Molecular Dissection of the Role of Two Methyltransferases in the Biosynthesis of Phenolglycolipids and Phthiocerol Dimycoserosate in the Mycobacterium tuberculosis Complex*

Esther Pérez‡, Patricia Constant, Françoise Laval, Anne Lemassu, Marie-Antoinette Lanéelle, Mamadou Daffe§, and Christophe Guilhot‡

From the Département “Mécanismes Moléculaires des Infections Mycobactériennes,” Institut de Pharmacologie et Biologie Structurale, CNRS and Université Paul Sabatier (Unité Mixte de Recherche 5089), 205 route de Narbonne, 31077 Toulouse Cedex, France

Received for publication, June 2, 2004, and in revised form, July 20, 2004
Published, JBC Papers in Press, August 3, 2004, DOI 10.1074/jbc.M406134200

A few mycobacterial species, most of which are pathogenic for humans, produce dimycocerosates of phthiocerol (DIM) and of glycosylated phenolphthiocerol, also called phenolglycolipid (PGL), two groups of molecules shown to be important virulence factors. The biosynthesis of these molecules is a very complex pathway that involves more than 15 enzymatic steps and has just begun to be elucidated. Most of the genes known to be involved in these pathways are clustered on the chromosome of M. tuberculosis. Based on their amino acid sequences, we hypothesized that the proteins encoded by \textit{Rv2952} and \textit{Rv2959c}, two open reading frames of this locus, are involved in the transfer of methyl group onto various hydroxyl functions during the biosynthesis of DIM, PGL, and related p-hydroxybenzoic acid derivatives (p-HBAD). Using allelic exchange and site-specific recombination, we produced three recombinant strains of \textit{Mycobacterium tuberculosis} carrying insertions in \textit{Rv2952} or \textit{Rv2959c}. Analysis of these mutants revealed that (i) the protein encoded by \textit{Rv2952} is a methyltransferase catalyzing the transfer of a methyl group onto the lipid moiety of phthiotriol and glycosylated phenolphthiocerol dimycoserosates to form DIM and PGL, respectively, (ii) \textit{Rv2959c} is part of an operon including the newly characterized \textit{Rv2958c} gene that encodes a glycosyltransferase also involved in PGL and p-HBAD biosynthesis, and (iii) the enzyme encoded by \textit{Rv2959c} catalyzes the O-methylation of the hydroxyl group located on carbon 2 of the rhamnosyl residue linked to the phenolic group of PGL and p-HBAD produced by \textit{M. tuberculosis}. These data further extend our understanding of the biosynthesis of important mycobacterial virulence factors and provide additional tools to decipher the molecular mechanisms of action of these molecules during the pathogenesis of tuberculosis.

Bacteria of the \textit{Mycobacterium tuberculosis} complex exhibit amazing capacities to infect their host, resist bactericidal responses, and subvert the host immune response. Thus, \textit{M. tuberculosis}, the causative agent of tuberculosis, has colonized one-third of the population of the world and kills more than two million people annually (1). This pathogenicity has been largely attributed to the unusual mycobacterial envelope. This complex structure has a high lipid content, up to 60% of the dry weight of the bacterium, and contains a large variety of lipids with unusual structures (2). Schematically, from the cytoplasm to the external side of the bacterium, the cell envelope is formed by (i) a plasma membrane, (ii) a cell wall core composed of three covalently attached macromolecules, i.e. the peptidoglycan, the arabinogalactan, and the mycolic acids, which are long chain (C60-C90) fatty acids, capped with a layer of non-covalently linked lipids and glycolipids, and (iii) a capsule of polysaccharides, proteins and lipids (2).

Among the extractable constituents of the cell envelope are two structurally related families of lipids, the diesters of phthiocerol and phenolglycolipids. The former class of lipids is a mixture of long chain β-diols, called phthiocerol and relatives, which are esterified by multimethyl-branched fatty acids (3). Depending on the stereochemistry of the chiral centers bearing the methyl branches, the fatty acids are called mycoseric or phthioceranic acids (4). To date, phthioceryl dimycocerosates (DIM) and diphthioceranates (DIP) have been identified in eight mycobacterial species; DIM have been found in \textit{M. tuberculosis}, \textit{Mycobacterium bovis}, \textit{Mycobacterium africanum}, \textit{Mycobacterium leprae}, \textit{Mycobacterium gastri}, and \textit{Mycobacterium kansasii}, whereas DIP have been found in \textit{Mycobacterium marinum} and \textit{Mycobacterium ulcerans} (4). With the exception of \textit{M. gastri}, all the DIM- or DIP-containing species are pathogenic. The mycobacterial species that produce DIM or DIP may also synthesize structurally related substances, called phenolphthiocerols and relatives, in which phthiocerol is α-terminated by an aromatic nucleus, usually glycosylated by a type- or species-specific mono-, tri-, or a tetrasaccharide unit leading to PGL (5, 6). More recently, we identified a new group of molecules related to PGL, the glycosylated \textit{p}-hydroxybenzoic acid methyl esters (p-HBAD), in the culture media of bacteria of the \textit{M. tuberculosis} complex (7).

The genes involved in the biosynthesis of DIM, PGL, and p-HBAD are clustered on the chromosome of bacteria of the \textit{M. tuberculosis} complex (Fig. 1) (8–11). Five of these genes, \textit{ppsA–E}, encode a type-I modular polyketide synthase responsible for the synthesis of phthiocerol and phenolphthiocerol by...

