tuberculosis* survival and persistence in host tissues, suggesting that iPhos concentration may serve as a trigger for the expression of many genes involved in TB pathogenesis (Rifat *et al.*, 2009). Because of RelA has previously been implicated in the *M. tuberculosis* transcriptional response to iPhos starvation and considered as a well characterized stringent response mediator which is required for mycobacterial persistence (Sureka *et al.*, 2008), we decided to examine the transcription level of *relA* and *pe_pgrs3* in low and high iPhos medium after 15 days of incubation at 37 C^{o} in the recombinant *M. smegmatis* strain expressing PE_PGRS3. By using Real time PCR we found that *pe_pgrs3* expression was upregulated by 3-4 times in low iPhos conditions respect to the high iPhos concentration and correlated with the expression of *relA*, which however showed a much more remarkable upregulation (up to 90 times) under low iPhos conditions (Fig 2. 9).

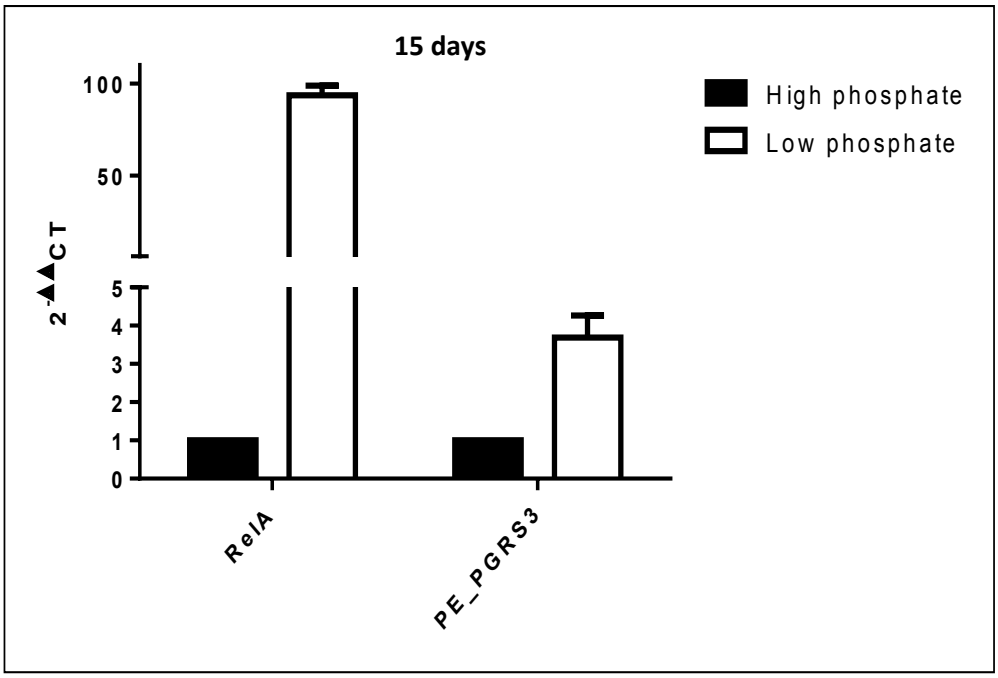
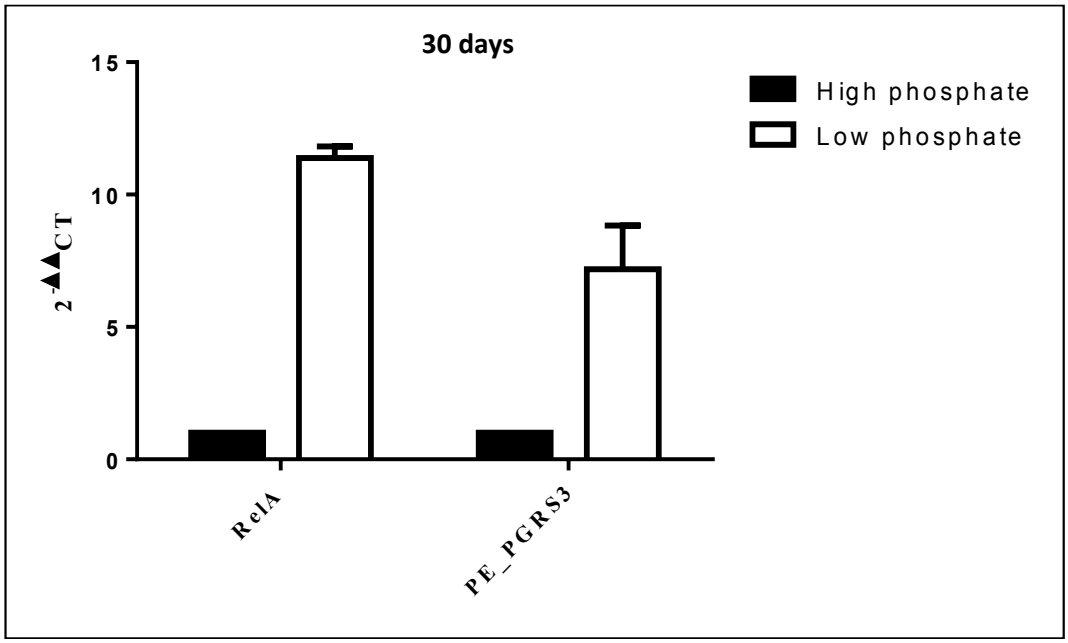


Figure 2.9: Shows Real time PCR results on *M. smegmatis* expressing PE_PGRS3 under its putative promoter, but using a pMV206 based vector. Strains were grown in Sauton medium with high iPhos and low iPhos (~50µM) for 15 days. *relA* and *pe_pgrs3* in phosphate starvation (low iPhos condition), were expressed respectively 93 times and 3-4 time more than the same genes in complete medium (high iPhos condition). Target cDNA was internally normalized to 16s cDNA.

Similar growth conditions were used for the *M. tuberculosis* H37Rv reference strain, cells were harvested at two time points 30 and 60 days of incubation, RNA isolated and real time RT-PCR performed. As shown in figure 2.10 a significant upregulation of *pe_pgrs3* in correlation with *relA* expression was observed at the two time points mentioned above (Figure 2.10).

A



B

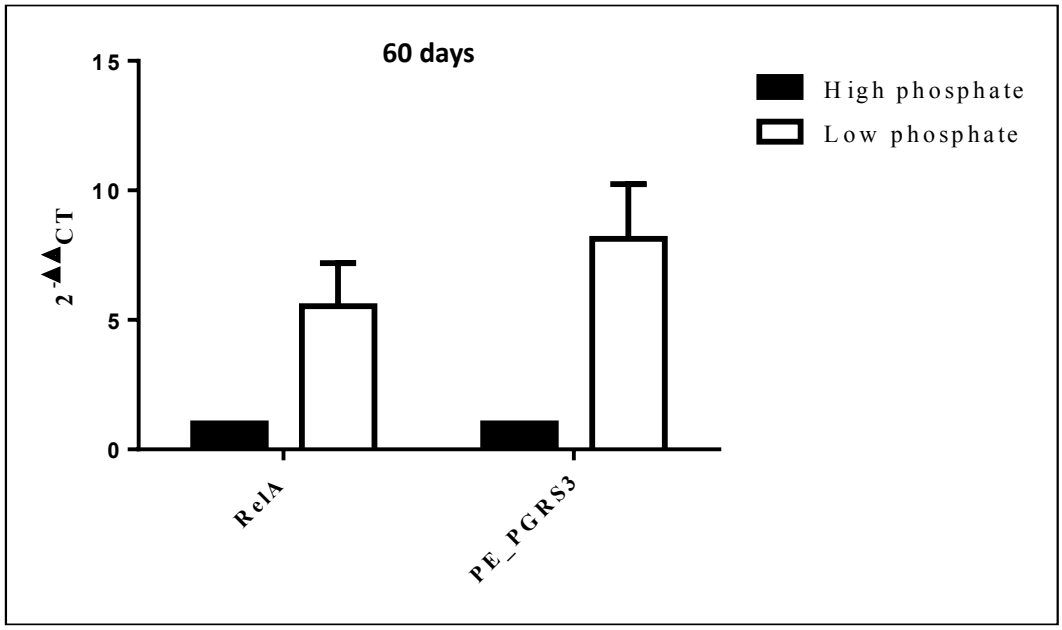


Figure 2.10: Shows Real time PCR results on *M. tuberculosis* H37Rv. *M. tuberculosis* strain were grown in Sauton medium with high iPhos and with low iPhos (~50µM) for 30 days (A) and 60 days (B). A significant increase in the expression of *relA* and *pe_pgrs3* is observed in low iPhos condition**.** Target cDNA was internally normalized to 16s cDNA.

4. *M. smegmatis* over expressing PE_PGRS3 were shorter in size than the strains expressing the functional domains

The findings that PE_PGRS3 is expressed only under low iPhos conditions, while of great interest, made challenging to further investigate the localization of this protein in the mycobacterial cell and more in general difficult to explore the role of the protein during infection *in vitro* models (as an example macrophages). Hence, to investigate the role of PE_PGRS3 in more detail, we expressed the full length protein and three functional deletion mutants (PE_PGRS3, PE_PGRS3ΔCt, PE_PGRS3ΔGRPLI, PE3), under the control of a constitutive promoter (*hbh*A promoter) (Fig 2.11A). All these chimeras were tagged with the HA epitope at the C-terminal and cloned in pMV multi copy vector expressing GFP (Fig 2. 11B).

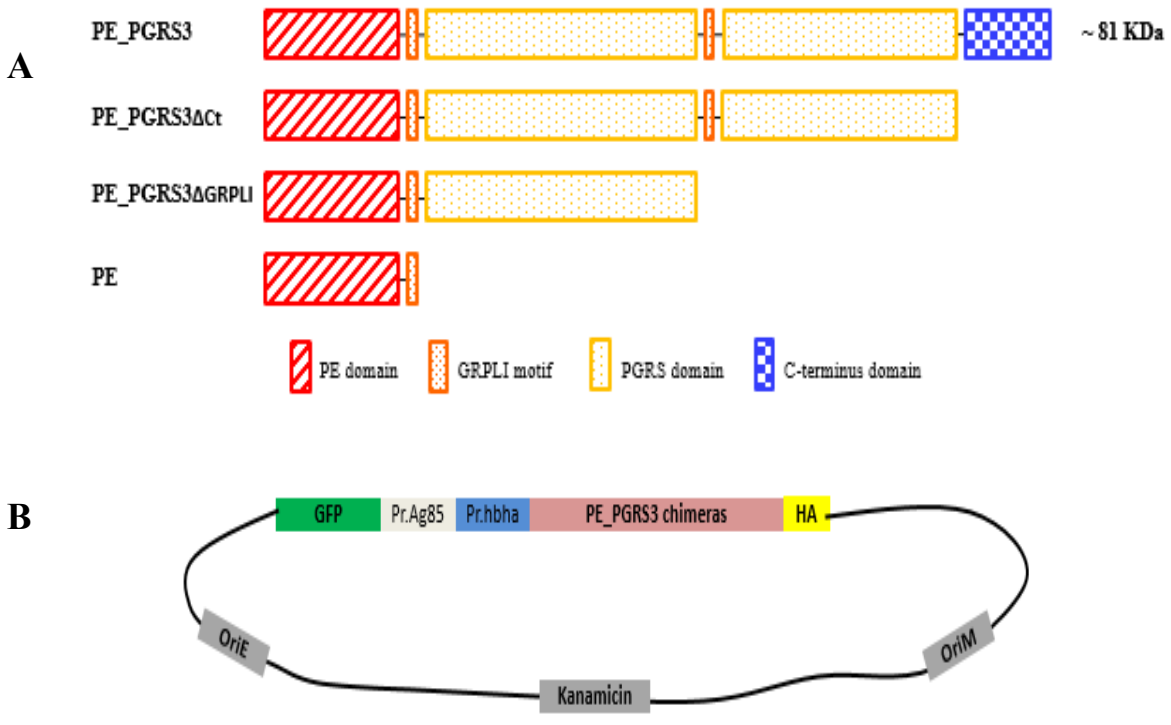


Figure 2.11: (**A)** Schematic representation of the PE_PGRS3 chimeras created; PE_PGRS3 and its functional different domains; PE_PGRS3ΔCt, PE_PGRS3ΔGRPLI, PE3. (**B)** schematic represents the episomal multicopy vector used in the cloning, tagging the GFP at the N-terminal and HA at the C-terminal end.

Then all these chimeras were electroporated in *M. smegmatis* mc^2 155. To confirm the expression of PE_PGRS3 and its different chimeras, all recombinant strains were grown in a normal condition and immunoblot on the bacterial whole cell lysates separated by SDS-PAGE was developed using anti HA antibody. The signals obtained in lane 2 is at ~ 90Kda corresponding to PE_PGRS3-HA; in lane 3 the band is ~50 Kda, corresponding to the expected MW of PE_PGRS3ΔGRPLI-HA; in lane 4 is about ~ 80Kda corresponding to PE_PGRS3ΔCt-HA (Fig 2.12). These results indicate that expression of *pe_pgrs3* under the control of the *hbh*A promoter warrant good level of expression.

