Table 1. Positivity rates of MMP-II and PGL-I tests in various groups of subjects by an enzyme-linked immunosorbent assay

	MMP-II				PGL-I				McNemar test (MMP-II vs. PGL)	Rate difference (95% CI)	Inter-rater agreement (κ)
	Tested	Positive	%	95% CI	Tested	Positive	%	95% CI			
Multi-bacillary leprosy	74	61	82.4	71.8–90.3	74	51	68.9	57.1–79.2	$P=0.063$	13.5% (−0.7 to 24.2)	0.14
Pauci-bacillary leprosy	77	30	39	28.8–50.1	77	15	19.5	12.2–29.7	$P=0.007$	19.5% (5.4 to 28.9)	0.189
Tuberculosis	55	9	16.4	8.9–28.3	ND	ND	ND	ND	–	–	–
Healthy subjects	81	8	9.9	5.1–18.3	81	8	9.9	5.1–18.3	–	–	–

ND, not detected.

Therefore, a systematic study was conducted, measuring the anti-MMP-II antibody IgG levels in the leprosy patients' sera. The study population consisted of multi-bacillary leprosy, pauci-bacillary leprosy, tuberculosis patients, and normal healthy BCG-vaccinated volunteers from Japan. Some of the leprosy patients were already under treatment, so that all patients were not active leprosy patients. The cut-off value of OD 0.130 was defined by an ROC curve analysis (MEDCALC software) using the OD titers from 81 normal individuals and 74 multi-bacillary leprosy patients. Using this cut-off value, it was observed that 61 (82.4%, 95% CI; 71.8–90.3) out of 74 multi-bacillary patients had positive results, and 30 (39.0%, 95% CI; 28.8–50.1) out of 77 pauci-bacillary patients had positive titers (Table 1). The only serological test for leprosy that is currently available is the detection of antibodies to PGL-I of *M. leprae*. The haptenic trisaccharide of PGL-I is known to be *M. leprae* specific, and this trisaccharide unit could be chemically synthesized (Fujiwara *et al.*, 1987). Several reports show the performance of PGL-I for serodiagnosis (Agis *et al.*, 1988; Cho *et al.*, 1991). When the anti-PGL-I IgM antibody levels were examined in the same Japanese leprosy patients by ELISA, it was found that only 68.9% (95% CI; 57.1–79.2) of multi-bacillary patients ($n=74$) and 19.5% (95% CI; 12.2–29.7) of pauci-bacillary patients ($n=77$) showed positive values (Table 1). These percentages were far lower than expected, which may be due to the influence of chemotherapy. However, the percent positivity for anti-MMP-II antibodies was significantly higher than that for anti-PGL-I antibodies of the same sera ($P=0.0008$, McNemar test, $n=152$) when both multi-bacillary and pauci-bacillary leprosy were considered together. When multi-bacillary sera were taken separately, the significance of the MMP-II test was statistically marginal compared with that of the PGL-I test ($P=0.06$), but there was a significant difference ($P=0.007$) between the two tests for pauci-bacillary leprosy (Table 1). The agreement between these two tests was low (κ value for multi-bacillary leprosy: 0.140, κ value for pauci-bacillary leprosy: 0.189), so that when positive titers for either or both PGL-I and MMP-II were considered, the percent positivity increased to 91.9% in multi-bacillary patients and 48.7% in pauci-bacillary patients. The specificity of both the MMP-II and PGL-I tests was 90.1%.

When the anti-MMP-II IgG levels were measured in the normal individuals (BCG-vaccinated), a low level of seropositivity was found, 9.9% (95% CI; 5.1–18.3) ($n=81$). As the amino acid homology between MMP-II protein from *M. leprae* and its homologue in *M. bovis* BCG is 90.6%, it was expected that a higher percentage of normal individuals would be positive. But this result was to the contrary and advantageous to the assay system. Also, tuberculosis patients showed a low positive result of 16.4% (95% CI; 9.9–28.3) ($n=55$). As, the genes encoding MMP-II were conserved between *M. leprae* and *M. tuberculosis*, the low positivity in tuberculosis was unexpected. The exact reason for the low positivity in tuberculosis patients is not clear, but might be due to slight conformational differences in the antibody recognition site on MMP-II between *M. leprae* and *M. tuberculosis*. More intensive studies need to be carried out using active tuberculosis patients and native MMP-II protein derived from *M. tuberculosis* to fully realize the significance of MMP-II homologue in the detection of tuberculosis.

There is only one report showing that the sera from leprosy patients had a higher IgG titer to MMP-II, regardless of the clinical type of leprosy (Deshpande *et al.*, 1995). However, their study was carried out with a limited number of individuals ($n=10$ in each group), and so, here, a more systematic study has been conducted for evaluating the expediency of measuring anti-MMP-II antibody levels. In addition, when monitoring of patients on multidrug treatment was conducted ($n=4$) for a period of 2 years, there was a definite decline in the MMP-II antibody levels after chemotherapy (data not shown). Further intensive studies have to be carried out with a larger cohort of patients to fully realize the significance of MMP-II in the monitoring of chemotherapy.

This report showed that MMP-II could contribute to the sero-detection of multi-bacillary as well as pauci-bacillary leprosy patients. Further study will be pursued to evaluate its efficacy for serodiagnosis in developing countries and for the development of user-friendly tools.

Acknowledgement

This work was supported by grants from Health Sciences Research Grants-Research on Emerging and Re-emerging

Infectious Diseases, from the Ministry of Health, Labour and Welfare, Tokyo, Japan.

References

Agis F, Schlich P, Cartel JL, Guidi C & Bach MA (1988) Use of anti-*M. leprae* phenolic glycolipid-I antibody detection for early diagnosis and prognosis of leprosy. *Int J Lepr Other Mycobact Dis* **56**: 527–535.

Buhrer-Sekula S, Smits HL, Gussenhoven GC, Leeuwen van J, Amador S, Fujiwara T, Klatser PR & Oskam L (2003) Simple and fast lateral flow test for classification of leprosy patients and identification of contacts with high risk of developing leprosy. *J Clin Microbiol* **41**: 1991–1995.

Cho SN, Cellona RV, Fajardo TT Jr, Abalos RM, dela Cruz EC, Walsh GP, Kim JD & Brennan PJ (1991) Detection of phenolic glycolipid-I antigen and antibody in sera from new and relapsed lepromatous patients treated with various drug regimens. *Int J Lepr Other Mycobact Dis* **59**: 25–31.

Deshpande RG, Khan MB, Bhat DA & Navalkar RG (1995) Immunoaffinity chromatographic isolation of a high molecular weight seroreactive protein from *Mycobacterium leprae* cell sonicate. *FEMS Immunol Med Microbiol* **11**: 163–169.

Fujiwara T, Hunter SW, Cho SN, Aspinall GO & Brennan PJ (1984) Chemical synthesis and serology of disaccharides and trisaccharides of phenolic glycolipid antigens from the leprosy bacillus and preparation of a disaccharide protein conjugate for serodiagnosis of leprosy. *Infect Immun* **43**: 245–252.

Fujiwara T, Aspinall GO, Hunter SW & Brennan PJ (1987) Chemical synthesis of the trisaccharide unit of the species-specific phenolic glycolipid from *Mycobacterium leprae*. *Carbohydr Res* **163**: 41–52.

Hunter SW & Brennan PJ (1981) A novel phenolic glycolipid from *Mycobacterium leprae* possibly involved in immunogenicity and pathogenicity. *J Bacteriol* **147**: 728–735.

Hunter SW, Rivoire B, Mehra V, Bloom BR & Brennan PJ (1990) The major native proteins of the leprosy bacillus. *J Biol Chem* **265**: 14065–14068.

Izumi S, Fujiwara T, Ikeda M, Nishimura Y, Sugiyama K & Kawatsu K (1990) Novel gelatin particle agglutination test for serodiagnosis of leprosy in the field. *J Clin Microbiol* **28**: 525–529.

Job CK (1989) Nerve damage in leprosy. *Int J Lepr Other Mycobact Dis* **57**: 532–539.

Maeda Y, Mukai T, Spencer J & Makino M (2005) Identification of an immunomodulating agent from *Mycobacterium leprae*. *Infect Immun* **73**: 2744–2750.