---

* This work was supported by the CNRS, France. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a Marie Curie Fellowship from the European Union.

¶ To whom correspondence may be addressed. Tel.: 33-561-175-845; Fax: 33-561-175-994; E-mail: Mamadou.Daffe@ipbs.fr.

The abbreviations used are: DIM, dimycocerosates of phthiocerol; DIP, diphthioceranates of phthiocerol; km, kanamycin; MALDI-TOF, matrix-assisted laser desorption-ionization time-of-flight; ORF, open reading frames; PGL, phenolglycolipid; PGL-th, PGL of \textit{M. tuberculosis}; p-HBAD, \textit{p}-hydroxybenzoic acid derivatives; TMS, trimethylsilyl; BCG, Bacille Calmette-Guérin; km, kanamycin.
the elongation of a C20-C22 fatty acyl chain or an acyl chain containing a phenol moiety with three malonyl-CoA and two methylmalonyl-CoA units (12). The pks151/p gene encodes an iterative type-I polyketide synthase catalyzing the elongation of p-HBA to form p-hydroxyphenylkanoates, which in turn are converted into phenolphthiophoric derivatives by the PpsA–E synthase (7). This gene is mutated in most M. tuberculosis clinical isolates, explaining why most strains of M. tuberculosis are unable to synthesize PGL even though they produce the structurally related DIM and P-HBAD. The mas gene encodes another iterative type-I polyketide synthase responsible for the synthesis of mycocerosic acids after 2–4 rounds of extension of C18-C20 fatty acids with methylmalonyl-CoA units (13). Two additional genes, fadD26 and fadD28 (also named acoa), encode two acyl-adenylate synthases involved in the formation of DIM and PGL presumably by activating the various polyketide synthase substrates (14, 15). Finally, four genes (drrA, drrB, and drrC encoding an ABC transporter and mmpL7 encoding a transporter of the RND superfam) are also located in this locus and are involved in the translocation of DIM from the cytoplasm to the bacterial cell surface (11). Close examination of this genomic region revealed five putative open reading frames (ORF) encoding proteins possibly involved in the glycosylation and methylation of DIM and PGL (7).

In the accompanying paper (16) we demonstrated that three of these genes (Rv2957, Rv2958c, and Rv2962c) encode proteins with similarities to glycosyltransferases and that these proteins are involved in the sequential elongation of both phenolphthiophoric dimycocerosates and p-hydroxybenzoic acid to form PGL and p-HBAD, respectively. In the present paper we show that the other two genes (Rv2959c and Rv2952) encode methyltransferases. We demonstrate (i) that the protein encoded by Rv2952 catalyzes the methylation of phthiotoic acid and phenolphthiolic acid to form phthiocerol and phenolphthiolic acid, respectively, and (ii) that the Rv2959c product is involved in the O-methylation of the hydroxyl group located at position 2 of the first rhamnosyl residue found in PGL-tb and P-HBAD of M. tuberculosis. We also show that Rv2959c and Rv2958c, which encodes the glycosyltransferase involved in the transfer of the terminal disaccharide moiety of PGL-tb and P-HBAD II (16), form an operon.

**EXPERIMENTAL PROCEDURES**

Most of the materials and methods used are described in the accompanying article (16). Only the experimental procedures specific to the present work are detailed.

**Construction of M. tuberculosis H37Rv Mutants—**PCR was carried out in a final volume of 50 µL containing M. tuberculosis genomic DNA, 1 µM primer A (Rv2959A or Rv2952A), 1 µM primer B (Rv2959B or Rv2952B) (Table I), 2.5 units of Taq DNA polymerase (Roche Applied Science), and 10% MgSO4. The amplification program consisted of 1 cycle of 10 min at 95 °C followed by 35 cycles of 30 s at 95 °C, 30 s at 57 °C, and 3 min at 72 °C and a final extension at 72 °C for 10 min. The PCR product was analyzed by electrophoresis in 0.8% agarose gel. The resulting fragments, 2805 bp for Rv2952 and 2860 bp for Rv2959c, were gel-purified using the Qiaquick gel extraction purification kit (Qiagen, Courtaboeuf, France). The Rv2952 fragment was digested with SacI and XbaI, and the Rv2959c fragment was digested with NotI and XhoI. The pBlueScript vector was digested with SpeI and KpnI and religated yielding pPET2. The Rv2952c and Rv2965c PCR fragments were inserted into the pPET2 vector between the XbaI and SacI or between the XbaI and NotI restriction sites, respectively, to give pPET2 (Rv2952c) and pPET5 (Rv2959c). A kanamycin resistance cassette (km) formed by the 1.4 km cassette from pHP45bm (17) flanked by two res sequences from transposon γ (18) was inserted between the KpnI and SalI sites of Rv2952c, generating a 605-bp deletion, and the two MfeI and BglII sites of pPET5, generating a 2860-bp deletion. The resulting plasmids were named pPET23 and pPET14, respectively. The PmeI fragments from the various plasmids, containing the disrupted gene constructs, were inserted at the XbaI site of pPR27 (19), generating pPET34 (Δr2952c::km) and pPET40 (Δr2959c::km). M. tuberculosis H37Rv was electrotransformed as previously described, and transformants were selected on 7H11 + oleic acid–albumin–dextrose–catalase + km at 32 °C (20). Two transformants obtained with each of the plasmids were grown in 5 ml of TH9 + albumin–dextrose–catalase + km + Tween at 32 °C until saturation. Dilutions of this culture were plated on 7H11 + oleic acid–albumin–dextrose–catalase + km + sucrose and incubated at 39 °C for 4 weeks. The colonies were screened by PCR using primers C, D, E, res1, or res2. The amplification program consisted of 1 cycle of 10 min at 95 °C followed by 35 cycles of 30 s at 95 °C, 30 s at 55 °C, and 3 min at 72 °C, and a final 10 min at 72 °C. Two clones giving the pattern corresponding to allelic exchange were retained for further analysis. These strains were renamed PMM28 (Δr2952c::km) and PMM28 (Δr2959c::km).

