HAL
open science

MmpS4 promotes glycopeptidolipids biosynthesis and export in Mycobacterium smegmatis

Caroline Deshayes, Horacio Bach, Daniel Euphrasie, Rodgoun Attarian, Mathieu Coureuil, Wladimir P Sougakoff, Françoise Laval, Yossef Av-Gay, Mamadou Daffe, Gilles Etienne, et al.

To cite this version:

Caroline Deshayes, Horacio Bach, Daniel Euphrasie, Rodgoun Attarian, Mathieu Coureuil, et al.. MmpS4 promotes glycopeptidolipids biosynthesis and export in Mycobacterium smegmatis. Molecular Microbiology, 2010, 78 (4), pp.989. 10.1111/j.1365-2958.2010.07385.x . hal-00580703

HAL Id: hal-00580703

https://hal.science/hal-00580703v1

Submitted on 29 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

molecular microbiology

MmpS4 promotes glycopeptidolipids biosynthesis and export in Mycobacterium smegmatis

Journal:	*Molecular Microbiology*
Manuscript ID:	MMI-2010-09792.R2
Manuscript Type:	Research Article
Date Submitted by the Author:	30-Aug-2010
Complete List of Authors:	Deshayes, Caroline; University Paris-Descartes, School of Medecine, Inserm U570 Bach, Horacio; University of British Columbia, Medicine Euphrasie, Daniel; INSERM, U570 Attarian, Rodgoun; University of British Columbia, Medicine Coureuil, Mathieu; University Paris-Descartes, School of Medecine, Inserm U570 Sougakoff, Wladimir; University Pierre et Marie Curie, Inserm UMRS-872 Laval, Françoise; CNRS, IPBS Av-Gay, Yossef; University of British Columbia, Medicine Daffe, Mamadou; CNRS-University Paul Sabatier Mixed Research Lab (UMR 5089), Department of Molecular Mechanisms of Mycobacterial Infections Etienne, Gilles; CNRS-University Paul Sabatier Mixed Research Lab (UMR 5089), Department of Molecular Mechanisms of Mycobacterial Infections Reyrat, Jean-Marc; University Paris-Descartes, School of Medecine, Inserm U570
Key Words:	mycobacteria, glycopeptidolipids, cell wall, membrane protein

SCHOLARONE™
Manuscripts

MmpS4 promotes glycopeptidolipids biosynthesis and export in *Mycobacterium smegmatis*

Caroline Deshayes[1, 2, 3,*], Horacio Bach[4], Daniel Euphrasie[2], Rodgoun Attarian[4], Mathieu Coureuil[1, 2], Wladimir Sougakoff[5], Françoise Laval[6, 7], Yossef Av-Gay[4], Mamadou Daffé[6, 7], Gilles Etienne[6, 7] and Jean-Marc Reyrat[1, 2, §]

[1] *INSERM-UMR 570, Unité de Pathogénie des Infections Systémiques, Paris Cedex 15, F-75730, France*
[2] *Université Paris Descartes, Faculté de Médecine, site Necker, Paris Cedex 15, F-75730, France*
[3] *Groupe d'Etude des Interactions Hôte-Pathogène, Université d'Angers, 4 rue Larrey, Angers, France*
[4] *Department of Medicine, Division of Infectious Diseases, University of British Columbia, 2733 Heather St., Vancouver, V5Z 3J5 British Columbia, Canada*
[5] *INSERM UMRS-872, équipe 12, Laboratoire de Recherche Moléculaire sur les Antibiotiques (LRMA), faculté de Médecine Pierre et Marie Curie, site Pitié-Salpêtrière, 91 bd de l'Hôpital, F-75634 Paris cedex 13, France*
[6] *CNRS, IPBS (Institut de Pharmacologie et de Biologie Structurale), Département Mécanismes Moléculaires des Infections Mycobactériennes. 205 route de Narbonne, F-31077 Toulouse, France*
[7] *Université de Toulouse, Université Paul Sabatier (Toulouse III), IPBS. F-31077 Toulouse, France*

***: corresponding author.** Dr Deshayes, C. Groupe d'Etude des Interactions Hôte-Pathogène, Université d'Angers, 4 rue Larrey, Angers, France. *deshayes@necker.fr*

§: this paper is dedicated to Dr Jean-Marc Reyrat, who deceased before the submission of the manuscript, for his excellent contribution to the mycobacterial genetics.

Key words: mycobacteria, glycopeptidolipids, cell wall, membrane protein

For Peer Review

Summary. The MmpS family (mycobacterial membrane protein small) includes over one hundred small membrane proteins specific to the genus *Mycobacterium* that have not yet been studied experimentally. The genes encoding MmpS proteins are often associated with *mmpL* genes, which are homologous to the RND (resistance nodulation cell division) genes of Gram-negative bacteria that encode proteins functioning as multidrug efflux system. We showed by molecular genetics and biochemical analysis that MmpS4 in *M. smegmatis* is required for the production and export of large amounts of cell surface glycolipids, but is dispensable for biosynthesis *per se*. A new specific and sensitive method utilizing single chain antibodies against the surface-exposed glycolipids was developed to confirm that MmpS4 was dispensable for transport to the surface. Orthologous complementation demonstrated that the MmpS4 proteins are exchangeable, thus not specific to a defined lipid species. MmpS4 function requires the formation of a protein complex at the pole of the bacillus, which requires the extracytosolic C-terminal domain of MmpS4. We suggest that MmpS proteins facilitate lipid biosynthesis by acting as a scaffold for coupled biosynthesis and transport machinery.

For Peer Review

Introduction

The number of complete genome sequences of high GC % Gram-positive bacteria available has increased considerably in the recent years, leading to the accumulation of unprecedented amounts of information. For example, 16 mycobacterial genomes have been sequenced to date and the species sequenced include both rapid and slow growers (Bernal *et al.*, 2001). *Mycobacterium* is a large genus containing over 150 species, including *M. tuberculosis*, the etiologic agent of tuberculosis (WHO), *M. ulcerans*, which causes a devastating necrotic disease of the skin, the so-called Buruli ulcer (WHO), and many opportunistic pathogens (De Groote & Huitt, 2006). Recent sequencing efforts have, through comparative and functional genomics, led to advancements in the characterization of the mycobacterial core genome and of genes involved in virulence. For example, genes encoding a new type of secretion system were characterized, the type VII (Abdallah *et al.*, 2007), in addition to genes required for the biosynthetic of arabino-galactan, a specific attribute restricted to this genus (Crick & Brennan, 2008) and related genera. Sequencing of the mycobacterial genome has also led to the identification of the MmpSL family of membrane proteins, the function of which remains unclear. The name of this protein family is short for mycobacterial membrane protein large (MmpL) and small (MmpS).

The genes encoding MmpS are often found closely associated with *mmpL* genes and nothing is currently known about the function of these proteins. MmpL proteins consist of approximately 1000 amino acids organized into twelve predicted transmembrane domains and two large periplasmic domains. MmpL are not restricted to mycobacteria, and are also found in related streptomycetes and *Rhodococcus*. This family of proteins is distantly related to the resistance nodulation cell division (RND) superfamily (Tekaia *et al.*, 1999), members of which are present in all domains of life and may serve as efflux pumps for toxic environmental molecules (Tseng *et al.*, 1999). Two of the fourteen MmpL members in *M. tuberculosis* have been shown to be involved in the biosynthesis and/or transport of mycobacterial lipid metabolites. MmpL7 is required together with an ABC transporter and a lipoprotein, for the specific transport of phthiocerol dimycoserate (PDIM) to the cell surface (Camacho *et al.*, 2001, Cox *et al.*, 1999). MmpL8 plays a role in the synthesis of sulfolipid-1 (SL-1), another cell surface polyketide, possibly in the transport of a SL-1 precursor from the cytoplasm to the

For Peer Review

periplasm (Converse *et al.*, 2003, Domenech *et al.*, 2004). In *M. smegmatis* and *M. abscessus*, the MmpL4a and MmpL4b proteins (previously known as TmtpB and TmtpC, respectively) are implicated in the biosynthesis of cell surface polyketides, the glycopeptidolipids (GPLs) (Recht *et al.*, 2000, Sonden *et al.*, 2005, Medjahed & Reyrat, 2009).

As previously mentioned, some *mmpL* genes are associated with *mmpS* genes, which encode proteins predicted to have only one N-terminal transmembrane domain with an extracytoplasmic C-terminus (Domenech *et al.*, 2005). Two (*M. leprae*) to twenty-seven (*M. abscessus*) such genes may be present in the genome, depending on the species. Most *mmpL* and *mmpS* genes are located close to genes involved in the synthesis or modification of polyketides. This close proximity suggests a possible role in the transport and/or biosynthesis of molecules synthesized by the neighbouring synthase gene products (Tekaia *et al.,* 1999). Little is known about MmpS function, but transcriptional regulation of these genes has been observed. For example, the *mmpS4* and *mmpS5* genes of *M. tuberculosis* are down regulated during nutrient starvation (Betts *et al.*, 2002). More recently, resistance to azole in a *M. tuberculosis* mutant strain was correlated with increased transcription of *mmpS5-mmpL5* genes (Milano *et al.*, 2009).

In *M. smegmatis*, *mmpS4* is organized into a putative operon with the *mmpL4a* and *mmpL4b* genes, which have been shown to be involved in GPL biosynthesis (Recht et al., 2000, Sonden et al., 2005). GPLs are the predominant glycolipids found at the surface of many non-tuberculous mycobacteria, including *M. avium* subsp. *avium* (Brennan *et al.*, 1981) and *M. abscessus* (Lopez-Marin *et al.*, 1994), an emerging pathogen predominantly infecting young cystic fibrosis patients. These surface polyketides are required for sliding motility and biofilm formation, and in some cases their level of production is correlated with strain virulence (Byrd & Lyons, 1999, Howard *et al.*, 2006). Upon engineering an in-frame unmarked deletion of the *mmpS4* gene in *M. smegmatis*, reduced levels of GPLs were observed. However, we found that MmpS4 is not required for GPLs localization. This finding was confirmed with a new immunofluorescent method for specific detection of GPLs at the cell surface. Our results suggest that MmpS are not specific for structurally defined lipid molecules. By studying fusions of red fluorescent protein (Rfp) with MmpS4 or two GPL biosynthesis enzymes, we show that MmpS4 is required for the formation of a protein complex at the pole of the

For Peer Review

bacillus, presumably *via* the C-terminal domain. Thus, we suggest that MmpS proteins promote efficient biosynthesis and secretion by acting as a scaffold for the enzymatic and export machinery. This report constitutes the first experimental investigation of a member of the MmpSL family of proteins.