Martinez AN, Britto CFPC, Nery JAC, Sampaio EP, Jardim MR, Sarno EN & Moraes MO (2006) Evaluation of real-time and conventional PCR targeting complex 85 genes for detection of *Mycobacterium leprae* DNA in skin biopsy samples from patients diagnosed with leprosy. *J Clin Microbiol* **44**: 3154–3159.

Pessolani MC, Smith DR, Rivoire B, McCormick J, Hefta SA, Cole ST & Brennan PJ (1994) Purification, characterization, gene sequence, and significance of a bacterioferritin from *Mycobacterium leprae*. *J Exp Med* **180**: 319–327.

Ridley DS & Jopling WH (1996) Classification of leprosy according to immunity. A five-group system. *Int J Lepr Other Mycobact Dis* **34**: 255–273.

Shepard CC & McRae DH (1968) A method for counting acid-fast bacteria. *Int J Lepr Other Mycobact Dis* **36**: 78–82.

Stoner GL (1979) Importance of the neural predilection of *Mycobacterium leprae* in leprosy. *Lancet* **2**: 994–996.

FEBS Letters 581 (2007) 3345–3350

Identification of trehalose dimycolate (cord factor) in *Mycobacterium leprae*

Masanori Kai[a,*], Yukiko Fujita[b], Yumi Maeda[a], Noboru Nakata[a], Shinzo Izumi[c], Ikuya Yano[b], Masahiko Makino[a]

[a] *Department of Microbiology, Leprosy Research Center, National Institute of Infectious Diseases, Higashimurayama-shi, Tokyo 189-0002, Japan*
[b] *BCG Laboratory, Kiyose, Tokyo 204-0022, Japan*
[c] *Leprosy Study Group, Tropical Disease Center, Airlangga University Kampus C UNAIR, Jl. Mulyoreijo 60115, Surabaya, Indonesia*

Received 15 May 2007; revised 11 June 2007; accepted 13 June 2007

Available online 21 June 2007

Edited by Sandro Sonnino

Abstract **Glycolipids of *Mycobacterium leprae* obtained from armadillo tissue nodules infected with the bacteria were analyzed. Mass spectrometric analysis of the glycolipids indicated the presence of trehalose 6,6′-dimycolate (TDM) together with trehalose 6-monomycolate (TMM) and phenolic glycolipid-I (PGL-I). The analysis showed that *M. leprae*-derived TDM and TMM possessed both α- and keto-mycolates centering at C78 in the former and at C81 or 83 in the latter subclasses, respectively. For the first time, MALDI-TOF mass analyses showed the presence of TDM in *M. leprae*.**

Keywords: Trehalose 6,6′-dimycolate; Trehalose 6-monomycolate; Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; Thin-layer chromatography; *Mycobacterium leprae*

1. Introduction

Mycolic acids and mycolyl glycolipids are unique and ubiquitous components of mycobacterial cell envelopes. Among such components, trehalose 6,6′-dimycolate (TDM) was first isolated as cord factor from highly virulent *Mycobacterium tuberculosis* showing cord-like growth on the surface culture in liquid media [1–3]. TDM of *M. tuberculosis* was recognized as one of the virulence factors capable of inhibiting fusion of phagosome with lysosome in infected macrophage [4]. However, on the other hand, the TDM was considered to be associated with host defence against the mycobacterial infection since it induced immune responses such as type 1 T cell activation and the formation of granuloma in the mycobacteria-infected lesion [5–7]. Wang et al. have reported that a high proportion of patients infected with *Mycobacterium leprae* possess IgG antibody against TDM of unknown origin as well as trehalose 6-monomycolate (TMM), a biosynthetic precursor of TDM [8]. These observations suggest the possible existence of both TDM and TMM in *M. leprae*. Previously, TDM has been isolated from almost all species of culturable mycobacteria [9,10], and also TMM was isolated from *M. leprae*, however, the search for TDM in *M. leprae* has been unsuccessful [11]. The possible reason for not being able to identify TDM, may be due to (1) inadequate supply of *M. leprae*, (2) negligible amount of the product, (3) technically inefficient to identify TDM. Recent development of newer techniques such as MALDI-TOF mass spectrometry has enabled us to identify even several pg amounts of products. Therefore, the present study was designed as an attempt to directly detect TDM in *M. leprae*, by use of newer technologies. In the process, higher amounts of *M. leprae* phenolic glycolipid-I (PGL-I) was obtained which was analyzed by MALDI-TOF mass spectrometry.

*Corresponding author. Fax: +81 42 391 8807.
E-mail address: mkai@nih.go.jp (M. Kai).

Abbreviations: TDM, trehalose 6,6′-dimycolate; TMM, trehalose 6-monomycolate; *M. leprae*, *Mycobacterium leprae*; *M. tuberculosis*, *Mycobacterium tuberculosis*; *M. bovis*, *Mycobacterium bovis*; TLC, thin-layer chromatography; MALDI-TOF mass, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

2. Materials and methods

2.1. Sources for extraction of glycolipids

M. tuberculosis Aoyama B (ATCC 31726) and *Mycobacterium bovis* BCG Connaught (ATCC 35745) were grown at 37 °C on Sauton's medium for four weeks as surface pellicles until early stationary phase. Cultivated mycobacterial strains were used for extraction of glycolipids. Because *M. leprae* cannot be cultivated in any artificial media, armadillo tissue nodules infected with *M. leprae* (Thai 53 strain) were used for the extraction of glycolipids.

2.2. Extraction of glycolipids and mycolic acid methyl esters

Glycolipids were extracted according to the methods described previously [12]. In brief, bacterial culture or tissues infected with *M. leprae* [13] were autoclaved at 121 °C for 15 min and collected by centrifugation. Lipids were extracted from homogenized tissue with 20 volumes of chloroform/methanol (2:1, v/v) three times with vigorous grinding. The two phases were separated in a funnel, the lower organic phase was collected, and the solvent was evaporated from the organic phase. The total lipids were separated by solvent fractionation and tetrahydrofuran-soluble fraction was further separated by thin-layer chromatography (TLC) on silica gel plates (Uniplate; Analtech Inc. Newark, DE) with the solvent system of chloroform/methanol/water (90:10:1, by vol.) or chloroform/methanol/acetone/acetic acid (90:10:6:1, by vol.). Glycolipid spots were visualized with a 9 M H_2SO_4 spray followed by charring at 200 °C for 10 min or with iodine vapor for preparative purposes.

To determine the subclass composition of the mycolic acids in each mycobacterial TMM and TDM, mycolic acid methyl esters were prepared by alkaline hydrolysis of glycolipids. The glycolipids were hydrolyzed with 1.25 M NaOH in 90% methanol at 70 °C for 1 h and the resultant mycolic acids were then extracted with *n*-hexane after acidification with HCl, followed by methylation with benzene/methanol/H_2SO_4 (10:20:1, by vol.) [14]. Mycolic acid methyl esters from each

0014-5793/$32.00
doi:10.1016/j.febslet.2007.06.029

glycolipid were fully separated into subclasses by TLC with the solvent system of benzene in a draft chamber under reduced pressure.

2.3. Mass spectrometry analysis

Analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF mass) was carried out on a Voyager DE-STR (Applied Biosystems, Tokyo, Japan) with pulsed UV light (337 nm) from an N_2 laser, essentially according to the method reported previously [14]. TDM and TMM were analyzed in the reflectron mode by the instrument operated at 20 kV in the positive ion mode. The 2,5-dihydroxybenzoic acid (2,5-DHB) matrix was used at a concentration of 10 mg/ml in chloroform/methanol (2:1 v/v). Typically, 5 μl of TDM or TMM samples (5 μg) in chloroform/methanol (2:1 v/v) solution and 5 μl of the matrix solution were mixed, and 1.5 μl of the mixture was applied on a sample plate. An external mass spectrum calibration was performed using calibration mixture 2 of the Sequazyme Peptide Mass Standards kit (Applied Biosystems), including known peptide standards in a mass range from 1290 to 5700 Da. The molecular mass of mycobacterial TMM, TDM, and PGL-I was determined based on the quasimolecular mass ions $[M+Na]^+$ by the reflectron mode. In general, nominal number of atomic mass is used for calculation of number of molecular mass. However, there is a slight difference between nominal mass number and accurate mass number read from spectrometry [15]. For instance, when both of the two molecules of mycocerosates ($R_1$and R_2) in PGL-I are C32, the nominal mass number $[M+Na]^+$ of PGL-I ($C_{124}H_{232}O_{19}$) is $[(C \times 124) + (H \times 232) + (O \times 19) + (Na \times 1)] = 2047$ (2024 + 23), but the accurate number is 2050.15 $[(12.0107 \times 124) + (1.00794 \times 232) + (15.9994 \times 19) + 22.9898]$. The nominal mass numbers are given in the text.