**Construction of M. tuberculosis H37Rv Δr2959c::res Unmarked Mutant—**To recover the res-Δkm-res cassette from M. tuberculosis PMM18, we transferred the plasmid pWM19 containing the resolvase gene of transposon γ under the control of the mycobacterial promotor pBlaF into this strain (18). The PMM16/pWM19 transformation mixture was resuspended in 5 ml of TH9 + albumin–dextrose–catalase and incubated for 48 h at 32 °C to allow the expression of hygromycin resistance. Transformants were selected directly in the liquid medium by adding hygromycin to the transformation mixture and incubating the culture at 32 °C for 12 days. Viable bacteria were then recovered by plating serial dilutions on 7H11 + oleic acid–albumin–dextrose–catalase plates without antibiotics and incubating at 39 °C, a non-permissive temperature for pWM19 replication. Twenty-one colonies picked randomly were then tested for their growth on km-containing plates. Thirteen of these colonies were unable to grow on km-containing plates but grew as the control on antibiotic-free plates. Three of these clones were analyzed by PCR using primers res1 + 2959c, 2959E + 2959C, and 2959F + 2959G. The amplification program consisted of 1 cycle of 10 min at 95 °C followed by 35 cycles of 30 s at 95 °C, 30 s at 55 °C, and 3 min at 72 °C, and a final extension of 10 min at 72 °C. The amplification pattern revealed that the res-Δkm-res cassette had been excised in these three clones, leaving a copy of the 132-bp res site within the Rv2959c gene. One clone, named PMM18res, was retained for further analysis.

### Table I

| Gene     | Oligonucleotide | Sequence                                      |
|----------|-----------------|-----------------------------------------------|
| Rv2959A  | 5'-GCTCTAGATTTTAAACGCCGGCGAACATGGTGATTT-3' |
| Rv2959B  | 5'-GCGGGCAACATTAAACGCCGGCGAACATGGTGATTT-3' |
| Rv2959C  | 5'-ACTGTTCTGATCCTTGAGCCGACATGGTGATTT-3'    |
| Rv2959D  | 5'-GGATGGGCGGATGTTGAGCCGACATGGTGATTT-3'    |
| Rv2959E  | 5'-TACCGCTGAGATCTCTCGTGATTT-3'             |
| Rv2959F  | 5'-GCCCGCTGAGATCTCTCGTGATTT-3'             |
| Rv2959G  | 5'-CCGCTGAGATCTCTCGTGATTT-3'               |
| Rv2952A  | 5'-GCGCGCTGAGATCTCTCGTGATTT-3'             |
| Rv2952B  | 5'-TGCAAGGTGTTAGTGGCCGACATGGTGATTT-3'      |
| Rv2952C  | 5'-TCCGCTGAGATCTCTCGTGATTT-3'              |
| Rv2952D  | 5'-CCTCGAAGGTGTTAGTGGCCGACATGGTGATTT-3'    |
| Rv2952E  | 5'-CGCGCGATCTCTCGTGATTT-3'                 |
| res1     | 5'-GCTCTAGATTTTAAACGCCGGCGAACATGGTGATTT-3' |
| res2     | 5'-GCTCTAGATTTTAAACGCCGGCGAACATGGTGATTT-3' |
RESULTS

Disruption of Rv2952 and Rv2959c Genes in M. tuberculosis H37Rv by Allelic Exchange—The biosynthesis of PGL, p-HBAD, and DIM in bacteria of the M. tuberculosis complex involves the transfer of several methyl groups onto the carbohydrate moieties of PGL and p-HBAD and onto the lipid domains of PGL and DIM. This implies that several unknown methyltransferases catalyze these reactions. Previous studies have shown that many genes involved in the biosynthesis of PGL and DIM are clustered on a 70-kilobase fragment of the M. tuberculosis chromosome; that is, the DIM and PGL locus (Fig. 1). Therefore, we looked for ORFs encoding putative proteins with similarities to methyltransferase in this locus. Two such ORFs, Rv2952 and Rv2959c, were identified downstream of pks15/1. Interestingly, the protein encoded by Rv2952, but not that encoded by Rv2959c, harbored an amino acid motif conserved in several other mycobacterial methyltransferases shown to be required for the transfer of methyl groups onto fatty acids (21). Thus, we hypothesized that the Rv2959c product is involved in the methylation of the glycosyl moiety of PGL-tb and p-HBAD and that the Rv2952 product methylates the lipid domains of both DIM and PGL-tb.