For Peer Review

Results

The *mmpS* family in mycobacteria. *mmpS* genes are specific to mycobacteria and are most often associated with *mmpL* genes. Indeed, our homology searches in genomes of closely related members of the *Actinomycetales*, such as *Streptomyces* and *Rhodococcus*, show that these bacteria contain *mmpL* homologues, but no *mmpS* genes (Table 1). The MmpS family is currently composed of more than one hundred proteins and new members are added with the release of each new mycobacterial genome sequence. The number of *mmpS* paralogs per genome varies from two to twenty seven and is correlated with the number of *mmpL* genes, rather than genome size (Table 1). These members can be organized into five classes, based on the clusters obtained by phylogenetic analysis, corresponding to the *M. tuberculosis mmpS* paralog-numbering (Fig. 1A).

MmpS proteins have an overall identity of 22 to 81 % within the same class and 8 to 37 % between members of different classes. Each *mmpS* gene encodes a predicted protein of approximately 150 amino acids in length, anchored in the membrane by a 15-amino-acid-transmembrane domain at its N-terminus. It has been suggested that the C-terminus of the protein is extracytosolic (Tekaia et al., 1999, Domenech et al., 2005) and this hypothesis has been supported by recent computer-assisted predicting of membrane protein topology (Fig. 1B). An alkaline phosphatase gene fusion *mmpS4-phoA* was used to confirm this predicted extracytosolic localization. The activity of bacterial alkaline phosphatase (PhoA) is indeed dependent on it being exported across the plasma membrane. *M. smegmatis* expressing the *mmpS4-phoA* fusion showed blue phenotype on plates containing X-P, the chromogenic substrate of PhoA, reflecting an activity of the alkaline phosphatase (Fig. 1C). In contrast, *M. smegmatis* carrying the empty pJEM11 plasmid was white, clearly confirming the bioinformatical prediction and indicating that the C-terminal domain of MmpS4 is extracytoplasmic.

The *mmpS3* genes differ from other *mmpS* genes in encoding predicted proteins of approximately 300 amino acids in size. These proteins are thus twice the size of other MmpSs proteins and likely arose from an ancient complete *mmpS3* copy through an event of duplication and fusion. *mmpS3* is present in both rapid and slow growing mycobacterial species, suggesting that the duplication event occurred in the common ancestor of these two phylogenetic clades. *M. smegmatis* has ten *mmpS* genes (Table 1), and is specifically enriched in paralogs of the *mmpS5* families (Table S1). By contrast, the *mmpS4*

For Peer Review

gene is present as only a single copy in all the sequenced mycobacterial genomes, with the exception of *M. abscessus* genome, which includes five paralogs.

MmpS4 is required for sliding motility and biofilm formation. We investigated the function of MmpS4 by generating a mutant strain for this gene in *M. smegmatis*. The *mmpS4*, *mmpL4a* and *mmpL4b* genes are probably transcribed as an operon, as these genes either overlap or are separated by only few nucleotides. We prevented possible polar effects on downstream genes by constructing an in frame unmarked deletion of the *mmpS4* gene by homologous recombination, using a two steps strategy and a counter-selectable marker (Pelicic *et al.*, 1996, Reyrat *et al.*, 1998). The mutant strain was verified by PCR, sequencing and Southern-blotting, all of which confirmed its identity and the occurrence of the anticipated event, a deletion of about 90% of the coding sequence of the *mmpS4* gene (Fig. S1). The resulting Δ*mmpS4* mutant had rough colony morphology readily visible by the naked eye. This mutant was unable to slide (Fig. 2A) or form biofilms (data not shown). Complementation of this mutant strain with the wild-type *mmpS4* allele from *M. smegmatis* restores the wild-type smooth phenotype and the ability to slide (Fig. 2A) and form biofilms. This excluded the possibility that the observed phenotype was due to secondary mutations or a polar effect on downstream genes.

In *M. smegmatis*, changes in sliding motility and biofilm formation are often associated with quantitative or qualitative changes in GPL biosynthesis, due to the lack of biochemical modifications by acetylation, methylation or glycosylation of the GPL backbone for instance (Deshayes *et al.*, 2005, Recht & Kolter, 2001, Sonden *et al.*, 2005). We therefore analyzed GPL production in the mutant strain. As shown in Figure 2B, the Δ*mmpS4* strain produced small amounts of GPL-like compounds. Only the major species remained detectable by TLC, whereas the complemented strain produced GPL compounds in amounts similar of those of the wild-type. For the precise quantification of GPL production in the Δ*mmpS4* strain, we grew large amounts of cells and purified GPLs, making use of their alkali-stable character (Belisle *et al.*, 1993). In these conditions the *mmpS4* mutant produced about 20% as much alkali-stable lipids as the wild-type and complemented strains did, accounting for

For Peer Review

the rough morphotype of the mutant strain (Fig. 2C). Complementation of the mutant strain with the *mmpS4* wild-type allele of *M. smegmatis* restored the wild-type level of GPL production (Fig. 2B and 2C). The levels of other lipids (triacyl glycerols, glycerol monomycolate, trehalose dimycolate, trehalose monomycolate, phosphatidylinositol mannosides) were similar (Fig. S2). Thus MmpS4 plays a specific role in controlling the level of GPLs biosynthesis in *M. smegmatis*.

MmpS4 is required for high level of GPL production. We checked for changes in GPL structure, by analyzing lipid extracts of the Δ*mmpS4* strain using MALDI-TOF mass spectrometry, a highly sensitive and accurate method of mycobacterial complex glycolipids detection and characterization (Villeneuve *et al.*, 2003). Total lipid extract prepared from the wild-type strain displayed a major peak at 1257 *m/z*, corresponding to the main diglycosylated (tri-*O*-Methyl Rhamnose and di-*O*-Acetylated 6-deoxyTalose) GPL of *M. smegmatis* mc²155 (Etienne *et al.*, 2005). A similar pattern of GPLs was also detected in the mutant strain, but in relatively smaller amounts and with the peaks 14 Da higher than those of the wild-type (Fig. S3). The major peak was observed at 1271 *m/z*. This 14 atomic mass-units increase is likely due to an additive *O*-methylation on the fatty acid moiety, as the only spot detectable by TLC (Fig. 2B) was the most apolar fully *O*-methylated GPL I (Patterson *et al.*, 2000). Complementation of the Δ*mmpS4* strain fully restored the wild-type profile with a major peak at 1257 *m/z*, demonstrating that this 14 atomic mass-units increase was dependent on the absence of the *mmpS4* gene.

MmpS4 is not involved in GPL subcellular localization. We previously showed that sliding motility requires not only the normal level of GPL production but also a localization of these molecules on the cell surface (Sonden et al., 2005). As the Δ*mmpS4* strain is not motile, we hypothesized that MmpS4 might be involved in GPL export to the cell surface. We tested this hypothesis, by preparing surface and cell-associated (after removal of the surface lipids with glass beads) fractions, extracting lipids from each fraction and analyzing them by TLC (Fig. 3A). As the Δ*mmpS4* strain produced less than 20% the amount of GPL produced by the wild-type strain, we used larger amounts of cells and

For Peer Review

concentrated lipid metabolites through solvent evaporation to visualize them by TLC. A Δ*gap* mutant strain was used as a control because this mutant produces wild-type levels of GPLs but does not export them to the surface (Sonden *et al.*, 2005). Consistent with quantification results, the Δ*mmpS4* strain produced much less GPL than did the wild-type and Δ*gap* strains, but small amounts of GPL were still present in the surface fraction of the mutant strain. Despite the low level of production of GPL by the mutant, the glycolipid could be visualized in the surface compartment, although barely, for the Δ*mmpS4* mutant but not for Δ*gap* strain, demonstrating the presence of GPL at the surface of the Δ*mmpS4* mutant (Fig. 3A). An overloaded TLC showed selective disappearance in the cellular extract of the lower spots, corresponding to polar GPLs in the Δ*mmpS4* mutant. Mass spectrometry analyses unambiguously confirmed the presence of GPLs in the cell surface of the Δ*mmpS4* strain while GPLs were not detectable on the cell surface of the Δ*gap* mutant (Fig. S3).

Development of a new sensitive immunofluorescent method for lipid localization. To further investigate the *in vivo* localization of GPLs, we developed a new and specific method. GPL localization in the Δ*mmpS4* mutant was characterized by an immunofluorescent method that selectively detected surface-exposed GPLs. Briefly, purified GPLs coupled to ovalbumin were used as antigens to screen a naïve human-derived antibody library using phage display technology. After four biopanning cycles and screening against ovalbumin, four clones with a high affinity for GPLs were selected. The gene encoding the antibody with the highest affinity was fused to the gene encoding the DSred fluorescent protein. The construct was expressed in *E. coli*, and the fused protein fully retained the fluorescent properties of the DSred as observed by epifluorescence microscopy (Figure 3B). The detection of GPLs was performed in a single step as secondary antibodies were not needed. Bacteria were first non-specifically labelled with FITC. The resultant green staining made it possible to visualize the bacterial bodies and served as a control for evaluating of the number of bacteria present on the slide. Two control strains were used in this experiment: a *pks* mutant strain producing no GPLs (pks^-) (Sonden *et al.*, 2005) used to evaluate labelling specificity, and the Δ*gap* mutant strain with its complemented companion strain (Sonden *et al.*, 2005). The DSred-antibody labelling perfectly co-

For Peer Review

localized with the FITC-labelled bacteria in the case of the wild-type and the complemented Δ*gap* strain, resulting in a yellow staining after image overlapping (Fig. 3B). The strain producing no GPL and the Δ*gap* strain were not immunolabelled with the anti-GPL antibody. This antibody was therefore highly specific for GPL and able to bind surface-exposed GPLs only, as the Δ*gap* strain, which produces wild-type level but does not export them on the surface, was not labelled by the antibody. Interestingly, a weak anti-GPL fluorescence signal was detected in the Δ*mmpS4* strain, corresponding to approximately 60% of the relative fluorescence compared to the wild-type strain. A clear signal was observed for the Δ*mmpS4* strain but not for the GPL- and Δ*gap* strains. These results confirm that the Δ*mmpS4* strain produces smaller amounts of GPL than the wild-type strain but also demonstrate that GPLs are present on the bacterial surface. The Δ*mmpS4* complemented strain gave a stronger signal, although not as strong as that of the wild-type strain, in approximately 95% of the bacteria. These experiments demonstrate that GPLs reach the cell surface in the Δ*mmpS4* strain but also show that that reintroduction of the *mmpS4* gene increases both the total and surface levels of GPLs.

