THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 49, pp. 40699–40706, December 9, 2005

p-Hydroxybenzoic Acid Synthesis in *Mycobacterium tuberculosis*[*]

Received for publication, July 29, 2005, and in revised form, October 6, 2005 Published, JBC Papers in Press, October 6, 2005, DOI 10.1074/jbc.M508332200

Gustavo Stadthagen[‡], Jana Korduláková[‡], Ruth Griffin[§1], Patricia Constant[¶], Iveta Bottová[‡2], Nathalie Barilone[‡], Brigitte Gicquel[‡], Mamadou Daffé[¶], and Mary Jackson[‡3]

From the [‡]Unité de Génétique Mycobactérienne, Institut Pasteur, 75015, Paris, France, [§]National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom, and [¶]Département Mécanismes Moléculaires des Infections Mycobactériennes, Institut de Pharmacologie et de Biologie structurale, CNRS, Toulouse 31077 cedex 04, France

Glycosylated *p*-hydroxybenzoic acid methyl esters and structurally related phenolphthiocerol glycolipids are important virulence factors of *Mycobacterium tuberculosis*. Although both types of molecules are thought to be derived from *p*-hydroxybenzoic acid, the origin of this putative biosynthetic precursor in mycobacteria remained to be established. We describe the characterization of a transposon mutant of *M. tuberculosis* deficient in the production of all forms of *p*-hydroxybenzoic acid derivatives. The transposon was found to be inserted in *Rv2949c*, a gene located in the vicinity of the polyketide synthase gene *pks15/1*, involved in the elongation of *p*-hydroxybenzoate to phenolphthiocerol in phenolic glycolipid-producing strains. A recombinant form of the Rv2949c enzyme was produced in the fast-growing non-pathogenic *Mycobacterium smegmatis* and purified to near homogeneity. The recombinant enzyme catalyzed the removal of the pyruvyl moiety of chorismate to form *p*-hydroxybenzoate with an apparent K_m value for chorismate of 19.7 μM and a k_{cat} value of 0.102 s^{-1}. Strong inhibition of the reaction by *p*-hydroxybenzoate but not by pyruvate was observed. These results establish Rv2949c as a chorismate pyruvate-lyase responsible for the direct conversion of chorismate to *p*-hydroxybenzoate and identify Rv2949c as the sole enzymatic source of *p*-hydroxybenzoic acid in *M. tuberculosis*.

The chorismate pathway, present only in bacteria, fungi, and plants, provides a wealth of compounds with diverse biological functions, including aromatic amino acids, folate cofactors, menaquinones, ubiquinones, pigments, and iron-chelating siderophores. *Mycobacterium tuberculosis*, the etiological agent of tuberculosis in humans, is no exception to the rule, and chorismate in this species is the probable common precursor for the biosynthesis of a series of products (see Fig. 1) that are important both in the physiology and in the pathogenicity of the bacterium. In addition to folate and aromatic amino acids (phenylalanine, tyrosine, and tryptophan), *M. tuberculosis* produces salicylic acid-derived siderophores known as mycobactins (1), isoprenoid quinones of the naphthalene series (menaquinones) (2), and glycosylated *p*-hydroxybenzoic acid methyl esters (*p*-HBAD[4]s) (3). A few strains of *M. tuberculosis* also produce species-specific phenolic glycolipids (PGL), complex lipids of the cell envelope that share with *p*-HBADs the same glycosylated aromatic nucleus. These lipids and their counterparts found in *Mycobacterium leprae* have been largely associated with pathogenicity (4–9). Likewise, preliminary data indicate that *p*-HBADs, which are essentially recovered from the secretion products of the tubercle bacillus, modulate the secretion of pro-inflammatory cytokines by murine macrophages (10) and are required for virulence in immunodeficient mice (11).

Because of the important roles played by PGL and *p*-HBADs in the pathogenesis of mycobacterial infections, their biosynthesis has stimulated some interest, and, during the last decade, several genes involved in the synthesis of the lipid core of PGL and in the methylation and glycosylation of PGL and *p*-HBADs have been described (3, 12–22). Most of these genes are clustered on a 70-kilobase region of the chromosome. The synthesis of both *p*-HBADs and PGL is thought to proceed from *p*-hydroxybenzoic acid. Comparison of the chemical structures of these compounds suggested that, in the biosynthetic route of *p*-HBADs, *p*-hydroxybenzoic acid would be first methylated to *p*-hydroxybenzoic acid methyl ester before being glycosylated and further methylated. In the case of PGL, *p*-hydroxybenzoic acid would be elongated by the specialized polyketide synthase Pks15/1 to give *p*-hydroxyphenylalkanoates, which would be in turn converted to phenolphthiocerol derivatives in catalytic reactions involving the PpsA-E synthases (3). The resulting molecules or their diesters, the phenolphthiocerol dimycocerosates, would be then glycosylated and methylated by the same enzymes as those involved in the production of *p*-HBADs to yield the final PGL molecules (18, 19). The origin of *p*-hydroxybenzoic acid (4-HB) in mycobacteria remains, however, to be established.

In *Escherichia coli*, 4-HB formation from chorismate is the first committed step in ubiquinone synthesis. The reaction consists of the removal of the pyruvyl moiety of chorismate and is catalyzed by the enzyme chorismate pyruvate-lyase (*p*-hydroxybenzoic acid synthase) encoded by the *ubiC* gene (23, 24). The apparent lack of production of isoprenoid quinones of the benzene series (ubiquinones) in *M. tuberculosis* (2) and the absence of proteins sharing significant sequence similarities with UbiC in the genome of this bacterium made it unclear whether 4-HB formation in mycobacteria proceeded through the same route as in *E. coli*. In the present study, we report the isolation and characterization of a transposon mutant of *M. tuberculosis* deficient in the production of all forms of 4-HB derivatives and provide evidence

[*] This work was supported by the Institut Pasteur, the European Commission, within the 6th Framework Program contract number LSHP-CT-2003-503367, the Heiser Program for Research in Leprosy and Tuberculosis (to J. K.), the CONACyT program from Mexico (to G. S.), and the Marie Curie Training Site program number CT-2000-00058 from the European Commission (to I. B.). 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.

[1] Present address: Centre for Molecular Microbiology and Infection, Imperial College, London, United Kingdom.

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

[3] To whom correspondence should be addressed: Unité de Génétique Mycobactérienne, Institut Pasteur, 25 rue du Dr. Roux, 75015, Paris, France. Tel.: 33-1-45-68-88-77; Fax: 33-1-45-68-88-43; E-mail: mjackson@pasteur.fr.

[4] The abbreviations used are: *p*-HBAD, *p*-hydroxybenzoic acid methyl ester; 4-HB, *p*-hydroxybenzoic acid; PGL, phenolic glycolipids; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; GC, gas chromatography; GC/MS, GC/mass spectrometry.

that *Rv2949c*, the gene interrupted by the transposon insertion, encodes a chorismate pyruvate-lyase responsible for the formation of 4-HB from chorismate.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions—Mt103, the clinical isolate of *M. tuberculosis* used in this study, and the mutant 66C7 were grown at 37 °C in Middlebrook 7H9 medium (Difco) supplemented with ADC (0.2% dextrose, 0.5% bovine serum albumin fraction V, 0.085% NaCl, 0.0003% beef catalase) and 0.05% Tween 80, in minimal Sauton's medium (25) or on solid Middlebrook 7H11 medium (Difco) supplemented with OADC (0.005% oleic acid, 0.2% dextrose, 0.5% bovine serum albumin fraction V, 0.085% NaCl, 0.0003% beef catalase). *E. coli* XL1-blue, the strain used for cloning experiments, was propagated in Luria Bertani (LB) broth (pH 7.5) (10 g liter^{-1} Bactotryptone, 5 g liter^{-1} Bacto^{TM} yeast extract, 5 g liter^{-1} NaCl) (Becton Dickinson, Sparks, MD) at 37 °C. *M. smegmatis* mc^{2}155 (26) was grown in Middlebrook 7H9 medium (Difco) supplemented with ADC and 0.05% Tween 80 or in LB broth supplemented with 0.025% tyloxapol at 30 and 37 °C. Where indicated, ampicillin, chloramphenicol, kanamycin, and hygromycin B were added to final concentrations of 100 μg ml^{-1}, 25 μg ml^{-1}, 20 μg ml^{-1}, and 50 μg ml^{-1}, respectively.

