Genetic Basis for the Biosynthesis of Methylglucose Lipopolysaccharides in Mycobacterium tuberculosis

Gustavo Stadthagen, Tounkang Sambou, Marcelo Guerin, Nathalie Barilone, Frédéric Boudou, Jana Korduláková, Patricia Charles, Pedro M. Alzari, Anne Lemassu, Mamadou Daffé, et al.

HAL Id: pasteur-03139205

https://pasteur.hal.science/pasteur-03139205

Submitted on 11 Feb 2021

Genetic Basis for the Biosynthesis of Methylglucose Lipopolysaccharides in *Mycobacterium tuberculosis**[S]

Gustavo Stadthagen[‡]**, Tounkang Sambou**[§]**, Marcelo Guerin**[¶]**, Nathalie Barilone**[‡]**, Frédéric Boudou**[‡]**,
Jana Korduláková**[‡1]**, Patricia Charles**[‡]**, Pedro M. Alzari**[¶]**, Anne Lemassu**[§]**, Mamadou Daffé**[§]**, Germain Puzo**[§]**,
Brigitte Gicquel**[‡]**, Michel Rivière**[§]**, and Mary Jackson**[‡2]

From the [‡]*Unité de Génétique Mycobactérienne and* [¶]*Unité de Biochimie Structurale, Institut Pasteur, 75015 Paris, France and the*
[§]*Département Mécanismes Moléculaires des Infections Mycobactériennes, Institut de Pharmacologie et de Biologie Structurale,
CNRS, 31077 Toulouse, France*

Mycobacteria produce two unusual polymethylated polysaccharides, the 6-*O*-methylglucosyl-containing lipopolysaccharides (MGLP) and the 3-*O*-methylmannose polysaccharides, which have been shown to regulate fatty acid biosynthesis *in vitro*. A cluster of genes dedicated to the synthesis of MGLP was identified in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. Overexpression of the putative glycosyltransferase gene *Rv3032* in *M. smegmatis* greatly stimulated MGLP production, whereas the targeted disruption of *Rv3032* in *M. tuberculosis* and that of the putative methyltransferase gene *MSMEG2349* in *M. smegmatis* resulted in a dramatic reduction in the amounts of MGLP synthesized and in the accumulation of precursors of these molecules. Disruption of *Rv3032* also led to a significant decrease in the glycogen content of the tubercle bacillus, indicating that the product of this gene is likely to be involved in the elongation of more than one α-(1→4)-glucan in this bacterium. Results thus suggest that *Rv3032* encodes the α-(1→4)-glucosyltransferase responsible for the elongation of MGLP, whereas *MSMEG2349* encodes the *O*-methyltransferase required for the 6-*O*-methylation of these compounds.

Mycobacteria produce two unusual polymethylated polysaccharides (PMPS),[3] the 3-*O*-methylmannose polysaccharides (MMP) (1–2) and the 6-*O*-methylglucosyl-containing lipopolysaccharides (MGLP) (3, 4). Both polysaccharides localize to the cytoplasm, where they have been postulated to regulate fatty acid synthesis by FAS-I as a consequence of their ability to form stable 1:1 complexes with long-chain fatty acids and acyl-coenzyme A derivatives (5–9) (for a review, see Ref. 10). PMPS were also proposed to protect fatty acid products from degradation and to serve as general lipid carriers facilitating the synthesis of the very large and insoluble mycolic acid esters while at the same time increasing the tolerance of mycobacteria to high cytoplasmic concentrations of long-chain acyl-CoA derivatives (10–12).

PMPS were first isolated from *Mycobacterium phlei*, *Mycobacterium smegmatis*, and *Mycobacterium tuberculosis* in the 1960s (2–4), and much of the information we have about the structure, biosynthesis, and biological activities of these molecules comes from this early work. Others then revised the structure of MGLP and extended the analysis of these molecules to other mycobacterial species (13–15). The structures of MGLP and MMP are shown in Fig. 1. MMP have been found in multiple nonpathogenic fast growing species of mycobacteria (16) and in *Streptomyces griseus* (17), whereas MGLP have been isolated from several *Nocardia* species as well as *M. phlei*, *M. smegmatis*, *Mycobacterium bovis* BCG, *M. tuberculosis*, *M. leprae*, and *M. xenopi* (3–4, 13–15, 18, 19).

Ballou and co-workers (2, 16, 20, 21) isolated precursors of MMP and characterized an α-(1→4)-mannosyltransferase and a 3-*O*-methyltransferase from cell-free extracts of *M. smegmatis*. These studies led to a biosynthetic model in which MMP is elongated by a linear alternating process of mannosylation followed by *O*-methylation, in which GDP-Man serves as the sugar donor for the mannosyltransferase and *S*-adenosylmethionine serves as the source of methyl groups. Termination of the elongation reaction occurs when the length of the chain is sufficient to confer on the polysaccharide fatty acid-binding properties (11–13 3-*O*-methylmannoses). At this stage, the chain is terminated with an unmethylated mannose because the acyl-CoA-bound oligosaccharides are no longer available as acceptors for the 3-*O*-methyltransferase.

Knowledge of the initiation, elongation, and termination reactions involved in the biosynthesis of MGLP is more limited. A membrane-associated acyltransferase activity responsible for the transfer of acetyl, propionyl, isobutyryl, octanoyl, and succinyl groups from their respective acyl-CoA derivatives onto

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–3.

[1] Present address: Dept. of Biochemistry, Faculty of Natural Sciences, Comenius University, Mlynska Dolina CH-1, 84215 Bratislava, Slovak Republic.

[2] To whom correspondence should be addressed: Dept. of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523-1682. Tel.: 970-491-3582; Fax: 970-491-1815; E-mail: Mary.Jackson@colostate.edu.