MmpS4 proteins assemble into a complex at the pole of the bacillus. We investigated the function and location of MmpS4 by constructing gene fusions based on the red fluorescent protein (RFP). The *rfp* gene was fused to the 5' or the 3' end of the *mmpS4* gene, yielding RFP-MmpS4 and MmpS4-RFP proteins, respectively. The fusion of RFP to the N-terminal domain did not affect the function of MmpS4. Indeed, the construct encoding RFP-MmpS4 fully complemented the □*mmpS4* mutant strain, resulting in the production of large amounts of GPLs (data not shown). In contrast, the fusion of RFP to the C-terminal domain of MmpS4 abolished its function, as this construct did not complement the □*mmpS4* mutant strain. This result suggests that the extracytosolic C-terminal domain of MmpS4 is crucial for its function and therefore does not tolerate fusion with RFP. Nevertheless, it is still possible that the formation of a fused protein may lead to misfolding of MmpS4 and this in turn affects its topology and membrane localization. To dissect the role of the MmpS4 extracytosolic C-terminal domain, truncated proteins were constructed. None of them were able to complement the Δ*mmpS4* strain (Fig. S4), strengthening our hypothesis that the C-terminal domain of MmpS4 is necessary for

For Peer Review

its function. Furthermore, this domain needs to be localized in the extracytosolic compartment. Indeed the intracellular expression of this domain could not restore the *mmpS4* function, since the MmpS4 protein deleted of its transmembrane region was not able to complement the mutant strain.

The strains expressing fusion proteins were observed by fluorescence microscopy and a control strain expressing cytosolic GFP was used (Barker *et al.*, 1998). In strains expressing MmpS4-RFP or the cytosolic GFP, the fluorescence was diffused and spread throughout the bacillus. In contrast, in the strain carrying the RFP-MmpS4 construct, the fluorescence was concentrated into small dot-like area (spot). This punctuate fluorescence was absent from strains expressing the MmpS4-RFP fusion and the cytosolic GFP, but present in approximately 25% of the bacilli expressing the RFP-MmpS4 fusion (Fig. 4A). There was never more than one spot per bacillus and the spots were located either at the site of division or at the pole of the cell (Fig. 4B). The presence of this complex at the division site may enable each of the two daughter cells to inherit half the complex, therefore becoming located at the pole after division, or can be correlated to the presence of the division machinery. Although we can not completely exclude that this apparent MmpS4 localization was an artefact of the RFP fusion, this seems unlikely since a polar or septum localization is rather expected for a complex involved in the synthesis of an envelope constituent. In conclusion, these findings strongly suggest that MmpS4 proteins are not randomly distributed in the membrane but instead assemble into a complex located at the pole of the bacterium. A free C-terminal extracytosolic domain is required for both complex formation and biological function.

MmpS4 proteins are necessary for the formation of a complex of GPL biosynthesis enzymes at the pole of the bacillus. Based on the data presented herein and previous studies reporting interaction between MmpL and enzymes required for the biosynthesis of small metabolites (Cox et al., 1999, Straight *et al.*, 2007), we suggest that MmpS4 acts as a scaffold for the assembly of the GPL biosynthesis megacomplex (Fig. 5A). Proteins fusions were constructed, fusing Rfp in the C-terminal domain of two GPL synthesis proteins (MbtH and FadD23) predicted to be cytosolic. Their localization was analyzed in the wild-type and Δ*mmpS4* mutant strains. Punctuate fluorescence was found at the poles only in the wild-type strain in roughly 25% and 35% of the bacilli for the MbtH-Rfp

For Peer Review

and FadD23-RFP constructs, respectively (Fig. 5B and C). These percentages were similar to the percentage obtained for the strain expressing the *rfp-mmpS4* fusion. In contrast, the fluorescence was mainly diffuse in the Δ*mmpS4* mutant strain. These results show that the absence of MmpS4 proteins affects the localization of other proteins involved in GPL synthesis and reinforce our megacomplex model.

MmpS4s are functionally exchangeable. MmpS4 orthologs are 50 to 60% identical and many of the conserved amino acids are located in the predicted extracytosolic region (Fig. S5). We assessed the activity of these MmpS4 orthologs, by transforming the Δ*mmpS4* strain with a plasmid carrying an ortholog of the *mmpS4* gene originating from other species (genes from several different species were used). For *M. abscessus*, which carries five paralogs of *mmpS4,* we used the copy (MAB4117c) most similar to *mmpS4* of *M. smegmatis*. This gene mapped to the GPL locus of *M. abcessus* (Ripoll *et al.*, 2007). We used orthologs from GPL-producing species such as *M. abscessus* and *M. avium* subsp. *avium* (MAV3247), but also from non-GPL-producing organisms such as *M. tuberculosis* (Rv0451c). In all cases (MAB4117c, MAV3247, and Rv0451c) complementation with the *mmpS4* gene restored a smooth wild-type phenotype. The mutant strain complemented with the *M. tuberculosis* ortholog exhibited wild-type sliding motility, consistent with functional complementation (Fig. 6A). As a control experiment we carried out complementation experiments with the *mmpS1* gene of *M. tuberculosis* (*mmpS1*Tb). Expression of this *mmpS1*Tb gene did neither restore the wild-type smooth phenotype, nor the sliding motility, consistent with the notion of orthology for the MmpS4 family (Fig. 6A). To avoid any possibility of non-expression of the *mmpS1* gene, its expression in the Δ*mmpS4/mmpS1*Tb strain was confirmed by reverse transcription (Fig. S6). Biochemical analyses of the Δ*mmpS4* mutant strain expressing the *mmpS4* gene of *M. tuberculosis* (*mmpS4*Tb) showed fully functional complementation. The *mmpS4* gene of *M. tuberculosis* therefore conferred to the mutant strain, upon complementation, the ability to produce more GPL, both in the cellular compartment and at the surface (Fig. 6B). Quantitative analyses showed that the *mmpS4*Tb-complemented strain produced significantly larger amounts of GPLs (2.09 ± 0.72 mg of alkali-stable lipids for 100 mg of

For Peer Review

cells, dry weight) than did the Δ*mmpS4* mutant strain (0.54 ± 0.33 mg). Furthermore, immunofluorescence experiments with the anti-GPL DSred antibody showed much more GPLs on the surface of the *mmpS4*Tb complemented strain (Fig. 6C) than on the Δ*mmpS4* mutant strain. Indeed, the *mmpS4*Tb complemented strain produced a stronger signal in approximately 95% of the bacilli. The signal obtained was similar to that for the *mmpS4*Ms complemented strain (Fig. 3B). As stated above, this complemented strain displayed greater sliding motility compared to the mutant, indicating a gain of function (Fig. 6A). In conclusion, these findings indicate that the *mmpS4* gene of a species that does not produce GPL can complement the *M. smegmatis* Δ*mmpS4* mutant strain, both in terms of production and localization of GPLs. Thus, MmpS4 are functionally exchangeable and constitute a true protein family. Our findings also suggest that MmpS proteins are not specific to the lipid compound itself.

For Peer Review

Discussion

The synthesis and export of complex polyketides present in the outermost structure of the cell envelope such as GPLs require the expression of large biosynthetic pathways, involving proteins encoded by grouped genes or genes dispersed throughout the genome, and the assembly of dedicated machineries (Gokhale *et al.*, 2007). The *mmpS* family was first identified a decade ago with the sequencing of the first mycobacterial genome (Cole *et al.*, 1998). These genes encode proteins involved in various specific pathways and are often associated with *mmpL* genes. *M. tuberculosis* has five *mmpS* genes and fourteen predicted *mmpL* genes. Consequently, not all biosynthetic pathways include proteins encoded by *mmpS* gene. This seems to be the case for PDIM and SL-1 loci for example (Camacho et al., 2001, Converse et al., 2003). The location of the *mmpS4* gene within the GPL biosynthesis locus suggests a role in GPL biosynthesis and/or transport to the surface, but our data do not support this hypothesis. Indeed, the Δ*mmpS4* mutant strain produced smaller amounts of GPLs than did the wild-type, but these compounds were still targeted to the bacterial surface as shown by biochemical fractionation and the new immunofluorescence sensitive technique for the specific tracking of GPLs at the bacterial surface. The low level of production of GPL in the Δ*mmpS4* strain accounts for its phenotype, lack of sliding motility and inability to form biofilms, as all these features are dependant on GPL production (Recht et al., 2000, Kocincova *et al.*, 2008). However, although produced in smaller amounts, the GPLs present in the mutant were chemically identical to those in the wild-type, excluding a possible role of MmpS4 in the synthesis, *per se*. The MmpS4 protein of *M. smegmatis* therefore seems to determine the amount of GPL produced. An alternative explanation would be that there was another protein, maybe another MmpS protein that provides a similar function in *M. smegmatis*, albeit not as efficient as MmpS4. The fact that functional complementation was observed when an orthologous gene (*mmpS4Tb*) from the same family was introduced into the Δ*mmpS4* strain, but not with a paralogous gene (*mmpS1Tb*), makes this last hypothesis unlikely, since only one *mmpS4* gene is present in the *M. smegmatis* genome. The functional complementation with an ortholog from *M. tuberculosis* was particularly surprising, as this species does not produce GPL. In *M. tuberculosis*, *mmpS4* is organized into an apparent operon with a single *mmpL4* gene and this

For Peer Review

operon is located close to genes encoding a probable cyclopropane fatty acyl phospholipid synthase and a putative enoyl-CoA hydratase. The lipid metabolite specified by this pathway has not yet been characterized, but a *M. tuberculosis mmpL4* mutant has been shown to be strongly attenuated *in vivo* (Domenech et al., 2005). One possible explanation for this orthologous complementation is that MmpS proteins are not specific for their lipid substrate but are, instead, specific for their cognate MmpL proteins, via molecular interactions involving amino acids conserved throughout the family (Fig. S5).