3. Results

3.1. TLC analysis of mycolic acids methyl ester

To determine the subclass composition of the mycolic acids in each mycobacterial TMM and/or TDM, mycolic acid methyl esters were analyzed by TLC. The TLC analysis indicated that fatty acid methyl esters had two spots corresponding to either α- and keto-mycolic acid (Fig. 1A). The same two spots pattern was observed for *M. bovis* BCG Connaught (BCG-C) and *M. leprae* while three spots were detected for *M. tuberculosis* corresponding to α-, methoxy-, and keto-mycolic acid [16,17]. The bottom spot in ML lane could be cholesterol from armadillo's tissues (data not shown).

3.2. MALDI-TOF-MS analysis of TMM

We separated the final solvent extracts of *M. leprae* into four fractions (M1–M4). Fig. 1B shows a thin-layer chromatogram of the TMM and TDM from *M. tuberculosis* and the solvent fractionated glycolipids from *M. leprae*. *M. leprae* exhibited bands, which were faint, but significantly reddish glycolipid-like, migrating close to bands of TDM or TMM of *M. tuberculosis*. The major bands of M2 and M4 migrated close to TDM and TMM positions of *M. tuberculosis*, respectively. Therefore, we tried to carefully analyze the bands which may correspond to TDM and TMM by mass spectrometry.

Major band in M4 in Fig. 1B was analyzed using BCG-C as a reference. The mass spectra of TMM from BCG-C showed a biphasic distribution of pseudomolecular ions, $[M+Na]^+$ (Fig. 2A). In the higher mass ranges of BCG-C TMM, dominant ions were detected at *m/z* 1555, 1583, 1597, 1611 and 1625 due to $[M+Na]^+$ of C82, C84, C85, C86 and C87 keto mycolyl TMM, and the major mass ions in the lower mass ranges were detected at *m/z* 1455, 1483, 1511 and 1539 due to $[M+Na]^+$ of α-mycolyl TMM centering at C78, respectively (Fig. 2A) [18]. On the other hand, *M. leprae* derived TMM showed in lower mass ranges at *m/z* 1427, 1455, 1483, 1511

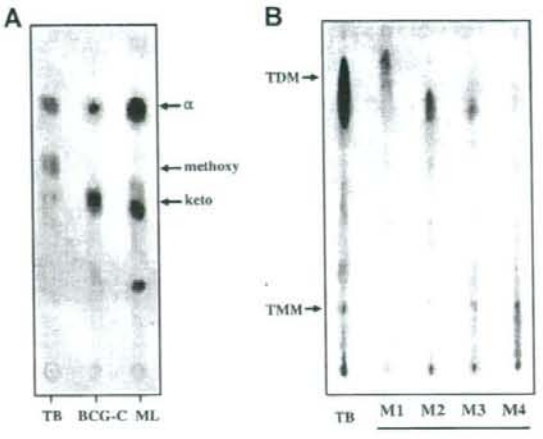


Fig. 1. Thin-layer chromatograms (TLC) of solvent extracts from tetrahydrofuran soluble fraction and mycolic acid methyl ester subclasses of *M. tuberculosis*, *M. bovis* BCG Connaught, and *M. leprae*. (A) TLC of mycolic acid methyl ester subclasses with the solvent system of benzene/methanol/H_2SO_4 (10:20:1, by vol.). Mycolic acid methyl ester subclasses of *M. tuberculosis* (TB); α-, methoxy- and keto-mycolic acid methyl esters are shown with the mycolic acid methyl ester mixture of glycolipids derived from armadillo tissue nodules infected with *M. leprae* (ML), and those of *M. bovis* BCG Connaught (BCG-C). (B) TLC of solvent extracts from *M. leprae* (ML) and *M. tuberculosis* (TB) with the solvent system of chloroform/methanol/acetone/acetic acid (90:10:6:1, by vol.). Trehalose monomycolate (TMM) and trehalose di-mycolate (TDM) bands of TB were identified previously and used as references in this TLC. Fractions 1–4 separated from the final extracts of *M. leprae* and designated M1–4.

and 1539 due to $[M+Na]^+$ of α-mycolyl TMM centering at C78 same to that of BCG-C (Fig. 2B). In the higher mass ranges, dominant ions were shifted lower and the major ions were detected at *m/z* 1541, 1569 and 1597 (Fig. 2B), indicating the major keto-mycolyl TMM consisted of C81, C83 or C85 mycolate, respectively. The molecular species of TMM from *M. leprae* and that from BCG-C are summarized in Table 1. These results indicate that *M. leprae* possess trehalose 6-monomycolate, with C78 α- and C83 keto-mycolates, as the major molecular species.

3.3. MALDI-TOF-MS analysis of TDM

TDM from BCG-C showed a diverse distribution of mass ions according to the combination of di α-, α- and keto-, and di keto-mycolic acid subclasses and each molecular species, that leads to a multiphasic distribution of mass ions due to the dominant combination of α–α, α–keto and keto–keto dimycolyl TDM (Fig. 2C). Given the small sample size, the identification of *M. leprae* TDM was achieved primarily on the thin-layer chromatographic behavior and MALDI-TOF mass analysis, in comparison with the analytical results from BCG-C possessing the same mycolic acid subclasses. Fig. 2D shows the positive MALDI-TOF mass spectra of *M. leprae* TDM (Fig. 1B, M2). In contrast to TDM from BCG-C, TDM from *M. leprae* showed a distinctive mass ion distribution shifted to lower mass ranges due to the major combinations of α–α dimycolyl TDM and α–keto dimycolyl TDM with a small shoulder due to keto–keto dimycolyl TDM. In *M. leprae* TDM, major

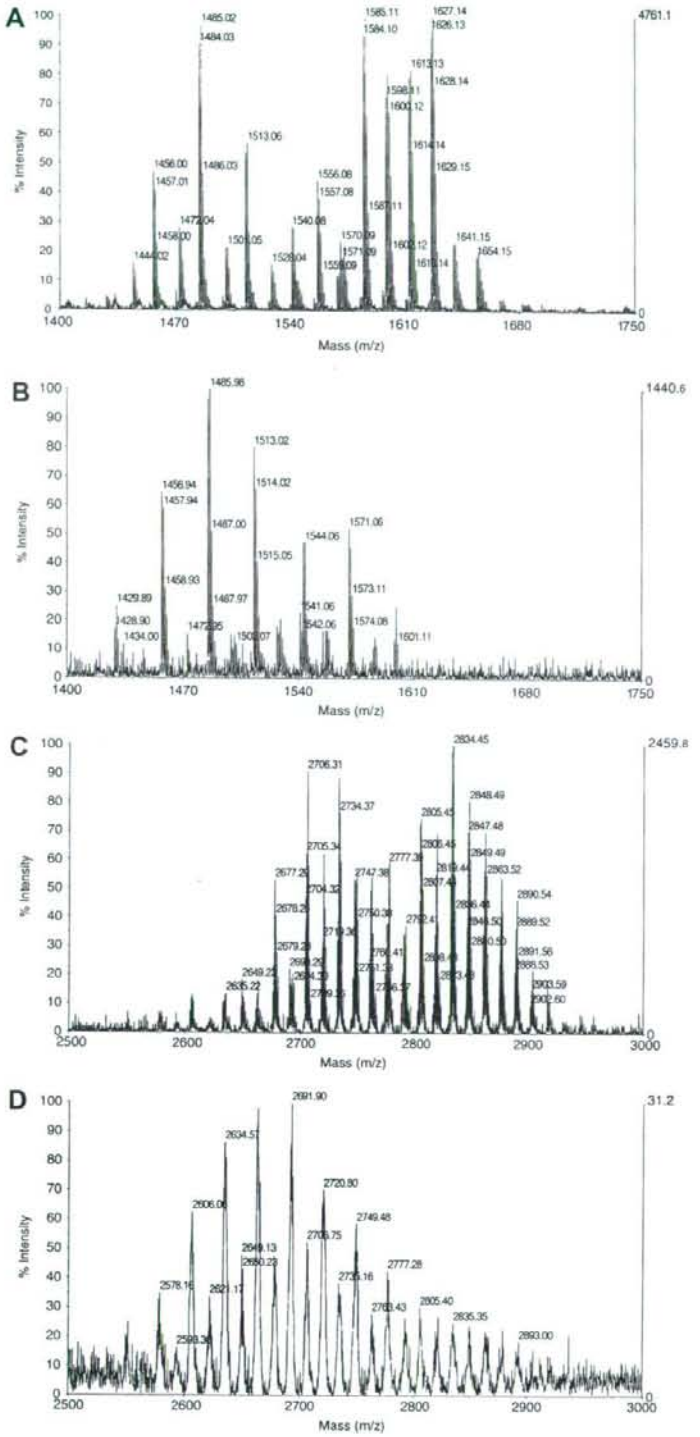


Fig. 2. MALDI-TOF-MS spectra of TMM and TDM. (A) TMM of *M. bovis* BCG Connaught, (B) TMM of *M. leprae* Thai 53, (C) TDM of *M. bovis* BCG Connaught, (D) TDM of *M. leprae* Thai 53.