[3] The abbreviations used are: PMPS, polymethylated polysaccharides; D-Glc*p*, D-glucopyranose; GT, glycosyltransferase; MALDI, matrix-assisted laser desorption ionization; TOF, time-of-flight; MS, mass spectrometry; MGLP, *O*-methylglucosyl-containing lipopolysaccharide(s); MGP, *O*-methylglucosyl-containing polysaccharide(s); MGP$_{x,y}$, MGP composed of x glucosyl residues, of which y are mono-*O*-methylated; MMP, methylmannose polysaccharide(s).
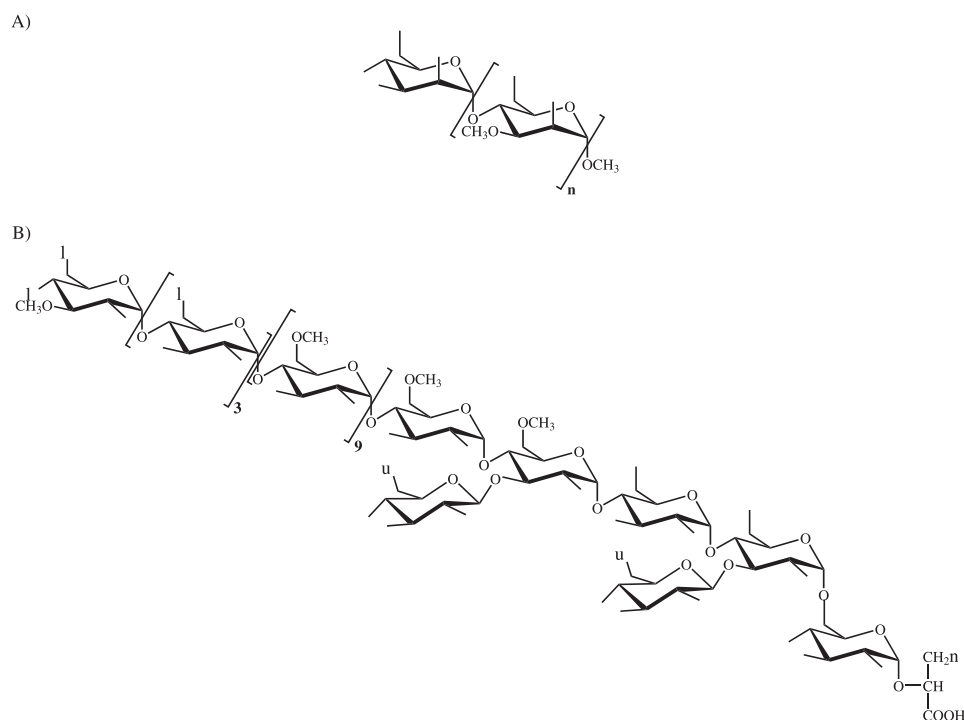
A)

B)

FIGURE 1. **Structures of the mycobacterial PMPS.** *A*, the MMP from *M. smegmatis* is shown. They are composed of 10–13 ($n = 10–13$) α-(1→4)-linked 3-*O*-methyl-D-mannoses terminated at the nonreducing end by a single α-(1→4)-linked unmethylated D-mannose and at the reducing end by an α-methyl aglycon. MMP occur in the cells as a mixture of at least four isomers due to differences in size and degree of *O*-methylation. *B*, MGLP from *M. bovis* BCG. MGLP are composed of 10 α-(1→4)-linked 6-*O*-methylglucosyl residues with a nonreducing end made of the tetrasaccharide 3-*O*-methyl-D-Glc*p*-(α-(1→4)-D-Glc*p*)$_3$-α(1→. The tetrasaccharide→4)-(α-(1→4)-D-Glc*p*)$_3$-α(1→6)-D-Glc*p*-α(1→ linked to the position 2 of D-glyceric acid constitutes the reducing end of the molecule. Position 3 of the second and fourth α-D-Glc*p* residues (closest to the reducing end) are substituted by single β-D-Glc*p* residues. The nonreducing end of the polymer is acylated by a combination of acetate, propionate, and isobutyrate (●), whereas octanoate (■) esterifies position 1 of glyceric acid and 0–3 succinate groups (♦) esterify the Glc residues of the reducing end. MGLP occurs as a mixture of four main components that differ in their content of esterified succinic acid.

purified MGLP and partially acetylated α-(1→4)-linked D-gluco-oligosaccharides has been described (22). A soluble protein fraction from *M. phlei* capable of catalyzing the transfer of methyl groups from *S*-adenosylmethionine to positions 6 and 3 of MGLP and partially acetylated α-(1→4)-D-gluco-oligosaccharides has also been reported (23, 24). Since the position of the *O*-methylation on the oligosaccharide acceptor was dependent on its degree of acetylation, it was suggested that acylation and *O*-methylation occurred together during the biosynthesis of the lipopolysaccharide, the former process exerting a control on the latter. The characterization of weakly acidic and partially *O*-methylated methylglucosyl-containing polysaccharide precursors from *M. smegmatis* led Kamisango *et al.* (25) to propose a model for the biosynthesis of MGLP in which the elongation of the chain proceeds stepwise, from the reducing end toward the nonreducing end, through a sequential glucosylation-methylation reaction. Although at least one α-(1→4)-glucosyltransferase is expected to be required for the elongation of the glucan backbone of MGLP, no such enzymatic activity had been reported.

Since these early studies, and despite the important roles PMPS might play in the regulation of fatty acid metabolism in mycobacteria, the biosynthesis of PMPS had not been reinvestigated in the postgenomic era.

EXPERIMENTAL PROCEDURES

Construction of the M. tuberculosis and M. smegmatis Glucosyltransferase and Methyltransferase Mutants—The *ts-sacB* method (26) was used to achieve allelic replacement at the *Rv3032* locus of *M. tuberculosis* (ATCC number 25618) and at the *MSMEG2349* locus of *M. smegmatis* mc²155. The *M. tuberculosis Rv3032* gene and flanking regions was PCR-amplified from *M. tuberculosis* H37Rv genomic DNA using primers Rv3032.1 (5′-gggctgc-agatcgccggcgcgctggcc-3′)/Rv3032.2 (5′-tgagccatgtcgcctccctgg-3′), and a disrupted allele, *Rv3032::kan*, was obtained by inserting the kanamycin resistance cassette from pUC4K (Amersham Biosciences) into the SmaI restriction site of *Rv3032*. *Rv3032::kan* was then cloned into the NotI-cut and blunt-ended pPR27-*xylE* (26) to obtain pPR27Rv3032KX, the construct used for allelic replacement in *M. tuberculosis*. The *M. smegmatis MSMEG2349* gene and flanking regions was PCR-amplified using the primers Ms3030f (5′-cacgttctgagcaacctacacc-3′) and Ms3030r (5′-tgatcagatgggaacggc-cctcgg-3′). A disrupted copy of *MSMEG2349* was obtained by substituting 152 bp of the coding region of this gene bracketed between two SalI sites by the Kan cassette from pUC4K. *MSMEG2349::kan* was then cloned into the XbaI-cut pJQ200-*xylE*, yielding pJQMSMEG2349KX.