The MmpL members are distantly related to the RND superfamily, but differ from members of this superfamily in several ways. For example, RND proteins have broad substrate specificity, whereas MmpL proteins seem to be much more specific, interacting with only one lipid metabolite. MmpL7 and MmpL8 affect specifically PDIM and SL-1 biosynthesis in *M. tuberculosis,* respectively (Cox *et al.,* 1999, Converse *et al.*, 2003), and MmpL4 affects GPL biosynthesis in *M. smegmatis* (Recht *et al.,* 2000). Furthermore, whereas RND family members from Gram-negative bacteria act purely as efflux pumps, the RND protein ActII-orf3 from the Gram-positive *Streptomyces coelicolor*, is involved in the biosynthesis of a blue pigment, γ-actinorhodin, a heterocyclic antibiotic synthesized by a polyketide synthase (Bystrykh *et al.*, 1996). In *M. tuberculosis,* MmpL8 plays a role in the synthesis of sulfolipid-1 (SL-1), possibly by transporting a precursor of this molecule to the surface (Converse *et al.,* 2003, Domenech *et al.*, 2004). In the same species, MmpL7 has been shown to be required for the transport of PDIM to the cell surface (Cox *et al.*, 1999, Camacho *et al.*, 2001). This transport process requires another protein, DrrC, a component of the putative ABC transporter, DrrABC (Camacho *et al*., 2001), and secretion into the medium requires LppX, a lipoprotein (Sulzenbacher *et al.*, 2006). Studies based on the two hybrid approach have shown that MmpL7 interacts with PpsE (Jain & Cox, 2005), which interacts with TesA (Rao & Ranganathan, 2004), both of them being enzymes required for PDIM synthesis. It has thus been suggested that MmpL7 interacts with the PDIM synthetic machinery to form a complex that coordinately synthesizes and transports PDIM across the cell membrane (Jain & Cox, 2005). In *Bacillus subtilis*, a polyketide synthase synthesizing an antibiotic, bacillaene, has been shown to assemble into a single megacomplex at the membrane (Straight *et al.*, 2007). The authors suggested that this organization is established *via* interaction with an unknown membrane protein.

For Peer Review

Based on such models, we suggest that MmpS4 proteins act as a scaffold with MmpL4 proteins for the assembly of the GPL biosynthesis megacomplex (Fig. 5A). Several enzymes needed for GPL synthesis/export are predicted to be membrane-bound or -inserted: Atf, a sugar acetyltransferase, Gap, needed for GPL surface localization, MmpL4a and MmpL4b, whose precise role is currently unknown. It is likely that these membrane proteins associate into a complex to maximize GPL biosynthesis and export. In this model, MmpS4 interacts with MmpL4a, MmpL4b and GPL biosynthetic enzymes such as the polyketide synthase (Pks, involved in the lipid moiety synthesis), non-ribosomal protein synthetases (Nrp, involved in the tetrapeptide moiety synthesis), acetyltransferase (Atf), methyltransferases (Mtf) or glycosyltransferases (Gtf), thereby increasing the synthesis and export of GPLs. Our data are consistent with the C-terminal extracytosolic domain of MmpS4 being essential for the function of this protein. The extracytosolic domain of MmpS4 could specifically interact with the extracytosolic domains of MmpL4 proteins, explaining the lack of specificity of MmpS4 towards the lipid moiety. MmpS4 is not randomly distributed throughout the membrane. Instead, it is found principally at the pole of the bacterium. This suggests that MmpS4 assembles into a complex that interacts with the biosynthetic machinery. In the absence of MmpS, the biosynthetic complex does not adopt the correct configuration and the enzymes are distributed throughout the cytoplasm (Fig. 5A), resulting in lower levels of GPL production than the wild-type strain. The absence of the scaffold would lead to inappropriate contacts with enzymes of the pathway, leading to the artefactual detection of GPL_{1271} as the main product. To strengthen our megacomplex model, we constructed fluorescent fusion proteins with two GPL biosynthesis enzymes, MbtH and FadD23. The role of the *mbtH* genes, whose homologs are found associated with *nrp* genes, is unclear to date but it has been shown that the *mbtH* gene is required for the clorobiocin biosyntheis in *Streptomyces coelicolor* (Wolpert *et al.*, 2007). The *FadD23* gene is a close homolog to the *M. tuberculosis fadD28* gene involved in acyl transfer of mycocerosic acid (Cox et al., 1999) and is probably involved in the lipid attachment to the GPL tetrapeptide moiety. Our results showed that MmpS4 is required for them to localize as a punctuate spot at the poles of the bacilli. Colocalization studies should be carried out to prove interactions between GPL biosynthesis/export enzymes. The megacomplex model is also supported by the finding that some enzymes predicted to be cytosolic and

For Peer Review

involved in PDIM (PapA5, Mas, FadD28) or mycolactone (MlsA1, MlsB) biosynthesis are found in the membrane fraction of *M. ulcerans* (Tafelmeyer *et al.*, 2008). It can be speculated that the localization of these proteins is due to their anchoring to a biosynthetic membrane megacomplex.

The presence of several paralogs, as in *M. abscessus*, is consistent with the diversification of lipid metabolism. The minimalist species *M. leprae* has only two *mmpS* members (*mmpS3* and *mmpS4*), suggesting that other pathways are not required for a strict intracellular lifestyle. The new sensitive immunofluorescent method developed for this study is clearly transposable to other lipid molecules. The method is therefore of considerable interest, because many lipid molecules are poorly immunogenic, making antibody-based detection difficult. This new method allows the sensitive and specific tracking of a single lipid only and is clearly of potential use for biochemical purposes and for investigating lipid trafficking during host cell infection. It is also potentially useful for diagnostic purposes, as some lipids are species- or even strain-specific.

In conclusion, according to our current working model of MmpS function, these small transmembrane proteins are not absolutely required for metabolite synthesis and export, but may promote interactions between the various partners involved in the pathway, possibly *via* their extracytosolic C-terminal domains, therefore stabilizing the complex and optimizing synthesis and export. MmpS proteins display no similarity to membrane fusion proteins of Gram-negative bacteria. However, the close association of *mmpS* and *mmpL* genes suggests that MmpS proteins may be the functional homologs of MFPs. Further biophysical studies are required to analyze the assembly and interactions required to stabilize this complex machinery consisting of a dozen enzymes.

For Peer Review

Experimental procedures

Bioinformatical analysis. The NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/) was searched with the BLASTP program, using *M. tuberculosis* MmpS proteins as queries. The sequences of MmpS proteins from the *M. tuberculosis* H37Rv, *M. bovis* AF2122/97, *M. ulcerans* Agy99 and *M. leprae* TN genomes were inspected at the Pasteur GenoList website (http://genolist.pasteur.fr/). The sequences of the MmpS proteins of *M. abscessus*, *M. avium subsp. avium* 104, *M. avium subsp. paratuberculosis* K-10, *M. gilvum* PYR-GCK, *M. smegmatis* mc^{2}155, *M. sp.* KMS and *M. vanbaalenii* PYR-1 were downloaded from the NCBI website. Protein sequences were aligned using the MultAlin program (Corpet, 1988), and a phylogenetic tree was inferred. Secondary structure prediction to identify membrane domains was carried out with Sosui software (Hirokawa *et al.*, 1998) and HMMTOP website (http://www.enzim.hu/hmmtop/).

Bacterial strains, medium and growth. *E. coli* DH5α was used for plasmids propagation and was grown in LB medium like *M. smegmatis* strain mc^{2}155 (Snapper *et al.*, 1990). When required, antibiotics were added to the medium at the following concentrations: hygromycin 50 µg/ml (200 µg/ml for *E. coli*) and kanamycin 25 µg/ml. For the alkaline phosphatase assay, LB plates supplemented with 25 µg/ml kanamycin and 50 µg/ml 5-bromo-4-chloro-3-indolyl phosphate (X-P), a chromogenic alkaline phosphatase substrate, were used (Lim *et al.*, 1995).

Disruption of the *mmpS4* gene of *M. smegmatis*. We used an in-frame deletion to disrupt the *mmpS4* gene by homologous recombination with the pPR27 vector to prevent polar effects (Pelicic et al., 1996). The primers used to amplify the upstream and downstream regions for each gene are listed in Table S2. The chromosomal DNA of clones obtained by these procedures was compared with the wild-type DNA by PCR and Southern blotting. The mutant was named Δ*mmpS4*.

Construction of the *mmpS4* expression plasmids. The wild-type *mmpS4* gene coding sequence (accession number AY439015) was amplified by PCR with *Pfu* Turbo DNA polymerase (Stratagene)

For Peer Review

using *M. smegmatis* mc^{2}155 genomic DNA as a template and the primers *mmpS4*_trans.5 and *mmpS4*_trans.3 (Table S2). PCR products were purified with a Qiagen PCR purification kit, digested with *Xba*I and inserted into the single *Xba*I site of the dephosphorylated integrative expression vector pNIP40b to generate pNIP*mmpS4* (de Mendonca-Lima *et al.*, 2001). Enzymatic digestions were used to select clones in which the *mmpS4* gene was inserted in the opposite orientation of the hygromycin resistant gene. One clone was selected, sequenced and named Δ*mmpS4*/*mmps4*Ms. A similar strategy was used to insert the *mmpS4* genes of *M. avium* (MAV3247), *M. abscessus* (MAB4117c), *M. tuberculosis* (*Rv0451c*) and the *mmpS1* gene (Rv0403c) of *M. tuberculosis* in pNIP40b yielding pNIP*mmpS4*Av, pNIP*mmpS4*Ab, pNIP*mmpS4*Tb, and pNIP*mmpS1*Tb respectively. These plasmids were used to electroporate *M. smegmatis* Δ*mmpS4* and transformants were selected on hygromycin and named Δ*mmpS4/mmpS4*Av, Δ*mmpS4/mmpS4Ab,* Δ*mmpS4/mmpS4*Tb and Δ*mmpS4/mmpS1*Tb respectively. For the construction of plasmid expressing the *mmpS4-phoA* fusion gene, the pJEM11, a *phoA* reporter shuttle plasmid that replicates in *E. coli* and in *M. smegmatis* (Lim *et al.*, 1995), was generously provided by B. Gicquel. The *M. smegmatis mmpS4* gene was PCR amplified (primers *mmpS4phoA*Cter.5 and *mmpS4phoA*Cter.3, Table S2) and inserted into the *Bam*HI site of pJEM11 plasmid to give pJEM11*mmpS4*. The pJEM11 and pJEM11*mmpS4* plasmids were then used to electoporate *M. smegmatis* strain mc^{2}155.