Table 1
MALDI-TOF mass spectrometry data of the individual types of TMM from *M. leprae* (Thai 53) and *M. bovis* BCG (Connaught)

Species (strain)	Mycolic acid type[a]	Total carbon number of TMM mycolate																	
		74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91
M. leprae (Thai 53)	I	1427		**1455**		**1483**	1497	**1511**	1525	1539	1553								
	II			1471		1499		1527	**1541**	1555	**1569**	1583	1597						
M. bovis BCG (Connaught)	I			**1455**	1469	**1483**		**1511**		1539									
	II									**1555**	1569	**1583**	**1597**	**1611**	**1625**	1639	1653	1667	1681

[a]I; α-dicyclopropanoic or dienoic, II; keto-monocyclopropanoic or monoenoic. Major homologues are shown in bold.

mass ions due to molecular species possessing C76, C78, C80 α–α dimycolic acids were detected at *m/z* 2573, 2601 and 2629, and those possessing α–keto dimycolic acids at *m/z* 2659, 2687 and 2715, and those possessing keto–keto dimycolic acids at *m/z* 2731, 2745, 2773 and 2801, respectively. The deduced molecular species of TDM from *M. leprae* and that from BCG-C are summarized in Table 2.

3.4. MALDI-TOF-MS analysis of PGL-I

The MALDI-TOF mass spectra of M2 in Fig. 1B showed the existence of TDM as described above, however, the major cluster ions in mass spectra of M2 were observed in lower mass ranges from around *m/z* 1990 to 2100. So, we analyzed the spectra in more detail. The result indicates that the cluster ions are derived from phenolic glycolipid-I (PGL-I) (Fig. 3A). The nominal mass number of the PGL-I, which is assumed to have two molecules of C32 mycocerosates, was *m/z* 2047 (Fig. 3A and B). Thus, the deduced combination of mycocerosic acids with different carbon numbers in PGL-I are shown in Fig. 3A and the general structure of PGL-I is shown in Fig. 3B.

4. Discussion

In the present study, we have directly detected TDM and TMM from armadillo tissues infected with *M. leprae*. The presence of TMM in *M. leprae* possessing C74-82 α-mycolic acids has been reported previously [8,18], however, the existence of keto-mycolate in *M. leprae* TMM was not clear. We identified keto-mycolate clearly in *M. leprae* TDM and TMM, and the chain length of keto-mycolate in *M. leprae* was found to be shorter than those from slow-growing culturable mycobacteria such as *M. tuberculosis* and *M. bovis* BCG [19].

Previously, no detectable TDM was identified by the analysis of the lipids obtained from *M. leprae* infected armadillo [9,20]. However, in our hands, we observed a meager but significant spot on TLC from *M. leprae* extract which migrated to the position of TDM of *M. tuberculosis* (Aoyama B strain). When this spot was analyzed by mass spectrometry in reference to TDM from *M. bovis* BCG Connaught (BCG-C), TDM having α–α dimycolates were observed in the lower mass ranges than *m/z* 2673 as seen in TDM from BCG-C (Fig. 2C). Therefore, both *M. leprae* and *M. bovis*

Table 2
Most probable combination of mycolic acids constructing TDM from *M. leprae* (Thai 53) and *M. bovis* BCG (Connaught)

Mass no. of TDM (*m/z*)	*M. leprae* (Thai 53)	*M. bovis* BCG (Connaught)
2573	α76:α78	
2601	α76:α80, α78:α78	
2617	α74:k82, α76:k80, α78:k78, α80:k76	
2629	α78:α80	
2645	α74:k84, α76:k82, α78:k80, α80:k78, α82:k76	
2659	α76:k83, α78:k81	
2673	α76:k84, α78:k82, α79:k81, α80:k80, α82:k78	α76:k84, α78:k82
2687	α78:k83, α80:k81	
2701	α78:k84, α79:k83, α80:k82, α81:k81, α82:k80	α76:k86, α78:k84, α80:k82
2715	α80:k83	α76:k87, α78:k85
2729		α78:k86, α80:k84
2731	k80:k83, k81:k82	
2743		α78:k87, α80:k85
2745	k81:k83	k82:k82
2757		α80:k86
2773	k83:k83	k82:k84
2787		k82:k85
2801	k83:k85, k84:k84	k82:k86, k84:k84
2815		k82:k87, k84:k85
2829		k84:k86, k85:k85
2843		k84:k87, k85:k86
2857		k85:k87, k86:k86
2871		k86:k87
2885		k87:k87

α,α-dicyclopropanoic or dienoic; k, keto-monocyclopropanoic or monoenoic mycolic acid. Molecular ions of TDM with intensities ⩾ 30% of the highest intensity observed are listed.

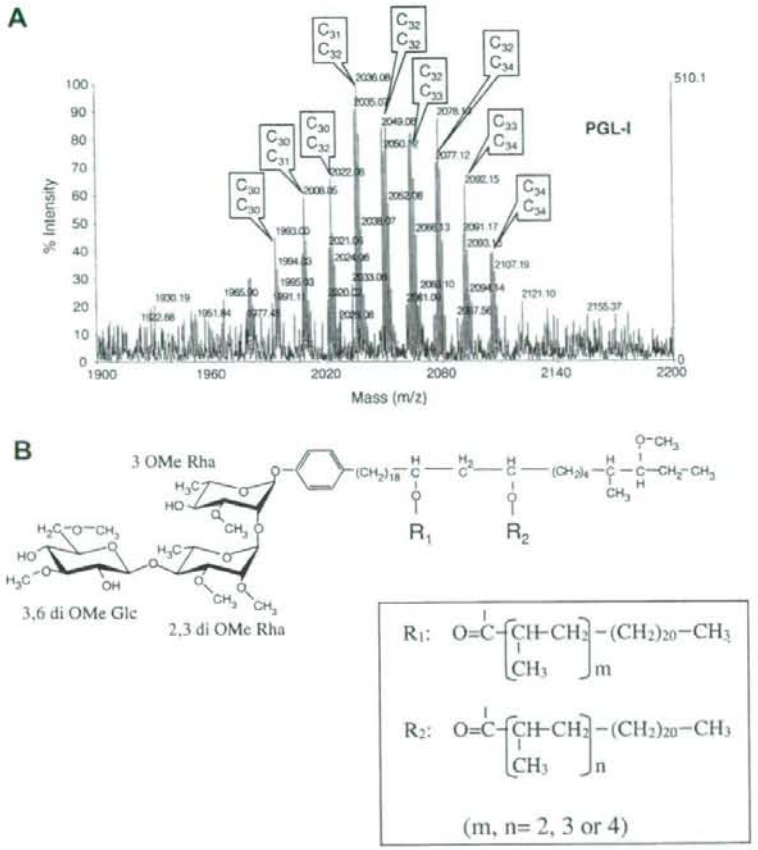


Fig. 3. MALDI-TOF mass spectra of *M. leprae* PGL-I and the deduced chemical structure. (A) The mass spectra of the major cluster ions ranging from around *m/z* 1997 (C30:C30) to 2103 (C34:C34) due to PGL-I separated by solvent fractionation, followed by TLC. C30–C34 shows carbon numbers of two sets of mycocerosate in PGL-I. (B) The chemical structure of *M. leprae* PGL-I. R indicates mycocerosate.