Overexpression of Rv3032 in M. smegmatis and Complementation Studies—The entire coding sequence of Rv3032 was PCR-amplified from *M. tuberculosis* H37Rv genomic DNA using the primers Rv3032.3 (5′-gcgcgcg<u>catatg</u>aggatcctcatggt-gtcgtg-3′) and Rv3032.4 (5′-ggg<u>aagctt</u>ccgatcgggaagagcgtgctc-3′) and cloned into the NdeI and HindIII restriction sites of the expression vector pVV16 (27), yielding pVVRv3032. The production of recombinant Rv3032 protein in *M. smegmatis* and *M. tuberculosis* was analyzed by immunoblotting with the monoclonal Penta-His antibody from Qiagen as described previously (27).

Whole Cell Radiolabeling Experiments—Radiolabeling of whole *M. smegmatis* and *M. tuberculosis* cells with [1,2-¹⁴C]acetic acid (0.5 μCi ml⁻¹; specific activity, 113 Ci mol⁻¹, MP Biomedicals Inc.) was performed at 37 °C in 7H9-ADC-Tween 80 broth for 14 h with shaking. Radiolabeling with [*methyl*-¹⁴C]L-methionine (0.5 μCi ml⁻¹; specific activity, 68 Ci

mol^{-1}; MP Biomedicals Inc.) was performed at 37 °C in Sauton medium for 48 h (*M. tuberculosis*) or 24 h (*M. smegmatis*).

Preparation, Purification, and Analysis of MGLP—For the preparation of PMPS from cold or radiolabeled cultures, *M. smegmatis* and *M. tuberculosis* were grown in Sauton's medium as surface pellicles. Total lipids and PMPS were extracted with chloroform/methanol (1:2) followed by two other extractions with chloroform/methanol (2:1). The pooled and dried organic extracts were then partitioned between chloroform and water (1:1), and the MGLP and MMP contained in the aqueous phase were further purified by reverse phase chromatography on Sep-Pak® Plus tC18 columns (Waters). Fractions were eluted with increasing concentrations of methanol in water, and the MMP and MGLP content of each fraction was determined by TLC and gas chromatography analysis of alditol acetates derived from these fractions (28). TLC analyses were performed on aluminum-backed silica gel 60-precoated plates F$_{254}$ (Merck) using chloroform/methanol/water (56:38:10) as the eluent.

PMPS were recovered in the fractions eluted with 40 and 70% methanol in water in which they were estimated to be 75–90% pure (supplemental Fig. 1). Dry MGLP were deacylated in 1 M NaOH at 37 °C for 3 h, yielding MGP. The mixture was neutralized with 0.2 M HCl and desalted using reverse phase SepPak® Plus tC18 (Waters) chromatography as described above.

Matrix-assisted Laser Desorption Ionization Time-of-flight (MALDI-TOF) Mass Spectrometry (MS) analysis of purified MGLP, MGP, and MMP samples was performed on a 4700 Proteomics Analyzer MALDI TOF/TOF system (Applied Biosystems, Framingham, MA) equipped with a pulsed Nd:YAG laser emitting at 355 nm and operating at 200 Hz. Typically, 0.3 μl of the sample in water was mixed with 0.3 μl of matrix solution (2-(4-hydroxyphenylazo)benzoic acid at a concentration of ~10 mg ml^{-1} in ethanol/water (1:1)) directly onto the target plate. All samples were analyzed in reflector mode using both positive and negative ion detection. Final spectra were obtained by accumulating 2,500 subspectra generated by 250 laser pulses on 10 distinct areas of the dried droplet and calibrated using the instrument's external calibration procedure.

Analysis of Glycogen and Glucan—For the analysis of glycogen and α-D-glucan, *M. tuberculosis* strains were grown in Sauton's medium as surface pellicles. High molecular weight carbohydrates and proteins were precipitated from 10-fold concentrated culture filtrates with six volumes of cold ethanol overnight at 4 °C, as described (29). Precipitates were recovered by centrifugation at 14,000 × *g* for 1 h, dissolved in distilled water, dialyzed for 1 day against water to eliminate traces of salts and glycerol, lyophilized, and weighted. The extracellular α-D-glucan content was determined by gas chromatography of sugar derivatives resulting from the acid hydrolysis of the lyophilized macromolecules. Further purification of α-D-glucan was carried out by anion-exchange chromatography on a DEAE-trisacryl gel, and the neutral fractions were rechromatographed on a Bio-Gel P-60 column, as previously described (30).

For the purification of glycogen, heat-killed cells were weighted, suspended in water, and centrifuged twice at 1,100 × *g* for 15 min to remove loosely attached compounds, among them the extracellular α-D-glucan. Washed cells were broken in a French press at 140 bars to liberate the intracellular glycogen. The material was centrifuged at 1,100 × *g* for 30 min to eliminate nonbroken cells, and the supernatant was recentrifuged at 27,000 × *g* for 15 min to remove cell walls. Glycogen was extracted as described (31) with only minor modifications.