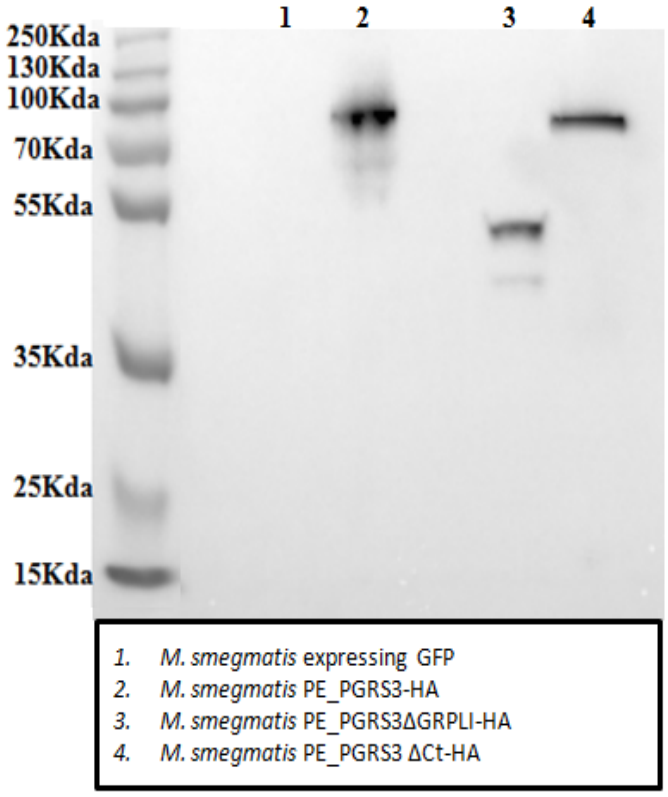


Figure 2.12: Shows SDS-PAGE and immunoblotting of the *M. smegmatis* expressing (GFP, PE_PGRS3-HA, PE_PGRS3ΔGRPLI-HA and PE_PGRS3ΔCt-HA) whole cell lysate by using anti HA as primary antibody**.** The blott represents PE_PGRS3HA, PE_PGRS3ΔGRPLI-HA and PE_PGRS3ΔCt-HA in lane 2, 3, 4 respectively.

Interestingly, analysis at confocal microscopy highlighted that the *M. smegmatis* expressing full length PE_PGRS3 were shorter in size (~50% less) than the control strain (*M. smegmatis* expressing GFP) and the strains expressing the functional domains (Fig 2. 13).

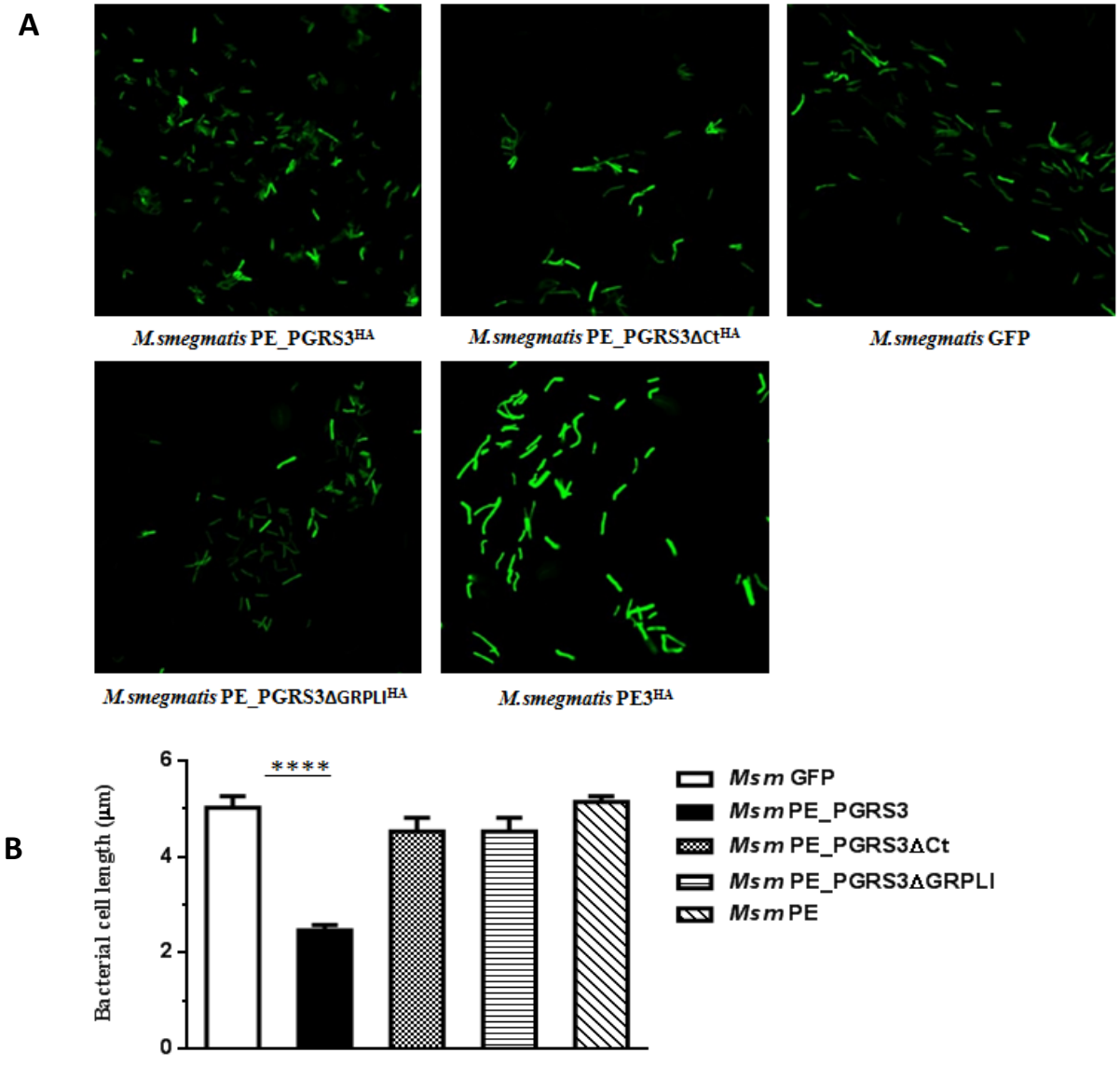


Figure 2.13: Confocal microscopy images were taken for *M. smegmatis* (*Msm*) expressing PE_PGRS3 and its functional deletion mutants under control *hbhA* promoter and *M. smegmatis* expressing GFP as a control with 60X objective (A), and bacterial cell length analyzed by Image J program and showed that *M. smegmatis* expression full length PE_PGRS3 was shorter in size (~50% less) than the control strain and the strains expressing PE_PGRS3 functional deletion mutants ****P<0.0001(B).

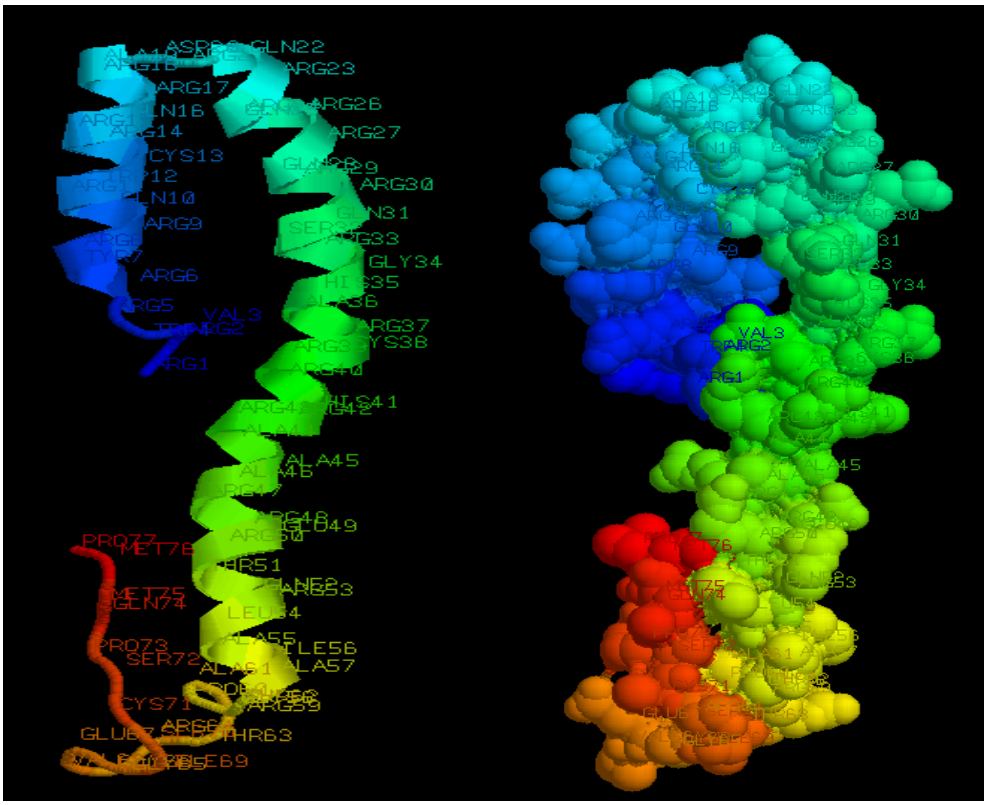
5. The *M. smegmatis* expressing PE_PGRS3 and its functional domains have a similar growth rates

M. smegmatis recombinant strains expressing PE_PGRS3, PE_PGRS3ΔCt, PE_PGRSΔGRPLI and *M. smegmatis* expressing GFP were also grown in MGIT to assess growth rate. The results obtained indicated that there is no difference between the growth rate of *M. smegmatis* expressing PE_PGRS3 and its functional domains (data not shown). Hence, overexpression of PE_PGRS3 in common media does not affect *M. smegmatis* replication.

6. Purified native C-terminal domain of PE_PGRS3 induces specific anti-serum in mice

The unique C-terminal domain of PE_PGRS3 is 77 aa in length and rich in arginine, which account for 30 aa residues (Fig 2.14). To develop a system that could be specifically used to detect PE_PGRS3 and to further investigate the role, we generated a system to express the recombinant C-terminal domain in *E. coli*. The gene fragment encoding the C-terminal domain of PE_PGRS3 was amplified by PCR as described in the material and methods and cloned in pET-SUMO vector, provided with 6XHis tag, and expressed into *E.coli* BL21 (DE3) cells. The induction of the protein expression in *E.coli* was carried out for 3 hours with 1mM IPTG, then the fermentation culture was pelleted and lysed. The soluble fraction of the cell lysate was applied to Ni-NTA affinity column. Non-specifically bound proteins were removed from the resin by increasing linear gradient of imidazole (from 50mM to 350mM) and recombinant C-terminal domain of the PE_PGRS3 was eluted at 50mM imidazole (Fig 2.15).

A



B

Figure 2.14: Shows in **(**A) the 3D structure of C-terminal domain of PE_PGRS3. Image coloured by rainbow N → C terminus, obtained by Phyre2 and analysed by RasMol. (B) a schematic representation of the PE_PGRS3 and its C-terminal domain with the amino acids sequences.

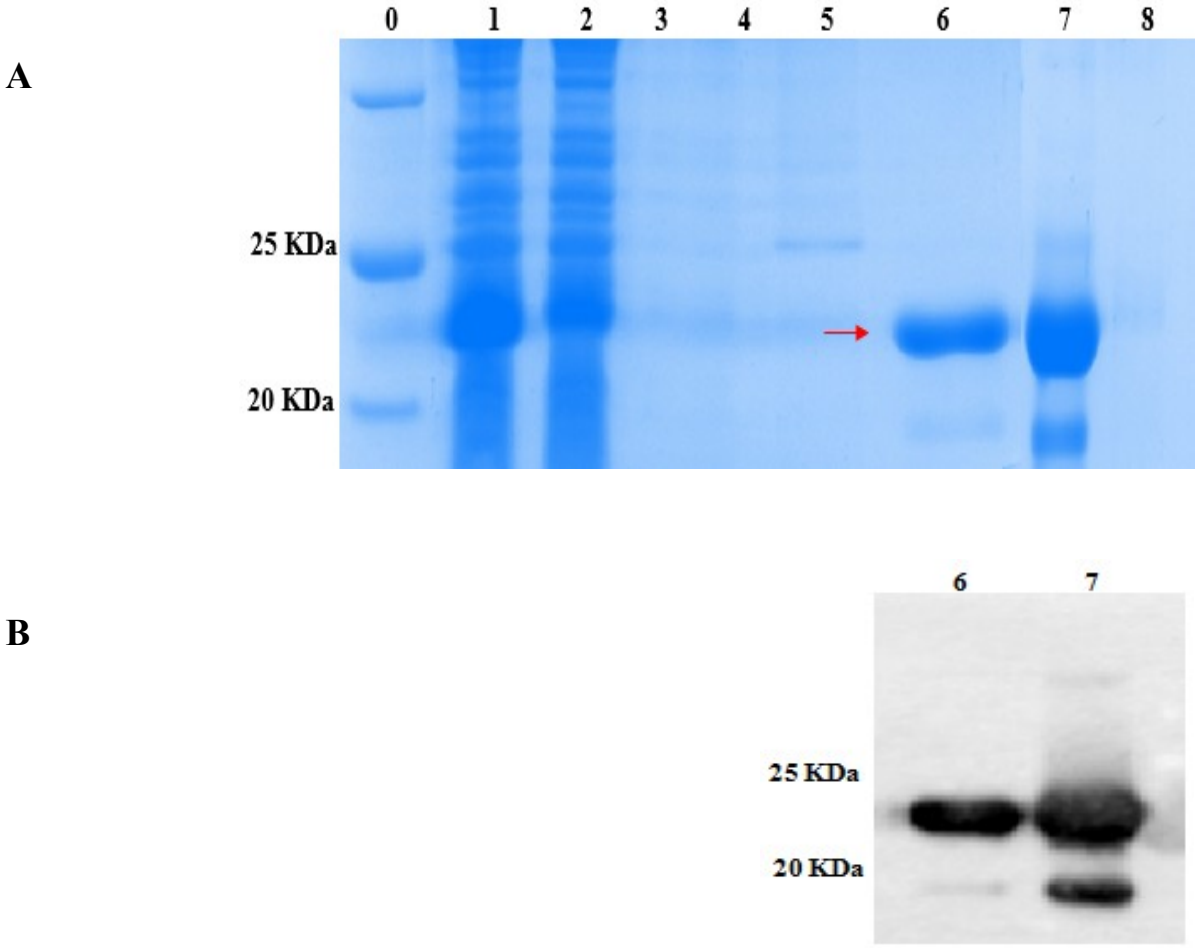


Figure 2.15: Shows SDS-PAGE, comassie brilliant blue stained gel (A) and immunoblot (B) to control the recombinant C-terminal domain of the PE_PGRS3 purification. **(**A) Lane 0: molecular weight markers, lane 1: soluble fraction of the *E.coli* cell lysate expressing C-terminal domain of the PE_PGRS3 protein, lane 2: flowthrough, lane 3,4,5: washing, lane 6,7,8: eluted fractions collected at increasing concentration of imidazole (50-200-350 mM) during purification by FPLC, eluted fraction at 50mM imidazole indicated by the arrow. (B) immunoblott with anti His 1:4000 represents a signal correspond to the molecular weight of the recombinant protein 23KDa (13KDa for the C-terminal domain of the PE_PGRS3 and 10KDa SUMO) in lane 6,7 eluted fraction 50, 200mM imidazole respectively.