Identification of the Transposon Insertion Site—The transposon insertion site in 66C7 was identified by ligation-mediated PCR as described previously (27). The primers used for amplification were Salgd (5′-tagcttattcctcaaggcacgagc-3′), ISada (5′-tttgagctctacaccgtcaagtgcgaagagc-3′), IS1 (5′-cttctgcagcaacgccaggtccacact-3′), and 66C7 (5′-ctatctgatcaagagatccg-3′). The products were sequenced on a capillary Applied Biosystems ABI 3100 Genetic Analyzer.

Complementation of M. tuberculosis Mutant Strain 66C7 and Overexpression of Rv2949c in M. smegmatis—Standard PCR strategies with Vent DNA polymerase (Biolabs) were used to amplify the *M. tuberculosis* H37Rv *Rv2949c* gene. PCR amplification consisted of predenaturation step (95 °C, 5 min) followed by 35 cycles of denaturation (95 °C, 1 min), annealing (60 °C, 1 min), and primer extension (72 °C, 1.5 min), and a final extension step at 72 °C for 10 min. The primers, Rv2949c.1 (5′-gggcgcccatatgaccgagtgttttctatctgatc-3′) and Rv2949c.2 (5′-ggaagcttgcgcgccagagtgatggcatttg-3′), incorporating NdeI and HindIII restriction sites (underlined), were designed to amplify the entire *Rv2949c* gene for direct cloning into the NdeI and HindIII sites of the expression vector pVV16 (28). The recombinant Rv2949c protein produced by the resultant vector, pVV2949c, carried a six-histidine tag at the carboxyl terminus. 66C7 and *M. smegmatis* mc^{2}155 were transformed with pVV2949c and transformants were selected for by plating on 7H11 agar containing kanamycin and hygromycin B. The production of recombinant Rv2949c protein in 66C7/pVV2949c and mc^{2}155/pVV2949c transformants was analyzed by immunoblotting with a mouse monoclonal anti-His antibody (Penta-His antibody, Qiagen) as previously described (28).

Extraction and Purification of Lipids—For the biochemical analyses of *p*-hydroxybenzoate derivatives, the *M. tuberculosis* strains Mt103, 66C7, and 66C7/pVV2949c were grown in Sauton's medium as surface pellicles. *p*-Hydroxybenzoate derivative production was investigated by thin-layer chromatography (TLC) analysis of total lipids extracted from bacterial cells and culture media. Culture filtrates were collected and filtered through 0.2-μm pore-size sterile filters to yield sterile extracellular materials. Exocellular lipids were then extracted by adding 2 volumes of CH_3OH and 1 volume of $CHCl_3$ to 0.8 volumes of extracellular materials to yield a homogeneous single-phase mixture. The mixture was incubated overnight at room temperature and then partitioned into two phases by adding 1 volume of $H_2O/CHCl_3$ (1:1, vol:vol). The organic phase was recovered, washed with 0.9% NaCl, and dried to yield the subcellular lipid extracts. Total lipids from bacterial cells were extracted with $CHCl_3/CH_3OH$ (1:2, vol:vol) for one night followed by two overnight extractions with $CHCl_3/CH_3OH$ (2:1, vol:vol). Mycolic acid methyl esters were prepared from extractable lipids and from delipidated cells by incubation with 15% tetrabutylammonium hydroxide (Aldrich) overnight at 100 °C followed by methylation with iodomethane (Aldrich) for 4 h at room temperature and extraction with dichloromethane. In some experiments, whole *M. tuberculosis* cells grown in 7H9 broth were radiolabeled overnight with [1,2-^{14}C]acetic acid (0.5 μCi ml^{-1}; specific activity, 113 Ci mol^{-1}, MP Biomedicals Inc.), a radiolabeled precursor incorporated in all classes of lipids. Lipids and fatty acid methyl esters were analyzed on silica gel 60-precoated TLC plates F_{254} (Merck) in the solvent systems $CHCl_3/CH_3OH/H_2O$ (90:10:1, vol:vol:vol) for *p*-HBADs, petroleum ether (60/80 °C)/ethyl acetate (98:2, vol:vol, three developments) for dimycocerosates of phthiocerol and *n*-hexane(s)/ethyl acetate (95:5, vol:vol, three developments) for fatty acid and mycolic acid methyl esters. An α-naphthol spray (1% α-naphthol in ethanol) and a cupric sulfate spray (10% $CuSO_4$ in a 8% phosphoric acid solution) were used to detect carbohydrate-containing lipids and all organic compounds, respectively. Radiolabeled lipids and fatty acids were visualized by exposure of TLC to Kodak BIOMAX MR films at −70 °C. Unmethylated *p*-HBAD-I standard was kindly provided by Drs. A. Liav and V. Vissa from Colorado State University (Fort Collins, CO). Cell-associated and exocellular mycobactins were extracted in the presence of ferric chloride as described previously (29) and analyzed by TLC using the solvent system ethanol/petroleum ether (60/80 °C)/ethyl acetate (1:4:6, vol:vol:vol).

p-HBADs were purified from the culture filtrate of the complemented strain 66C7/pVV2949c by preparative TLC using $CHCl_3/CH_3OH/H_2O$ (90:10:1, vol:vol:vol) as the eluent. The products of interest were scraped off and eluted from the silica with $CHCl_3/CH_3OH$ (2:1, vol:vol).

Matrix-assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry—MALDI mass spectrometry was performed using a voyager DE-STR MALDI-TOF instrument (PerSeptive Biosystems) equipped with a pulse nitrogen laser emitting at 337 nm. Samples were analyzed in the Reflector mode using an extraction delay time set at 100 ns and an accelerating voltage operating in positive ion mode of 20 kV. The mass spectra were mass-assigned by external calibration. Samples (1 μl of a 1 mg ml^{-1} solution in $CHCl_3$) were directly applied onto the sample plate. The matrix solution (0.5 μl of 2,5-dihydroxybenzoic acid at 10 mg ml^{-1} in $CHCl_3/CH_3OH$ (1:1, vol:vol)) was added. The samples were then allowed to crystallize at room temperature.

Chemical Analysis—4-HB derivatives were methanolized using CH_3OH/HCl prepared by the acetyl chloride reaction on methanol in anhydrous conditions. Samples were dissolved in 500 μl of CH_3OH/HCl (1 N) and incubated overnight at 80 °C under nitrogen atmosphere. The solvent was evaporated under nitrogen and then co-evaporated three times with anhydrous CH_3OH.

For trimethylsilyl derivatization, samples were dissolved in 200 μl of anhydrous pyridine, followed by the addition of 100 μl of hexamethyldisilazane and 50 μl of trimethylchlorosilane. The reaction was incubated at room temperature for 30 min. The mixture was dried under nitrogen, and the trimethylsilyl derivatives were solubilized in either petroleum ether, for gas chromatography (GC) and GC-mass spectrometry (GC/MS) or in $CHCl_3$ for MALDI-mass spectrometry.