RESULTS

Identification of a Cluster of Genes Potentially Involved in the Synthesis of MGLP—Glucosyltransferase(s) likely to be involved in the elongation of MGLP were searched within the CAZy classification of glycosyltransferases (GTs) (Carbohydrate-Active enZymes; available on the World Wide Web). The cytosolic localization of MGLP suggested that the glucosyltransferase(s) probably used ADP-D-Glc or UDP-D-Glc as D-Glc donor. The fact that these lipopolysaccharides essentially consist of α-(1→4)-linked D-Glc*p* residues further suggested that the glucosyltransferase(s) used a retaining type of catalytic mechanism for sugar transfer, leading to a glycosidic bond in α-configuration. The CAZy GT-4 family is the largest family of NDP-sugar-dependent retaining GTs. Enzymes from this family carry a conserved motif (D/E)X_7E proposed to be involved in the binding of the sugar-nucleotide donor. *M. tuberculosis* H37Rv contains seven representatives of this family, among which three (PimA, PimB, and MshA) have been functionally characterized in mycobacteria or related species and shown to participate in the biosynthesis of phosphatidylinositol mannosides, glycosylated diacylglycerols, and mycothiol (27, 32–34). Among the four remaining GTs, Rv1212c and Rv3032 showed the greatest level of sequence similarity with the glycogen synthase from *E. coli* (GlgA) (46.7 and 46% similarity at the amino acid level, respectively). *Rv1212c* is orthologous to the *Corynebacterium glutamicum* glycogen synthase gene, *glgA*, with which it shares 71.8% similarity (60.3% identity) at the amino acid level (35). In contrast, no functional information was available for *Rv3032*. Interestingly, this gene displayed a limited distribution within prokaryotes, having orthologs only in mycobacterial spp. (*M. tuberculosis* H37Rv and CDC1551, *M. bovis*, *M. leprae*, *M. smegmatis*, *Mycobacterium avium*, *M. avium* subsp. *paratuberculosis*, *Mycobacterium ulcerans*, *M. flavescens*) and *Nocardia farcinica*. A closer examination of the genomic region encompassing *Rv3032* in *M. tuberculosis* H37Rv revealed that, adjacent or in close vicinity to this putative glucosyltransferase gene, lay a putative acetyltransferase gene (*Rv3034c*), two putative *S*-adenosyl-methionine-dependent-methyltransferase genes (*Rv3030* and *Rv3037c*), and a putative α-amylase/glucoside hydrolase/GH-57 family branching enzyme gene (*Rv3031*) (Fig. 2). Given the likely involvement of some if not all of these genes in MGLP biosynthesis, efforts were focused on this cluster of genes and, more particularly, on *Rv3032* and *Rv3030*.

Construction of M. smegmatis and M. tuberculosis Glucosyltransferase and Methyltransferase Knock-out Mutants and Phenotypic Analyses—The putative glucosyltransferase gene, *Rv3032*, and methyltransferase gene, *MSMEG2349* (orthologous to *Rv3030* of *M. tuberculosis*), were disrupted by homologous recombination in *M. tuberculosis* H37Rv and *M. smegmatis*, respectively, using the *ts-sacB* method (26). Allelic replacements at the *Rv3032* and *MSMEG2349* loci were con-
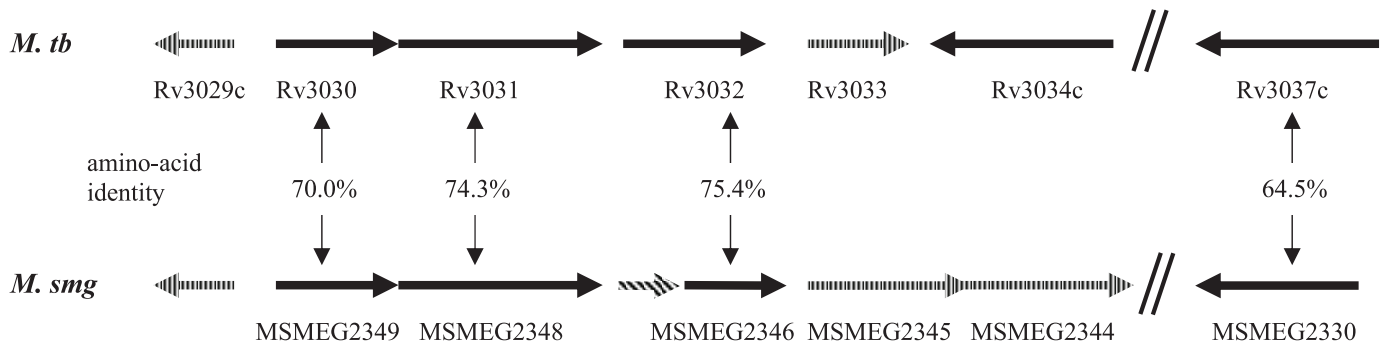
FIGURE 2. **Organization of the MGLP gene cluster in *M. tuberculosis* H37Rv (*M. tb*) and *M. smegmatis* mc²155 (*M. smg*).** *Rv3032/MSMEG2346*, glucosyltransferase genes; *Rv3031/MSMEG2348*, putative α-amylase/glucoside hydrolase/GH-57 family branching enzyme genes; *Rv3030/MSMEG2349*, putative *S*-adenosylmethionine-dependent methyltransferase genes; *Rv3034c*, putative acetyltransferase gene (two *M. smegmatis* mc²155 orthologs of *Rv3034c* exist at a different location of the chromosome, *MSMEG6949* and *MSMEG7005*); *Rv3037c/MSMEG2330*, putative *S*-adenosylmethionine-dependent methyltransferase gene. The genes likely to be involved in MGLP synthesis are in *black*; genes of unknown function are indicated in *hatched bars*. In *M. tuberculosis*, the start codon of *Rv3031* overlaps the stop codon of *Rv3030*, and *Rv3031* and *Rv3032* are separated by an intergenic region of 32 bp.
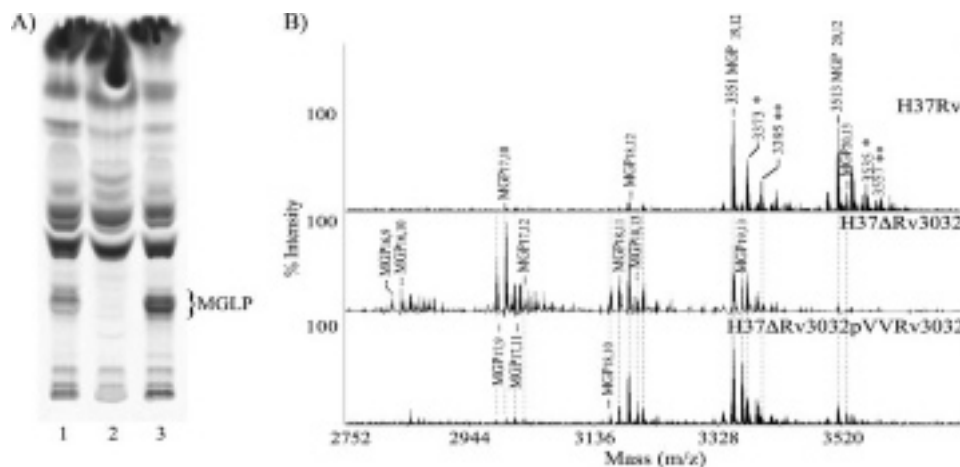