As shown in figure 2.15, the Comassie brilliant blue staining of the polyacrylamide gel highlighted that recombinant C-terminal of the PE_PGRS3 was eluted without other contaminating proteins. The molecular weight of the recombinant protein was around 23KDa (13 KDa for the C-terminal domain of the PE_PGRS3 and 10 KDa SUMO). To ensure that the purified protein was the C-terminal domain of the PE_PGRS3, an immunoblot with monoclonal

anti-polyHistidine antibody was developed (Fig 2.15B). Selected fraction containing recombinant C-terminal of the PE_PGRS3 were dialyzed for 24 hours against PBS to remove salts, detoxified (removing LPS) and concentrated. About 0.5 mg/ml was obtained from 250ml of the fermentation culture.

6.1 Sera from immunized mice specifically recognize C-terminal of PE_PGRS3

Purified recombinant C-terminal domain of the PE_PGRS3 was used to immunize BALB/c mice to obtain specific serum against the protein. The anti C-terminal polyclonal serum was collected after the third immunization, and tested by ELISA to assess antibody titer, which was found to be 1:2000. The serum was found specific for the C-terminal of PE_PGRS3 and no signal was observed when other mycobacterial recombinant proteins were tested (Fig 2.16). The result of both immunoblots assessed that the antiserum raised against the native C-terminal of the PE_PGRS3 specifically recognized this protein.

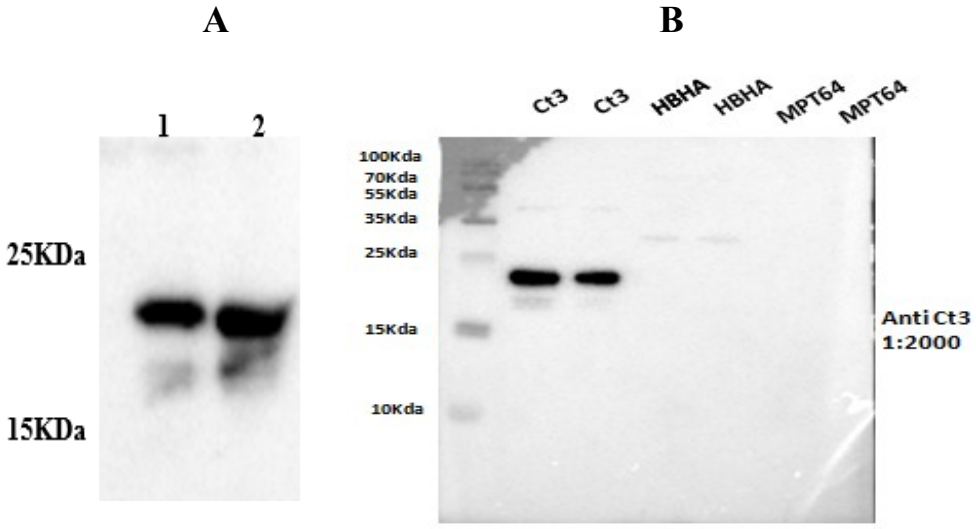


Figure 2.16: Shows in (A) SDS-PAGE and immunoblott with anti Ct polyclonal serum 1:2000 against dialyzed purified recombinant C-terminal of the PE_PGRS3 (rCt3) in lane 1 and dialyzed LPS free in lane 2. (B) immunoblot with anti Ct polyclonal serum against rCt3, rHBHA and rMPT64 other mycobacterial recombinant proteins.

7. PE_PGRS3 could be cleaved at the C-terminal domain secreted or surface exposed

M. smegmatis recombinant strains expressing PE_PGRS3 and PE_PGRS4 were grown in a high and low iPhos medium and culture supernatant were harvested and analyzed in immunoblot using the anti recombinant C-terminal domain of the PE_PGRS3 specific serum. A band at about 42 kDa was observed only on the lane containing the secreted proteins expressed by *M. smegmatis* PE_PGRS3 grown in low iPhos medium. The MW of the observed band (~40KDa) could correspond to the C-terminal ~13KDa + GFP~26KDa (Fig 2. 17). Hence, these results suggest that the arginine-rich C-terminal domain of PE_PGRS3 may be cleaved by mycobacteria following expression under low iPhos conditions.

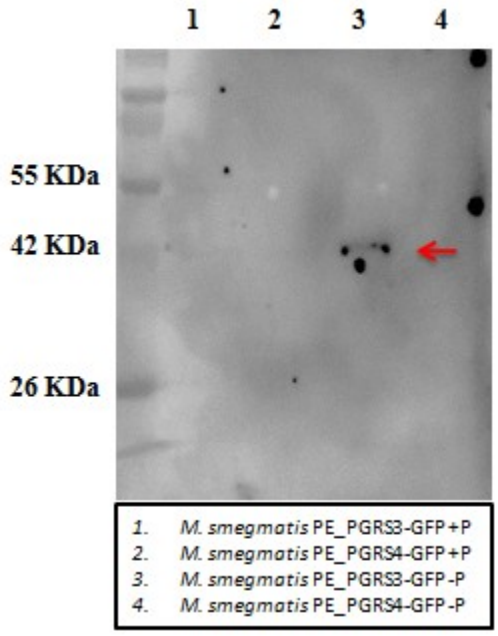


Figure 2.17: Shows SDS-PAGE and immunoblott of the culture supernatant of *M. smegmatis* expressing PE_PGRS3 and PE_PGRS4 were analysed after protein precipitating according to the standard protocol, the immunoblot was developed by using anti recombinant C-terminal domain of the PE_PGRS3 specific serum produced in mice (1:2000). A signal was observed in the secreted protein fraction obtained from the culture in low iPhos medium. The fragment may represent the peptide Ct+GFP (C-terminal domain of the PE_PGRS3 with GFP).

8. The recombinant purified C-terminal of PE_PGRS3 has no cytotoxic effect on the murine macrophages and human alveolar epithelial cells

It has been suggested that the arginine rich peptide could be a cell penetrating peptide and play a role in disrupting the plasma membrane of the eukaryotic cells (Schmidt *et al.*, 2010). Since the C-terminal domain of PE_PGRS3 may be cleaved following expression, we decided to study the possible cytotoxic effect of the C-terminal domains on macrophages and epithelial cells. Different concentration of the purified LPS-free recombinant C-terminal domain of the PE_PGRS3 protein was incubated with J774 murine macrophages and A549 human alveolar epithelial cells, and Alamar blue was used to evaluate the cytotoxic effect by measuring the absorbance at different time points (4 hours, 1day, 4 and 7days) following incubation at 37 C^{o}. No cytotoxic effect due to the recombinant C-terminal domain was observed on J774 murine macrophages and A549 human alveolar epithelial cells.

9. *M. smegmatis* expressing PE_PGRS3 enhanced entry in macrophages and alveolar epithelial cells

To investigate the role of PE_PGRS3 in TB pathogenesis, the recombinant *M. smegmatis* strains expressing PE_PGRS3, or its functional deletion mutant PE_PGRS3ΔCt, under the control of the *hbh*A promoter and *M. smegmatis* expressing GFP as a control strain, all were used to infect J774 murine macrophages at MOI 10:1. Four hours post infection cells were collected and lysed with Triton-X100 to count intracellular CFU. As shown in figure 2.18, *M. smegmatis* expressing full length PE_PGRS3 shows enhanced ability to entry into macrophages compared with the *M. smegmatis* expressing PE_PGRS3ΔCt and *M. smegmatis* expressing GFP ($P<0.01$).

Because of the importance of the alveolar epithelium in the pathogenic processes of *M. tuberculosis* and the role of the alveolar epithelial cells in active TB progression and *M. tuberculosis* persistence (Scordo *et al.*, 2016), we decided to infect the human alveolar epithelial cells (A549) at MOI 10:1. Intracellular CFUs were evaluated at 4h post infection, as shown in figure 2.18. The *M. smegmatis* expressing full length PE_PGRS3 shows enhanced ability to entry in alveolar epithelial cells (pneumocytes) compared with the *M. smegmatis* expressing PE_PGRS3ΔCt and *M. smegmatis* expressing GFP (P< 0.0001).

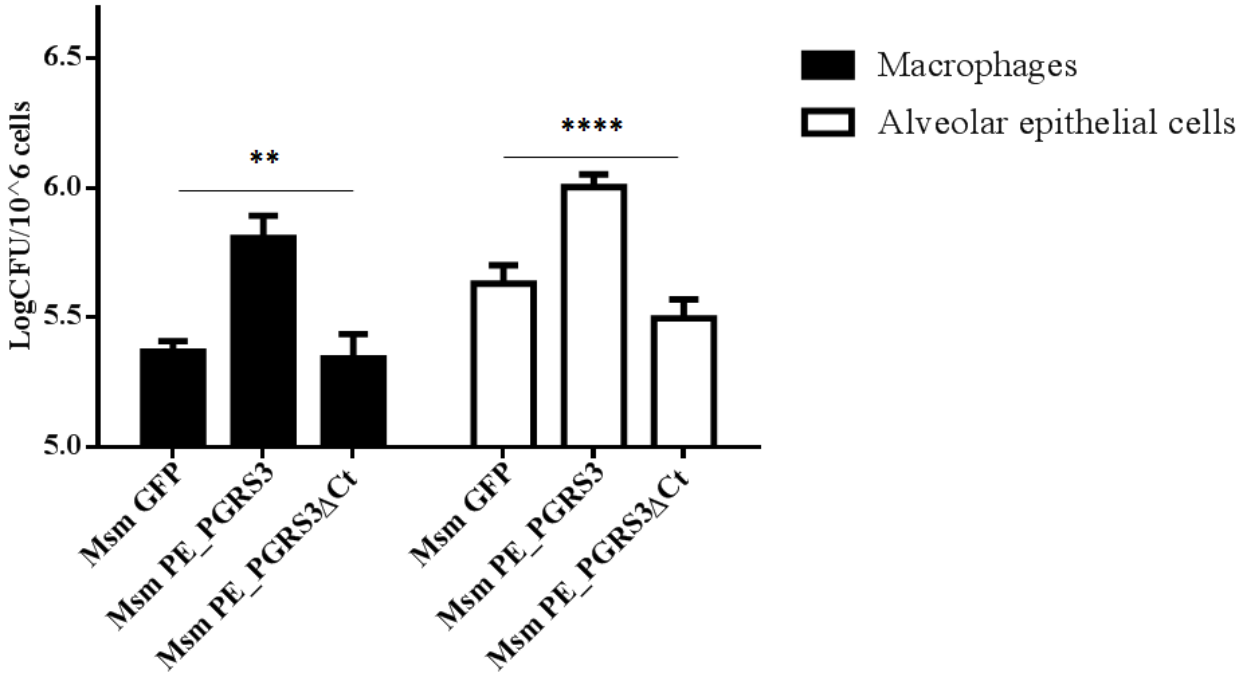


Figure 2.18: Shows CFU at 4h post infection where recombinant *M. smegmatis* (*Msm*) strains expressing PE_PGRS3 or its functional deletion mutant PE_PGRS3ΔCt, under control of the *hbha* promoter and *Msm* expressing GFP as a control strain, all were used to infect murine macrophages cells (J774) and human alveolar epithelial cells (A549) MOI 10:1. CFU were evaluated at 4 hours post infection. Significant result was obtained for *M. smegmatis* expressing PE_PGRS3 in murine macrophages **P<0.01 and human alveolar epithelial cells (pneumocytes) ****P<0.0001.

To assess the ability of *M. smegmatis* above mentioned recombinant strains to survive and persist intracellularly, J774 macrophages and A549 epithelial cells were infected with each strain at MOI 10:1, then 4h hours later cells were washed three times with PBS to remove extracellular bacteria. After washing the cells were incubated in fresh medium for 4 days and then the cells were lysed with Triton X-100 and intracellular bacteria determined by CFU counting. As shown in figure 2.19A, there is no significant difference in the intracellular CFU at 4 days in macrophages with the different *M. smegmatis* recombinant strains. Similarly, the results obtained by CFU after 4 days with A549 epithelial cells indicate that *M. smegmatis* expressing PE_PGRS3, while showing increased cell entry at 4 h, shows similar intracellular replication compared with the other two strains tested (figure 2.19B). Hence, expression of PE_PGRS3 enhances mycobacterial entry in macrophages and epithelial cells and this effect is due to the arginine-rich domain.