GC and GC/MS Analysis—GC analyses were performed using a Girdel series 30 instrument equipped with an OV1 capillary column

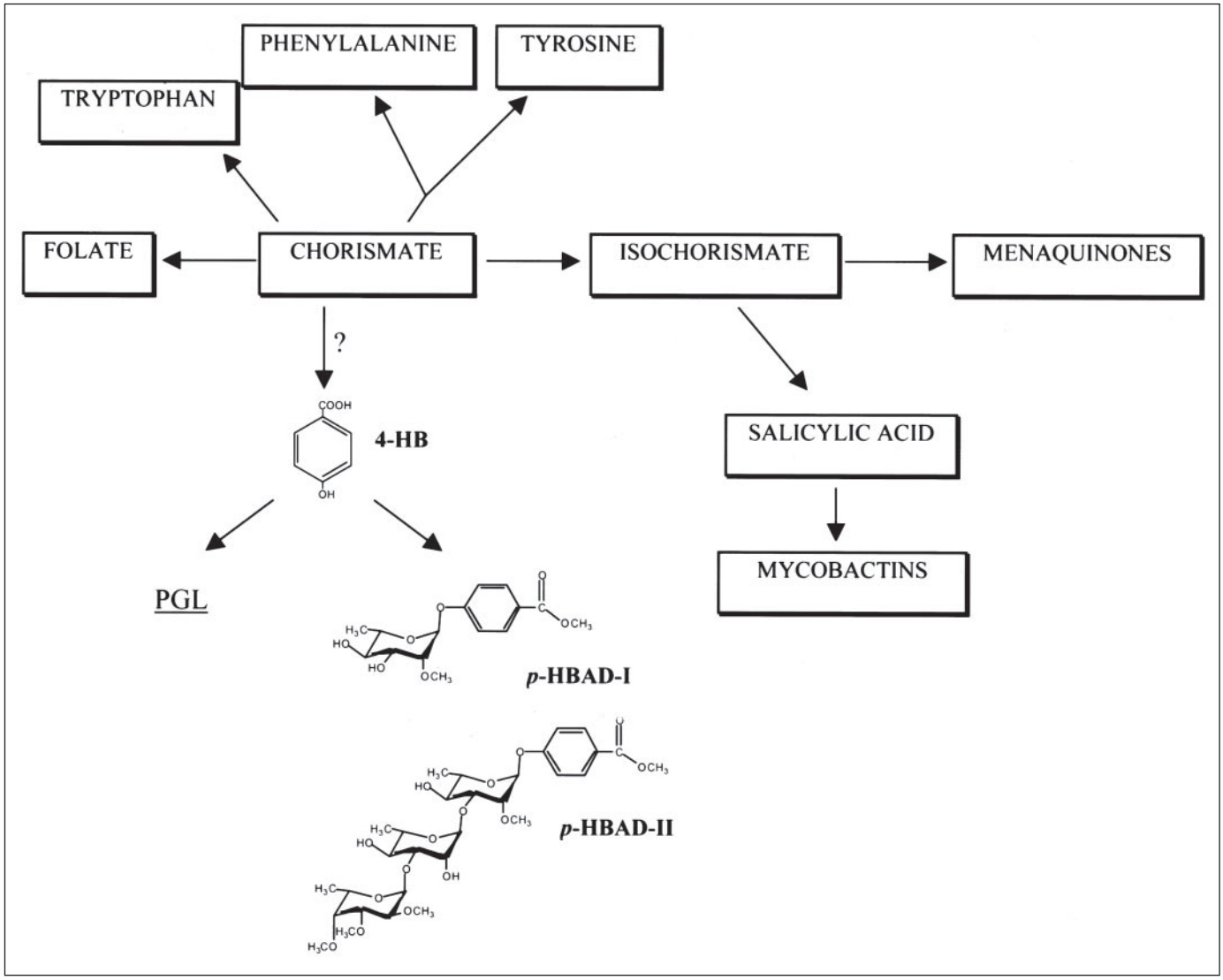

FIGURE 1. **Chorismate-utilizing pathways in *M. tuberculosis*.** The trisaccharide substituent found in *p*-HBAD-II, consists of 2,3,4-tri-*O*-methyl-α-L-Fuc*p*-(1→3)-α-L-Rha*p*-(1→3)-2-*O*-methyl-α-L-Rha*p*. The monosaccharide substituent found in *p*-HBAD-I consists of 2-*O*-methyl-α-L-Rha*p*.

(0.30 mm × 25 m) using helium gas (0.7 bar) with a flame ionization detector at 310 °C. The temperature program was from 60 to 310 °C, at 5 °C min^{-1}. GC/MS analyses were performed on a Hewlett-Packard 5889 X mass spectrometer (electron energy, 70 eV) working in electron impact (EI) mode, coupled with a Hewlett-Packard 5890 series II gas chromatograph fitted with a similar OV1 column (0.30 mm × 12 m).

Production and Purification of Recombinant Enzymes from E. coli and M. smegmatis—The *M. tuberculosis* H37Rv *Rv2949c* gene was amplified by PCR from genomic DNA using the pair of primers Rv2949c.tb1 (5′-ggccgcccatatgaccgagtgttttctatctgatc-3′) and Rv2949c.tb2 (5′-gccggatcctagcgcgccagagtgatggc-3′). Primers were designed to create NdeI and BamHI restriction sites (underlined) enabling direct cloning of the PCR product into the corresponding restriction sites of the expression vector pET14b (Novagen). The resulting expression plasmid, pETRv2949c was used to transform *E. coli* BL21(DE3)pLysS cells. Recombinant bacteria were grown overnight in LB-chloramphenicol-ampicillin broth at 37 °C and then transferred to 200 ml of the same medium at 30 °C. When A_{600} reached 0.7–0.8, isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 0.05 mM, and incubation was continued for an additional 3 h at 30 °C. Production of the recombinant protein was confirmed by SDS-PAGE with Coomassie Brilliant Blue 250R staining and by Western blot using a mouse monoclonal anti-His antibody (Penta-His antibody, Qiagen).

E. coli cells producing recombinant enzyme were washed and resuspended in buffer A (50 mM Tris-HCl, pH 7.5, 10 mM β-mercaptoethanol). Cells were disrupted by probe sonication on ice (4 cycles of 30 s on and 90 s off), and the resulting sonicate was centrifuged at 1,200 × *g* for 20 min. The pellet was discarded, and the supernatant containing the soluble *M. tuberculosis* His-tagged protein was loaded onto a cobalt-based immobilized metal affinity BD TALON™ Spin Column (Clontech). Unbound proteins were removed by washing the resin with buffer A containing 10 mM imidazole, and His-tagged protein bound to the resin was then gradually eluted with buffer A containing increasing concentrations of imidazole (50, 150, 300, 500, and 1,000 mM). The recombinant Rv2949c protein was detected by SDS-PAGE in the fractions eluted with 150, 300, 500, and 1,000 mM imidazole. Fractions that were estimated to be at least 90% pure by Coomassie Blue staining were pooled, desalted using a PD-10 column (Amersham Pharmacia Biotech), concentrated in an Amicon Ultra-15 (10,000 MWCO) (Millipore), and used for further characterization.

mc^2155/pVV2949c, the *M. smegmatis* strain overexpressing *Rv2949c*, was obtained upon transformation of mc^2155 with the expression vec-

tor pVV2949c. Cells were grown in LB-kanamycin-tyloxapol broth at 30 °C and collected by centrifugation, and the production of recombinant Rv2949c protein was analyzed by Western blotting as described above. For the purification of recombinant His-tagged Rv2949c protein, mc^2155/pVV2949c cells (1 g of wet weight) were washed and resuspended in 2 ml of buffer A before probe sonication for 9 min at 4 °C in the form of nine 60-s pulses with 90-s cooling intervals between pulses. The unbroken cells and bacterial debris were removed by centrifugation of the sonicate at 700 × *g* for 10 min, and the recombinant His-tagged Rv2949c protein was purified from the supernatant of this centrifugation using a BD TALON™ Spin Column as described for the recombinant protein produced in *E. coli*. Recombinant Rv2949c protein was detected in the fractions eluted with 150, 300, 500, and 1,000 mM imidazole. These fractions were pooled, desalted, and concentrated as described above.