To establish the exact roles of the proteins encoded by Rv2952 and Rv2959c in the biosynthesis of DIM, PGL-tb, and p-HBAD, we constructed two M. tuberculosis H37Rv mutants in which these genes were disrupted by allelic exchange. Briefly, chromosomal fragments overlapping the 5′ or 3′ ends of Rv2952 or Rv2959c were cloned flanking a km resistance cassette into the vector pPR27 (19) yielding plasmids pPET40 (Rv2959c) and pPET34 (Rv2952). These constructs were independently transferred by electroporation into M. tuberculosis H37Rv. Two transformants obtained with each plasmid were selected on the counterselective plates using C or E allele (Fig. 2). Two clones, named PMM18 (ΔRv2959c::km) and PMM28 (ΔRv2952::km), were retained for further studies.

Biochemical Phenotypes of Methyltransferase Mutants—To examine the effects of the mutations in the Rv2952 and Rv2959c genes on p-HBAD, DIM, and PGL production, we first transferred the plasmid pPET1 into the H37Rv strain and in the two mutants, PMM18 and PMM28. This plasmid carries a functional pks15/1 gene from M. bovis BCG because the wild-type H37Rv strain is unable to produce PGL-tb due to a frameshift mutation in pks15/1 (7).

The p-HBAD compounds are mostly found in the culture supernatant of the H37Rv strain of M. tuberculosis, whereas DIM and PGL-tb remain associated with the bacterial cells. Therefore, the mutants and the wild-type strain (either transformed with the plasmid pPET1 or not transformed) were grown in liquid culture, and lipids were extracted from both the culture supernatant and the bacterial pellets.

TLC analysis of the organic solvent extracts from the culture supernatants obtained with the various strains revealed the presence of one glycoconjugate spot in the lipid extract of H37Rv. It corresponds to the previously described p-HBAD II (tri-O-methyl-fucosyl-(α-1→3)-rhamnosyl-(α-1→3)-2-O-methyl-rhamnosyl-α-p-hydroxybenzoic acid methyl ester) (7). This compound is the major p-HBAD produced by M. tuberculosis H37Rv. A minor compound p-HBAD I corresponding to 2-O-methyl-rhamnosyl-α-p-hydroxybenzoic acid methyl ester is sometimes visible but was undetectable in this experiment. The same pattern of glycoconjugate production was observed with the PMM28 mutant strain (Fig. 3A). In contrast, no glycoconjugates were detected in the organic-solvent extracts from the PMM18 mutant (Fig. 3A). Hence, the putative methyltransferase encoded by Rv2959c, but not that encoded by Rv2952, plays a role in the biosynthesis of p-HBAD.

To determine the roles of the putative methyltransferase in the biosynthesis of PGL-tb, lipids from the bacterial cells were examined. TLC analysis of these lipids showed that the PMM28::pPET1 mutant produced two glycoconjugates, products B and C (Fig. 3B). These molecules have slightly different mobilities to PGL from M. bovis BCG (2-O-methyl-rhamnosyl-phenolphthiocerol dimycocerosates) and from M. tuberculosis (tri-O-methyl-fucosyl-(α-1→3)-rhamnosyl-(α-1→3)-2-O-methyl-rhamnosyl-phenolphthiocerol dimycocerosates). In the case of PMM18::pPET1 mutant, a glycoconjugate (product A, Fig. 3B) with a lower mobility than that of PGL-tb was observed. These results strongly suggest that both Rv2952 and Rv2959c are involved in PGL-tb biosynthesis.

Because PGL and DIM are structurally related, we investigated the roles of the putative methyltransferases in the biosynthesis of DIM. As expected, the control strain (H37Rv::pPET1) produced both DIM A (phenolphthiocerol dimycocerosates) and DIM B (phthiodiolone dimycocerosate) (Fig. 3B). The mutation in the Rv2952 gene did not affect the production of DIM, as both forms were detected in similar amounts to those found in the wild-type strain (data not shown). In contrast, the mutation in the Rv2952 gene seemed to abolish the synthesis of DIM A completely but did not affect the production of DIM B both in PMM28 and PMM28::pPET1 (Fig. 3C). When large amounts of lipid extracts were loaded on TLC plates, an additional spot (product D) was detected in the organic solvent extracts of PMM28::pPET1 in comparison with that of the wild-type strain.
Thus, these preliminary analyses demonstrated that mutations in \( \text{Rv2959c} \) and \( \text{Rv2952} \) affected the production of \( p\)-HBAD, PGL-tb, and DIM but in different ways. The protein encoded by \( \text{Rv2959c} \) seems to be required for the biosynthesis of both \( p\)-HBAD and PGL-tb but not DIM, whereas that encoded by \( \text{Rv2952} \) appears to be required for the biosynthesis of PGL-tb and DIM in \( M. \) tuberculosis but not \( p\)-HBAD. Because \( p\)-HBAD and PGL-tb share a common glycosyl-phenolic domain absent from DIM, whereas DIM and PGL-tb exhibit a common lipid domain absent from \( p\)-HBAD, these results support our hypothesis, based on the amino acid sequences of the two putative methyltransferases, that the \( \text{Rv2959c} \) product is involved in the methylation of the glycosyl moiety and that the \( \text{Rv2952} \) product is involved in the methylation of the lipid domain.