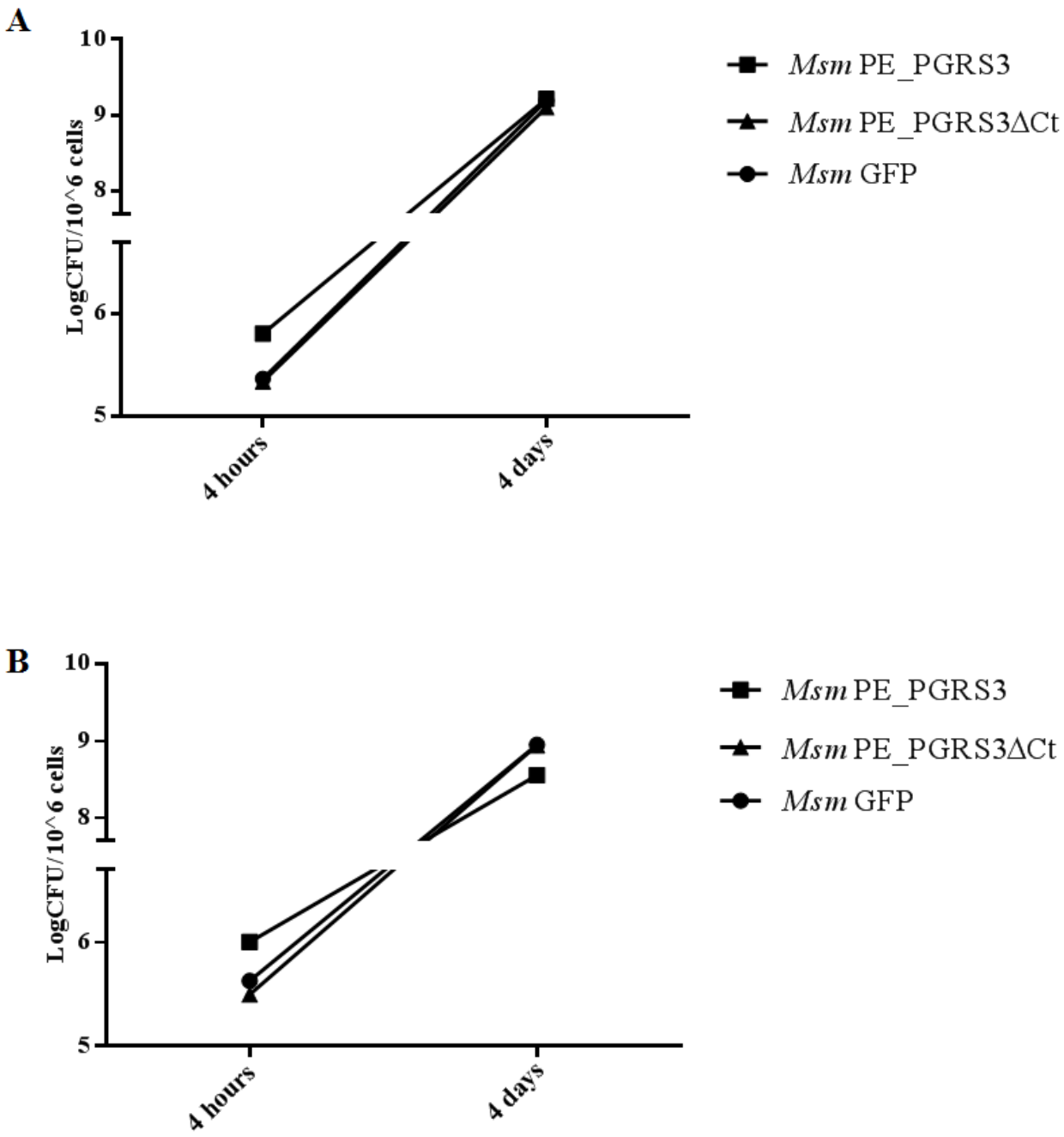


Figure 2.19: Shows CFU at 4h and 4 days post infection of macrophages (A) and alveolar epithelial cells (B). Recombinant *M. smegmatis* strain (*Msm*) expressing PE_PGRS3, and *Msm* expressing PE_PGRS3ΔCt under control of the *hbhA* promoter, and *Msm* expressing GFP as a control strain were used to infect (J774) murine macrophages and human alveolar epithelial cells (A549) at MOI 10:1. CFU were evaluated at 4 hours and 4 days post infection, no significant difference in the intracellular CFU at 4 days post infection was observed with the different *M. smegmatis* recombinant strains.

To investigate whether the purified recombinant C-terminal domain of the PE_PGRS3 could restore the ability of entry of *M. smegmatis* expressing PE_PGRS3 lacking the C-terminal domain, we added the purified C-terminal domain of the PE_PGRS3 (rCt $^{PE\_PGRS3}$) at concentration 0.5µg/ml, to the infecting solution containing the *M. smegmatis* PE_PGRS3ΔCt and to the infecting solution containing *M. smegmatis* GFP. As shown in figure 2.20, addition of recombinant C-terminal domain to *M. smegmatis* expressing PE_PGRS3ΔCt and to *M. smegmatis* expressing cytoplasmic GFP, enhanced mycobacterial entry in macrophages and alveolar epithelail cells ($P<0.01$ for *M. smegmatis* PE_PGRS3ΔCt with added rCt $^{PE\_PGRS3}$ versus PE_PGRS3ΔCt in macrophages and alveolar epithelial cells, $P<0.0001$ for *M. smegmatis* GFP with added rCt $^{PE\_PGRS3}$ versus *M. smegmatis* GFP in macrophages and $P<0.001$ in alveolar epithelial cells). These results further implicate the C-terminal domain of PE_PGRS3 in the entry in mammalian cells.

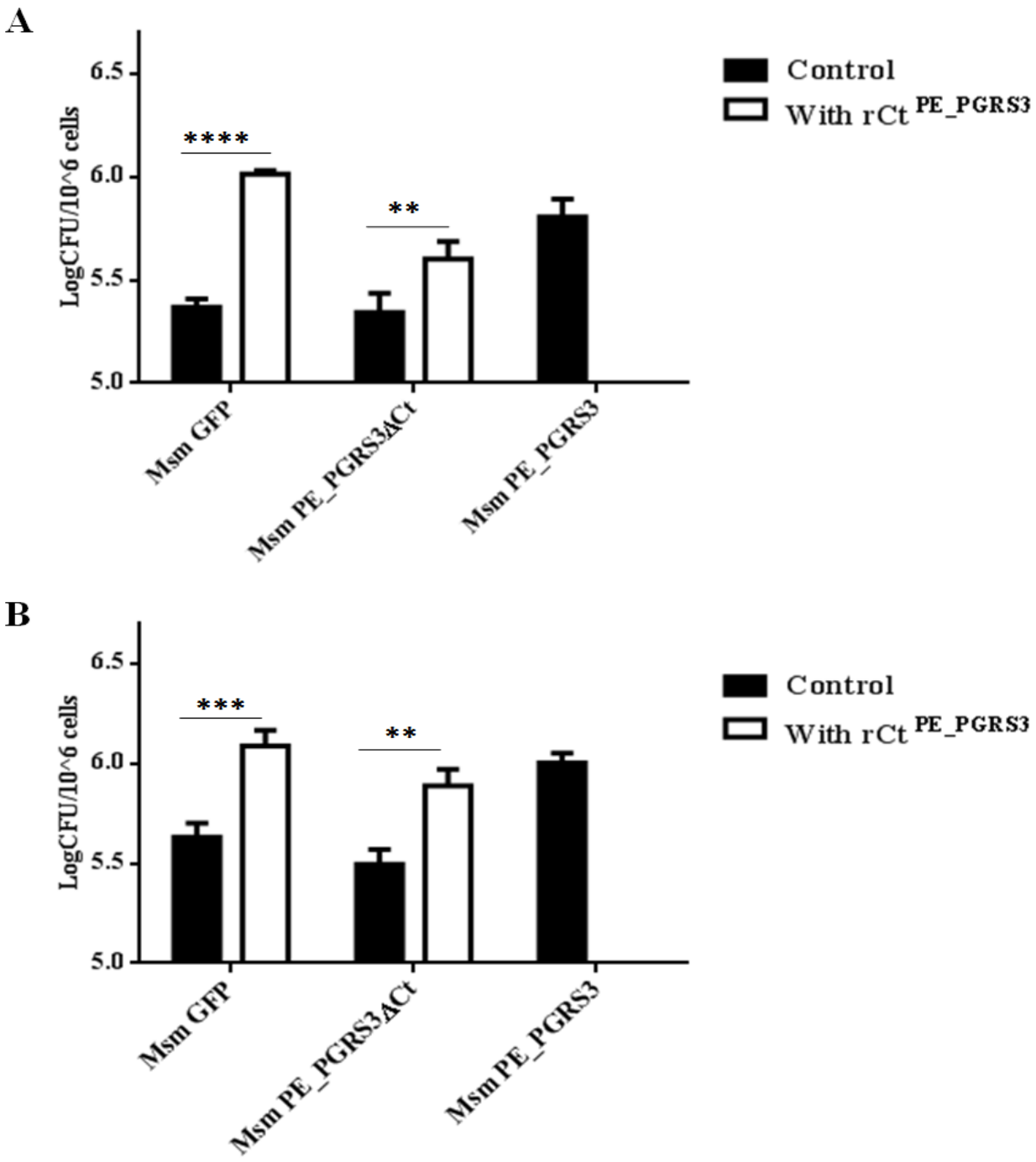


Figure 2.20: Shows the co-infection with the recombinant C-terminal domain of the PE_PGRS3(rCt^{PE_PGRS3}) of the murine macrophages (J774) (A) and human alveolar epithelial cells (A549) (B). Recombinant *M. smegmatis* (*Msm*) strains expressing PE_PGRS3 and *Msm* expressing PE_PGRS3ΔCt, under control of the *hbha* promoter and *Msm* expressing GFP were used to infect (J774) murine macrophages and human alveolar epithelial cells (A549) at MOI 10:1 after adding the rCt $^{PE\_PGRS3}$ at 0.5µg/ml to the infecting solution containing *Msm* PE_PGRS3ΔCt and to the infecting solution containing *Msm* GFP. CFU were evaluated at 4 hours post infection.****P<0.0001 for *Msm* GFP + rCt $^{PE\_PGRS3}$ vs *Msm* GFP in macrophages and ***P<0.001 in alveolar epithelial cells, **P<0.01 for *Msm* PE_PGRSΔCt + rCt $^{PE\_PGRS3}$ vs *Msm* PE_PGRSΔCt in macrophages and alveolar epithelial cells.

Chapter III

Material and methods

1. Construction of gene reporter vectors

1.1 Construction of plasmids expressing PE_PGRS3 and PE_PGR4 fused with green fluorescent protein (GFP) and with haemagglutinin (HA) epitope

The *pe_pgrs3* and *pe_pgrs4* full length genes (*Rv0278c-Rv279c* respectively) and their putative promoter sequences were amplified from *M. tuberculosis* H37Rv genomic DNA (Cole *et al.*, 1998) using a set of primer indicated in table.1. Briefly, the forward primers were designed to anneal to 250bp upstream of start codon of *Rv0278c* and *Rv0279c* so to amplify their putative promoters sequence and contained the HindIII restriction site adaptor sequence. Reverse primers were designed to anneal to the end of the gene sequences, but without stop codon and contained the XbaI or NheI adaptor sequence. Cloning was performed using standard procedures; PCR products were amplified using Expand High Fidelity PCR system (Roche) polymerase and cloned in PCR 2.1 Topo T/A cloning (Life Technology). Both gene sequences were controlled by sequencing before to continue the next cloning. *pe_pgrs3* and *pe_pgrs4* with their own promoters were then inserted in pMV206 medium copy episomal plasmid in frame and upstream of GFP coding sequence. The same genes were also inserted in the integrative plasmid pMV306 in frame and upstream of the sequence encoding haemagglutinin (HA) epitope.

Table 1: list of primer used to amplify *Rv0278c-Rv279c* genes and the C-terminal domain of the *Rv0278c* of the *M. tuberculosis* genome.

Primer	sequence	restriction enzyme	note	construct
P3prHn-5	**5`-**ACCAAGCTTACCGCGAACCGGCCATCAGAC-3`	HindIII	forward	pMVPE_PGRS3
P3Xb-3	**5`-**ACCTCTAGACGGCATCATCTGCGGTGAGCA-3`	Xba1	reverse	pMVPE_PGRS3
P4prHn-5	**5`-** ACCAAGCTTTTACCACTATTCCACACCACG-**3`**	HindIII	forward	pMVPE_PGRS4
P4Nh-3	**5`-** ACCGCTAGCCAGGCCGTTGAGCCCGTT-3`	NdeI	reverse	pMVPE_PGRS4
P3CtEXP-5	**5`-** CGGCGGGTTTGGCGCCGGTAC-3`	none	forward	pET_SUMO-Ct3
P3EXP-3	**5`-** CTACGGCATCATCTGCGGTGAGCA-3`	none	reverse	pET_SUMO-Ct3

1.2 Construction of plasmid expressing PE_PGRS3 under control of the *hbh*A promoter and of its functional deletion mutants in a plasmid expressing green fluorescence protein (GFP)

The *pe_pgrs3* full length and its selected fragments were amplified from pCR vector previously created, using primers indicated in table. 2 and cloned using standard procedure. Briefly, the forward primer was designed to anneal to the start codon *Rv0278c* region so to amplify the gene sequence without its putative promoter and contained *NdeI* restriction site adaptor sequence. Reverse primers were designed to anneal to different positions of the *Rv0278c* coding sequence and contained *Nhe*I adaptor sequence. PCR products were amplified using Expand High Fidelity PCR system (Roche) polymerase and cloned in PCR 2.1 topo T/A cloning (life technology), controlled by sequencing as previously described and purified (QIAGEN). Finally, *pe_pgrs3* and its gene chimeras were cloned in a modified episomal pMV206 vector that contained *hbhA* promoter upstream the cloning site and the sequence coding haemagglutinin (HA) epitope downstream the cloning site. Furthermore, the modified pMV vector carried also the GFP sequence under control of mycobacterial antigen 85 promoter as gene reporter (Fig 2.11B).