Sliding motility assay. A 3 µl aliquot of liquid culture was dispensed onto plates containing 7H9 medium plus 0.3% agar with no added carbon source. These plates were then incubated at 37°C for one week.

Extraction and purification of mycobacterial lipids. Lipids were extracted from cell pellets with a mixture of chloroform and methanol as previously described (Villeneuve *et al.,* 2003). The extracts were dried under vacuum and partitioned between water and chloroform (1:1, vol/vol). The organic phases was extensively washed with distilled water and evaporated to dryness. The lipid extracts were dissolved in chloroform and analyzed by thin-layer chromatography (TLC) on plates coated with

For Peer Review

Durasil 25- silica gel (0.25 mm thickness; Macherey-Nagel). The GPLs (250 µg each deposit) were resolved by TLC run in chloroform-methanol (9:1, vol/vol) and visualised by spraying the plates with 0.2% anthrone in concentrated sulfuric acid, and heating at 110°C. Identification of GPLs was achieved by matrix-assisted laser-desorption/ionization-time-of-flight (MALDI-TOF) mass spectrometry analysis (Villeneuve et al., 2003).

Quantification of GPL production. Three 200 mL independent cultures of each strain were grown in LB medium, centrifuged and the pellets were weighted. Lipids were extracted and GPLs were purified as described above. GPL purification was completed by deacylating the lipids with 0.1 M KOH as described by Brennan and Goren (Brennan & Goren, 1979). We expressed the amount of GPLs produced as a function of bacterial pellet weight.

Surface-exposed material preparation. The surface-exposed material was recovered from mycobacteria cells treated with 10 g of glass beads as previously described (Sonden *et al.*, 2005). The surface-exposed and the cell-associated lipids were extracted with chloroform and methanol from the surface-exposed material and from the residual cells, respectively, and their GPLs components identified by TLC as described above. The amount of lipid spotted on the TLC plates corresponded to the amount of lipid produced by 5 mg of cells, dry weight.

Immobilization of GPL and selection of single-chain antibodies. We coupled 1 mg of purified GPL to 6 mg of ovalbumin using 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (Pierce) according to the manufacturer's instructions. A Tomlinson I single-chain Fv antibodies library was kindly supplied by Geneservice, Cambridge, UK. Fv antibodies were screened according to the instructions supplied with the library. GPL-coupled ovalbumin was used as the antigen for screening. After screening, the gene encoding the antibody identified (sc051) was inserted into the pMal-2X vector (NEB) (Bach *et al.*, 2001) containing the *ds-red2* gene (Clontech) downstream from the multiple cloning site. This construct was used to produce a fusion protein consisting of the antibody fused to the fluorescent DS-Red2 protein.

For Peer Review

Immunofluorescence microscopy. Bacteria were labelled with 10 µg/ml FITC (Sigma) by incubating with gentle shaking for 1h at 37°C. The bacteria were washed three times with PBS and three times with double distilled water and immobilized on cover slip by flaming. The cover slip was covered with 100 □g of sc051 in PBS and incubated for 30 min. Unbound antibody was then washed away, and the cover slip was washed three times, for three minutes each with double distilled water, and mounted on a glass slide containing FluorSave (Calbiochem). Samples were analyzed by epifluorescence microscopy, as previously described (Sendide *et al.*, 2004). The averages of intensities were measured for each panel and the Δ*mmpS4* strain relative fluorescence was calculated compared to the wild-type strain fluorescence.

Construction of the red fluorescent fusion expression plasmids. For construction of the chimeric *rfp-mmpS4* gene, the wild-type *mmpS4* gene coding sequence was amplified by PCR with *Pfu* Turbo DNA polymerase (Stratagene) using *M. smegmatis* mc^{2}155 genomic DNA as the template and the primers *mmpS4*rfpNter.5 and *mmpS4*_trans.3 (Table S2). The *rfp* gene was amplified using the plasmid pMal-2X as a template and the primers rfpNter.5 and rfpNter.3. The PCR products were purified with a Qiagen PCR purification kit, digested with *Nde*I and *Xba*I and inserted into the dephosphorylated integrative expression vector pNip40b at the single *Xba*I site to generate pNip*Rfp-mmpS4*. Enzymatic digestion was used to select clones on which the *rfp-mmpS4* fusion gene was inserted in the opposite orientation to the hygromycin resistant gene. One clone was selected and sequenced. A similar strategy was used to clone the *mmpS4-rfp* (*mmpS4*_trans.5 and *mmpS4rfp*Cter.3 primers), *mbtH-rfp* (*mbtH*_trans.5 and *mbtH_rfp*Cter.3) and *fadD23-rfp* (*fadD23*_trans.5 and *fadD23_rfp*Cter.3) chimeric genes. In this case, *rfp*Cter.5 and *rfp*Cter.3 primers were used for the amplification of the *rfp* gene. The resulting plasmids were named pNip*mmpS4-rfp*, pNip*mbtH-rfp* and pNip*fadD23rfp*. Plasmids were used to electroporate the wild-type and mutant *M. smegmatis* Δ*mmpS4* strains and transformants were selected on hygromycin.

For Peer Review

Construction of the truncated *mmpS4* expression plasmids. For construction of the truncated *mmpS4* genes, the *M. smegmatis* mc^{2}155 *mmpS4* gene coding sequence was amplified by PCR as described above using the primers *mmpS4*_trans.5 and *mmpS4*-48.3 or *mmpS4*-82.3 or *mmpS4*-116.3 (Table S2) to generate respectively the plasmids pNip*mmpS4*-48, pNip*mmpS4*-82 or pNip*mmpS4*-116. For the amplification of the extracytosolic domain of *mmpS4* gene the primers *mmpS4*cyto.5 and *mmpS4*cyto.3 were used.

RNA isolation and RT-PCR assay. Total RNA was extracted from 10 ml of log phase cultures of *M. smegmatis* strains grown in LB medium as previously described (Deshayes et al., 2005). Contaminating DNA was removed by digestion with DNase I according to the manufacturer's instructions (Promega). The DNase I enzyme was removed with two phenol-chloroform–isoamylalcohol extractions, followed by ethanol precipitation. Reverse transcription was carried out with both primers *mmpS4*-R and *mmpS1*-R. A PCR run on cDNA was then performed by using the primers *mmpS4*-F/*mmpS4*-R and *mmpS1*-F/*mmpS1*-R. The PCR products were resolved by horizontal electrophoresis on a 1.5% agarose gel.

For Peer Review

Acknowledgements.

We thank C. Jeanneau and M. Bertili for bacterial medium preparation. CD is funded by a doctoral grant from "FRM" and a post-doctoral grant form Conseil Général Maine-et-Loire and Angers Loire Métropole. We gratefully acknowledge "CEFIPRA" (grant 3403-B) and INSERM for funding this project. We thank V. Pelicic for critical reading of the manuscript. JMR was Directeur de Recherches at INSERM. We also thank Jeffrey Helm for technical assistance.

For Peer Review

Legends to Figures.

Table 1. Number of *mmpS* and *mmpL* orthologs per bacterial species considered. The asterisk indicates the presence of a gene containing frameshift (the *mmpL13a* and *mmpL13b* genes in *M. tuberculosis*, *mmpL1a*, *mmpL1b*, *mmpL9a* and *mmpL9b* in *M. bovis*, *mmpL1_1, mmpL13_1, mmpL2, mmpL4, mmpl4_5, mmpL4_6, mmpL4_7, mmpL4_8* and *mmpL4_9* in *M. ulcerans*).

Figure 1. A- Phylogenetic tree of the *mmpS* paralogs of *M. smegmatis* (MSMEG), *M. avium* subsp. *avium* (MAV), *M. bovis* (MB), *M. ulcerans* (MUL), *M. leprae* (ML) and *M. tuberculosis* H37Rv (Rv). **B-** Topology of MmpS4 from *M. smegmatis* generated though the prediction obtained by the HMMTOP software. The hydrophobic residues are indicated in black, the polar residues are indicated in light blue, the positively charged residues are indicated in dark blue and the negative residues are indicated in red. **C-** Phenotype on X-P containing LB plates of wild-type *M. smegmatis* strains carrying pJEM11 or pJEM11*mmpS4* reflecting the alkaline phosphatase activity.

Figure 2. A- Sliding motility of the wild-type and Δ*mmpS4 M. smegmatis* strains on low agar containing plate. **B-** TLC analyses of GPLs (250 µg each deposit) produced by the wild-type, Δ*mmpS4* and Δ*mmpS4/mmpS4*Ms strains. **C-** Quantification of alkali-stable lipids in wild-type, mutant and complemented strains. The percentages of alkali-stable lipids differ significantly between the mutant and the wild-type and complemented strains ($p < 0.01$).

Figure 3: **A-** TLC analyses of the lipid extracts of the wild-type, Δ*mmpS4*, and Δ*mmpS4/mmpS4*Ms strains C, cell-associated (non-surface-exposed) GPLs; S, surface-exposed GPLs. Each deposit corresponds to the estimated amount of lipid produced by 5 mg of cells, dry weight. Apolar and polar GPLs are indicated with open and closed boxes, respectively. **B-** Immunofluorescence microscopy analyses. Bacteria were labelled with FITC (green) and incubated with purified anti-GPL antibodies coupled to DSred. The merge image is indicated on the panel on the right.