**Structural Analysis of the Compounds Accumulated in the**
**\( M. \) tuberculosis H37Rv \( \Delta \text{Rv2952::km} \) Mutant**—The mutations in both genes resulted in the accumulation of new compounds (Fig. 3) that may be biosynthetic intermediates of \( p\)-HBAD, PGL-tb, and DIM. To test this hypothesis, we purified these substances by chromatography on Florisil and analyzed them further.

First, products B and C, which accumulated in the \( M. \) tuberculosis \( \Delta \text{Rv2952::km} \) mutant, were analyzed by MALDI-TOF mass spectrometry. The spectrum of glycoconjugate B showed a series of pseudomolecular ion \((M+H)^+\) peaks at 1850, 1864, 1878, 1892, 1906, 1920, 1934, 1948, 1962, 1976, 1990, and 2004 \( m/z \) (Fig. 4B) (the major peaks are underlined). The same peaks were observed in the mass spectrum of the purified PGL-tb.
Biosynthesis of Phenolglycolipids in M. tuberculosis

from the *M. tuberculosis* H37Rv:pPET1 strain (Fig. 4A), but the mass values of the major pseudomolecular ion peaks corresponding to glycolipid B from the PMM28:pPET1 mutant were 14 mass units lower than those in the spectrum of PGL-tb. This suggests that the PMM28:pPET1 mutant produces a PGL-like substance that may differ from PGL-tb by the absence of a methyl group. This hypothesis was supported by the \(^1\)H NMR analysis of the glycolipid B, which showed that the \(^1\)H NMR spectrum of the purified glycoconjugate B from PMM28:pPET1 was very similar to that of PGL-tb (Fig. 5A). Two deshielded doublets were observed at 6.97 and 7.10 ppm and corresponding to proton resonances of the phenolic group of PGLs (7, 22). Three anomic proton resonances were seen in the \(^1\)H NMR spectrum of the glycolipid B from PMM28:pPET1 at 5.50 ppm (1H) and 5.15 ppm (2H) (Fig. 5A). Their chemical shift values were identical to those of PGL-tb (7, 22). In addition, four singlets were observed in the region of the resonances of sugar-linked methoxyl (OCH\(_3\)) protons at 3.5–3.7 ppm. Again, the chemical shift values of the methoxyl proton resonances of product B (Fig. 5A) were identical to those found located on the trisaccharide portion of PGL-tb (7, 22). These data strongly suggest that compound B has the same carbohydrate moiety as PGL-tb. The occurrence of polymethylene (CH\(_2\)) units in the glycolipid was deduced from the presence of a broad signal resonance at 1.25 ppm. The resonances of the expected terminal methyl (CH\(_3\)) protons were observed at 0.8–1.0 ppm, whereas those attributable to methyl branches located on carbons 2 of fatty acyl residues were seen at 1.14 ppm. The resonance of the methine (CH) proton of the esterified \(\beta\)-glycol was seen at 4.83 ppm (Fig. 5A, signal a). Interestingly, the proton resonance typical of the methoxyl groups of the phthiocerol and phenolphthiocerol moiety of DIM A and PGL, respectively, expected at 3.32 ppm (singlet 3H) (7, 22, 23), was absent from the \(^1\)H NMR spectrum of glycolipid B (Fig. 5A). This observation was consistent with the absence of a signal at 2.85 ppm (multiplet, 1H) assignable to the resonance of the methine proton of the carbon bearing this methoxyl group (7, 22, 23). Instead, a broad signal was observed at 3.31 ppm (1H) in the \(^1\)H NMR spectrum of glycolipid B. This signal may correspond to a proton resonance of the carbon bearing a hydroxyl group (Fig. 5A, signal c). Altogether these data indicate that the glycolipid B from the PMM28:pPET1 mutant strain corresponds to a tri-O-methyl-fucoyl-(\(\alpha\)-1,3)-rhamnosyl-(\(\alpha\)-1,3)-2-O-rhamnosyl-\(\alpha\)-phenolphthiotriol dimycocerosates.

The same structural analyses were repeated with the glycoconjugate C purified from the PMM28:pPET1 mutant. MALDI-TOF mass spectrometry analysis of the product revealed a series of pseudomolecular ion (M+Na\(^+\)) peaks, at 2088, 2102, 2116, 2130, 2144, 2158, 2172, 2186, 2200, 2214, 2228, 2242, and 2256 m/z (Fig. 4C). These mass peaks were 238 mass units higher than those observed in the mass spectrum of the glycolipid B (Fig. 4B). Furthermore, the analysis of the \(^1\)H NMR spectrum of glycoconjugate C revealed the PGL nature of the glycolipid C in that the spectrum (Fig. 5B) was identical to that of glycoconjugate B (Fig. 5A) except for an additional broad signal at 4.7 ppm (1H), which may correspond to the proton resonance of a carbon bearing an ester group (Fig. 5B, signal c), and the absence of the signal assigned to the resonance of the methine proton of the carbon bearing the hydroxyl group of the phenolphthiotriol moiety of glycolipid B (Fig. 5A, signal c). These data indicate that the free OH group of phenolphthiotriol found in glycoconjugate B is esterified in compound C by a fatty acid the acyl moiety of which has 239 mass units, e.g. a palmityl residue.