FIGURE 3. **MGLP composition of the wild-type and *Rv3032* mutant strains of *M. tuberculosis* H37Rv.** *A*, TLC analysis of the MGLP from wild-type H37Rv (1), H37RvΔRv3032 (2), and H37RvΔRv3032/pVVRv3032 (3). *B*, partial high mass range (*m/z* 2750–3700) MALDI mass spectra in the negative ion detection mode of the purified MGP (deacylated MGLP) from the *M. tuberculosis* wild-type, mutant, and complemented mutant strains. All annotated peaks correspond to [M − H]⁻ pseudomolecular ions with the exception of those labeled with *one* or *two* *asterisks* attributed to [M − 2H + Na]⁻ and [M − 3H + 2Na]⁻ ions, respectively.
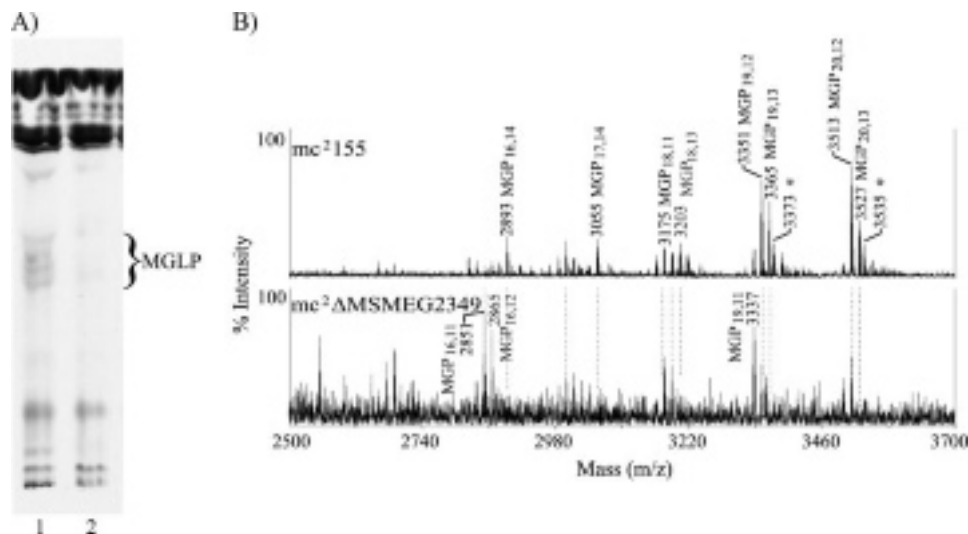

FIGURE 4. **MGLP composition of the wild-type and *MSMEG2349* mutant strains of *M. smegmatis*.** *A*, TLC analysis of the MGLP from mc²155 (1) and mc²ΔMSMEG2349 (2). *B*, partial high mass range (*m/z* 2750–3700) MALDI mass spectra in the negative ion detection mode of the purified MGP (deacylated MGLP) from the *M. smegmatis* wild-type and mutant strains. All annotated peaks correspond to [M − H]⁻ pseudomolecular ions with the exception of those labeled with an *asterisk* attributed to [M − 2H + Na]⁻ ions.

firmed by Southern hybridization (data not shown). Interestingly, the growth of H37RvΔRv3032 was slightly slower than that of wild-type H37Rv in 7H9 broth at 37 °C and completely abolished at a higher temperature (39 °C) (supplemental Fig. 2). Likewise, the mc²ΔMSMEG2349 mutant did not grow at high temperature (42 °C) in Sauton's medium, although its growth was comparable with that of mc²155 at 30 and 37 °C in this medium (data not shown). An analysis of the MGLP content of the wild-type and mutant strains upon metabolic labeling with [*methyl*-¹⁴C]L-methionine revealed that the amount of radioactivity incorporated into the MGLP of the H37RvΔRv3032 and mc²ΔMSMEG2349 mutants was only 17 and 38%, respectively, of that incorporated by their wild-type parent (Figs. 3*A* and 4*A*). Comparison of the yields of MGLP from the different strains also supported these conclusions in that, depending on the growth phase, the amounts of lipopolysaccharides recovered from H37RvΔRv3032 were 20–60% less than that recovered from wild-type *M. tuberculosis* H37Rv. Recovery from mc²ΔMSMEG2349 was significantly less than that from wild-type *M. smegmatis*. However, quantification was not possible due to the important amounts of MMP co-eluting with MGLP in this *Mycobacterium* species (see supplemen-

tal Fig. 1). It is noteworthy that MMP could not be detected in the extracts or in the purified PMPS fractions of *M. tuberculosis*, neither by MALDI-MS nor by TLC (supplemental Fig. 1). Consistent with this observation, gas chromatography analyses of the alditol acetates derived from purified PMPS revealed the presence of 3-*O*-methylmannose in the *O*-methylated polysaccharide fractions from *M. smegmatis* but not in those from *M. tuberculosis* (strains H37Rv and Mt103) (data not shown).

Complementation of H37RvΔRv3032 with a wild-type copy of *Rv3032* carried by the multicopy plasmid pVVRv3032 restored the production of MGLP in the mutant, thereby confirming the involvement of Rv3032 in MGLP synthesis (Fig. 3*A*). Attempts to complement the mc²ΔMSMEG2349 mutant with *Rv3030* from *M. tuberculosis* carried by the multicopy expression plasmid pVV16 were not successful due to the apparent toxic effect of overexpressing this gene in *M. smegmatis*. However, the fact that *MSMEG2348* transcripts were detected by reverse transcription-PCR in mc²ΔMSMEG2349 (data not shown) suggests that inactivation of the methyltransferase gene, rather than a polar effect on the expression of the downstream putative branching enzyme gene, is responsible for the phenotype.