Chorismate Pyruvate-lyase Assay—Chorismate pyruvate-lyase activity was monitored using a coupled photometric assay protocol similar to that of Siebert *et al*. (30). The release of pyruvate from chorismate, which occurs in the course of the reaction, was monitored by following NADH oxidation in a second reaction that was not rate-limiting. Decrease in absorbance was measured over time at 340 nm in a Beckman Coulter DU-800 spectrophotometer. The standard assay was performed at 37 °C and contained in a final volume of 1 ml: chorismate (25–500 μM) (Sigma), 50 mM Tris-HCl (pH 7.5), 200 μM NADH, 5.5 units of lactate dehydrogenase (Sigma), and 0.472 μM (10 μg) of purified recombinant Rv2949c enzyme. For the determination of kinetic parameters, incubations were carried out for 2 min to measure accurately initial reaction velocities and to minimize product-inhibition effects.

In some experiments (to analyze pH dependence and inhibition by pyruvate), NADH and lactate dehydrogenase were omitted from the reaction mixture, and the reactions were stopped at different time points by adding 600 μl of sodium acetate (0.75 M, pH 4). The *p*-hydroxybenzoate formed in the reaction was directly detected and quantitated by HPLC. 100 μl of each reaction mixture was applied to an HPLC system (Shimadzu) equipped with a reverse phase Hypersil ODS column (250 × 4.6 mm, particle size 5 μm) (Thermo). Elution was isocratic at 1 ml min^{-1} using water/acetonitrile/acetic acid (80:10:5, vol:vol:vol) as the solvent system. Analytes were detected at 240 nm.

RESULTS

Isolation of a Rv2949c Mutant of M. tuberculosis Mt103 from a Transposon Mutant Library—Mutants of *M. tuberculosis* Mt103 carrying transposon insertions in a region of the chromosome associated with the synthesis of PGL and structurally related glycosylated *p*-HBADs were recently isolated (10). One of these mutants, 66C7, which carried an insertion 237 base pairs downstream from the predicted start codon of the previously uncharacterized gene *Rv2949c*, was selected for further studies.

Rv2949c is located upstream from the *pks15/1* genes (*Rv2947c*/*Rv2946c*) in the genome of *M. tuberculosis* H37Rv and has orthologs in other PGL-producing species of mycobacteria such as *M. leprae*, *M. bovis*, and *M. marinum*. In fact, in the latter species, two orthologs of *Rv2949c* (sharing 75 and 67% similarity with Rv2949c at the amino acid level) were identified. At the amino acid level, Rv2949c (199 amino acids) shows some sequence similarities with putative chorismate pyruvate-lyases (*p*-hydroxybenzoic acid synthases) from Archaea (48% similarity on a 155-amino-acid overlap with the protein of *Archaeoglobus fulgidus*, 44% similarity on a 174-amino-acid overlap with the protein of *Methanococcus jannaschii*), and *Streptomyces capreolus* (44% similarity on a 180-amino-acid overlap) but no significant similarities with the

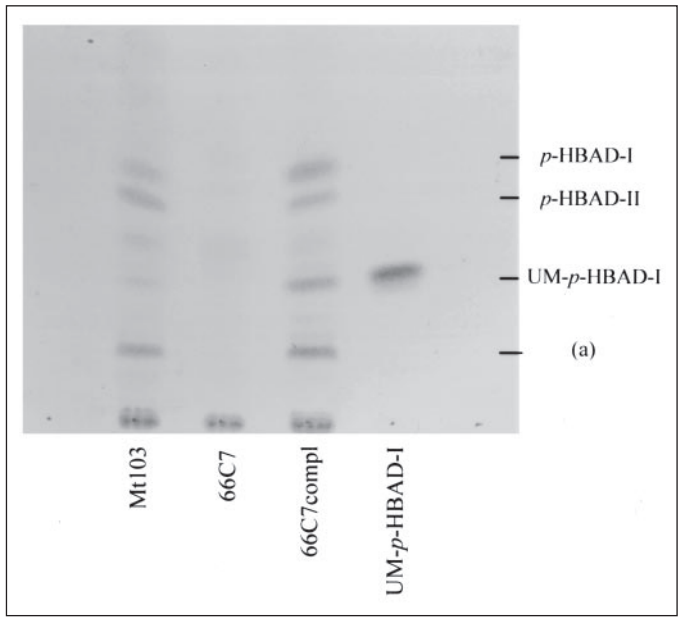

FIGURE 2. **Thin-layer chromatography analysis of lipids extracted from the culture supernatants of *M. tuberculosis* strains Mt103, 66C7, and 66C7/pVV2949c.** Equal volumes of lipid extracts prepared from culture filtrates of *M. tuberculosis* Mt103, 66C7, and 66C7/pVV2949c (*66C7compl*) were applied to a TLC plate and developed in $CHCl_3$/CH_3OH/H_2O (90:10:1). The plate was revealed with α-naphthol. 1 μg of unmethylated *p*-HBAD-I (*UM-p-HBAD-I*) was used as the migration standard.

chorismate pyruvate-lyase of *E. coli*, UbiC. As the location of *Rv2949c* on the chromosome suggested a possible involvement of this gene in the metabolism of *p*-HBADs and PGL, we sought to determine whether it participated in the synthesis of the aromatic nucleus of these important biological molecules.

Biochemical Phenotype of the Rv2949c Mutant—Important differences were found between the *p*-HBAD content of Mt103 and 66C7 (Fig. 2). Although wild-type Mt103 released into the culture medium the mono- and triglycosylated forms of *p*-hydroxybenzoic acid methyl esters, *p*-HBAD-I and *p*-HBAD-II (Fig. 1), there was no evidence for the production of either of these two glycosylated 4-HB derivatives in the culture medium or in bacterial cells of 66C7 (Fig. 2). 66C7 also clearly lacked two other glycoconjugates. One of these, UM-*p*-HBAD-I, was identified by co-migration with an authentic standard as being rhamnosyl-α-*p*-hydroxybenzoic acid methyl ester, a truncated form of *p*-HBAD-I carrying an unmethylated rhamnosyl residue. The second compound (*a*) was purified and first analyzed by MALDI-TOF mass spectrometry. The mass spectrum showed a pseudomolecular ion $(M + Na)^+$ peak at *m*/*z* 641 Da (Fig. 3*B*). This mass value is 28 mass units lower than that observed for *p*-HBAD-II (Fig. 3*A*), suggesting that compound *a* corresponds to an undermethylated form of *p*-HBAD-II. To determine the number of free hydroxyl groups in compound *a*, the substance was trimethylsilylated and reanalyzed by MALDI-TOF mass spectrometry. The mass spectrum of the trimethylsilylated compound *a* showed pseudomolecular ion peaks at *m*/*z* 713, 785, 857, 929, and 1,001 Da (Fig. 3*C*). The peak at *m*/*z* 1,001 Da corresponds to the pertrimethylsilylated compound *a* bearing five trimethylsilyl groups; the remaining pseudomolecular ion peaks correspond to undertrimethylsilylated forms of compound *a* that contain either one, two, three, or four trimethylsilyl groups. These data indicated that the native compound *a* contains five free hydroxyl groups. Thus, compared with *p*-HBAD-II, and in agreement with the 28 mass units default, compound *a* corresponds to a substance lacking two *O*-methyl substituents on the sugar moiety. Acid methanolysis of the purified compound *a*, followed by GC