As well as producing two new glycoconjugates, the PMM28:pPET1 mutant synthesized a new apolar lipid, product D, the structure of which was solved by both MALDI-TOF mass spectrometry and \(^1\)H NMR spectroscopy analyses. The mass spectrum revealed a series of pseudomolecular ion (M+Na\(^+\)) peaks at 1334, 1348, 1362, 1376, 1390, 1404, 1418, 1432, and 1446 m/z (data not shown). These mass values were 14 mass units lower than those observed in the mass spectrum of DIM A from wild-type strain (11). Comparison of the \(^1\)H NMR spectra of product D and DIM A from the wild-type H37Rv strain showed that all the signal resonances detected with DIM A were present (11) in product D accumulated in PMM28:pPET1 (Fig. 5C), with the notable exception of two signals at 3.32 ppm (singlet, 1H) and 2.85 ppm (multiplet, 1H). These latter signals correspond, respectively, to the proton resonances of the methoxyl group and the carbon bearing this group (11). Remarkably, a broad signal at 3.31 ppm was observed in the spectrum of product D (1H). This signal was absent from the spectrum of DIM A (11) but present in that of the triglycosylphenolphthiotriol dimycocerosates (product B) (Fig. 5A). This proton resonance may correspond to that of a hydroxyl group in place of the methoxyl group of DIM A. Altogether these results indicate that PMM28:pPET1 synthesizes a new apolar lipid corresponding to phthiotriol dimycocerosates. Accordingly, structural analyses of the three compounds that accumulated in the PMM28:pPET1 strain revealed that the insertional mutation in the \(Rv2952\) gene affected the conversion of the phenolphthiotriol and phthiotriol into their methoxylated forms usually found in PGL-tb and DIM A.

**Structural Characterization of the PGL from *M. tuberculosis* H37Rv \(\Delta\)Rv2952::km Mutant Strain**—We used first MALDI-TOF mass spectrometry to determine the structure of the ad-
ditional compound that accumulated in PMM18:pPET1 strain, compound A. The MALDI-TOF mass spectrum showed a series of pseudomolecular ion (M+Na+) peaks at 1516, 1530, 1544, 1558, 1572, 1586, 1600, 1614, 1628, and 1642 m/z (Fig. 6A). The spectrum was similar to that of the PGL-tb from the H37Rv: pPET1 strain (Fig. 4A), but the mass values of the major pseudomolecular ion peaks were 348 mass units lower. This mass difference corresponds to that calculated for the tri-O-methyl-fucosyl-rhamnosyl moiety (334 mass units) plus that corresponding to the loss of a methyl group (14 mass units). Consistently, the mass values observed for compound A correspond to those obtained with the 2-O-methyl-rhamnosyl-phe- nolphthiocol dimycocerosates produced by PMM24:pPET1 (16 minus 14 mass units (corresponding to the mass of a methyl group). Accordingly, we speculated that the sugar moiety of the glycoconjugate A accumulated in PMM18:pPET1 is an unmethylated rhamnosyl residue. To firmly establish the nature of the sugar residue, we performed an acid hydrolysis of the glycolipid A followed by trimethylsilylation and gas chromatography-mass spectrometry analysis. The mass spectra of the persilylated sugar derivative showed a (M+NH4+) mass peak at 470 m/z (in chemical ionization mode) and fragmentation peaks at 217, 204, 191, 147, and 73 m/z (in the electron impact mode). Identical mass spectra were obtained with a
standard perillylated 6-deoxyxugar, e.g. rhamnose. Finally, gas chromatography analysis of the sugar derivative from the acid hydrolysis of compound A identified rhamnose as the only sugar constituent of this new glycoconjugate (data not shown). These results suggest that the methyltransferase encoded by the \( \text{Rv2959c} \) gene catalyzes the transfer of the methyl group on rhamnol-syl-phenolphthiocerol dimycocerosates to form 2-O-methyl-rhamnol-syl-phenolphthiocerol dimycocerosate (also called mycoside B).

Production of a \( \text{M. tuberculosis} \) \( \text{H37Rv} \) Mutant Strain with an Unmarked Mutation within \( \text{Rv2959c} \) — The detection of rhamnosyl-phenolphthiocerol dimycocerosates in the lipid extract of \( \text{PMM18:pPET1} \), but not that of other glycoconjugates, such as the tri-O-methyl-fucosyl-di-rhamnosyl phenolphthiocerol found as a minor constituent of \( \text{M. tuberculosis} \) Canetti (24), suggests that the transfer of the terminal disaccharide unit found in PGL-tb was also impaired in the mutant strain. We reasoned that either the glycosyltransferase(s) involved in the transfer of the two terminal sugars (16) do not recognize the unmethylated rhamnosyl-phenolphthiocerol dimycocerosates as a substrate, or alternatively, the enzyme(s) is not produced within the \( \text{Rv2959c} \) gene. Therefore, these clones contained a deletion of 290 nucleotides within the \( \text{Rv2959c} \) gene and leaving behind a single \( \text{res} \) site. Thirteen of the 21 clones tested were sensitive to \( \text{km} \). PCR analysis of these clones using primers \( \text{res}1 \) and \( \text{res}2 \) (18) into \( \text{PMM18} \). The transformants were plated on antibiotic-free plates and incubated at 39 °C to cure the thermosensitive plasmid. Thirteen of the 21 clones tested were sensitive to \( \text{km} \). PCR analysis of these clones using primers \( \text{res}1 \) and \( \text{res}2 \) (18) into \( \text{PMM18} \) revealed that recombination occurred between the two \( \text{res} \) sites of the \( \text{res-Okm-res} \) inserted within the \( \text{Rv2959c} \) gene in \( \text{PMM18} \) (Fig. 2). Therefore, these clones contained a deletion of 290 nucleotides within the \( \text{Rv2959c} \) gene and an insertion of one \( \text{res} \) sequence at the same locus (Fig. 2). One of these clones, named \( \text{PMM18res} \), was retained for further analysis. Because no transcription terminator has been found within the \( \text{res} \) sequence, the occurrence of this sequence within the \( \text{Rv2959c} \) gene should not prevent the transcription of the \( \text{Rv2958c} \) gene.