BCG-C can synthesize α-mycolic acids with similar sizes (C74 to C82 centering at C78), however, the subclass composition of keto-mycolate differed. *M. bovis* BCG-C synthesizes multi sub-classes of ketomycolic acids such as *cis*-monoenoic, *cis*-monocyclopropanoic, *cis*-monoenoic *cis*-monocyclopropanoic, *trans*-monoenoic and *trans*-monocyclopropanoic mycolates [21], while *M. leprae* synthesize only *cis*-monoenoic ketomycolate centering at C81 or C83 (Tables 1 and 2) [22]. Recently, the immunological role of mycolic acid has become a point of focus. The *trans*-, but not *cis*-cyclopropanation of mycolic acids in TDM induce massive inflammation in the lesion infected with *M. tuberculosis* and, in this respects, the *trans*-cyclopropane structure may be contributing to virulence [23,24]. Furthermore, recently, it was reported that the Δ*kasB* (a β-keto acyl-ACP synthase gene) mutant strain of *M. tuberculosis* synthesized shorter chain mycolate than the wild ones, and resulted in the loss of keto-mycolic acid *trans*-cyclopropanation. The *kasB* deletion induced the ability of the mutant strain to persist in infected mice for up to 600 days without causing disease [25]. Thus, the lack of *trans*-cyclopropanoic keto-mycolic acid in *M. leprae* may contribute to the low virulency in contrary to substantial quantities found in most of the slow growing pathogenic mycobacteria.

In the search for TDM, we also identified a major cluster ions in M2 fraction. Detailed analyses showed that the ions were derived from phenolic glycolipid (PGL-I), which is a major species specific glycolipid of *M. leprae*. Previously, combined GLC-mass spectroscopic analysis of the trisaccharide of PGL-I had been achieved, but TOF-mass spectrometric data have not been reported [26]. We analyzed the M1 fraction by MALDI TOF mass spectrometry and detected some clusters. Mass spectra of major clusters look like monoglycosyl phthiocerol diester and diglycosyl phthiocerol diester (data not shown), which differ from PGL-I in lacking the terminal sugars [27]. The co-existence of both PGL-I and TDM in armadillo tissue nodules infected with *M. leprae* indicates that these lipids were originated actually from *M. leprae*. Thus, this paper identifies TDM structure of *M. leprae* for the first time and shows that *M. leprae* TDM contains only α-mycolic acids and shorter chain *cis*-monoenoic keto-mycolic acids, exclusively.

Acknowledgments: We thank Dr. M. Matsuoka, and Dr. M. Gidoh, for providing armadillo tissues infected with *M. leprae*. This work was supported in part by a Health Science Research Grant for Research on Emerging and Re-emerging Infectious Diseases, Ministry of Health, Labour and Welfare, Japan.

References

[1] Minnikin, D. E. (1982) Presented at the Part I. Physiology of the mycobacteria. In The Biology of the Mycobacteria.
[2] Rastogi, N., Legrand, E. and Sola, C. (2001) The mycobacteria: an introduction to nomenclature and pathogenesis. Rev. Sci. Tech. 20, 21–54.
[3] Ryll, R., Kumazawa, Y. and Yano, I. (2001) Immunological properties of trehalose dimycolate (cord factor) and other mycolic acid-containing glycolipids – a review. Microbiol. Immunol. 45, 801–811.
[4] Fujiwara, N. (1997) Distribution of antigenic glycolipids among *Mycobacterium tuberculosis* strains and their contribution to virulence. Kekkaku 72, 193–205.
[5] Coelho, E.A., Tavares, C.A., de Melo Lima, K., Silva, C.L., Rodrigues Jr., J.M. and Fernandes, A.P. (2006) *Mycobacterium* hsp65 DNA entrapped into TDM-loaded PLGA microspheres induces protection in mice against *Leishmania* (*Leishmania*) *major* infection. Parasitol. Res. 98, 568–575.
[6] Oiso, R., Fujiwara, N., Yamagami, H., Maeda, S., Matsumoto, S., Nakamura, S., Oshitani, N., Matsumoto, T., Arakawa, T. and Kobayashi, K. (2005) Mycobacterial trehalose 6,6′-dimycolate preferentially induces type 1 helper T cell responses through signal transducer and activator of transcription 4 protein. Microb. Pathog. 39, 35–43.
[7] Hunter, R.L., Olsen, M.R., Jagannath, C. and Actor, J.K. (2006) Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease. Ann. Clin. Lab. Sci. 36, 371–386.
[8] Dhariwal, K.R., Yang, Y.-M., Fales, H.M. and Goren, M.B. (1987) Detection of trehalose monomycolate in *Mycobacterium leprae* grown in armadillo tissues. J. Gen. Microbiol. 133, 201–209.
[9] Prome, J.C., Lacave, C.A., Ahibo-Coffy and Savagnac, A. (1976) Separation of structural studies of the molecular species of monomycolates and dimycolates of alpha-D-trehalose present in *Mycobacterium phlei*. Eur. J. Biochem. 63, 543–552.
[10] Wang, L., Izumi, S., He, H., Fujiwara, N., Saita, N., Yano, I., Kobayashi, K. and Tatsumi, N. (1999) Serodiagnosis of Hansen's disease/leprosy by enzyme-linked immunosorbent assay using cord factor (trehalose 6,6′-dimycolate) as an antigen. Nihon Hansenbyo Gakkai Zasshi 68, 165–174.
[11] Goren, M.B., Brokl, O. and Roller, P. (1979) Cord factor (trehalose-6,6′-dimycolate) of in vivo-derived *Mycobacterium lepraemurium*. Biochim. Biophys. Acta 574, 70–78.
[12] Fujita, Y., Ogata, H. and Yano, I. (2005) Clinical evaluation of serodiagnosis of active tuberculosis by multiple-antigen ELISA using lipids from *Mycobacterium bovis* BCG Tokyo 172. Clin. Chem. Lab. Med. 43, 1253–1262.
[13] Sasaki, N., Kawatsu, K., Tsutsumi, S., Gidoh, M., Nakagawa, H., Kashiwabara, Y., Matsuki, G. and Endo, H. (1997) Pathological investigation of armadillos infected with *Mycobacterium leprae*. Nihon Hansenbyo Gakkai Zasshi 66, 227–235.
[14] Fujita, Y., Naka, T., Doi, T. and Yano, I. (2005) Direct molecular mass determination of trehalose monomycolate from 11 species of mycobacteria by MALDI-TOF mass spectrometry. Microbiology 151, 1443–1452.
[15] Fujita, Y., Naka, T., McNeil, M.R. and Yano, I. (2005) Intact molecular characterization of cord factor (trehalose 6,6′-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. Microbiology 151, 3403–3416.
[16] Minnikin, D.E. and Polgar, N. (1967) Synthesis of methyl 3,12,15-trimethyldocosanoate and 3,15-dimethyldocosanoate. J. Chem. Soc. 7, 575–577.
[17] Yuan, Y., Crane, D.C., Musser, J.M., Sreevatsan, S. and Barry 3rd., C.E. (1997) MMAS-1, the branch point between *cis*- and *trans*-cyclopropane-containing oxygenated mycolates in *Mycobacterium tuberculosis*. J. Biol. Chem. 272, 10041–10049.
[18] Draper, P., Dobson, G., Minnikin, D.E. and Minnikin, S.M. (1982) The mycolic acids of *Mycobacterium leprae* harvested from experimentally infected nine-banded armadillos. Ann. Microbiol. 133, 39–47.
[19] Takayama, K., Wang, C. and Besra, G.S. (2005) Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. Clin. Microbiol. Rev. 18, 81–101.
[20] Minnikin, D.E., Dobson, G., Goodfellow, M., Draper, P. and Magnusson, M. (1985) Quantitative comparison of the mycolic and fatty acid compositions of *Mycobacterium leprae* and *Mycobacterium gordonae*. J. Gen. Microbiol. 131, 2013–2021.
[21] Watanabe, M., Aoyagi, Y., Mitome, H., Fujita, T., Naoki, H., Ridell, M. and Minnikin, D.E. (2002) Location of functional groups in mycobacterial meromycolate chains; the recognition of new structural principles in mycolic acids. Microbiology 148, 1881–1902.
[22] Watanabe, M., Aoyagi, Y., Ridell, M. and Minnikin, D.E. (2001) Separation and characterization of individual mycolic acids in representative mycobacteria. Microbiology 147, 1825–1837.
[23] Glickman, M.S., Cahill, S.M. and Jacobs Jr., W.R. (2001) The *Mycobacterium tuberculosis* cmaA2 gene encodes a mycolic acid *trans*-cyclopropane synthetase. J. Biol. Chem. 276, 2228–2233.
[24] Yuan, Y., Zhu, Y., Crane, D.D. and Barry 3rd., C.E. (1998) The effect of oxygenated mycolic acid composition on cell wall function and macrophage growth in *Mycobacterium tuberculosis*. Mol. Microbiol. 29, 1449–1458.
[25] Bhatt, A., Fujiwara, N., Bhatt, K., Gurcha, S.S., Kremer, L., Chen, B., Chan, J., Porcelli, S.A., Kobayashi, K., Besra, G.S. and Jacobs Jr., W.R. (2007) Deletion of *kasB* in *Mycobacterium tuberculosis* causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. Proc. Natl. Acad. Sci. USA 104, 5157–5162.
[26] Hunter, S.W. and Brennan, P.J. (1981) A novel phenolic glycolipid from *Mycobacterium leprae* possibly involved in immunogenicity and pathogenicity. J. Bacteriol. 147, 728–735.
[27] Daffe, M. and Laneelle, M.A. (1989) Diglycosyl phenol phthiocerol diester of *Mycobacterium leprae*. Biochim. Biophys. Acta 1002, 333–337.