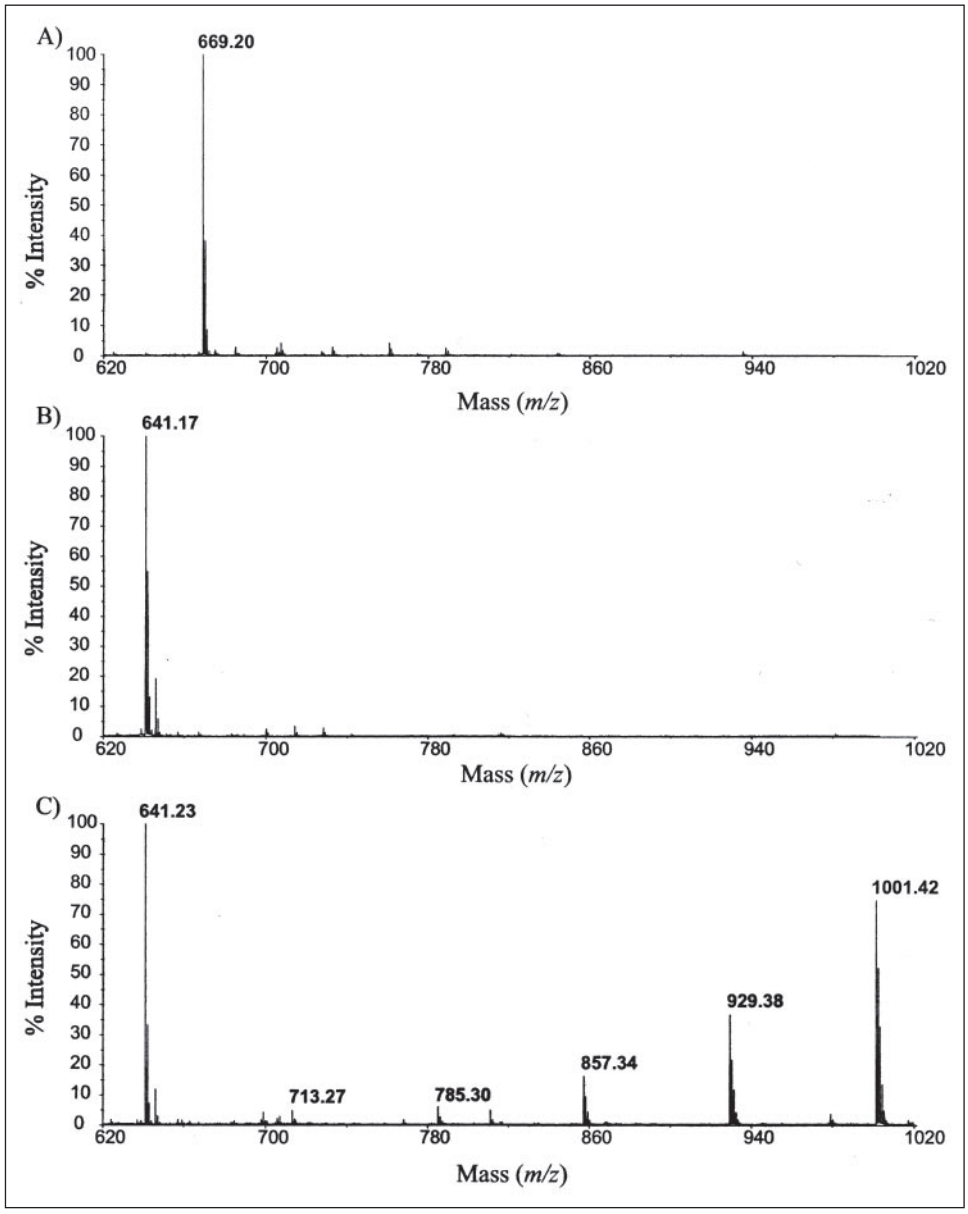

FIGURE 3. **MALDI-TOF mass spectra of *p*-HBAD-II (*A*), native (*B*), and trimethylsilylated (*C*) compound a.** Samples (1 μl of a 1 mg ml^{-1} solution in $CHCl_3$) were directly applied onto the sample plate and then allowed to crystallize at room temperature. Samples were analyzed in the Reflector mode using an extraction delay time set at 100 ns and an accelerating voltage operating in positive ion mode of 20 kV. The matrix solution used consisted of 2,5-dihydroxybenzoic acid (10 mg ml^{-1} in $CHCl_3$/CH_3OH (1:1, vol:vol)).

and GC-MS analyses of the trimethylsilyl derivatives of the methanolysate and of various standards, identified 2-*O*-methylfucose, 2-*O*-methylrhamnose, and rhamnose as the sugar constituents of compound *a* (data not shown). The non-carbohydrate moiety of the compound was identified by GC-MS as *p*-hydroxybenzoic acid methyl ester (data not shown). We thus concluded that compound *a* corresponds to the triglycosylated form of *p*-HBAD in which the trisaccharide part consists of 2-*O*-methyl-α-L-Fuc*p*-(1→3)-α-L-Rha*p*-(1→3)-2-*O*-methyl-α-L-Rha*p*. The 1→3 α-linkages between the three sugar units were assumed by analogy with the structure of *p*-HBAD-II (3). Compound *a* is either an intermediate product in the biosynthetic pathway leading to *p*-HBAD-II from *p*-HBAD-I or a degradation product of *p*-HBAD-II.

Complementation of 66C7 with a wild-type copy of *Rv2949c* placed under control of the *hsp60* promoter and carried on a multicopy plasmid clearly restored the production of all glycoconjugates in the mutant, including that of *p*-HBAD-I and *p*-HBAD-II (Fig. 2). Their synthesis, especially that of unmethylated *p*-HBAD-I (Fig. 2), appeared to be slightly increased in 66C7/pVV2949c as compared with Mt103, most likely because of the strong expression of *Rv2949c* in the complemented strain. Interestingly, only a tiny amount, if any, of free 4-HB could be detected by TLC in the different lipid batches prepared from the wild-type and complemented strains (data not shown) suggesting that this product is immediately metabolized upon synthesis. The impact of disrupting *Rv2949c* on PGL production could not be tested in the context of this study as the clinical isolate *M. tuberculosis* Mt103 is naturally devoid of these molecules (3).

TLC analyses of non radiolabeled and [1,2-^{14}C]acetate-derived lipids from Mt103 and 66C7 revealed no other quantitative or qualitative difference in the cell envelope composition or in the secreted products of the two strains (including their phthiocerol dimycocerosate content) (data not shown). Likewise, the same types and amounts of mycolates esterifying arabinogalactan and outer membrane glycolipids (predominantly trehalose mono- and dimycolates) were recovered from Mt103 and 66C7 (data not shown).

Finally, as chorismate pyruvate-lyase (involved in the formation of 4-HB from chorismate) and isochorismate pyruvate-lyase (responsible for the formation of salicylic acid from isochorismate) catalyze similar reactions, we verified whether the disruption of *Rv2949c* in 66C7 had

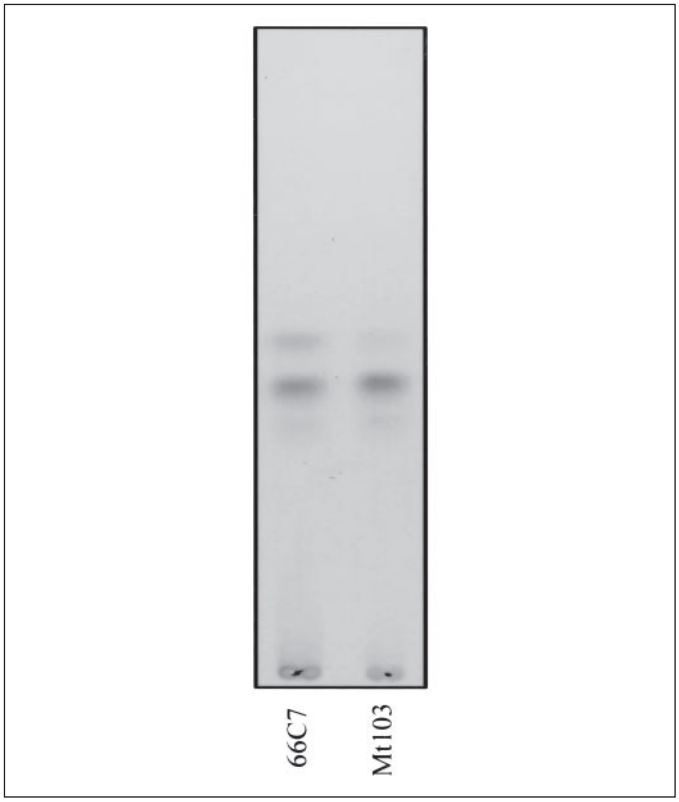

FIGURE 4. **Mycobactin production in Mt103 and 66C7.** Mycobactins were extracted from bacteria grown in low-iron Sauton's medium as described under "Experimental Procedures". They were then applied to a TLC plate and developed in the solvent system ethanol/petroleum ether (60°/80 °C)/ethyl acetate (1:4:6). Upon fixation of iron from ferric chloride during the extraction process, mycobactins develop a rusty color allowing their detection on the plate. Only the cell-associated mycobactins are shown.