Table 2: list of primers used to amplify *Rv0278c* gene without promoter and its domains.

Primer	sequence	restriction enzyme	note	construct
P3Nd-F	5`-ACCCATATGATGTCGTTTGTGATTGCGGCGCCA -3`	NdeI	forward	pMVPE_PGRS3Δpr
P3Nh-R	5`-ACCGCTAGCCGGCATCATCTGCGGTGAGCA-3`	NheI	reverse	pMVPE_PGRS3Δpr
P3NhdCT-R	5'-ACCGCTAGCGTCACCGCCGTTGCCGAACAC-3`	NheI	reverse	pMVPE_PGRS3Δct
P3NhPG1-R	5`-ACCGCTAGCGGTGAGCGTCTGGAAGGGCTC-3`	NheI	reverse	pMVPE_PGRS3ΔGRPLI
P3NhPE-R	5`-ACCGCTAGCGATCAACGGGCGCCCGGTATTGCC-3`	NheI	reverse	pMVPE3

2. Construction of 6xHis-SUMO fusion expression vector

C-terminus domain of the gene *Rv0278* were amplified starting from the pCR vector previously generated using the set of primers indicated in table.1, to obtain a PCR product suitable to be ligated into pET-SUMO vector (Life Technology), according to manufacturer's instruction. PCR was performed using primer in table.1 were generated to amplify the last 234bp of the gene with stop codon , with Expand High Fidelity PCR system (Roche) polymerase which adds A (adenine) nucleotide at each 3′-ends of the amplified fragment. pET_SUMO vectors has TA-cloning system, so that linear pET-SUMO vector 5′-T ends and PCR fragment 3′ -A can easily pairing. PCR amplified fragment of about 234bp for the C-terminal of the PE_PGRS3 was purified (QUIGEN) and ligated into pET_SUMO vector. Ligation mixture was transformed into *E.coli* TOP10 chemically competent cells (Life Technology) and positives clones were screened by double digestion with HindIII and Kpn1. Finally, the C-terminus sequence was also verified by sequencing.

3. Bacterial strains media and growth conditions

In order to obtain plasmids that were used to transform *M. smegmatis* strain, competent *E. coli* Top 10 (Life technologies) was used. *E. coli* was grown using Lauria Bertani broth medium (Sigma Aldrich) or Lauria Bertani agar medium (Sigma Aldrich) following the manufacture standard protocol. To select single colonies that had acquired the pCR 2.1 T/A cloning vector, ampicillin and kanamycin (Life technologies) were added at final concentration of 100 µg/ml and 40µg/ml respectively. Transformants for pMV vectors were selected on Lauria Bertani agar medium (Sigma Aldrich) containing 150µg/ml hygromycin B (Sigma Aldrich) or 40µg/ml kanamycin for the strains transformed with pMV206 vector

and with pMV306 or pMV206 modified vector respectively. *M. tuberculosis* and *M. smegmatis* mc^2 155 were grown at 37c° in Middelbrook 7H9 broth or on 7H11 solid medium (Difco Becton-Dickinson), supplemented with 0.2% glycerol (Sigma-Aldrich), ADC 10% (Becton-Dickinison), and 0.05% v/v Tween 80 (Sigma-Aldrich) (Delogu *et al.*, 2004) (De *et al.*, 2014). *M. smegmatis* transformants were selected on 7H11 agar media supplemented with 10% OADC (Microbiol) containing 50µg/ml hygromycin B (Sigma Aldrich) or 40µg/ml kanamycin for the strains transformed with pMV206 vector and with pMV306 or pMV206 modified vector respectively. Single individual antibiotic-resistant colonies were isolated and subcultured in a 7H9 media supplemented with 10% ADC (Becton-Dickinson) and 0.05% Tween 80 containing hygromycin B (50 µg/ml) and incubated at 37°C. Mycobacteria cultures were stocked at -80°C in 20% glycerol. Strains processed for cell fractionation were grown after inoculation at 1:100 in Sauton's medium (Difco) for additional 24h at 37c°. The low phosphate growth condition was obtained by subinoculation in Sauton's medium without inorganic phosphate (iPhos) (1:100) to arrive ~50µM iPhos and the bacteria incubated for 15 days for *M. smegmatis* and 30-60 days for *M. tuberculosis*. For washing, the cells were centrifuged 3500rpm for 15minutes at 4C° and washed one time with complete Sauton medium and resuspended with the same medium. Test of different growth conditions were achieved by growing the *M. smegmatis* recombinant strains in 48 wells plates using a different modified media containing different concentration of inorganic phosphate or metals. The media used for growing the bacteria under different stress condition are : 7H9/ADC/tween standard medium; Sauton standard medium; Sauton iPhos depleted; 7H9/ADC/tween pH5; Sauton ferric iron depleted Fe^{+3}; Sauton magnesium depleted Mg^{+2}; Phosphate-buffered saline (PBS); low oxygen was achieved by adding oil to the medium to mimic anaerobic conditions. Then the

fluorescence of the mycobacterial recombinant strains which correspond to the protein expression was observed by fluorescence microscopy (NIKON, phase-contrast 2, ELWD 0.3).

4. Electroporation

M. smegmatis mc^2 155 grown to mid-exponential (log) phase were extensively washed in 10% glycerol and concentrated approximately 40-fold. 50-100µl concentrated cells were mixed with 1µg of DNA, and transferred to 0.25 cm gap cuvettes (eppendrof). Samples were electroporated using a Bio-RAD GenePulser X Cell™ electroporation system (voltage 1250 V, capacitance 25 µF, resistance 800 Ω, cuvette 2 mm). After the pulse, the cells were diluted in 900 µl of liquid medium, incubated for 4h then plated on 7H11/OADC/tween80 selective solid medium.

5. Expression of C-terminal domain of the PE_PGRS3

5.1 Protein purification

To purify the C-terminal domain of the PE_PGRS3, *E. coli* BL21 (DE3) chemically competent cells (life technologies) was transformed with pET_SUMO-CT, then a single colony of pET_SUMO-Ct BL21 selected on Lauria Bertani agar medium (Sigma Aldrich) 50µg/ml Kanamycin (Sigma Aldrich), was grown in 5 ml of Lauria Bertani broth medium (Sigma Aldrich) 50µg/ml Kanamycin, 1% glucose over night at 37C°. Then a 250ml culture was prepared and incubated until the mid-log phase (OD_{600}: 0.5), when IPTG (3V chimica) was added at final concentration of 1mM. After 3h of induction the culture was pelleted (14000 rpm for 20 minutes at 4Cº), the pellet was washed one time with cold PBS buffer, resuspended 1:10 in lysis buffer (PBS, 350mM Nacl, 1:100 protease inhibitor cocktail –SIGMA Aldrich, pH: 7.4) and finally lysated by sonication (five times

per 5 minutes, with 1 minute in ice between each sonication). Supernatant was separated from bacterial cells by centrifugation (25000rpm, for 15 minutes at 4Cº) and collected for purification. Protein purification was carried out by Fast Protein Liquid Chromatography (FPLC- ÄKTA, GE health care life sciences). Soluble lysate containing 6XHis-tagged recombinant C-terminal domain of the PE_PGRS3 was purified by ion metal affinity chromatography with Ni-NTA agarose 1ml bed volume column (GE-health Care) with flow rate 1ml/min. The column was properly equilibrated with buffer A (PBS-350mM NaCl). Protein elution was obtained mixing automatically buffer A with buffer B (PBS-350mM NaCl, 500mM Imidazole, 3V chimica) in order to have a linear increasing of the salt gradient (50-350mM).

5.2 Purified recombinant C-terminal domain of the PE_PGRS3 LPS free preparation

We started with dialyzation of 5ml of the purified protein using dialysis device (Float-A-Lyzer G2) in PBS at 4 C°, the procedure performed according to standard protocols. Then 2ml of the dialyzed purified C-terminal of the PE_PGRS3 was incubated with 2ml PBS Triton X-114 (0.2%) for 1hour at 4C° in agitation, followed by a centrifuge 7000g for 20minutes using Amicon filter 10KDa, then washing several times with 3ml PBS Triton X-114 0.1% and then 3ml of PBS with centrifugation (7000g for 20minutes). Finally, the protein was collected with 1ml of PBS and the sample was prepared for measuring the LPS according to standard protocols (Limulus Amebocyte Lystae –QCL-1000™ LONZA). The endotoxin free was considered as a concentration < 0.1 EU/ml and the reaction absorbance was measured using filter 415nm (BIO-RAD).

6. Mice immunization

Pathogen – free female BALB/C mice obtained from enclosure labs of the Catholic University of Sacred Heart (Rome) and housed in a temperature controlled environment with 12h light/dark cycle, and received food and water *ad libitum*. Group of animal, 4-8 weeks old, were injected subcutaneously in two sites at the base of the tail with 100µl of antigen recombinant C-terminal domain of the PE_PGRS3 emulsified with TiterMax Gold Adjuvant (Sigma-Aldrich) at final concentration of 100µg/ml. Immunization was repeated three times, 15 days apart, and the serum was collected at the 15th day after the third immunization by the tail vien (De *et al.*, 2014). All animal experiments were authorized by Ethical Committe of the Catholic University of Sacred Heart (Rome) and performed in compliance with legislative decree of the italian goverment 27 january 1992, n. 6. All manipulation were performed under isoflurane anesthesia and all efforts were made to minimize suffering.

7. SDS_PAGE, Western blotting and Immunoblotting

Whole cell lysates were prepared from the bacterial culture which were pelleted by centrifuge (4000 rpm for 20minutes at 4C°) (De *et al.*, 2014), then the pellet were washed three times with a cold PBS and lysed in lysis buffer pH=7.5 (20mM Tris, 150mM Nacl with 1:100 protease inhibitor cocktail –SIGMA Aldrich) by sonication five times 5 minutes each with 1 minute in ice between each sonication. The secreted fraction was prepared by precipitating the proteins in the supernatant of the bacterial culture after centrifuge 3500 rpm for 15 mintes at 4C°, using acetone precipitation of proteins standard protocol. Briefly, secreted proteins were obtained after incubation of the supernatant with 4 volumes of acetone for 1 hour at -20C° and centrifugation 10000 rpm for 10 minutes at 4C°,

secreted proteins were re-suspended in cold PBS. SDS_PAGE was performed according to the standard protocols. The proteins were separated on 12% and 10% polyacrylamide gel. Samples separated by SDS_PAGE were transferred to nitrocellulose membrane (Bio-Rad) by Western blotting. Non-specific protein binding was blocked by incubating membrane into Tris-5% MILK for one hour at room temperature. Once removed blocking buffer, membranes were incubated with primary antibody diluted into Tris-2% milk at 4Cº overnight. Primary antibodies used were monoclonal anti-polyhistidine clone HIS-1 antibody (Sigma-Aldrich) diluted 1:6000, polyclonal anti C-terminal PE_PGRS3 1:2000, polyclonal anti-GFP antibody (abcam) 1:7000 or monoclonal anti-HA (biolegand) 1:500. Then the membranes were washed three times for 10 minutes with Tris-Tween 0.05% and incubated for 1h with 1:4000 anti-IgG horseradish peroxidase (HRP) conjugated antibody (Sigma-Aldrich) as a secondary antibody, after washing three times with Tris-tween 0.05% the membranes were incubated with SuperSignal West Dura extended duration substrate (Thermoscientific) and chemiluminescence detected by ChemiDoc ^{TM} XRS+ system (Bio-Rad).

8. FACS analysis

Recombinant *M. smegmatis* strains expressing PE_PGRS3, PE_PGRS4 and PE_PGRS33 under their own promoter fused with GFP, *M. smegmatis* expressing GFP and *M. smegmatis* mc^{2}155 were grown in 7H9 medium as described above and sub-inoculated in Sauton without inorganic phosphate medium (sub-inoculation in a dilution 1:100 to arrive ~50µM iPhos) supplemented with HygromycineB (50µg/ml). *Msm* expressing GFP and wild type strains were used as a control to set the instrument, while *Msm* expressing PE_PGRS4 also grown in Sauton complete medium and the fluorescence was analyzed to confirm the confocal microscope results. The fluorescence analyzed at different time points on

FACS - Canto (Bioscience) flow cytometer, The fluorescence intensity of 50.000 ungated events was measured as described (Sureka *et al.*, 2008). A side scatter threshold was set to get out inappropriate heterogeneities. It is difficult to analyze single mycobacteria by FACS because its ability to clump. Carefully gating procedures to minimize the effect of clumping. The data files were analyzed using CellQuestPro (Nioscience) and WINMIDI. The flow cytometer data were represented in figure 2.8, where x-axis is the days of incubations and y-axis represent the fluorescent bacteria %.