JOURNAL OF BACTERIOLOGY, Aug. 2007, p. 5515–5522 Vol. 189, No. 15
0021-9193/07/$08.00+0 doi:10.1128/JB.00344-07

Characterization of the Fucosylation Pathway in the Biosynthesis of Glycopeptidolipids from *Mycobacterium avium* Complex[▽]

Yuji Miyamoto,[1] Tetsu Mukai,[1]* Yumi Maeda,[1] Noboru Nakata,[1] Masanori Kai,[1] Takashi Naka,[2] Ikuya Yano,[2] and Masahiko Makino[1]

Department of Microbiology, Leprosy Research Center, National Institute of Infectious Diseases, 4-2-1 Aobacho, Higashimurayama, Tokyo 189-0002, Japan,[1] and Japan BCG Central Laboratory, 3-1-5 Matsuyama, Kiyose, Tokyo 204-0022, Japan[2]

Received 3 March 2007/Accepted 2 May 2007

The cell envelopes of several species of nontuberculous mycobacteria, including the *Mycobacterium avium* complex, contain glycopeptidolipids (GPLs) as major glycolipid components. GPLs are highly antigenic surface molecules, and their variant oligosaccharides define each serotype of the *M. avium* complex. In the oligosaccharide portion of GPLs, the fucose residue is one of the major sugar moieties, but its biosynthesis remains unclear. To elucidate it, we focused on the 5.0-kb chromosomal region of the *M. avium* complex that includes five genes, two of which showed high levels of similarity to the genes involved in fucose synthesis. For the characterization of this region by deletion and expression analyses, we constructed a recombinant *Mycobacterium smegmatis* strain that possesses the *rtfA* gene of the *M. avium* complex to produce serovar 1 GPL. The results revealed that the 5.0-kb chromosomal region is responsible for the addition of the fucose residue to serovar 1 GPL and that the three genes *mdhtA*, *merA*, and *gtfD* are indispensable for the fucosylation. Functional characterization revealed that the *gtfD* gene encodes a glycosyltransferase that transfers a fucose residue via 1→3 linkage to a rhamnose residue of serovar 1 GPL. The other two genes, *mdhtA* and *merA*, contributed to the formation of the fucose residue and were predicted to encode the enzymes responsible for the synthesis of fucose from mannose based on their deduced amino acid sequences. These results indicate that the fucosylation pathway in GPL biosynthesis is controlled by a combination of the *mdhtA*, *merA*, and *gtfD* genes. Our findings may contribute to the clarification of the complex glycosylation pathways involved in forming the oligosaccharide portion of GPLs from the *M. avium* complex, which are structurally distinct.

Mycobacteria have a unique multilayer cell envelope composed of peptidoglycan, arabinogalactan, mycolic acids, and an outer layer that contains abundant species-specific glycolipids. It is thought that these cell wall characteristics allow mycobacteria to survive in host cells (8, 11). Among the glycolipids present in the outer layer of the cell envelope, glycopeptidolipids (GPLs) are recognized as highly antigenic surface molecules in several species of nontuberculous mycobacteria, including the *Mycobacterium avium* complex (29). The common structure in which the fatty acyl-tetrapeptide core is glycosylated with both 6-deoxy-talose (6-d-Tal) and *O*-methyl-rhamnose (*O*-Me-Rha), termed core GPLs, is present in all types of GPLs (2, 4, 12). The core GPLs are further glycosylated with a Rha residue or various oligosaccharides linked to the 6-d-Tal residue. These glycosylation events give rise to structural variations, especially in the GPLs from the *M. avium* complex, that are responsible for the creation of serotypes, and they are also known as serovar-specific GPLs (ssGPLs). The ssGPLs define approximately 30 serotypes, some of which are found predominantly in isolates from patients. For example, the serotypes 1, 4, 6, and 8 are frequently isolated from AIDS patients coinfected with the *M. avium* complex (17, 30). In addition, ssGPLs are associated with host responses to infection and function as agonists of cell surface receptors such as Toll-like receptor 2 (28).

The biosynthesis of GPLs, including methylation, acetylation, and peptide synthesis, has been investigated mainly with *M. smegmatis*, but it has not been fully elucidated (6, 15, 21, 23). We have recently identified the glycosyltransferase genes involved in the formation of core GPLs in *M. smegmatis* and *M. avium* (19), but the glycosylation pathways of ssGPLs have not been clarified, except for that of serovar 1 GPL (14). The presence of a Rha residue linked to 6-d-Tal, as observed with serovar 1 GPL, is structurally required for the formation of all types of ssGPLs. The Rha residue is further extended by the subsequent addition of one of the three different sugars, Rha, fucose (Fuc), or glucose, to form the structural group of ssGPLs (9). In these groups, the Fuc-containing ssGPLs, such as serovar 2, 3, 4, and 9 GPLs, shown in Table 1, comprise the major group in the ssGPLs of the *M. avium* complex, and serovar 2 GPL is the basic structure for these ssGPLs (9). The biosynthetic pathway for the Fuc-containing ssGPLs is still unclear. In the *M. avium* complex, a 22- to 25-kb chromosomal region with the ability to synthesize the serovar 2 GPL has been identified and cloned from two strains (4, 5). However, the key genes responsible for the formation of the Fuc residue of serovar 2 GPL have not been identified in this chromosomal region. In this study, we have focused on a 5.0-kb segment that includes the genes predicted to be associated with Fuc synthesis and have characterized their functions in the biosynthesis of GPLs in the *M. avium* complex.

* Corresponding author. Mailing address: Department of Microbiology, Leprosy Research Center, National Institute of Infectious Diseases, 4-2-1 Aobacho, Higashimurayama, Tokyo 189-0002, Japan. Phone: 81-42-391-8211. Fax: 81-42-394-9092. E-mail: tmukai@nih.go.jp.

▽ Published ahead of print on 25 May 2007.

TABLE 1. Oligosaccharide structures of Fuc-containing ssGPLs

Serotype	Oligosaccharide[a]
2	2,3-di-*O*-Me-α-L-Fuc-(1→3)-α-L-Rha-(1→2)-L-6-d-Tal
4	4-*O*-Me-α-L-Rha-(1→4)-2-*O*-Me-α-L-Fuc-(1→3)-α-L-Rha-(1→2)-L-6-d-Tal
3	2,3-di-*O*-Me-α-L-Fuc-(1→4)-β-D-GlcA-(1→4)-2,3-di-*O*-Me-α-L-Fuc-(1→3)-α-L-Rha-(1→2)-L-6-d-Tal
9	4-*O*-Ac-2,3-di-*O*-Me-α-L-Fuc-(1→4)-β-D-GlcA-(1→4)-2,3-di-*O*-Me-α-L-Fuc-(1→3)-α-L-Rha-(1→2)-L-6-d-Tal

[a] The oligosaccharide structures are from reference 9. GlcA, glucuronic acid; Ac, acetyl.