For Peer Review

Figure 4. A- Percentage of punctuate fluorescence in *M. smegmatis* producing cytosolic GFP (pG13), the MmpS4-RFP or the RFP-MmpS4 fusion protein. **B**- Immunofluorescence of *M. smegmatis* producing the MmpS4-RFP and RFP-MmpS4 fusion proteins. The bar indicates 2.5 µm. The merge image is indicated on the panel on the right.

Figure 5. A- Hypothetical model of the GPL biosynthesis enzymatic megacomplex stabilized and anchored to the bacterial membrane by MmpS and MmpL proteins. **B**- Immunofluorescence of *M. smegmatis* producing the MbtH-Rfp and FadD23-Rfp fusion proteins. The merge image is indicated on the right panel. The bar indicates 2.5 µm. **C**- Percentage of punctuate fluorescence in the wild-type or Δ*mmpS4* strains producing the MbtH-RFP or FadD23-Rfp fusion proteins.

Figure 6. Orthologous complementation**. A-** Sliding motility of Δ*mmpS4/mmpS4*Tb *and* Δ*mmpS4/mmpS1*Tb strains on low agar containing plate. **B-** TLC analysis of the lipid extracts from the Δ*mmpS4/mmpS4*Tb strain. C, cell-associated (non-surface-exposed) GPLs; S, surface-exposed GPLs. Apolar and polar GPLs are indicated with open and closed boxes, respectively. **C-** Immunofluorescent microscopy analyses of the Δ*mmpS4/mmpS4*Tb strain. Bacteria were labelled with FITC (green) and incubated with purified anti-GPL antibodies coupled to DSred. The merge image is indicated on the panel on the right.

For Peer Review

Table 1. Number of *mmpS* and *mmpL* orthologs per considered bacterial species. The asterisk indicates the presence of gene containing frameshift (The *mmpL13a* and *mmpL13b* in *M. tuberculosis*, *mmpL1a*, *mmpL1b*, *mmpL9a* and *mmpL9b* in *M. bovis*, *mmpL1_1, mmpL13_1, mmpL2, mmpL4, mmpl4_5, mmpL4_6, mmpL4_7, mmpL4_8 and mmpL4_9* in *M. ulcerans*).

Bacterial strains	*mmpS*	*mmpL*	Genome size (Mb)
M. smegmatis mc^{2}155	10	17	6.9
M. abscessus ATCC 19977	27	31	5.1
M. gilvum PYR-GCK	6	16	5.6
M. sp. KMS	6	15	6.3
M. vanbaalenii PYR-1	5	14	6.4
M. avium subsp. avium ATCC25291	10	15	5.5
M. avium subsp. paratuberculosis K-10	12	18	4.8
M. bovis AF2122/97	6	16 *	4,3
M. leprae TN	2	7	3.3
M. tuberculosis H37Rv	5	14 *	4.4
M. ulcerans Agy99	6	16 *	5.8
Rhodococcus jostii RHA1	0	9	9,7
Streptomyces coelicolor A3	0	13	8.6

For Peer Review

References

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

Bach, H., Y. Mazor, S. Shaky, A. Shoham-Lev, Y. Berdichevsky, D. L. Gutnick & I. Benhar, (2001) Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *J Mol Biol* **312**: 79-93.

Barker, L. P., D. M. Brooks & P. L. Small, (1998) The identification of Mycobacterium marinum genes differentially expressed in macrophage phagosomes using promoter fusions to green fluorescent protein. *Mol Microbiol* **29**: 1167-1177.

Belisle, J. T., M. R. McNeil, D. Chatterjee, J. M. Inamine & P. J. Brennan, (1993) Expression of the core lipopeptide of the glycopeptidolipid surface antigens in rough mutants of Mycobacterium avium. *J Biol Chem* **268**: 10510-10516.

Bernal, A., U. Ear & N. Kyrpides, (2001) Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. *Nucleic Acids Res* **29**: 126-127.

Betts, J. C., P. T. Lukey, L. C. Robb, R. A. McAdam & K. Duncan, (2002) Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. *Mol Microbiol* **43**: 717-731.

Brennan, P. J., G. O. Aspinall & J. E. Shin, (1981) Structure of the specific oligosaccharides from the glycopeptidolipid antigens of serovars in the Mycobacterium avium-Mycobacterium intracellulare-Mycobacterium scrofulaceum complex. *J Biol Chem* **256**: 6817-6822.

Brennan, P. J. & M. B. Goren, (1979) Structural studies on the type-specific antigens and lipids of the *Mycobacterium avium.Mycobacterium intracellulare.Mycobacterium scrofulaceum* serocomplex. *Mycobacterium intracellulare* serotype 9. *J Biol Chem* **254**: 4205-4211.

Byrd, T. F. & C. R. Lyons, (1999) Preliminary characterization of a *Mycobacterium abscessus* mutant in human and murine models of infection. *Infect Immun* **67**: 4700-4707.

Bystrykh, L. V., M. A. Fernandez-Moreno, J. K. Herrema, F. Malpartida, D. A. Hopwood & L. Dijkhuizen, (1996) Production of actinorhodin-related "blue pigments" by *Streptomyces coelicolor* A3(2). *J Bacteriol* **178**: 2238-2244.

Camacho, L. R., P. Constant, C. Raynaud, M. A. Laneelle, J. A. Triccas, B. Gicquel, M. Daffe & C. Guilhot, (2001) Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. *J Biol Chem* **276**: 19845-19854.

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels & B. G. Barrell, (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**: 537-544.

Converse, S. E., J. D. Mougous, M. D. Leavell, J. A. Leary, C. R. Bertozzi & J. S. Cox, (2003) MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc Natl Acad Sci U S A* **100**: 6121-6126.

Corpet, F., (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**: 10881-10890.

Cox, J. S., B. Chen, M. McNeil & W. R. Jacobs, Jr., (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402**: 79-83.

Crick, D. C. & P. J. Brennan, (2008) Biosynthesis of the arabinogalactan peptidoglycan complex of Mycobacterium tuberculosis.; Daffe M, Reyrat JM, editors: ASM press. 395 p.

De Groote, M. A. & G. Huitt, (2006) Infections due to rapidly growing mycobacteria. *Clin Infect Dis* **42**: 1756-1763.

de Mendonca-Lima, L., M. Picardeau, C. Raynaud, J. Rauzier, Y. O. de la Salmoniere, L. Barker, F. Bigi, A. Cataldi, B. Gicquel & J. M. Reyrat, (2001) Erp, an extracellular protein family specific to mycobacteria. *Microbiology* **147**: 2315-2320.

Deshayes, C., F. Laval, H. Montrozier, M. Daffe, G. Etienne & J. M. Reyrat, (2005) A glycosyltransferase involved in biosynthesis of triglycosylated glycopeptidolipids in *Mycobacterium smegmatis*: impact on surface properties. *J Bacteriol* **187**: 7283-7291.

Domenech, P., M. B. Reed & C. E. Barry, 3rd, (2005) Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. *Infect Immun* **73**: 3492-3501.

Domenech, P., M. B. Reed, C. S. Dowd, C. Manca, G. Kaplan & C. E. Barry, 3rd, (2004) The role of MmpL8 in sulfatide biogenesis and virulence of *Mycobacterium tuberculosis*. *J Biol Chem* **279**: 21257-21265.

For Peer Review

Etienne, G., F. Laval, C. Villeneuve, P. Dinadayala, A. Abouwarda, D. Zerbib, A. Galamba & M. Daffe, (2005) The cell envelope structure and properties of *Mycobacterium smegmatis* mc(2)155: is there a clue for the unique transformability of the strain? *Microbiology* **151**: 2075-2086.

Gokhale, R. S., P. Saxena, T. Chopra & D. Mohanty, (2007) Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. *Nat Prod Rep* **24**: 267-277.

Hirokawa, T., S. Boon-Chieng & S. Mitaku, (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**: 378-379.

Howard, S. T., E. Rhoades, J. Recht, X. Pang, A. Alsup, R. Kolter, C. R. Lyons & T. F. Byrd, (2006) Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. *Microbiology* **152**: 1581-1590.

Jain, M. & J. S. Cox, (2005) Interaction between Polyketide Synthase and Transporter Suggests Coupled Synthesis and Export of Virulence Lipid in *M. tuberculosis*. *PLoS Pathog* **1**: e2.

Kocincova, D., A. K. Singh, J. L. Beretti, H. Ren, D. Euphrasie, J. Liu, M. Daffe, G. Etienne & J. M. Reyrat, (2008) Spontaneous transposition of IS1096 or ISMsm3 leads to glycopeptidolipid overproduction and affects surface properties in Mycobacterium smegmatis. *Tuberculosis (Edinb)* **88**: 390-398.

Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel & D. Portnoi, (1995) Identification of mycobacterium tuberculosis DNA sequences encoding exported proteins by using phoA gene fusions. *J Bacteriol* **177**: 59-65.

Lopez-Marin, L. M., N. Gautier, M. A. Laneelle, G. Silve & M. Daffe, (1994) Structures of the glycopeptidolipid antigens of *Mycobacterium abscessus* and *Mycobacterium chelonae* and possible chemical basis of the serological cross-reactions in the *Mycobacterium fortuitum* complex. *Microbiology* **140 (Pt 5)**: 1109-1118.

Medjahed, H. & J. M. Reyrat, (2009) Construction of Mycobacterium abscessus defined glycopeptidolipid mutants: comparison of genetic tools. *Appl Environ Microbiol* **75**: 1331-1338.

Milano, A., M. R. Pasca, R. Provvedi, A. P. Lucarelli, G. Manina, A. L. Ribeiro, R. Manganelli & G. Riccardi, (2009) Azole resistance in Mycobacterium tuberculosis is mediated by the MmpS5-MmpL5 efflux system. *Tuberculosis (Edinb)* **89**: 84-90.