To further compare the nature of the MGLP produced by the wild-type and mutant strains, PMPS were purified by reverse phase chromatography, and the MGLP-containing fractions, either native or deacylated, were analyzed by MALDI-MS. The partial negative ion MALDI mass spectra of the MGLP-containing fractions of H37Rv, H37RvΔRv3032, and H37RvΔRv3032/pVVRv3032 are presented in supplemental Fig. 3. The mass spectrum of the MGLP from wild-type H37Rv showed a group of intense ions between 3500 and 4000 atomic mass units, potentially attributable to MGLP $[M + Na]^-$ or $[M - H + 2Na]^-$ pseudomolecular ions. A similar pattern was seen in the spectrum of the complemented mutant but not in that of H37RvΔRv3032, in good agreement with the quantification results presented above. Interpretation of these clustered ions indicates that they may result from the overlap of ions differing by 162, 100, and 14 mass units reflecting, respectively, the variability of MGLP structures in terms of their degree of glycosylation, succinylation, acylation, and/or *O*-methylation. Therefore, to best perceive and analyze potential qualitative differences in the oligosaccharide backbones of the MGLP from the wild-type and mutant strains, the purified MGLP-containing fractions were deacylated by alkali treatment and reanalyzed by MALDI-MS. The MGP fraction of wild-type H37Rv mainly showed two series of ions, separated by 162 mass units and differing by 22 mass units attributable to sodium adducts of pseudomolecular ions of MGP (Fig. 3*B*). The two most intense ions at *m/z* 3351 and *m/z* 3513 were respectively assigned to the $[M - H]^-$ pseudomolecular ions of $MGP_{19,12}$ and $MGP_{20,12}$ (MGP containing 19 and 20 glucose units, among which 12 are *O*-methylated). This attribution was further confirmed by the fact that these ions were found, respectively, at *m/z* 3397 and *m/z* 3559 ($[M - H + 2Na]^+$) in the MALDI mass spectrum run in the positive ion detection mode (data not shown). This result thus confirmed that the complex MGLP pattern described above resulted from variable combinations of acylation of mainly two oligosaccharide back-

bones, $MGP_{19,12}$ and $MGP_{20,12}$, and, to a much lesser extent, $MGP_{20,11}$ and $MGP_{20,13}$.

As shown in Figs. 3*B* and 4*B*, the disruptions of *Rv3032* and *MSMEG2349* had significant impacts on the MGLP patterns of *M. tuberculosis* and *M. smegmatis*. H37RvΔRv3032 accumulated polysaccharidic compounds with lower masses, consisting essentially of 17 Glc residues ($MGP_{17,9}$, $MGP_{17,10}$, $MGP_{17,11}$), with lesser quantities of $MGP_{16,10}$, $MGP_{18,10}$, $MGP_{18,11}$, $MGP_{18,12}$, and $MGP_{19,12}$. More glucosylated forms of MGLP could not be detected in this strain. Complementation of the mutant partially restored MGP synthesis, resulting in a clear shift of its MGP profile toward the mature forms of these polysaccharides ($MGP_{19,12}$ and $MGP_{20,12}$) (Fig. 3*B*). MS analysis of the MGP from the *MSMEG2349* mutant also revealed a shift in the types of MGP produced from the mature to the less glycosylated forms of these molecules, with a clear accumulation of $MGP_{16,11}$ and $MGP_{16,12}$ (*m/z* 2851 and *m/z* 2865, respectively) (Fig. 4*B*).

In conclusion, the analyses performed on the two mutant strains indicate that *Rv3032* in *M. tuberculosis* and *MSMEG2349* in *M. smegmatis* participate in the biosynthesis of MGLP, although compensatory glucosyltransferase and *O*-methyltransferase activities exist in these species, allowing the knock-out mutants to produce residual quantities of glucosylated PMPS.

Disruption of Rv3032 Also Affects the Glycogen Content of M. tuberculosis—The likely involvement of *Rv3032* in the formation of α-(1→4)-glucosidic bonds prompted us to further analyze the glycogen (36) and capsular α-D-glucan (29) content of H37RvΔRv3032. Analyses performed on two independent cultures of *M. tuberculosis* wild-type and *Rv3032* mutant strains revealed an important decrease (44.5 ± 9.0%) in the glycogen content of the mutant relative to the wild-type strain. This tendency was partially reversed upon complementation of the mutant with pVVRv3032 (the glycogen content of the complemented mutant was 74 ± 0.6% of that of the wild-type strain). To analyze the impact of the inactivation of *Rv3032* on the production of glucan, which represents the major extracellular polysaccharide of *M. tuberculosis* (29), the ratios of Glc to the monosaccharide constituents of the other extracellular polysaccharides (*i.e.* arabinomannan and mannan) (29) (Glc *versus* Ara + Man + Glc) were determined in the culture filtrates of three or four independent batches of the wild-type and mutant strains. Very similar ratios were obtained (0.56 ± 0.07 and 0.58 ± 0.09, respectively) for the different batches and strains, indicating that the synthesis of the capsular glucan was unaffected in the mutant. This suggests either that Rv3032 is not involved in the synthesis of the capsular polysaccharide or that residual α-(1→4)-glucosyltransferase activity (or activities) in the mutant cells is preferentially utilized to synthesize glucan. The glycogen and glucan produced by the mutant were otherwise structurally identical to those of *M. tuberculosis* H37Rv as determined by ¹H NMR analyses (data not shown). Altogether, our results implicated Rv3032 in the synthesis of two different α-(1→4)-linked glucans in *M. tuberculosis*: MGLP and glycogen.