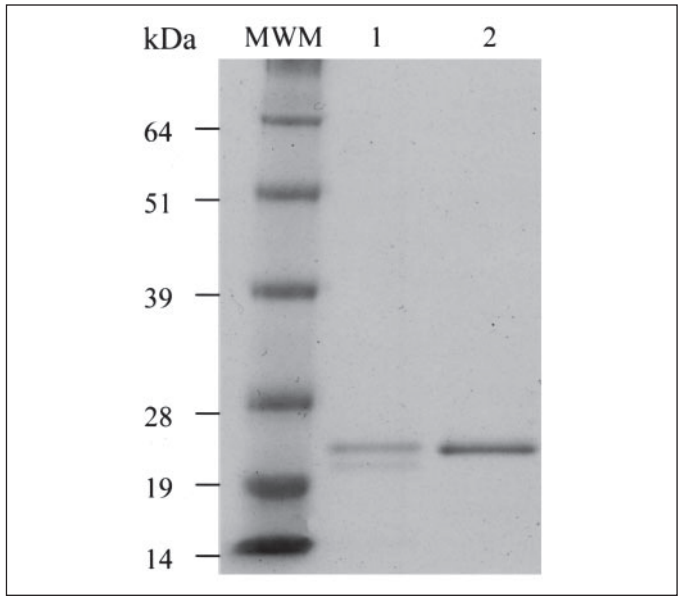

FIGURE 5. **Purification of recombinant forms of Rv2949c from *M. smegmatis* and *E. coli* BL21(DE3)pLysS.** 1 μg of purified recombinant proteins from *M. smegmatis* (*lane 1*) and from *E. coli* BL21DE3pLysS (*lane 2*) were submitted to SDS-PAGE and visualized by staining with Coomassie Brilliant Blue 250R. *MWM*, molecular weight marker.

any effect on the production of salicylic acid-derived mycobactins (Fig. 1). This analysis was further prompted by the fact that, to date, the genes involved in the formation of the salicylate moiety in mycobactin siderophores have not been clearly defined (1, 31). To conduct this analysis, wild-type Mt103 and 66C7 were grown in low iron Sauton's medium (in which ferric ammonium citrate was omitted) to stimulate siderophore production. Both the cell-associated and the secreted forms of mycobactin were extracted and analyzed by TLC (Fig. 4). No difference was found between the two strains indicating that *Rv2949c* is not required for mycobactin formation.

66C7 grew at the same rate as its parent strain Mt103 in 7H9 and in Sauton's medium at 37 °C (data not shown). In contrast, compared with the other strains, the growth of the complemented mutant, 66C7/pVV2949c, in Sauton's medium was delayed by ~5 weeks suggesting that overexpression of *Rv2949c* had some toxic effects on the cells. This observation was not unexpected with regard to the well known antibacterial and antifungal activities of 4-HB esters (32). Together, these results suggest that *Rv2949c* specifically encodes the enzyme responsible for the production of the aromatic nucleus of all *p*-HBADs and that it is the sole enzymatic source of 4-HB in *M. tuberculosis*.

Production of Recombinant Forms of Rv2949c in M. smegmatis and E. coli—To characterize the function of Rv2949c, recombinant forms of this protein were produced and purified. Because some studies aimed at analyzing the structure, post-translational modifications, and immunogenicity of mycobacterial proteins have highlighted the superiority of recombinant proteins purified from mycobacterial hosts compared with *E. coli*-derived products (33–36), both an *E. coli* and a *M. smegmatis* expression system were used and compared.

For the production of Rv2949c in *M. smegmatis*, the expression plasmid for complementation experiments described above was used. *M. smegmatis* mc^{2}155 was transformed with pVV2949c or with the empty pVV16 vector and recombinant clones were selected on LB-kanamycin plates at 30 °C. mc^{2}155/pVV2949c transformants grew very poorly in LB-tyloxapol broth at 37 °C but grew at about the same rate as mc^{2}155/pVV16 at 30 °C (data not shown). Production of recombinant carboxyl-terminal hexahistidine-tagged Rv2949c protein of the expected size (22.6 kDa) was confirmed by Western blot (data not shown); however, the production of the protein was found to be quite unstable as it was lost after more than two passages of the recombinant clones in liquid broth. Interestingly, when 7H9-ADC-Tween 80 broth was used instead of LB-Tween 80, mc^{2}155/pVV2949c transformants were able to grow similarly to the control strain at 30 and 37 °C, but the production of recombinant protein in this medium was much less than in LB broth (data not shown). Therefore, consistent with the observations made on 66C7/pVV2949c, overexpression of *Rv2949c* had some toxic effects on *M. smegmatis*, which could be partially reversed by growing the cells at 30 °C. mc^{2}155/pVV2949c grown in LB broth at 30 °C was, thus, used to produce Rv2949c, and the recombinant protein was purified to near homogeneity from these cells as described under "Experimental Procedures" (Fig. 5, *lane 1*).

Recombinant amino-terminal hexahistidine-tagged Rv2949c protein was also produced in *E. coli* using the pET14b expression system. Although more than 80% of the protein produced was in the insoluble fraction (under the form of inclusion bodies) (data not shown), sufficient amounts of the protein were recovered in the soluble extract to allow the purification of native recombinant Rv2949c from *E. coli* BL21(DE3)pLysS (Fig. 5, *lane 2*). Both forms of recombinant Rv2949c were used in enzymatic assays.

Characterization of Rv2949c as a Chorismate Pyruvate-lyase—Chorismate pyruvate-lyase produces 4-HB by releasing pyruvate from chorismate. To analyze and compare the catalytic activities of the recombinant proteins produced in *M. smegmatis* and in *E. coli*, an activity assay based on that developed for the UbiC protein of *E. coli* was used (30). The release of pyruvate from chorismate was monitored over time

using a coupled photometric assay that was not rate-limiting. 4-HB formation was also directly detected by HPLC.

No spontaneous (*i.e.* non-enzymatic) degradation of chorismate to 4-HB occurred under the conditions used in the assay over a 30-min incubation (data not shown). However, when either form of purified recombinant Rv2949c (amino-terminal or carboxyl-terminal His-tagged) was added to the reaction mixture, pyruvate was released from chorismate to form 4-HB in a time-dependent fashion. The substrate saturation curves were hyperbolic for both enzymes indicating that they followed Michaelis Menten's kinetics. Non-linear regression analysis was conducted on the curves, and steady-state kinetic parameters were calculated (Fig. 6). The apparent K_m values for chorismate were determined as 19.7 μM for the recombinant Rv2949c produced in *M. smegmatis* (Fig. 6*A*) and as 39.6 μM for the enzyme produced in *E. coli* (Fig. 6*B*). k_{cat} values were determined as 0.102 and 0.215 s^{-1}, respectively. These k_{cat} values are about 10-fold lower than those reported for the *E. coli* UbiC protein (1.5 s^{-1}) (37). Comparable kinetic parameters were obtained when pyruvate release was monitored using the coupled photometric assay or when 4-HB formation was directly quantified by HPLC (data not shown).