Patterson, J. H., M. J. McConville, R. E. Haites, R. L. Coppel & H. Billman-Jacobe, (2000) Identification of a methyltransferase from *Mycobacterium smegmatis* involved in glycopeptidolipid synthesis. *J Biol Chem* **275**: 24900-24906.

Pelicic, V., J. M. Reyrat & B. Gicquel, (1996) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol Microbiol* **20**: 919-925.

Rao, A. & A. Ranganathan, (2004) Interaction studies on proteins encoded by the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. *Mol Genet Genomics* **272**: 571-579.

Recht, J. & R. Kolter, (2001) Glycopeptidolipid acetylation affects sliding motility and biofilm formation in *Mycobacterium smegmatis*. *J Bacteriol* **183**: 5718-5724.

Recht, J., A. Martinez, S. Torello & R. Kolter, (2000) Genetic analysis of sliding motility in *Mycobacterium smegmatis*. *J Bacteriol* **182**: 4348-4351.

Reyrat, J. M., V. Pelicic, B. Gicquel & R. Rappuoli, (1998) Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infect Immun* **66**: 4011-4017.

Ripoll, F., C. Deshayes, S. Pasek, F. Laval, J. L. Beretti, F. Biet, J. L. Risler, M. Daffe, G. Etienne, J. L. Gaillard & J. M. Reyrat, (2007) Genomics of glycopeptidolipid biosynthesis in *Mycobacterium abscessus* and *M. chelonae*. *BMC Genomics* **8**: 114.

Sendide, K., A. E. Deghmane, J. M. Reyrat, A. Talal & Z. Hmama, (2004) Mycobacterium bovis BCG urease attenuates major histocompatibility complex class II trafficking to the macrophage cell surface. *Infect Immun* **72**: 4200-4209.

Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser & W. R. Jacobs, Jr., (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**: 1911-1919.

Sonden, B., D. Kocincova, C. Deshayes, D. Euphrasie, L. Rhayat, F. Laval, C. Frehel, M. Daffe, G. Etienne & J. M. Reyrat, (2005) Gap, a mycobacterial specific integral membrane protein, is required for glycolipid transport to the cell surface. *Mol Microbiol* **58**: 426-440.

Straight, P. D., M. A. Fischbach, C. T. Walsh, D. Z. Rudner & R. Kolter, (2007) A singular enzymatic megacomplex from *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **104**: 305-310.

Sulzenbacher, G., S. Canaan, Y. Bordat, O. Neyrolles, G. Stadthagen, V. Roig-Zamboni, J. Rauzier, D. Maurin, F. Laval, M. Daffe, C. Cambillau, B. Gicquel, Y. Bourne & M. Jackson, (2006) LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *Embo J* **25**: 1436-1444.

For Peer Review

Tafelmeyer, P., C. Laurent, P. Lenormand, J. C. Rousselle, L. Marsollier, G. Reysset, R. Zhang, A. Sickmann, T. P. Stinear, A. Namane & S. T. Cole, (2008) Comprehensive proteome analysis of Mycobacterium ulcerans and quantitative comparison of mycolactone biosynthesis. *Proteomics* **8**: 3124-3138.

Tekaia, F., S. V. Gordon, T. Garnier, R. Brosch, B. G. Barrell & S. T. Cole, (1999) Analysis of the proteome of *Mycobacterium tuberculosis* in silico. *Tuber Lung Dis* **79**: 329-342.

Tseng, T. T., K. S. Gratwick, J. Kollman, D. Park, D. H. Nies, A. Goffeau & M. H. Saier, Jr., (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol* **1**: 107-125.

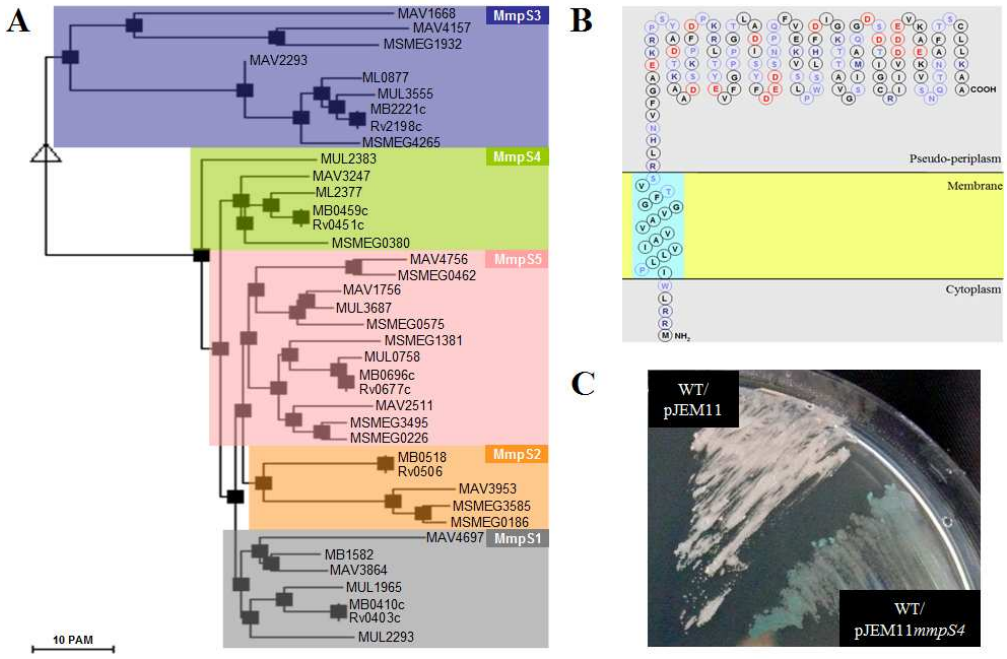
Villeneuve, C., G. Etienne, V. Abadie, H. Montrozier, C. Bordier, F. Laval, M. Daffe, I. Maridonneau-Parini & C. Astarie-Dequeker, (2003) Surface-exposed glycopeptidolipids of *Mycobacterium smegmatis* specifically inhibit the phagocytosis of mycobacteria by human macrophages. Identification of a novel family of glycopeptidolipids. *J Biol Chem* **278**: 51291-51300.

WHO, Buruli Ulcer Fact Sheets. [http://www.who.int/mediacentre/factsheets/fs199/en/].

WHO, Tuberculosis Fact Sheets. [http://www.who.int/mediacentre/factsheets/fs104/en/].

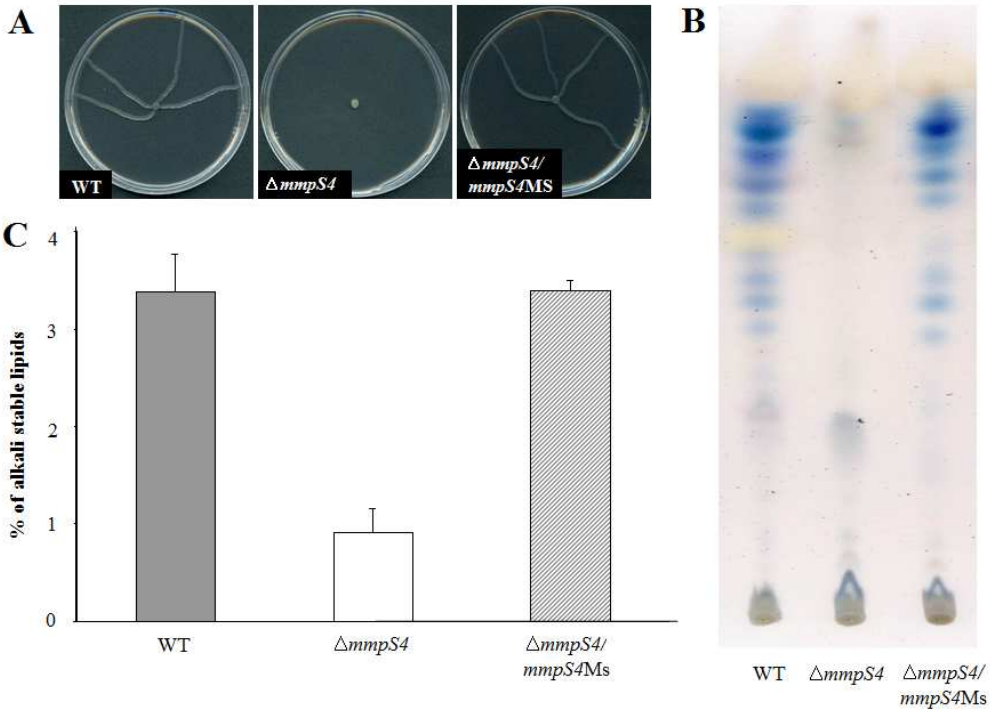
Wolpert, M., B. Gust, B. Kammerer & L. Heide, (2007) Effects of deletions of mbtH-like genes on clorobiocin biosynthesis in *Streptomyces coelicolor*. *Microbiology* **153**: 1413-1423.