MATERIALS AND METHODS

Bacterial strains, culture conditions, and DNA manipulation. The bacterial strains and vectors used and constructed in this study are listed in Table 2. *M. avium*, used for the isolation of chromosomal DNA, was grown in Middlebrook 7H9 broth (Difco) with 0.05% Tween 80 supplemented with 10% Middlebrook ADC enrichment broth (BBL). Recombinant *M. smegmatis* strains for GPL production were cultured in Luria-Bertani broth with 0.2% Tween 80. The isolation of DNA, transformation, and PCR were carried out as previously described (20). *Escherichia coli* strain DH5α was used for the routine manipulation and propagation of plasmid DNA. The following antibiotics were added as required: kanamycin, 50 μg/ml for *E. coli* and 25 μg/ml for *M. smegmatis*; hygromycin B, 150 μg/ml for *E. coli* and 75 μg/ml for *M. smegmatis*. Oligonucleotide primers used for PCR are available on request.

Construction of the integrating mycobacterial vector (pYM301). The site-specific integrating mycobacterial vector pYM301 was constructed from parts of pYUB854, pMV306kan, and pMV261 (3, 24, 27). To replace the *oriM* region of pMV261 with the region needed for integration, a fragment containing *attP* and *int* was amplified from pMV306kan DNA using the primers INT-S and INT-A, digested with the respective restriction enzyme, and cloned into the XbaI-MluI site of pMV261 to give pMV301kan. The hygromycin-resistant cassette, a selective marker of integration, was excised from pYUB854 with XbaI and NheI and inserted into the NheI-SpeI site of pMV301kan to obtain *hyg* instead of *kan*. Because the resulting plasmid had restriction sites for EcoRI, BamHI, and PstI outside of the multicloning site, it was disrupted in turn by PCR to create pYM301 by using the following primers: 301ECO-D and 301ECO-U for disruption of the EcoRI site, 301PST-D and 301PST-U for disruption of the PstI site, and 301BAM-D and 301BAM-U for disruption of the BamHI site.

Construction of expression vectors. Previous studies have shown that the clustered *gtfC-gtfD* region is about 5.0 kb long and contains five genes, designated *gtfC*, *mdhtA*, *merA*, *mtfF*, and *gtfD* (13). To express these five genes as one operon, the 5.0-kb segment was obtained as a PstI-EcoRI fragment to be inserted into the expression cassette of pMV261. Prior to cloning into pMV261, it was necessary to clone the 5.0-kb segment into pUC18 to confirm the DNA sequences. Since we could not directly clone the 5.0-kb segment as one PCR-amplified fragment, three divided fragments were amplified from the genomic DNA of *M. avium* JATA51-01 by using following primers: GTFC-S and HA for the 2.0-kb PstI-HindIII fragment, HS and KA for the 1.0-kb HindIII-KpnI fragment, and KS and GTFD-A for the 2.0-kb KpnI-EcoRI fragment (Fig. 1A). These three fragments then were combined in pUC18 as a PstI-EcoRI fragment to give pUCgtfCD (Fig. 1A). A 5.0-kb PstI-EcoRI fragment was excised from pUCgtfCD and inserted into the PstI-EcoRI site of pMV261 to create pMVgtfCD, which expressed the above-described five genes (Fig. 1B).

Deletion of each gene from the *gtfC-gtfD* region was performed as follows. The expression vectors were constructed from pUCgtfCD by PCR using following primers: 18PST-U and MDHTA-S for the deletion of *gtfC*, MDHTA-U and MDHTA-D for the deletion of *mdhtA*, MERA-U and MERA-D for the deletion of *merA*, MTFF-U and MTFF-D for the deletion of *mtfF*, and MTFF-A and 18ECO-D for the deletion of *gtfD*. The PCR products were digested with each restriction enzyme and were ligated. The PstI-EcoRI fragment was excised from each resulting plasmid and was cloned into the same restriction sites of pMV261 to give pMVΔgtfC, pMVΔmdhtA, pMVΔmerA, pMVΔmtfF, and pMVΔgtfD (Fig. 1B).

The two *M. avium* genes *rtfA* and *gtfD* were amplified from genomic DNA of *M. avium* JATA51-01 by using the following primers: RTFA-S and RTFA-A for *rtfA* and GTFD-S and GTFD-PA for *gtfD*. The PCR products were digested with each restriction enzyme and were cloned into the corresponding site of pYM301 and pMV261 to give pYMrtfA-int and pMVgtfD, respectively.

To construct the vector for the simultaneous expression of *rtfA*, *mdhtA*, and *merA*, the HpaI site of pYMrtfA-int was replaced with an AflII site by PCR using AFL-U and AFL-D to give pYMrtfA-int-Afl. The *mdhtA* and *merA* genes were amplified as one operon from genomic DNA of *M. avium* JATA51-01 by using the primers MDHTA-S2 and MERA-A. The PCR product was digested with

TABLE 2. Bacterial strains and vectors used in this study

Strain or vector	Characteristic(s)	Source or reference
Bacterial strains		
E. coli DH5α	Cloning host	
M. smegmatis mc^{2}155	Expression host	25
M. avium JATA51-01 (ATCC 25291)	Source of the 5.0-kb region containing *gtfC*, *mdhtA*, *merA*, *mtfF*, and *gtfD*	
Vectors		
pUC18	*E. coli* cloning vector	
pBluescript II SK(+)	*E. coli* cloning vector	
pYM301	Site-specific integrating mycobacterial vector carrying the *hsp60* promoter cassette	This study
pYUB854	Source of pYM301	3
pMV306kan	Source of pYM301	24
pMV261	*E. coli-Mycobacterium* shuttle vector carrying the *hsp60* promoter cassette	27
pMVgtfCD	pMV261 with *gtfC*, *mdhtA*, *merA*, *mtfF*, and *gtfD*	This study
pMVΔgtfC	pMV261 with *mdhtA*, *merA*, *mtfF*, and *gtfD*	This study
pMVΔmdhtA	pMV261 with *gtfC*, *merA*, *mtfF*, and *gtfD*	This study
pMVΔmerA	pMV261 with *gtfC*, *mdhtA*, *mtfF*, and *gtfD*	This study
pMVΔmtfF	pMV261 with *gtfC*, *mdhtA*, *merA*, and *gtfD*	This study
pMVΔgtfD	pMV261 with *gtfC*, *mdhtA*, *merA*, and *mtfF*	This study
pMVgtfD	pMV261 with *gtfD*	This study
pYMrtfA-int	pYM301 with *rtfA*	This study
pYMrtfA-mdhtA-merA-int	pYM301 with *rtfA*, *mdhtA*, and *merA*	This study

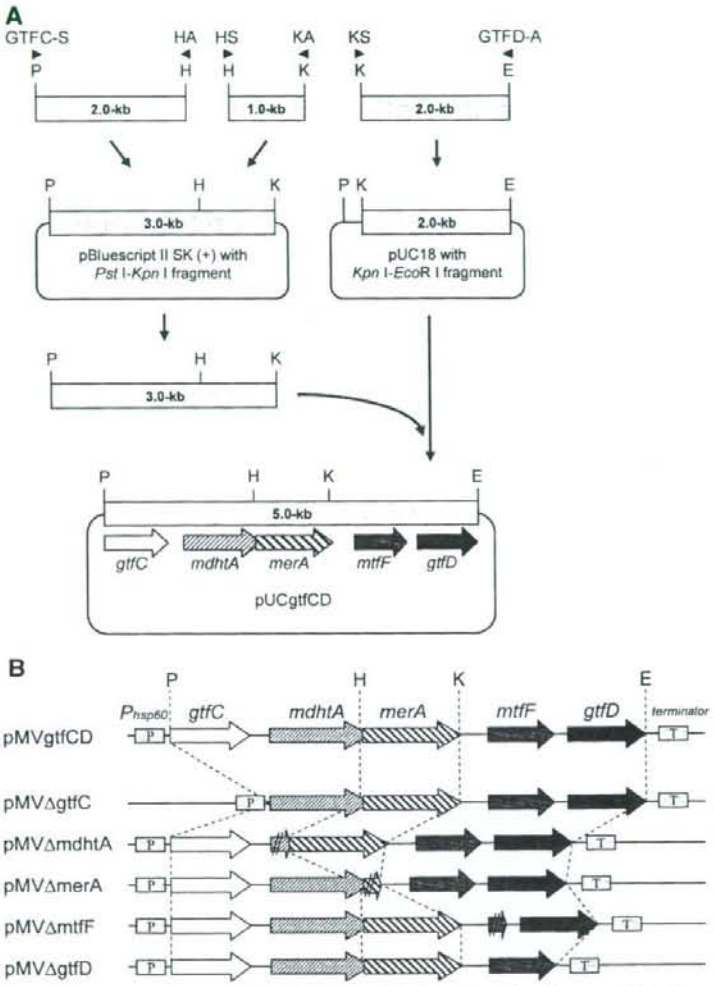


FIG. 1. Schematic presentation of the cloning procedure for the 5.0-kb *gtfC-gtfD* region (A) and its gene-deleted constructs, which were inserted into an expression cassette made up of the *hsp60* promoter and the terminator of pMV261 (B). The primers used for PCR amplification of the three fragments are indicated by filled triangles. In pMVΔgtfC and pMVΔgtfD, the genes *gtfC* and *gtfD* were completely deleted from the *gtfC-gtfD* region of pMVgtfCD. In-frame deletions were designed for the construction of the expression cassettes of pMVΔmdhtA, pMVΔmerA, and pMVΔmtfF to prevent the polar effect on each downstream gene. P, PstI; H, HindIII; K, KpnI; E, EcoRI.

each restriction enzyme and cloned into the PstI-AflII site of pYMrtfA-int-Afl to give pYMrtfA-mdhtA-merA-int.