The addition of 10 mM $MgCl_2$ to the reaction buffer had neither a stimulatory nor an inhibitory effect on the activity of Rv2949c produced in *M. smegmatis* suggesting that this enzyme, like the *E. coli* UbiC protein, has no requirement for Mg^{2+} (data not shown). pH dependence was also studied using buffers containing Tris-HCl and glycine-NaOH. The enzyme was found to be active over a range of pH between 6.8 and 10 with an optimum pH at 7.5 (data not shown). As reported for the UbiC enzyme, a strong product inhibition was observed in the presence of 4-HB. Using a chorismate concentration close to the apparent K_m value, *i.e.* 25 μM, the addition of 100 μM 4-HB to the reaction mixture reduced pyruvate formation by 86% after 10 min of incubation. Pyruvate formation after the same incubation time was reduced by 65 and 46% when chorismate concentration was raised to 250 and 500 μM, respectively. This partial reversion of inhibition in the presence of excess chorismate is suggestive of a competitive inhibition. The initial rate inhibition because of 4-HB was measured, and the inhibition constant (K_i) for this product was determined to be 6 μM, close to the K_i value for 4-HB of UbiC (2.1 μM) (37). In contrast, using a chorismate concentration of 100 μM, pyruvate did not show any inhibitory effect at concentrations up to 1 mM (data not shown). Finally, in the presence of 10% glycerol, the purified recombinant enzymes could be stored at −20 °C for at least 3 weeks without significant loss of activity.

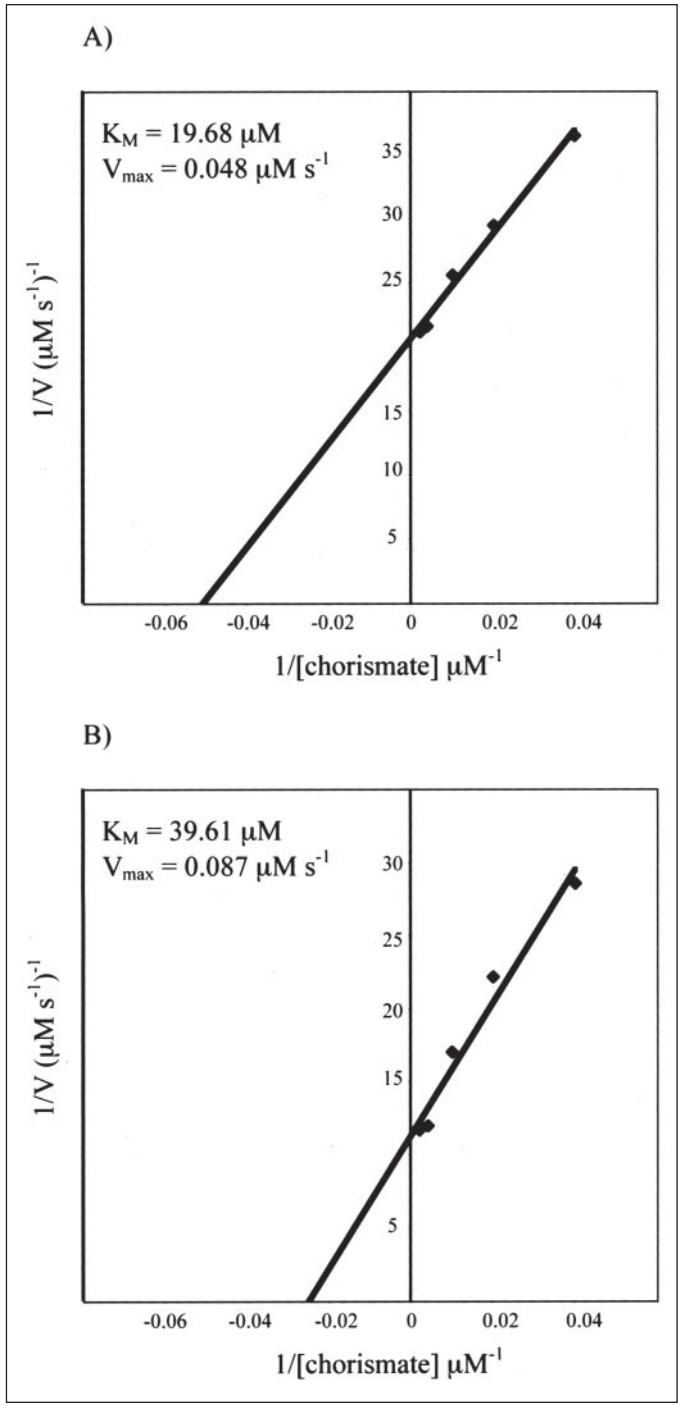

FIGURE 6. **Double reciprocal plots of initial velocity *versus* chorismate concentration.** Purified recombinant forms of Rv2949c produced in *M. smegmatis* (*A*) and *E. coli* (*B*) were used. Incubations were carried out for 2 min to measure accurately initial reaction velocities and to minimize product-inhibition effects. The chorismate concentration ranged from 25 to 500 μM.

DISCUSSION

This study establishes that *Rv2949c* encodes a chorismate pyruvate-lyase (4-HB synthase) responsible for the production of 4-HB from chorismate and that it is the sole enzymatic source of this product in *M. tuberculosis*. Interestingly, Rv2949c shares some sequence similarities with putative chorismate pyruvate-lyases from Archaea and *Streptomyces capreolus* but shows no significant similarity with the chorismate pyruvate-lyase of *E. coli*, UbiC. Although the three-dimensional structure of Rv2949c remains to be determined, and compared with that of other chorismate-utilizing enzymes, this finding tends to support the earlier observation that these enzymes are examples of convergent evolution toward similar reaction capabilities rather than enzymes arising from divergent evolution (31).

From our complementation experiment, it seems that in *M. tuberculosis* Mt103, 4-HB is immediately metabolized into *p*-HBADs. Although one cannot totally exclude from our analyses that 4-HB is the precursor of multiple end products, it seems likely that *p*-HBADs (and PGL in phenolglycolipid-producing strains) are the major metabolic end products derived from 4-HB produced by tubercle bacilli. *M. tuberculosis* strains are apparently devoid of ubiquinones (2), and, to our knowledge, no other complex compound sharing the same aromatic nucleus as *p*-HBADs has yet been described in this species. It is, therefore, puzzling to observe that all *M. tuberculosis* strains have retained the ability to produce such complex molecules. The fact that the 66C7 mutant did not exhibit any growth defect when cultured under axenic conditions argues against a major role of 4-HB derivatives in the *in vitro* physiology

of the tubercle bacillus. The structure of phenolic glycolipids and the observation that some of the PGL species produced by *M. marinum* were *O*-acylated with mycoloyl residues led Gastambide-Odier (38) to propose that these molecules acted as carriers of key constituents of the cell envelope. Accordingly, a thorough analysis of the lipid and mycolic acid content of 66C7 was undertaken in the present study to investigate this possibility, and, as no qualitative or quantitative difference was found between the mutant and the wild-type strain, we conclude that *p*-HBADs are unlikely to be carriers of key envelope constituents.

A more likely explanation for the conservation of *p*-HBADs and the occurrence of PGL in some *M. tuberculosis* strains is related to their role in pathogenesis. Remarkably, glycosylated phenolphthiocerol diesters are only found in a few *Mycobacterium* species, namely *M. ulcerans*, *M. marinum*, *M. tuberculosis*, *M. bovis*, *M. leprae*, *M. kansasii*, *M. gastri*, *M. microti*, and *M. hemophilum* (39). With the exception of *M. gastri*, all of these species are pathogenic for humans. Although PGL are well known virulence factors of *M. leprae* playing pleiotropic roles in the course of leprosy (5–8), evidence for the contribution of 4-HB derivatives to the pathogenesis of tuberculosis has been provided only recently. It was shown that the low innate immunogenicity and hypervirulence of a Beijing isolate of *M. tuberculosis* was associated with the presence of PGL in this strain (9). Moreover, a *Rv2958c* mutant of *M. tuberculosis* H37Rv, likely to be affected in the production of the di- and triglycosylated forms of *p*-HBADs (18), exhibited a decreased virulence in immunodeficient mice (11). Finally, our data indicate that a *Rv2958c* mutant constructed in *M. tuberculosis* Mt103 induces mouse bone marrow-derived macrophages to secrete more proinflammatory cytokines than the wild-type strain (10). Thus, it is clear that both *p*-HBADs and PGL play important roles in the modulation of the host immune response and that the glycosylated aromatic nucleus of these molecules is directly involved in some of their biological activities. The availability of 66C7, a *M. tuberculosis* mutant devoid of all forms of *p*-HBADs, now provides a unique opportunity to measure the contribution of these molecules to immunopathogenesis.