For Peer Review



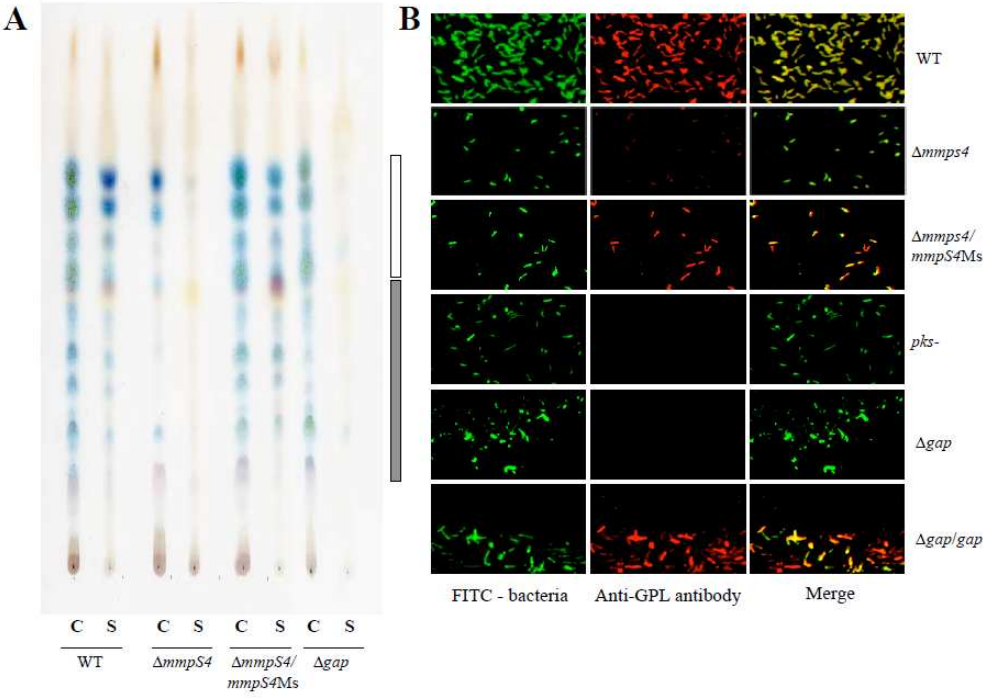
348x230mm (72 x 72 DPI)

Review



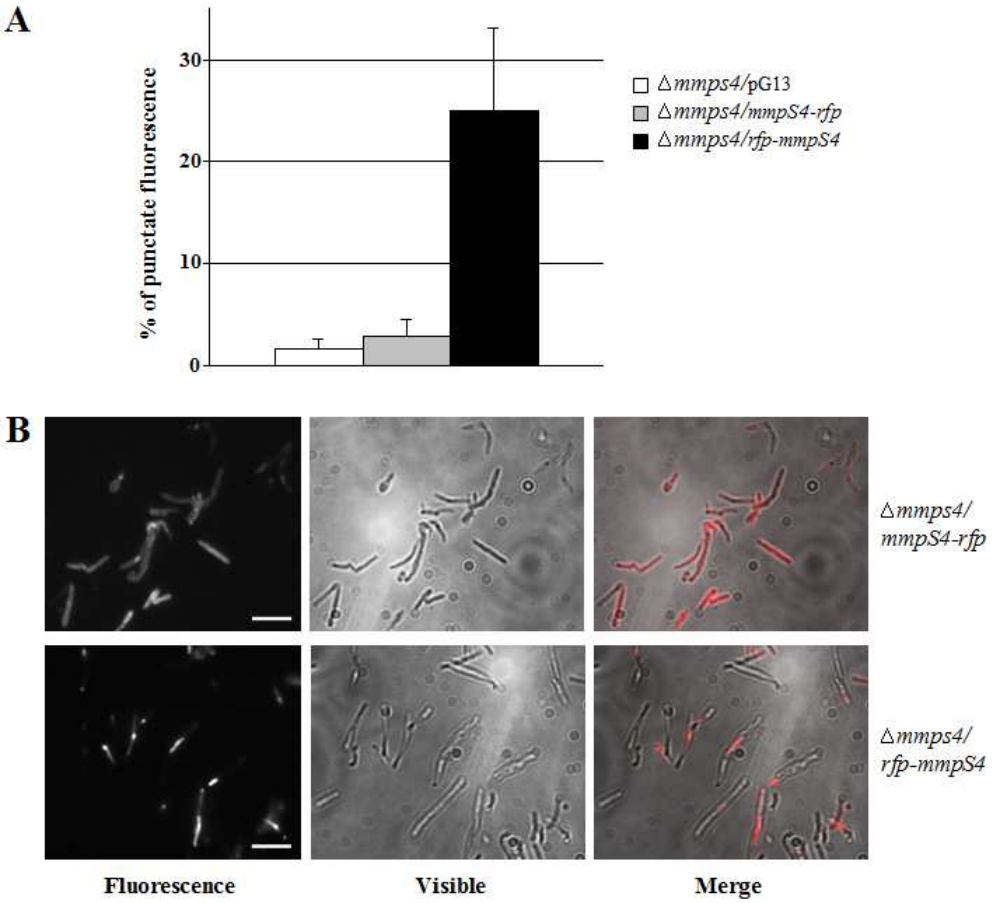
309x222mm (72 x 72 DPI)

Review

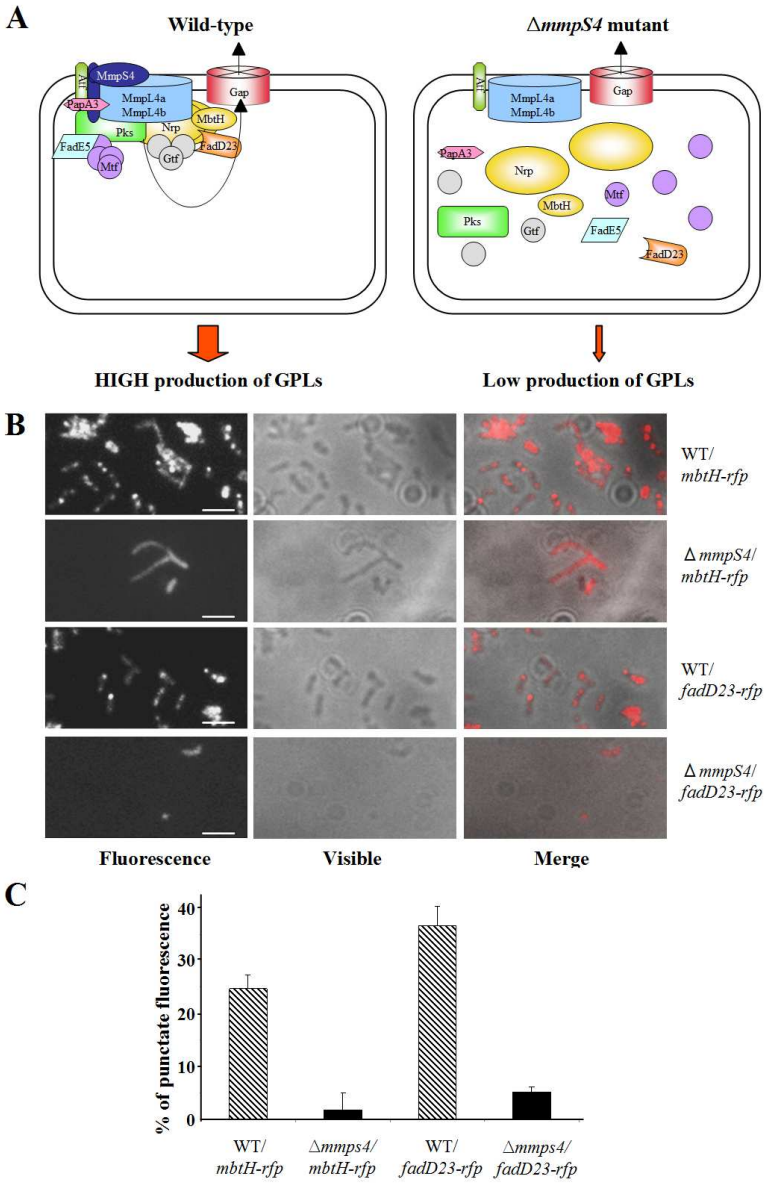


321x228mm (72 x 72 DPI)

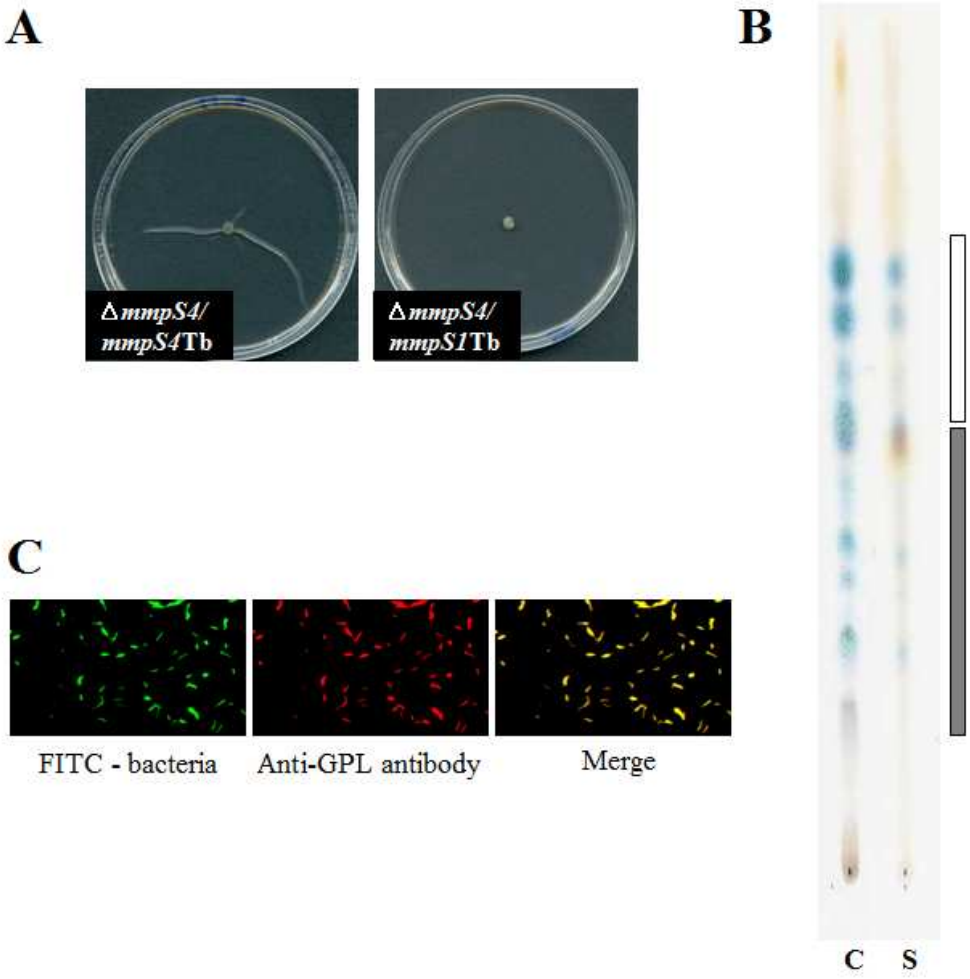
Review



251x231mm (72 x 72 DPI)



376x567mm (72 x 72 DPI)



219x221mm (72 x 72 DPI)