9. Quantitative reverse transcription - Real time PCR (qRT-PCR)

For confirmation of transcriptional profiling results. Bacteria were grown to mid-exponential phase (OD_{600} of 0.5) in 7H9/ADC/tween medium. To assess the response to inorganic Phosphate starvation, bacteria were grown to mid-exponential phase in 7H9 /ADC/tween medium and then resuspended (1:100) in Sauton medium without iPhos and in Sauton complete medium . Culture were incubated at 37C^{o} and bacteria were collected for RNA extraction at 15 days for *M. smegmatis* recombinant strains expressing PE_PGRS3 and at two time points 30 days and 60days for *M. tuberculosis* H37Rv reference strains. Cells were collected by centrifugation (4000 rpm for 15 minutes at 4C^{o}), and RNA extracted in 1ml of Trizol (Invitrogen) by beat beating with 0.1-mm zirconia-silica bead (BioSpec) then the extraction was processed using (RNase miniKit QUIGEN). The RNA was eluted in 30-50µl RNase free water and transcribed to cDNA using high capacity cDNA reverse transcription Kit (Applied Biosystem). The reverse transcription reaction used following standard manufacturer's protocol. cDNA was stored at -20C^{o} until Real time quantitative PCRs were performed. Primers for Real time quantitative reverse transcription-PCR (qRT-PCR) specific for the *16S* rRNA, *Rv2583c*, *Rv0278c* genes were designed using primer express software

(Applied Bio System) with similar melting temperature (62-64C^{o}), are listed in table. 3. Primers were tested in PCRs using *M. tuberculosis* genome equivalents as a template, and products were analysed by gel electrophoresis. Real time quantitative PCRs were prepared with 2X Sybr master mix (Applied Biosystems), 2µl cDNA, 0.5 µM primers and were run in absolute quantification mode on a 7300 real time machine (Applied Biosystems), using standard cycling conditions. Mock reactions (no RT) were performed on each sample to confirm the absence of genomic DNA contamination. Cycle threshold C_T values were concerted to copy numbers using standard curves for each genes. Target cDNA was internally normalized to 16s cDNA.

Table 3: list of primers used for Real time quantitative reverse transcription PCR (qRT-PCR) specific for the *16S* rRNA, *Rv2583c*, *Rv0278c* genes.

Primer	sequence	note	gene
PE_PGRS3-Fw	CGTGTTGATCGGCAATGG	forward	*pe_pgrs3*
PE_PGRS3-Rev	TTCACCACATTGAGCACAT	reverse	*pe_pgrs3*
RelA-Fw	TACCAGAAGATGATCGTTAAG	forward	*relA*
RelA-Rev	GCGATGTAGTCCTTGAAC	revers	*relA*
16S-Fw	GTATTCACCGCAGCGTTGC	forward	*16S*
16S-Rev	CCACTGGCTTCGGGTGTTA	reverse	*16S*

10. Confocal microscopy and image analysis

The recombinant mycobacterial strains *M. smegmatis* mc^{2}155 expressing: PE_PGRS4-GFP, PE_PGRS3-GFP, *M. smegmatis* expressing PE_PGRS33- GFP, *M. smegmatis* expressing GFP, all strains were grown in 7H9/ADC/tween80 at mid-log phase then sub inoculated in Sauton complete medium and Sauton with low inorganic phosphate and incubated for 15 days at 37C°. Furthermore recombinant mycobacterial strains expressing PE_PGRS3 and its chimeras under constitutive promoter *hbhA* promoter all tagged with GFP, were grown in

7H9/ADC/tween 80 and at mid-log phase at 37C°. All strains were plated on chamber slides pre-treated with polylisine (Sigma, Alrich, Sant Louis, Mo). Subsequently chamber slides were incubated for 24h at 37C°, then fixed with 4% paraformaldehyde and washed with PBS. Chamber slides were closed and observed with a confocal microscope. Confocal and transmission images were collected using an inverted confocal microscope (DMIRE2, Leica microsystem, wetzlar Germany) equipped with a 60X oil immersion objective (NA 1.25). For GFP excitation a He/Ne laser at 476nm was used. Internal photon multiplier tubes collected 8-bit unsigned images at 400Hz scan speed in an emission range comprised between 500 and 550. Imaging was performed at room temperature. Image processing was performed with image J software; image background value (defined as intensities below 7% of the maximum intensity) were set in zero and colored in black. Confocal images for measuring bacterial size was collected using confocal microscope (Nikon A1MP) with a 60X oil immersion objective (1 pixel = 0.21µm). For GFP excitation (laser at 488.3nm, HV-saturation : 30, laser potency :5, offset : -2) were used. The bacterial cell length was measured using images j program. At least 100 bacteria were analysed per 3field for each strain.

11. Cell culture and mycobacteria infection

J774- murine macrophages and A549 human alveolar epithelial cells were grown in DMEM medium (Euroclone) enriched with 10% fetal bovine serum (FBS), 2mM glutamine (Euroclone), 100 µg/ml streptomycine and penicillin (Euroclone), were kept in humidified atmosphere containing 5% CO2 at 37C°. Before infection, cells were collected and suspended in complete medium with 2% FBS without antibiotics (maintaining medium). Cells were plated at 1.2 x10^6 cell per ml in 48-wells plates infected with *M. smegmatis* expressing PE_PGRS3, *M. smegmatis* expressing PE_PGRS3ΔCt and *M. smegmatis* expressing GFP as a

control strain at multiplicity of infection MOI 10:1 for 4h and incubated in a 5% CO2 atmosphere at 37C°. Once removed infecting solution, cells were washed three times with PBS and incubated with maintaining medium for the second time points 4 days post infection. Intracellular colony forming units (CFU) were obtained at 4h post infection and 4d post infection, where the infected cells were washed three times with PBS to remove any remained extracellular bacteria, lysed in 0.1% Triton X-100 and intracellular bacteria determined by colony forming units (CFU) counting determined as described (Zumbo *et al.*, 2013). The co-infection with the purified C-terminal domain of the PE_PGRS3 was prepared by adding the purified protein at concentration 0.5µg/ml to, *M. smegmatis* expressing PE_PGRS3ΔCt and *M. smegmatis* expressing GFP infecting solution, CFUs at 4h post infection were evaluated as previously described.

12. *M. smegmatis* recombinant strains growth rate measurements

A dilution of the *M. smegmatis* strains expressing: PE_PGRS3, PE_PGRSΔCt, PE_PGRSΔGRPLI, PE3, *M. smegmatis* expressing GFP was prepared from the conserved stock in PBS to arrive a concentration 10^6 cell/ml. Then were subinoculated in MGIT (mycobacterial growth indicating tube) tube containing 7ml Middelbrook 7H9 broth base supplemented with OADC enrichment, BBL™ MGIT™ PANTA™ antibiotic mixture to arrive a concentration 10^3 cell/ml then all tubes were incubated in the instrument (BACTEC MGIT 960) until give a positive corresponding to the fluorescence detected depending on the oxygen consume.

13. Cytotoxicity assay

The cytotoxicity effect of the purified recombinant C-terminal domain of the PE_PGRS3 were measured using Alamar Blue Assay which detects the colorimetric changes due to oxidation and reduction of the dye. A549 human alveolar epithelial cells and J774 murine macrophages were used to evaluate the cytotoxic effect of the purified recombinant C-terminal domain of the PE_PGRS3. Two 96-wells plates were prepared at concentration $5x10^5$ cell/well and incubated with different concentration of the purified LPS free protein (0.1, 1, 5, 10µg/ml), 10% Alamar Blue was added to the fresh media. The cytotoxic effect was evaluated at four time points (4h, 1d, 4d, 7d), visible as a media colour change, was evaluated by measuring in microplate reader (BIORAD), the absorbance at OD (570-630nm) after 3h from incubation.

14. Multiple sequence alignments

Phylogenetic tree obtained of multiple sequence alignment of PE_PGRS3 between different MTBC strains; *M. tuberculosis H37Rv*, *M. africanum*, *M. tuberculosis EAI*, *M. bovis*, *M. bovis BCG*, *M. canetti*, using multi-alignments tools (Clustel Omega and Jail view software) for nucleotides sequence and amino acids sequence multi-alignments.

15. Statistical analysis

All experiment were performed at least in triplicate and replicated at least three times. Graphpad Prism software 7.01(GraphPad software, CA, USA) was used for statistical analysis. All data were expressed as mean with SD and analyzed by one-way and two-way ANOVA followed by bonferroni's and Tukey multiple comparison test, as specified in the caption under each figure, family wise significance and confidence levels 0.05 (95% confidence intervals).

Chapter IV

Discussion

TB is still one of the most important infectious diseases at global level, causing high mortality and morbidity particularly in low income countries. The association of TB with the HIV pandemic and the spread of multidrug-resistance strains (MDR-) and more recently extensively drug resistant (XDR-) *M. tuberculosis* strains is a cause of major concern for public health authorities. There is a need for a new and improved vaccine, for new drugs and diagnostics, though there is a wide consensus that to design and develop these new tools we need a better understanding of *M. tuberculosis* pathogenesis and of the biology of the tubercle bacilli. For example, the ability of *M. tuberculosis* to remain dormant after primary infection and cause a latent infection remains still poorly understood as well as the mechanism responsible for disease reactivation in an immunocompetent host. Among the peculiar features of *M. tuberculosis* is certainly the presence of two large protein families (PE and PPE), which are restricted to mycobacteria causing tuberculosis or disease in mammals. The genes coding for these proteins account for almost 8% of the coding capacity and were shown to be responsible for most of the genetic variability of *M. tuberculosis* (Delogu G and Cole S.T and borsch R, 2008). The PE family is further divided in two subfamilies PE and PE_PGRS and several studies on the latter group suggested that these proteins could be directly implicated in TB pathogenesis and may be involved in the evasion of host immune responses (Cole *et al.*, 1998), (Brennan and Delogu, 2002), (Banu *et al.*, 2002). Until now, our understanding on the expression, function and host response to PE_PGRS proteins remains limited to few published articles and several questions remains unanswered.

In this study, for the first time we studied PE_PGRS3 and PE_PGRS4 which are found in the same gene locus and show high level of genetic homology. Both PE_PGRSs are among the few of the family with two GRPLI motifs and

PE_PGRS3 has a unique arginine rich C-terminal domain. We demonstrated that PE_PGRS3 and PE_PGRS4 have a different expression pattern. Moreover PE_PGRS3 is expressed in low iPhos condition and we confirmed for the first time that the inorganic phosphate is a specific regulator for the expression of a PE_PGRS protein. Overexpression of the *pe_pgrs3* gene in low phosphate conditions correlated with *relA* upregulation in *M. smegmatis* recombinant strains and *M. tuberculosis*. To investigate the role of PE_PGRS3, we expressed the protein under the control of the *hbhA* promoter and analysis at the confocal microscope of the *M. smegmatis* strain overexpressing PE_PGRS3 and its functional deletion mutant demonstrated shorter bacterial cells, which were not observed in the *M. smegmatis* strain overexpressing the PE_PGRS3 functional deletion mutant lacking the C-terminal domain of the protein.When the *M. smegmatis* strains overexpressing PE_PGRS3 were used to infect *in vitro* murine macrophages or alveolar epithelial cells, a superior ability to entry in macrophages and epithelial cells was observed for the strain expressing full length PE_PGRS3, while no significance in the intracellular survival were measured. These results provide new and original insights on the functional characterization of PE_PGRS3 and support the role of PE_PGRS proteins in TB pathogenesis. PE_PGRS3 is 957 amino acids protein and belong to the group of PE_PGRSs containing, within the PGRS domain, the second GRPLI motif (Delogu G and Cole S.T and borsch R, 2008) and characterized by the presence of a C-terminal unique domain rich in arginine residues. The different expression pattern that we observed for the two neighboring genes *pe_pgrs3* and *pe_pgrs4* confirm that *pe_pgrs* genes are differentially regulated during TB pathogenesis. Among the different conditions tested, only the presence of low iPhos was able to induce expression of PE_PGRS3.