Effect of Overexpressing Rv3032 in M. smegmatis—To further study the role of Rv3032 in MGLP synthesis, we analyzed the
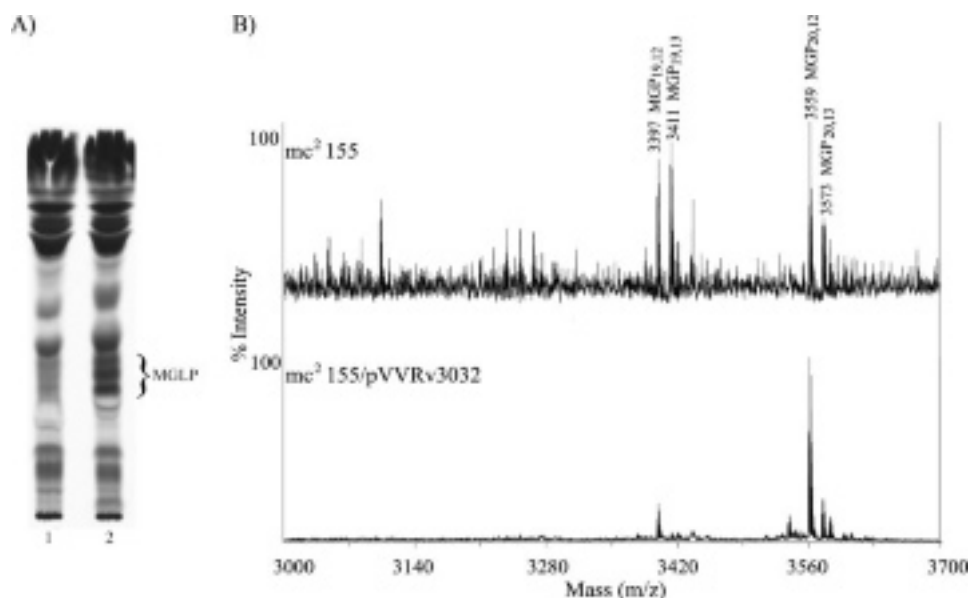
FIGURE 5. **Effect of overexpressing *Rv3032* on the MGLP composition of *M. smegmatis*.** *A*, TLC analysis of the MGLP from wild-type mc²155 (*lane 1*) and mc²155/pVVRv3032 (*lane 2*). *B*, MALDI mass spectra in the positive ion detection mode of the purified MGP (deacylated MGLP) from *M. smegmatis* mc²155 and mc²155/pVVRv3032.

among which the glucosyltransferase gene *Rv3032* and the ortholog of the *O*-methyltransferase gene *Rv3030* in *M. smegmatis* were studied. Also present within the gene cluster is *Rv3031*, annotated as a putative α-amylase in the genome of *M. tuberculosis* H37Rv, *Rv3034c*, annotated as a putative acetyltransferase, and *Rv3037c*, encoding a putative *S*-adenosylmethionine-dependent methyltransferase of unknown function. Sequence similarities between Rv3031 and a GH-57 family branching enzyme from *Thermococcus kodakaraensis* (38) suggest that this enzyme is involved in generating the α(1→6) glycosidic bond linking the first and second D-Glc*p* residues at the reducing end of the molecule.

In *M. tuberculosis*, Rv3032 participates in the production of glycogen and MGLP. It also stimu-

effects of overexpressing *Rv3032* on the production of these glycoconjugates in *M. smegmatis*. Transformation of mc²155 with pVVRv3032, the multicopy plasmid used in the complementation studies described above, resulted in the production of a carboxyl-terminal hexahistidine-tagged recombinant protein of the expected size (~45 kDa) (data not shown). TLC analyses of the MGLP contents from [*methyl*-¹⁴C]L-methionine-labeled mc²155 and mc²155/pVVRv3032 cells (Fig. 5*A*) revealed a clear increase in the synthesis of the lipopolysaccharides in the overexpressor. Moreover, MALDI-MS analyses of the purified MGP fractions from the same strains indicated a stimulatory effect of the overexpression of *Rv3032* on the synthesis of the most glucosylated (mature) forms of MGLP (Fig. 5*B*).

These results, together with the fact that the inactivation of *Rv3032* in *M. tuberculosis* resulted in a general decrease in the production of glycogen and all forms of MGLP, suggested that Rv3032 is the main α-(1→4)-glucosyltransferase responsible for the elongation of the lipopolysaccharides. Supporting this assumption, preliminary cell-free assays, which will be presented elsewhere, indicated that a purified recombinant form of Rv3032 catalyzes *in vitro* the transfer of α-(1→4)-linked Glc residues from UDP-D-Glc onto short-chain gluco-oligosaccharides, capsular α-D-glucan and glycogen (data not shown).

DISCUSSION

The availability of a growing number of mycobacterial genome sequences and the rapidly growing knowledge of the processes underlying sugar transfer in these bacteria (37) have opened the way to the elucidation of the biosynthetic pathways of complex glycoconjugates in *M. tuberculosis*. This work was undertaken with the aims of investigating the biosynthetic pathway of MGLP. A cluster of genes relatively conserved within the *Mycobacterium* genus was identified,

lates the synthesis of MGLP when overexpressed in *M. smegmatis*. The general decrease in all forms of MGLP that followed the disruption of *Rv3032* in *M. tuberculosis* and the stimulatory effect the overexpression of *Rv3032* had on the production of mature MGLP in *M. smegmatis* strongly suggest that Rv3032 is the main α-(1→4)-glucosyltransferase committed in the elongation of those lipopolysaccharides. Likewise, the similar phenotype that arose from the genetic disruption of *MSMEG2349* in *M. smegmatis* suggests that this gene encodes the main *O*-methyltransferase of the pathway (*i.e.* the one required for the 6-*O*-methylation of the polysaccharide backbone of MGLP). Important information derived from the analysis of the *M. smegmatis* mutant is that a defect in *O*-methylation abolishes MGLP synthesis. Thus, despite both the α-(1→4)-glucosyltransferase and the 6-*O*-methyltransferase being active on unmethylated α-(1→4)-D-gluco-oligosaccharides *in vitro* (23–24), the elongation of MGLP in whole bacterial cells appears to proceed with glucosylation and *O*-methylation occurring hand in hand. This observation is consistent with the biosynthetic model proposed by Kamisango *et al.* (25) based on the structural analysis of MGLP precursors.

The residual quantities and types of MGLP found in H37RvΔRv3032 and mc²ΔMSMEG2349 imply that compensatory glucosyltransferase and *O*-methyltransferase activities exist allowing *M. tuberculosis* and *M. smegmatis* to synthesize basal amounts of wild-type MGLP. According to our bioinformatics analyses, an obvious α-(1→4)-glucosyltransferase gene candidate in *M. tuberculosis* H37Rv is *Rv1212c*, the ortholog of the glycogen synthase gene from *Corynebacterium glutamicum* (35). Work is in progress in our laboratories to study the involvement of this gene in glycogen and MGLP synthesis and to generate a double *Rv3032/Rv1212c* knock-out mutant of *M. tuberculosis*.