Biochemical analysis of the \( \text{M. tuberculosis} \) \( \text{H37Rv} \) \( \Delta \text{Rv2959c-:res} \) Mutant Strain — To examine the effect of the unmarked mutation within \( \text{Rv2959c} \) on PGL-tb production, the plasmid \( \text{pPET1} \) was transferred into the \( \text{PMM18res} \) strain. Lipids were then extracted with organic solvents from the bacterial cells of a \( \text{PMM18res:pPET1} \) liquid culture. TLC analysis of these lipids revealed two glycoconjugates spots (Fig. 7). The compound that exhibited the lowest mobility on TLC corresponded to the unmethylated rhamnosyl-phenolphthiocerol dimycocerosates (product A) previously characterized in the lipids from the \( \text{PMM18:pPET1} \) mutant. The second glycoconjugate, product E, exhibited a higher mobility than product A (Fig. 7). This product was purified and analyzed by MALDI-TOF mass spectrometry and \(^1\)H NMR spectroscopy. The mass spectrum showed a series of
Biosynthesis of Phenolglycolipids in M. tuberculosis

Concerning the methylation of the saccharide domain of PGL-tb, we showed that the enzymes catalyzing the O-methylation of position 2 of the rhamnosyl residue linked to the phenol group in PGL-tb are encoded by \(Rv2959c\). We also demonstrated that this gene is part of an operon including a gene, \(Rv2958c\), encoding the glycosyltransferase involved in the transfer of additional sugar residues onto position 3 of this rhamnosyl unit (16). These results are also consistent with the enzyme encoded by \(Rv2959c\) being involved in the methylation of \(p\)-HBAD. Indeed, the \(\Delta Rv2959c::km\) mutant strain no longer synthesizes \(p\)-HBAD I or \(p\)-HBAD II. The unmethylated \(p\)-HBAD are probably too polar to be recovered from the culture supernatant in the organic phase after extraction using the protocol described by Bligh and Dyer (26), and as a consequence, are probably lost during the extraction procedure. However, this result supports the model proposed in the accompanying paper, according to which the same enzymatic machinery is involved in the formation of the saccharide moieties of \(p\)-HBAD and PGL-tb. The accumulation of unmethylated rhamnosyl-phenolphthiocerol dimycocerosates and tri-\(O\)-methylfucosyl-di-rhamnosyl-phenolphthiocerol dimycocerosates in the \(\Delta Rv2959c::res\) mutant demonstrates that other methyltransferases are involved in the O-methylation of the terminal fucose residue of PGL-tb. This was not unexpected because it was unlikely that the same methyltransferase enzyme would be involved in the O-methylation of various positions of different sugar units. Indeed, different methyltransferases are required for the methylation of three positions of the same rhamnosyl residue during the biosynthesis of glycopeptidolipids in \(M. smegmatis\) (27). The novel glycoconjugates that accumulated in the \(\Delta Rv2959c::res\) mutant are probably biosynthetic intermediates in the formation of mycoside B and PGL-tb and were previously identified in a \(M. tuberculosis\) wild-type strain as minor products (23, 24). Because no obvious gene candidate that may encode the methyltransferase(s) responsible for the O-methylation of the terminal fucosyl residue is present on the cluster of genes studied here, further studies are needed to identify this gene(s).

The two methyltransferases encoded by the \(Rv2952\) and \(Rv2959c\) genes are highly conserved in \(M. leprae\) (83 and 77% identity, respectively, over the entire protein with their orthologs \(ML0130\) and \(ML0127\)). For the enzyme involved in the methylation of the lipid domain of PGL and DIM, this was not unexpected because this domain is similar in these molecules in both species (4). However, this is not the case with the saccharide domain, which is a 3,6-di-\(O\)-methyl-glucosyl-di-\(O\)-methylrhamnosyl-3-\(O\)-methylrhamnossidase in \(M. leprae\), i.e. the methyl group at position 3 of the first rhamnosyl residue is located at a position different in PGL-tb. This suggests that, despite the very high sequence similarity between the two proteins encoded by \(Rv2959c\) and \(ML0130\), the two enzymes have different specificities in terms of positions of the hydroxyl groups of the rhamnosyl residue that would be methylated.