A phosphate-limited environment is encountered by *M. tuberculosis* in macrophages and in host tissues, when the tubercle bacilli copes with the harsh environment triggered by the host immune response by activating the stringent response (Rifat *et al.*, 2009). The finding that PE_PGRS3 expression follows the same pattern of RelA, a well characterized mediator of the *M. tuberculosis* stringent response, (Ghosh *et al.*, 2011). RelA has been demonstrated that catalyses the synthesis of Guanosine tetra phosphate which regulates the expression of several genes involved in the biogenesis of the cell wall and more in general regulates the bacterial stringent response, that is known to be essential for long term survival of *M. tuberculosis* and persistence infection in mice (Dahl *et al.*, 2003). The results of our study suggest that RelA could be a mediator for PE_PGRS3 expression and as such PE_PGRS3 may have a specific role in *M. tuberculosis* biology during persistence in host tissues. Phosphate starvation has been shown to induce expression of the genes of the ESX5 secretion system (Elliott and Tischler, 2016). The ESX-5 secretion system plays an important role in *M. tuberculosis* pathogenesis, promoting cell death and facilitating bacterial spread (Abdallah *et al.*, 2011) (Abdallah *et al.*, 2008). ESX5 is also considered the major secretion pathway for PE and PPE proteins, including PE_PGRSs (Abdallah *et al.*, 2006), (Abdallah *et al.*, 2009). Our findings indicate that *M. tuberculosis* specifically tunes expression of at least one PE_PGRS (PE_PGRS3) under conditions classically associated with the stringent response support the involvement of PE_PGRS3 in the stringent response.
To investigate the role of PE_PGRS3, we developed an expression system that warranted good expression of the protein in experimental conditions normally used to investigate bacterial physiology or *in vitro* models of infections. Overexpression of PE_PGRS3 did not affect bacterial replication though we observed shorter cells in the *M. smegmatis* strain expressing PE_PGRS3 compared

to the control or the *M. smegmatis* strain expressing PE_PGRS3 lacking the arginine rich C-terminal domain (PE_PGRS3ΔCt). Interestingly, *M. smegmatis* strains expressing full length PE_PGRS3 showed enhanced tropism for macrophages and alveolar pneumocytes compared to *M. smegmatis* control and *M. smegmatis* expressing PE_PGRS3ΔCt, clearly implicating the arginine rich C-terminal domain in the process. We hypothesize that the cationic C-terminal domain may serve to promote adhesion of *M. tuberculosis* to host cells, which may be important for the tubercle bacilli to create a protected niche during *M. tuberculosis* persistence in macrophages. It was suggested that MTBC strain similar to *M. canetti*, also named smooth tubercle bacilli, could be the common ancestors of *M. tuberculosis* (Brosch *et al.*, 2002). Interestingly, multiple sequence alignment of PE_PGRS3 between different MTBC strains belonging to different phylogeographic lineages indicated that not all strains have full length PE_PGRS3. Instead, full length PE_PGRS3 is observed only in *M. tuberculosis* strains (new and ancient) that cause disease in human but not in those that cause disease in animal (*M. bovis*). The analysis that we obtained from the phylogenetic tree of the multiple sequence alignment suggests that the *pe_pgrs3* as a single gene exists only in *M. tuberculosis*. In the ancestral *M. canetti* strain the *pe_pgrs3* exists as a two genes *pe_pgrs3* (2700bp), shorter than the *Rv0278c* (2874bp), and *pe_pgrs3a* (2184bp), both of them lacking of region encoding the C-terminal domain. Interestingly, also the *M. africanum* and animal adapted strains *M. bovis* strains present two genes: *pe_pgrs3* and *pe_pgrs3a* (Fig. 4.1), and in both cases expression of the arginine-rich C-terminal domain is not expected. While further studies are required to dissect these events, we speculate that the expression of the full length PE_PGRS3 is restricted to *M. tuberculosis* causing disease in humans, highlighting the specific role of this protein in the immunopathogenesis of human TB.

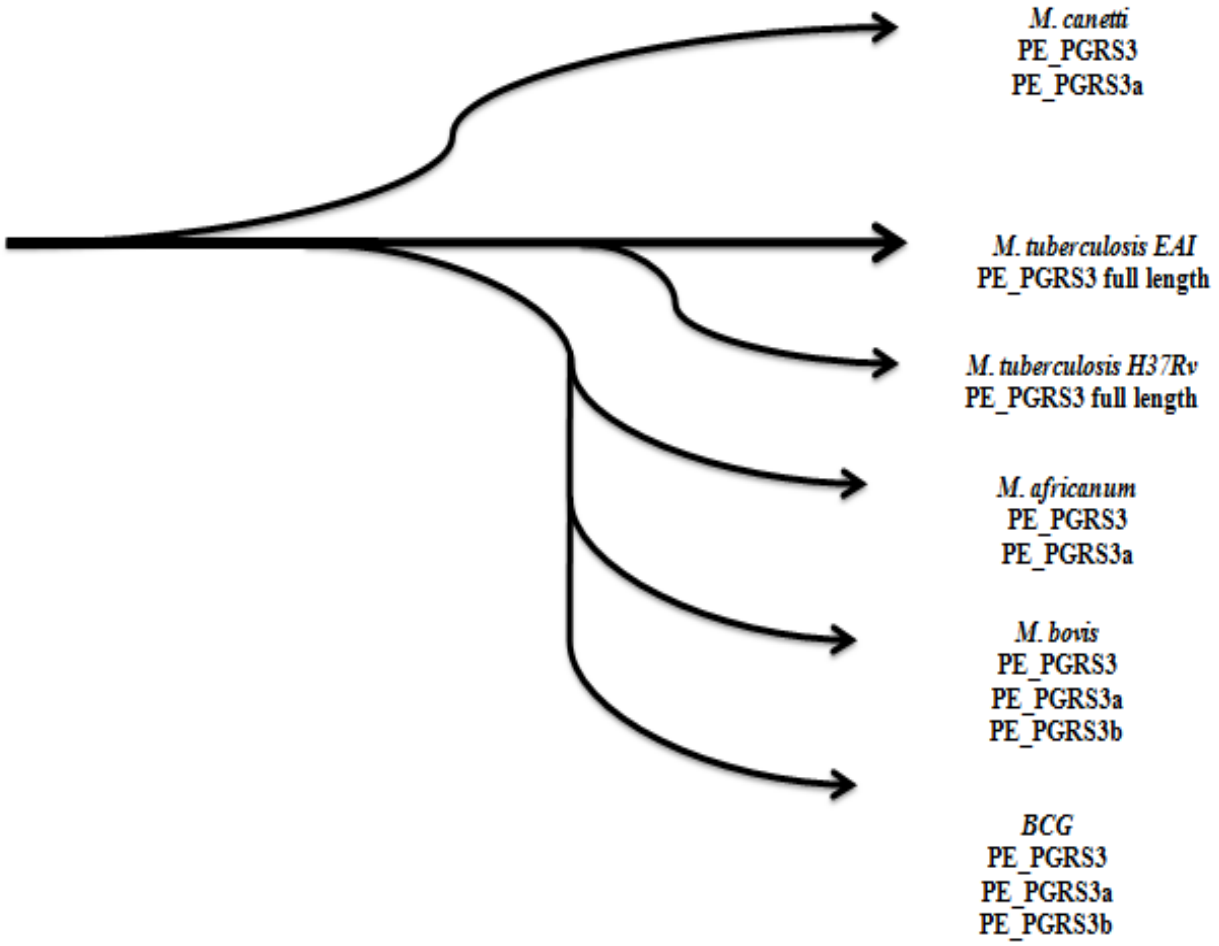


Figure 4.1: Scheme of the proposed evolutionary pathway of the tubercle bacilli illustrating evolutionary genetic events of the *Rv0278c*.

These results of this study implicate for the first time PE_PGRS3 in TB pathogenesis and highlight the role of the arginine rich C-terminal domain in this process. PE_PGRS proteins could be considered multifunctional proteins sharing a similar architecture with different functions resulting from the different domains,

which may play specific roles in different steps of the infectious process. Understanding their role and function in *M. tuberculosis* biology and TB pathogenesis may open new avenues to fight.

References

Abdallah, A. M., Bestebroer, J., Savage, N. D., de, P. K., van, Z. M., Wilson, L. *et al.* (2011) Mycobacterial secretion systems ESX-1 and ESX-5 play distinct roles in host cell death and inflammasome activation. *J Immunol* 187: 4744-4753.

Abdallah, A. M., Gey van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C. M. *et al.* (2007) Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* 5: 883-891.

Abdallah, A. M., Savage, N. D., van, Z. M., Wilson, L., Vandenbroucke-Grauls, C. M., van der Wel, N. N. *et al.* (2008) The ESX-5 secretion system of Mycobacterium marinum modulates the macrophage response. *J Immunol* 181: 7166-7175.

Abdallah, A. M., Verboom, T., Hannes, F., Safi, M., Strong, M., Eisenberg, D. *et al.* (2006) A specific secretion system mediates PPE41 transport in pathogenic mycobacteria. *Mol Microbiol* 62: 667-679.

Abdallah, A. M., Verboom, T., Weerdenburg, E. M., Gey van Pittius, N. C., Mahasha, P. W., Jimenez, C. *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.

Ahmed, A., Das, A., and Mukhopadhyay, S. (2015) Immunoregulatory functions and expression patterns of PE/PPE family members: Roles in pathogenicity and impact on anti-tuberculosis vaccine and drug design. *IUBMB Life* 67: 414-427.

Bansal, K., Elluru, S. R., Narayana, Y., Chaturvedi, R., Patil, S. A., Kaveri, S. V. *et al.* (2010) PE_PGRS antigens of Mycobacterium tuberculosis induce maturation and activation of human dendritic cells. *J Immunol* 184: 3495-3504.

Banu, S., Honore, N., Saint-Joanis, B., Philpott, D., Prevost, M. C., and Cole, S. T. (2002) Are the PE-PGRS proteins of Mycobacterium tuberculosis variable surface antigens? *Mol Microbiol* 44: 9-19.

Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S. *et al.* (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284: 1520-1523.

Berthet, F. X., Rasmussen, P. B., Rosenkrands, I., Andersen, P., and Gicquel, B. (1998) A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). *Microbiology* 144 (Pt 11): 3195-3203.

Bottai, D. and Brosch, R. (2009) Mycobacterial PE, PPE and ESX clusters: novel insights into the secretion of these most unusual protein families. *Mol Microbiol* 73: 325-328.

Boutte, C. C. and Crosson, S. (2013) Bacterial lifestyle shapes stringent response activation. *Trends Microbiol* 21: 174-180.

Brandt, L., Elhay, M., Rosenkrands, I., Lindblad, E. B., and Andersen, P. (2000) ESAT-6 subunit vaccination against Mycobacterium tuberculosis. *Infect Immun* 68: 791-795.

Brennan, M. J. and Delogu, G. (2002) The PE multigene family: a 'molecular mantra' for mycobacteria. *Trends Microbiol* 10: 246-249.

Brennan, P. J. and Besra, G. S. (1997) Structure, function and biogenesis of the mycobacterial cell wall. *Biochem Soc Trans* 25: 188-194.

Brennan, P. J. and Nikaido, H. (1995) The envelope of mycobacteria. *Annu Rev Biochem* 64: 29-63.

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K. *et al.* (2002) A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A* 99: 3684-3689.

Cardona, P. J. (2006) RUTI: a new chance to shorten the treatment of latent tuberculosis infection. *Tuberculosis (Edinb)* 86: 273-289.

Champion, P. A. and Cox, J. S. (2007) Protein secretion systems in Mycobacteria. *Cell Microbiol* 9: 1376-1384.

Chao, M. C. and Rubin, E. J. (2010) Letting sleeping dos lie: does dormancy play a role in tuberculosis? *Annu Rev Microbiol* 64: 293-311.

Cohen, I., Parada, C., Acosta-Gio, E., and Espitia, C. (2014) The PGRS Domain from PE_PGRS33 of Mycobacterium tuberculosis is Target of Humoral Immune Response in Mice and Humans. *Front Immunol* 5: 236.

Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V. *et al.* (1994) Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *JAMA* 271: 698-702.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D. *et al.* (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 393: 537-544.

Converse, S. E. and Cox, J. S. (2005) A protein secretion pathway critical for Mycobacterium tuberculosis virulence is conserved and functional in Mycobacterium smegmatis. *J Bacteriol* 187: 1238-1245.

Dahl, J. L., Arora, K., Boshoff, H. I., Whiteford, D. C., Pacheco, S. A., Walsh, O. J. *et al.* (2005) The relA homolog of Mycobacterium smegmatis affects cell appearance, viability, and gene expression. *J Bacteriol* 187: 2439-2447.

Dahl, J. L., Kraus, C. N., Boshoff, H. I., Doan, B., Foley, K., Avarbock, D. *et al.* (2003) The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of Mycobacterium tuberculosis in mice. *Proc Natl Acad Sci U S A* 100: 10026-10031.

Daniel, T. M. (2006) The history of tuberculosis. *Respir Med* 100: 1862-1870.

De, M. F., Maulucci, G., Minerva, M., Anoosheh, S., Palucci, I., Iantomasi, R. *et al.* (2014) Impact of protein domains on PE_PGRS30 polar localization in Mycobacteria. *PLoS One* 9: e112482.

Delogu G and Cole S.T and borsch R (2008) The PE and PPE protein families of Mycobacteriun tuberculosis. In *Hand book of tuuberculosis*. Kaufmann S H and Rubin E (ed.) Weinheim: Willy-VCH, pp. 131-150.

Delogu, G. and Brennan, M. J. (2001) Comparative immune response to PE and PE_PGRS antigens of Mycobacterium tuberculosis. *Infect Immun* 69: 5606-5611.

Delogu, G., Pusceddu, C., Bua, A., Fadda, G., Brennan, M. J., and Zanetti, S. (2004) Rv1818c-encoded PE_PGRS protein of Mycobacterium tuberculosis is surface exposed and influences bacterial cell structure. *Mol Microbiol* 52: 725-733.

Delogu, G., Sali, M., and Fadda, G. (2013) The biology of mycobacterium tuberculosis infection. *Mediterr J Hematol Infect Dis* 5: e2013070.

Dheenadhayalan, V., Delogu, G., Sanguinetti, M., Fadda, G., and Brennan, M. J. (2006) Variable expression patterns of Mycobacterium tuberculosis PE_PGRS genes: evidence that PE_PGRS16 and PE_PGRS26 are inversely regulated in vivo. *J Bacteriol* 188: 3721-3725.

Elliott, S. R. and Tischler, A. D. (2016) Phosphate starvation: a novel signal that triggers ESX-5 secretion in Mycobacterium tuberculosis. *Mol Microbiol* 100: 510-526.

Feltcher, M. E., Sullivan, J. T., and Braunstein, M. (2010) Protein export systems of Mycobacterium tuberculosis: novel targets for drug development? *Future Microbiol* 5: 1581-1597.

Forrellad, M. A., Klepp, L. I., Gioffre, A., Garcia, J., Morbidoni, H. R., de la Paz, S. M. *et al.* (2013) Virulence factors of the Mycobacterium tuberculosis complex. *Virulence* 4: 3-66.