An intriguing preliminary observation that arose from our study is the fact that, contrary to *M. smegmatis*, *M. tuberculosis* (strains H37Rv and Mt103) does not produce detectable amounts of MMP. Although more isolates will have to be tested before a definitive conclusion can be drawn, these results and the fact that MMP have so far only been reported in fast growing nonpathogenic species of mycobacteria (16) raise the interesting possibility that mannosylated PMPS may not be produced by slow growing mycobacterial species.

Due to the existence of compensatory enzymatic activities and to the dual involvement of Rv3032 in MGLP and glycogen synthesis, it is at present not possible to draw any conclusions as to the potential essentiality and physiological role(s) of MGLP in mycobacteria. Answers to these questions will require the construction of an *M. tuberculosis* mutant totally deficient in their synthesis, a goal that is likely to be achievable through the individual or combined inactivation of the branching enzyme, *O*-methyltransferase and acetyltransferase genes, *Rv3031*, *Rv3030, Rv3037c*, and *Rv3034c*.

REFERENCES

1. Gray, G. R., and Ballou, C. E. (1971) *J. Biol. Chem.* **246,** 6835–6842
2. Maitra, S. K., and Ballou, C. E. (1977) *J. Biol. Chem.* **252,** 2459–2469
3. Lee, Y. C., and Ballou, C. E. (1964) *J. Biol. Chem.* **239,** PC3602–PC3603
4. Lee, Y. C. (1966) *J. Biol. Chem.* **241,** 1899–1908
5. Yabusaki, K. K., and Ballou, C. E. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75,** 691–695
6. Yabusaki, K. K., Cohen, R. E., and Ballou, C. E. (1979) *J. Biol. Chem.* **254,** 7282–7286
7. Maggio, J. E. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 2582–2586
8. Hindsgaul, O., and Ballou, C. E. (1984) *Biochemistry* **23,** 577–584
9. Tuffal, G., Tuong, A., Dhers, C., Uzabiaga, F., Rivière, M., Picard, C., and Puzo, G. (1998) *Anal. Chem.* **70,** 1853–1858
10. Bloch, K. (1977) *Adv. Enzymol. Relat. Areas Mol. Biol.* **45,** 1–84
11. Yabusaki, K. K., and Ballou, C. E. (1979) *J. Biol. Chem.* **254,** 12314–12317
12. Forsberg, L. S., Dell, A., Walton, D. J., and Ballou, C. E. (1982) *J. Biol. Chem.* **257,** 3555–3563
13. Hunter, S. W., Gaylord, H., and Brennan, P. J. (1986) *J. Biol. Chem.* **261,** 12345–12351
14. Tuffal, G., Ponthus, C., Picard, C., Rivière, M., and Puzo, G. (1995) *Eur. J. Biochem.* **233,** 377–383
15. Tuffal, G., Albigot, R., Rivière, M., and Puzo, G. (1998) *Glycobiology* **8,** 675–684
16. Weisman, L. S., and Ballou, C. E. (1984) *J. Biol. Chem.* **259,** 3457–3463
17. Harris, L. S., and Gray, G. R. (1977) *J. Biol. Chem.* **252,** 2470–2477
18. Lornitzo, F. A., and Goldman, D. S. (1968) *Biochem. Biophys. Acta* **158,** 329–335
19. Pommier, M. T., and Michel, G. (1986) *J. Gen. Microbiol.* **132,** 2433–2441
20. Yamada, H., Cohen, R. E., and Ballou, C. E. (1979) *J. Biol. Chem.* **254,** 1972–1979
21. Weisman, L. S., and Ballou, C. E. (1984) *J. Biol. Chem.* **259,** 3464–3469
22. Tung, K.-K., and Ballou, C. E. (1973) *J. Biol. Chem.* **248,** 7126–7133
23. Ferguson, J. A., and Ballou, C. E. (1970) *J. Biol. Chem.* **245,** 4213–4223
24. Grellert, E., and Ballou, C. E. (1972) *J. Biol. Chem.* **247,** 3236–3241
25. Kamisango, K., Dell, A., and Ballou, C. E. (1987) *J. Biol. Chem.* **262,** 4580–4586
26. Jackson, M., Camacho, L. R., Gicquel, B., and Guilhot, C. (2001) in *Mycobacterium tuberculosis Protocols* (Parish, T., and Stocker, N. G., eds) Vol. 54, pp. 59–75, Humana Press, Totowa, NJ
27. Korduláková, J., Gilleron, M., Mikusová, K., Puzo, G., Brennan, P. J., Gicquel, B., and Jackson, M. (2002) *J. Biol. Chem.* **277,** 31335–31344
28. McNeil, M., Chatterjee, D., Hunter, S. W., and Brennan, P. J. (1989) *Methods Enzymol.* **179,** 215–242
29. Lemassu, A., and Daffé, M. (1994) *Biochem. J.* **297,** 351–357
30. Dinadayala, P., Lemassu, A., Granovski, P., Cérantola, S., Winter, N., and Daffé, M. (2004) *J. Biol. Chem.* **279,** 12369–12378
31. Bueding, E., and Orrell, S. (1964) *J. Biol. Chem.* **239,** 4018–4020
32. Schaeffer, M. L., Khoo, K.-H., Besra, G. S., Chatterjee, D., Brennan, P. J., Belisle, J. T., and Inamine, J. M. (1999) *J. Biol. Chem.* **274,** 31625–31631
33. Newton, G. L., Koledin, T., Gorovitz, B., Rawat, M., Fahey, R. C., and Av-Gay, Y. (2003) *J. Bacteriol.* **185,** 3476–3479
34. Tatituri, R. V. V., Illarionov, P. A., Dover, L. G., Nigou, J., Gilleron, M., Hitchen, P. G. H., Krumbach, K., Morris, H. R., Spencer, N., Dell, A., Eggeling, L., and Besra, G. S. (2007) *J. Biol. Chem.* **282,** 4561–4572
35. Tzvetkov, M., Klopprogge, C., Zelder, O., and Liebl, W. (2003) *Microbiology* **149,** 1659–1673
36. Elbein, A. D., and Mitchell, M. (1973) *J. Bacteriol.* **113,** 863–873
37. Berg, S., Kaur, D., Jackson, M., and Brennan, P. J. (2007) *Glycobiology* **17,** 35R–56R
38. Murakami, T., Kanai, T., Takata, H., Kuriki, T., and Imanaka, T. (2006) *J. Bacteriol.* **188,** 5915–5924