In this work and in the accompanying study (16) we have provided new insight into the biosynthesis of DIM, PGL, and \(p\)-HBAD. We have characterized the roles of five new enzymes involved in the formation of the saccharide domains of PGL-tb and \(p\)-HBAD and in the modification of the lipid moieties of PGL-tb and DIM A. However, the enzymes involved in several other biosynthetic steps remain to be identified. For instance, it has been suggested that \(p\)-hydroxybenzoic acid is the precursor of PGL (2), but the source of this molecule in mycobacteria remains to be established. Additional work is clearly required to complete our understanding of these biosynthetic pathways. Nevertheless, it is noteworthy that more than 1.5% of the \(M. tuberculosis\) genome is devoted to the synthesis and trans-
location of this group of molecules. The maintenance of such complex biosynthetic pathways, even in \textit{M. leprae}, which has undergone massive gene decay, strongly suggests that these molecules play essential roles for the biology of mycobacteria. These roles remain to be clarified, but there is considerable evidence suggesting that they are related to pathogenicity. The various recombinant \textit{M. tuberculosis} strains producing defined variants of DIM, PGL, and \(p\)-HBAD constructed during the course of the present study and the accompanying work (16) provide valuable tools for the characterization of the molecular mechanisms of action of these molecules in pathogenicity.

**Acknowledgments**—We are grateful to Dr. Henri Montrozier (Institut de Pharmacologie et Biologie Structurale (IPBS), Toulouse) for help with the biochemical analyses of the various glycoconjugates and to Françoise Viala (IPBS, Toulouse) for technical assistance in the preparation of the figures.

**REFERENCES**

1. World Health Organization (2002) *Fact Sheet 104*, World Health Organization, Geneva
2. Daffe, M., and Draper, P. (1998) *Adv. Microb. Physiol.*** **39**, 131–203
3. Asselineau, J. (1982) *Indian J. Chest. Dis.* **24**, 143–157
4. Daffe, M., and Lanéelle, M. A. (1988) *J. Gen. Microbiol.* **134**, 2049–2055
5. Brennan, P. J. (1988) in *Microbial Lipids* (Ratledge, C., and Wilkinson, S. G., eds.) pp. 203–298, Academic Press, Inc., London
6. Daffe, M., and Lemassu, A. (2000) in *Glycomicrobiology* (Doyle, R. J., ed) pp. 225–273, Plenum Press, New York
7. Constant, P., Perez, E., Malaga, W., Lanéelle, M.-A., Saurel, O., Daffe, M., and Guilhot, C. (2002) *J. Biol. Chem.* **277**, 38148–38158
8. Cole, S. T., Boshch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gao, S., Barrey, C. E., III, Tekapia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devin, K., Feltwell, T., Gentle, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krögh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajaedream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537–544
9. Camacho, L. R., Eusereguex, D., Perez, E., Gicquel, B., and Guilhot, C. (1999) *Mol. Microbiol.* **34**, 257–267
10. Cox, J. S., Chen, B., McNeil, M., and Jacobs, W. R., Jr. (1999) *Nature* **402**, 79–83
11. Camacho, L. R., Constant, P., Raynaud, C., Lanéelle, M.-A., Triccas, J.-A., Gicquel, B., Daffe, M., and Guilhot, C. (2001) *J. Biol. Chem.* **276**, 19845–19854
12. Azad, A. K., Sirakova, T. D., Fernandes, N. D., and Kolattukudy, P. E. (1997) *J. Biol. Chem.* **272**, 16741–16745
13. Azad, A. K., Sirakova, T. D., Rogers, L. M., and Kolattukudy, P. E. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4787–4792
14. Fitzmaurice, A. M., and Kolattukudy, P. E. (1998) *J. Biol. Chem.* **273**, 8033–8039
15. Trivedi, O. A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D., and Gokhale, R. S. (2004) *Nature* **428**, 441–445
16. Perez, E., Constant, P., Lemassu, A., Leval, F., Daffe, M., and Guilhot, C. (2004) *J. Biol. Chem.* **279**, 42574–42583
17. Fellay, R., Frey, J., and Krish, H. M. (1987) *Gene (Amst.)* **52**, 147–154
18. Malaga, W., Perez, E., and Guilhot, C. (2003) *FEBS Microbiol. Lett.* **219**, 261–268
19. Pelicic, V., Jackson, M., Reyrat, J. M., Jacobs, W. R., Gicquel, B., and Guilhot, C. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10955–10960
20. Pelicic, V., Reyrat, J. M., and Gicquel, B. (1996) *Mol. Microbiol.* **20**, 919–925
21. Jeevarajah, D., Patterson, J. H., McConville, M. J., and Billman-Jacobe, H. (2002) *Microbiology* **148**, 3079–3087
22. Daffe, M., Lacave, C., Lanéelle, M.-A., and Lanéelle, G. (1987) *Eur. J. Biochem.* **167**, 155–169
23. Daffe, M., Lanéelle, M.-A., Lacave, C., and Lanéelle, G. (1988) *Biochem. Biophys. Acta* **956**, 443–449
24. Daffe, M. (1989) *Biochim. Biophys. Acta* **1002**, 257–260
25. Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devin, K., Feltwell, T., Gentle, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajaedream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537–544
26. Bligh, E., and Dyer, W. J. (1959) *Can. J. Biochem. Physiol.* **37**, 911–917
27. Patterson, J. H., McConville, M. J., Haines, R. E., Coppell, R. L., and Billman-Jacobe, H. (2000) *J. Biol. Chem.* **275**, 24960–24966