Gao, L. Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J. N., and Brown, E. J. (2004) A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol* 53: 1677-1693.

Gengenbacher, M. and Kaufmann, S. H. (2012) Mycobacterium tuberculosis: success through dormancy. *FEMS Microbiol Rev* 36: 514-532.

Gey van Pittius, N. C., Gamieldien, J., Hide, W., Brown, G. D., Siezen, R. J., and Beyers, A. D. (2001) The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria. *Genome Biol* 2: RESEARCH0044.

Gey van Pittius, N. C., Sampson, S. L., Lee, H., Kim, Y., van Helden, P. D., and Warren, R. M. (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.

Ghosh, S., Sureka, K., Ghosh, B., Bose, I., Basu, J., and Kundu, M. (2011) Phenotypic heterogeneity in mycobacterial stringent response. *BMC Syst Biol* 5: 18.

Gillespie, S. H. (2002) Evolution of drug resistance in Mycobacterium tuberculosis: clinical and molecular perspective. *Antimicrob Agents Chemother* 46: 267-274.

Glover, R. T., Kriakov, J., Garforth, S. J., Baughn, A. D., and Jacobs, W. R., Jr. (2007) The two-component regulatory system senX3-regX3 regulates phosphate-dependent gene expression in Mycobacterium smegmatis. *J Bacteriol* 189: 5495-5503.

Groschel, M. I., Prabowo, S. A., Cardona, P. J., Stanford, J. L., and van der Werf, T. S. (2014) Therapeutic vaccines for tuberculosis--a systematic review. *Vaccine* 32: 3162-3168.

Hett, E. C. and Rubin, E. J. (2008) Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev* 72: 126-56, table.

Huitric, E., Verhasselt, P., Koul, A., Andries, K., Hoffner, S., and Andersson, D. I. (2010) Rates and mechanisms of resistance development in Mycobacterium tuberculosis to a novel diarylquinoline ATP synthase inhibitor. *Antimicrob Agents Chemother* 54: 1022-1028.

Iantomasi, R., Sali, M., Cascioferro, A., Palucci, I., Zumbo, A., Soldini, S. *et al.* (2012) PE_PGRS30 is required for the full virulence of Mycobacterium tuberculosis. *Cell Microbiol* 14: 356-367.

Karsdal, M. A., Nielsen, M. J., Sand, J. M., Henriksen, K., Genovese, F., Bay-Jensen, A. C. *et al.* (2013) Extracellular matrix remodeling: the common denominator in connective tissue diseases. Possibilities for evaluation and current understanding of the matrix as more than a passive architecture, but a key player in tissue failure. *Assay Drug Dev Technol* 11: 70-92.

Kenneth Todar (2017) Text of Bacteriology-*Mycobacterium tuberculosis* and tuberculosis. [.

Koh, K. W., Lehming, N., and Seah, G. T. (2009) Degradation-resistant protein domains limit host cell processing and immune detection of mycobacteria. *Mol Immunol* 46: 1312-1318.

Lalita Ramakrishnan, N. A. F. S. F. granuloma specific expression of mycobacterium virulence proteins from the Glycin-Rich PE_PGRS family. 288, 1463-1439. 2000. USA, science. Lalita Ramakrishnan, Nancy A. Federspiel Stanley Falkow.
Ref Type: Generic

Laurenzo, D. and Mousa, S. A. (2011) Mechanisms of drug resistance in Mycobacterium tuberculosis and current status of rapid molecular diagnostic testing. *Acta Trop* 119: 5-10.

levinsin W (2008) Mycobacteria review. In *medical microbiology and immunology*. warren levinson (ed.) san francisco: Mc Graw Hill, pp. 161-168.

Luca, S. and Mihaescu, T. (2013) History of BCG Vaccine. *Maedica (Buchar)* 8: 53-58.

Majlessi, L., Prados-Rosales, R., Casadevall, A., and Brosch, R. (2015) Release of mycobacterial antigens. *Immunol Rev* 264: 25-45.

Malen, H., Pathak, S., Softeland, T., de Souza, G. A., and Wiker, H. G. (2010) Definition of novel cell envelope associated proteins in Triton X-114 extracts of Mycobacterium tuberculosis H37Rv. *BMC Microbiol* 10: 132.

Marisol Ocampo C (2015) vaccine-recent advances and clininal trails. In *tuberculosis-expanding knowledge*. Wellman Ribon (ed.) INTECH, pp. 103-115.

Mayra Silva Miranda, A. B. S. A. F. D. a. F. A. (2012) The Tuberculous Granuloma: An Unsuccessful Host Defence Mechanism Providing a Safety Shelter for the Bacteria? *journal of immunology research* 1-14.

Mohareer, K., Tundup, S., and Hasnain, S. E. (2011) Transcriptional regulation of Mycobacterium tuberculosis PE/PPE genes: a molecular switch to virulence? *J Mol Microbiol Biotechnol* 21: 97-109.

Ojha, A. K., Mukherjee, T. K., and Chatterji, D. (2000) High intracellular level of guanosine tetraphosphate in Mycobacterium smegmatis changes the morphology of the bacterium. *Infect Immun* 68: 4084-4091.

Pai, M., Behr, M. A., Dowdy, D., Dheda, K., Divangahi, M., Boehme, C. C. *et al.* (2016) Tuberculosis. *Nat Rev Dis Primers* 2: 16076.

Palomino, J. C. and Martin, A. (2014) Drug Resistance Mechanisms in Mycobacterium tuberculosis. *Antibiotics (Basel)* 3: 317-340.

Palucci, I., Camassa, S., Cascioferro, A., Sali, M., Anoosheh, S., Zumbo, A. *et al.* (2016) PE_PGRS33 Contributes to Mycobacterium tuberculosis Entry in Macrophages through Interaction with TLR2. *PLoS One* 11: e0150800.

Parish, T., Smith, D. A., Roberts, G., Betts, J., and Stoker, N. G. (2003) The senX3-regX3 two-component regulatory system of Mycobacterium tuberculosis is required for virulence. *Microbiology* 149: 1423-1435.

Pym, A. S., Brodin, P., Brosch, R., Huerre, M., and Cole, S. T. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol Microbiol* 46: 709-717.

Ramaswamy, S. and Musser, J. M. (1998) Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tuber Lung Dis* 79: 3-29.

Reva, O., Korotetskiy, I., and Ilin, A. (2015) Role of the horizontal gene exchange in evolution of pathogenic Mycobacteria. *BMC Evol Biol* 15 Suppl 1: S2.

Rifat, D., Bishai, W. R., and Karakousis, P. C. (2009) Phosphate depletion: a novel trigger for Mycobacterium tuberculosis persistence. *J Infect Dis* 200: 1126-1135.

Rodriguez, G. M., Voskuil, M. I., Gold, B., Schoolnik, G. K., and Smith, I. (2002) ideR, An essential gene in mycobacterium tuberculosis: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infect Immun* 70: 3371-3381.

Romagnoli, A., Etna, M. P., Giacomini, E., Pardini, M., Remoli, M. E., Corazzari, M. *et al.* (2012) ESX-1 dependent impairment of autophagic flux by Mycobacterium tuberculosis in human dendritic cells. *Autophagy* 8: 1357-1370.

Russell, D. G., Barry, C. E., III, and Flynn, J. L. (2010) Tuberculosis: what we don't know can, and does, hurt us. *Science* 328: 852-856.

Saini, N. K., Baena, A., Ng, T. W., Venkataswamy, M. M., Kennedy, S. C., Kunnath-Velayudhan, S. *et al.* (2016) Suppression of autophagy and antigen presentation by Mycobacterium tuberculosis PE_PGRS47. *Nat Microbiol* 1: 16133.

Sampson, S. L. (2011) Mycobacterial PE/PPE proteins at the host-pathogen interface. *Clin Dev Immunol* 2011: 497203.

Sandor, M., Weinstock, J. V., and Wynn, T. A. (2003) Granulomas in schistosome and mycobacterial infections: a model of local immune responses. *Trends Immunol* 24: 44-52.

Sani, M., Houben, E. N., Geurtsen, J., Pierson, J., de, P. K., van, Z. M. *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.

Sanyal, S., Banerjee, S. K., Banerjee, R., Mukhopadhyay, J., and Kundu, M. (2013) Polyphosphate kinase 1, a central node in the stress response network of Mycobacterium tuberculosis, connects the two-component systems MprAB and SenX3-RegX3 and the extracytoplasmic function sigma factor, sigma E. *Microbiology* 159: 2074-2086.

Schmidt, N., Mishra, A., Lai, G. H., and Wong, G. C. (2010) Arginine-rich cell-penetrating peptides. *FEBS Lett* 584: 1806-1813.

Scordo, J. M., Knoell, D. L., and Torrelles, J. B. (2016) Alveolar Epithelial Cells in Mycobacterium tuberculosis Infection: Active Players or Innocent Bystanders? *J Innate Immun* 8: 3-14.

Serafini, A., Boldrin, F., Palu, G., and Manganelli, R. (2009) Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc. *J Bacteriol* 191: 6340-6344.

Singh, P. P., Parra, M., Cadieux, N., and Brennan, M. J. (2008) A comparative study of host response to three Mycobacterium tuberculosis PE_PGRS proteins. *Microbiology* 154: 3469-3479.

Singh, R., Singh, M., Arora, G., Kumar, S., Tiwari, P., and Kidwai, S. (2013) Polyphosphate deficiency in Mycobacterium tuberculosis is associated with enhanced drug susceptibility and impaired growth in guinea pigs. *J Bacteriol* 195: 2839-2851.

Singh, V. K., Berry, L., Bernut, A., Singh, S., Carrere-Kremer, S., Viljoen, A. *et al.* (2016) A unique PE_PGRS protein inhibiting host cell cytosolic defenses and sustaining full virulence of Mycobacterium marinum in multiple hosts. *Cell Microbiol* 18: 1489-1507.

Smith, I. (2003) Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev* 16: 463-496.

Solans, L., Gonzalo-Asensio, J., Sala, C., Benjak, A., Uplekar, S., Rougemont, J. *et al.* (2014) The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in Mycobacterium tuberculosis. *PLoS Pathog* 10: e1004183.

Sorensen, A. L., Nagai, S., Houen, G., Andersen, P., and Andersen, A. B. (1995) Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis. *Infect Immun* 63: 1710-1717.

Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. (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.

Sureka, K., Dey, S., Datta, P., Singh, A. K., Dasgupta, A., Rodrigue, S. *et al.* (2007) Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria. *Mol Microbiol* 65: 261-276.

Sureka, K., Ghosh, B., Dasgupta, A., Basu, J., Kundu, M., and Bose, I. (2008) Positive feedback and noise activate the stringent response regulator rel in mycobacteria. *PLoS One* 3: e1771.

Thayil, S. M., Morrison, N., Schechter, N., Rubin, H., and Karakousis, P. C. (2011) The role of the novel exopolyphosphatase MT0516 in Mycobacterium tuberculosis drug tolerance and persistence. *PLoS One* 6: e28076.

Tischler, A. D., Leistikow, R. L., Kirksey, M. A., Voskuil, M. I., and McKinney, J. D. (2013) Mycobacterium tuberculosis requires phosphate-responsive gene regulation to resist host immunity. *Infect Immun* 81: 317-328.

Tortoli, E. (2006) The new mycobacteria: an update. *FEMS Immunol Med Microbiol* 48: 159-178.

Tufariello, J. M., Chapman, J. R., Kerantzas, C. A., Wong, K. W., Vilcheze, C., Jones, C. M. *et al.* (2016) Separable roles for Mycobacterium tuberculosis ESX-3 effectors in iron acquisition and virulence. *Proc Natl Acad Sci U S A* 113: E348-E357.

Vallecillo, A. J. and Espitia, C. (2009) Expression of Mycobacterium tuberculosis pe_pgrs33 is repressed during stationary phase and stress conditions, and its transcription is mediated by sigma factor A. *Microb Pathog* 46: 119-127.

van, C. R., Ottenhoff, T. H., and van der Meer, J. W. (2002) Innate immunity to Mycobacterium tuberculosis. *Clin Microbiol Rev* 15: 294-309.

Weiss, L. A. and Stallings, C. L. (2013) Essential roles for Mycobacterium tuberculosis Rel beyond the production of (p)ppGpp. *J Bacteriol* 195: 5629-5638.

WHO. Global Tuberculosis Report. 5-23. 2016. switzerland.
Ref Type: Report

William A.Strohl, H. R. B. D. F. (2001) Mycobacteria and Actinomycetes. In *Microbiology*. Richard A.Harvey, P. A. c. (ed.) Lippincot William and Wilkins, pp. 245-258.

Zhang, Y., Garbe, T., and Young, D. (1993) Transformation with katG restores isoniazid-sensitivity in Mycobacterium tuberculosis isolates resistant to a range of drug concentrations. *Mol Microbiol* 8: 521-524.

Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffe, M. (2008) Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* 190: 5672-5680.

Zumbo, A., Palucci, I., Cascioferro, A., Sali, M., Ventura, M., D'Alfonso, P. *et al.* (2013) Functional dissection of protein domains involved in the immunomodulatory properties of PE_PGRS33 of Mycobacterium tuberculosis. *Pathog Dis* 69: 232-239.

Acknowledgments

Universita` di Sassari

Prof. Leonardo Sechi

Prof. Salvatore Rubino

Universita` Cattolica del Sacro Cuore, Roma

Dott. Flavio De Maio

Dott.ssa Ivana Palucci

Dott.ssa Serena Camassa

Dott.ssa Maria Chiara Minerva

Prof.ssa Michela Sali

Prof. Maurizio Sanguinetti

Prof. Giovanni Delogu