Acknowledgments—We thank C. Chaput and I. Boneca for their assistance with HPLC, H. Munier-Lehman for helpful discussion, and K. Davenport and K. Mikusová for critical reading of the manuscript.

REFERENCES

1. De Voss, J. J., Rutter, K., Schroeder, B. G., and Barry, C. E., III (1999) *J. Bacteriol.* **181,** 4443–4451
2. Collins, M. D., and Jones, D. (1981) *Microbiol. Rev.* **45,** 316–354
3. Constant, P., Pérez, E., Malaga, W., Lanéelle, M.-A., Saurel, O., Daffé, M., and Guilhot, C. (2002) *J. Biol. Chem.* **277,** 38148–38158
4. Brennan, P. J. (1988) in *Microbial Lipids* (Ratledge, C., and Wilkinson, S. G., eds) Vol. 1, pp. 203–298, Academic Press Ltd., London
5. Chan, J., Fujiwara, T., Brennan, P. J., McNeil, M., Turco, S. J., Sibille, J.-C., Snapper, M., Aisen, P., and Bloom, B. R. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 2453–2457
6. Puzo, G. (1990) *Critic. Rev. Microbiol.* **17,** 305–327
7. Ng, V., Zanazzi, G., Timpl, R., Talts, J. F., Salzer, J. L., Brennan, P. J., and Rambukkana, A. (2000) *Cell* **103,** 511–524
8. Schlesinger, L. S., and Horwitz, M. A. (1991) *J. Exp. Med.* **174,** 1031–1038
9. Reed, M. B., Domenech, P., Manca, C., Su, H., Barczak, A. K., Kreiswirth, B. N., Kaplan, G., and Barry, C. E., III (2004) *Nature* **431,** 84–87
10. Griffin, R., Stadthagen Gomez, G., Young, D., Coade, S., Bottová, I., Nigou, J., Brando, T., Puzo, G., Gicquel, B., Lowrie, D., Jackson, M., and Tascon, R. E. (2005) *Keystone Symposium, British Columbia, April 2–7, 2005, Bill and Melinda Gates Foundation*, Whistler, British Columbia, Canada
11. Hisert, K. B., Kirksey, M. A., Gomez, J. E., Sousa, A. O., Cox, J. S., Jacobs Jr., W. R., Nathan, C. F., and McKinney, J. D. (2004) *Infect. Immun.* **72,** 5315–5321
12. Azad, A. K., Siraková, T. D., Rogers, L. M., and Kolattukudy, P. E. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 4787–4792
13. Azad, A. K., Siraková, T. D., Fernandes, N. D., and Kolattukudy, P. E. (1997) *J. Biol. Chem.* **272,** 16741–16745
14. Fitzmaurice, A. M., and Kolattukudy, P. E. (1998) *J. Biol. Chem.* **273,** 8033–8039
15. Camacho, L. R., Ensergueix, D., Pérez, E., Gicquel, B., and Guilhot, C. (1999) *Mol. Microbiol.* **34,** 257–267
16. Cox, J. S., Chen, B., McNeil, M., and Jacobs Jr., W. R. (1999) *Nature* **402,** 79–83
17. Camacho, L. R., Constant, P., Raynaud, C., Lanéelle, M.-A., Triccas, J. A., Gicquel, B., Daffé, M., and Guilhot, C. (2001) *J. Biol. Chem.* **276,** 19845–19854
18. Pérez, E., Constant, P., Lemassu, A., Laval, F., Daffé, M., and Guilhot, C. (2004) *J. Biol. Chem.* **279,** 42574–42583
19. Pérez, E., Constant, P., Laval, F., Lemassu, A., Lanéelle, M.-A., Daffé, M., and Guilhot, C. (2004) *J. Biol. Chem.* **279,** 42584–42592
20. Onwueme, K. C., Ferreras, J. A., Buglino, J., Lima, C. D., and Quadri, L. E. N. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 4608–4613
21. Trivedi, O. A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D., and Gokhale, R. S. (2004) *Nature* **428,** 441–445
22. Trivedi, O. A., Arora, P., Vats, A., Ansari, M. Z., Tickoo, R., Sridharan, V., Mohanty, D., and Gokhale, R. S. (2005) *Mol. Cell* **17,** 631–643
23. Siebert, M., Bechthold, A., Melzer, M., May, U., Berger, U., Schröder, G., Schröder, J., Severin, K., and Heide, L. (1992) *FEBS Lett.* **307,** 347–350
24. Nichols, B. P., and Green, J. M. (1992) *J. Bacteriol.* **174,** 5309–5316
25. Sauton, M. B. (1912) *C. R. Acad. Sci.* **155,** 860–861
26. Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., and Jacobs, W. R. (1990) *Mol. Microbiol.* **4,** 1911–1919
27. Prod'hom, G., Lagier, B., Pelicic, V., Hance, A. J., Gicquel, B., and Guilhot, C. (1998) *FEMS Microbiol. Lett.* **158,** 75–81
28. Kordulaková, J., Gilleron, M., Mikusová, K., Puzo, G., Brennan, P. J., Gicquel, B., and Jackson, M. (2002) *J. Biol. Chem.* **277,** 31335–31344
29. De Voss, J. J., Rutter, K., Schroeder, B. G., Su, H., Zhu, Y., and Barry III, C. E. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 1252–1257
30. Siebert, M., Severin, K., and Heide, L. (1994) *Microbiology* **140,** 897–904
31. Dosselaere, F., and Vanderleyden, J. (2001) *Crit. Rev. Microbiol.* **27,** 75–131
32. Soni, M. G., Taylor, S. L., Greenberg, N. A., and Burdock, G. A. (2002) *Food Chem. Tox.* **40,** 1335–1373
33. Garbe, T., Harris, D., Vordermeier, M., Lathigra, R., Ivanyi, J., and Young, D. (1993) *Infect. Immun.* **61,** 260–267
34. Triccas, J. A., Roche, P. W., Winter, N., Feng, C. G., Butlin, C. R., and Britton, W. J. (1996) *Infect. Immun.* **64,** 5171–5177
35. Roche, P. W., Winter, N., Triccas, J. A., Feng, C., and Britton, W. J. (1996) *Clin. Exp. Immunol.* **103,** 226–232
36. Daugelat, S., Kowall, J., Mattow, J., Bumann, D., Winter, R., Hurwitz, R., and Kaufmann, S. H. (2003) *Microbes Infect.* **5,** 1082–1095
37. Holden, M. J., Mayhew, M. P., Gallagher, D. T., and Vilker, V. L. (2002) *Biochim. Biophys. Acta* **1594,** 160–167
38. Gastambide-Odier, M. (1973) *Eur. J. Biochem.* **33,** 81–86
39. Daffé, M., and Draper, P. (1998) *Adv. Microb. Physiol.* **39,** 131–203

p-Hydroxybenzoic Acid Synthesis in *Mycobacterium tuberculosis*

Gustavo Stadthagen, Jana Korduláková, Ruth Griffin, Patricia Constant, Iveta Bottová, Nathalie Barilone, Brigitte Gicquel, Mamadou Daffé and Mary Jackson

J. Biol. Chem. 2005, 280:40699-40706.
doi: 10.1074/jbc.M508332200 originally published online October 6, 2005