Isolation and purification of GPLs. Whole-lipid extracts were isolated from harvested bacterial cells that had been mixed with $CHCl_3$-CH_3OH (2:1, vol/vol) for several hours at room temperature. The extracts obtained from the organic phase were separated from the aqueous phase and evaporated to dryness. To remove the lipid components except for GPLs, the whole-lipid extracts were subjected to mild alkaline hydrolysis as previously described (20, 21). For analytical thin-layer chromatography (TLC), crude GPLs from equal amounts of harvested bacterial cells were spotted on silica gel 60 plates (Merck) using $CHCl_3$-CH_3OH (9:1, vol/vol) as the solvent and were visualized by spraying with 10% H_2SO_4 and then charring. Purified GPLs were prepared from crude GPLs by preparative TLC on the same plates and were extracted from the bands corresponding to each GPL. Perdeuteriomethylation for determination of the linkage position of sugar moieties was carried out as previously described (7, 10, 14).

GC-MS and MALDI-TOF analyses. For the monosaccharide analysis, crude GPLs from equal amounts of harvested cells were hydrolyzed in 2 M trifluoroacetic acid (2 h, 120°C), and the released sugars were reduced with sodium tetradeuterborate and then were acetylated with pyridine-acetic anhydride (1:1, vol/vol) at room temperature overnight. The resulting alditol acetates were separated and analyzed by gas chromatography-mass spectrometry (GC-MS) on a TRACE DSQ (Thermo Electron) equipped with an SP-2380 column (SUPELCO) using helium gas. The temperature program was from 52 to 172°C with 40°C/min increments and then from 172 to 250°C with 3°C/min increments. To determine the total mass of the purified GPLs, matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectra (in the positive mode) were acquired on a QSTAR XL (Applied Biosystems) with a pulse laser emitting at 337 nm. Samples mixed with 2,5-dihydroxybenzoic acid as the matrix were analyzed in the reflectron mode with an accelerating voltage of 20 kV and operating in positive-ion mode.

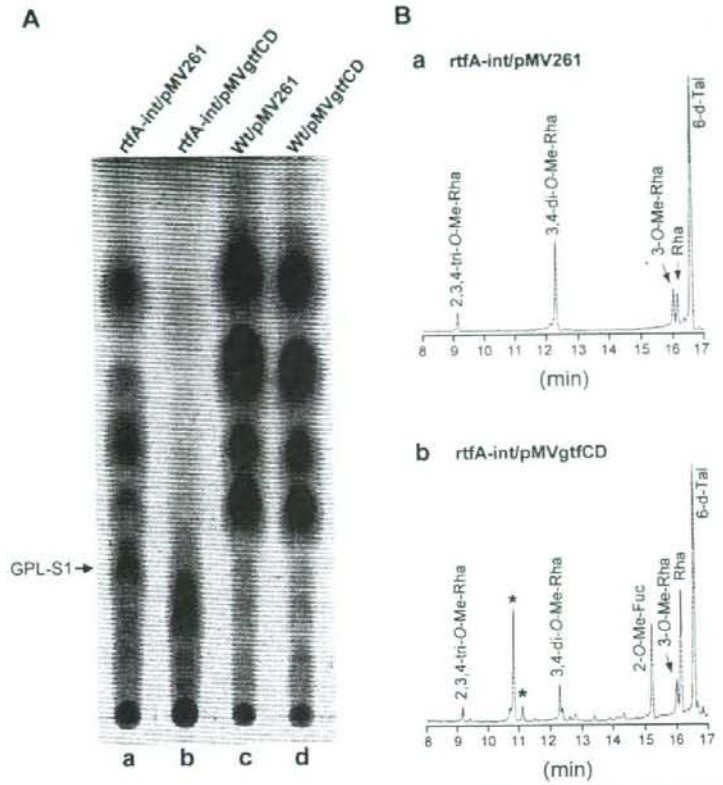


FIG. 2. Functional analyses of the *gtfC-gtfD* region. (A) TLC analysis of crude GPL extracts from rtfA-int/pMV261 (lane a), rtfA-int/pMVgtfCD (lane b), Wt/pMV261 (lane c), and Wt/pMVgtfCD (lane d). The total lipid fraction after mild alkaline hydrolysis was spotted on plates and was developed in $CHCl_3$-CH_3OH (9:1 [vol/vol]). (B) GC-MS analyses of alditol acetates of sugars released from crude GPL extracts of rtfA-int/pMV261 (a) and rtfA-int/pMVgtfCD (b). Alditol acetate derivatives were prepared from the total lipid fraction after mild alkaline hydrolysis. Asterisks indicate noncarbohydrates.

RESULTS

Identification of the genes involved in the formation of the Fuc residue in serovar 2 GPL. To reveal the genes responsible for the fucosylation pathway that lead to the formation of serovar 2 GPL, we focused on the 5.0-kb segment designated the *gtfC-gtfD* region (GenBank accession no. AF125999.1). This region contains five genes: *mdhtA* and *merA*, whose deduced amino acid sequences show a high level of similarity to those of enzymes involved in Fuc synthesis; *mtfF*, previously identified as the fucosyl 2-*O*-methyltransferase gene; and *gtfC* and *gtfD*, putative glycosyltransferase genes whose functions remain unknown (13, 18). Since *M. smegmatis* only produces core GPLs, we introduced the chromosomal integrating vector pYMrtfA-int possessing the *M. avium* gene *rtfA*, whose gene product transfers the Rha residue to 6-d-Tal of core GPLs, and obtained the recombinant strain rtfA-int, which produces GPL with a terminal Rha residue (termed GPL-S1) that could be a substrate for the synthesis of serovar 2 GPL. The expression vector pMVgtfCD, harboring the *gtfC-gtfD* region (Fig. 1B), then was introduced into the GPL-S1-producing strain (rtfA-int) and wild-type mc^2155, and GPL production was analyzed by TLC (Fig. 2A). Although there were no differences between the TLC profiles of Wt/pMV261 and Wt/pMVgtfCD (Fig. 2A, lanes c and d), new spots appeared in rtfA-int/pMVgtfCD (Fig. 2A, lane b), indicating that the *gtfC-gtfD* region contains genes with the ability to convert GPL-S1 into structurally different compounds, but it can do so only in the presence of the *rtfA* gene. To confirm that the products formed in rtfA-int/pMVgtfCD contained Fuc, a characteristic of serovar 2 GPL, alkaline-stable extracts from rtfA-int/pMVgtfCD were hydrolyzed, and the released monosaccharides were analyzed by GC-MS (Fig. 2B). The results showed that 2-*O*-Me-Fuc, which is structurally related to serovar 2 GPL, was present together with Rha, 6-d-Tal, 3-*O*-Me-Rha, 3,4-di-*O*-Me-Rha, and 2,3,4-tri-*O*-Me-Rha in rtfA-int/pMVgtfCD (Fig. 2B, graph b). 2-*O*-Me-Fuc was not detected in strain rtfA-int/pMV261 (vector control) (Fig. 2B, graph a) or recombinant wild-type mc^2155 (data not shown). These results indicated that the *gtfC-gtfD* region is responsible for the transfer of the Fuc residue to serovar 1 GPL. Additionally, for identification of the individual genes involved in this fucosylation, we constructed various plasmids that have one of the genes deleted from the *gtfC-gtfD* region, as shown in

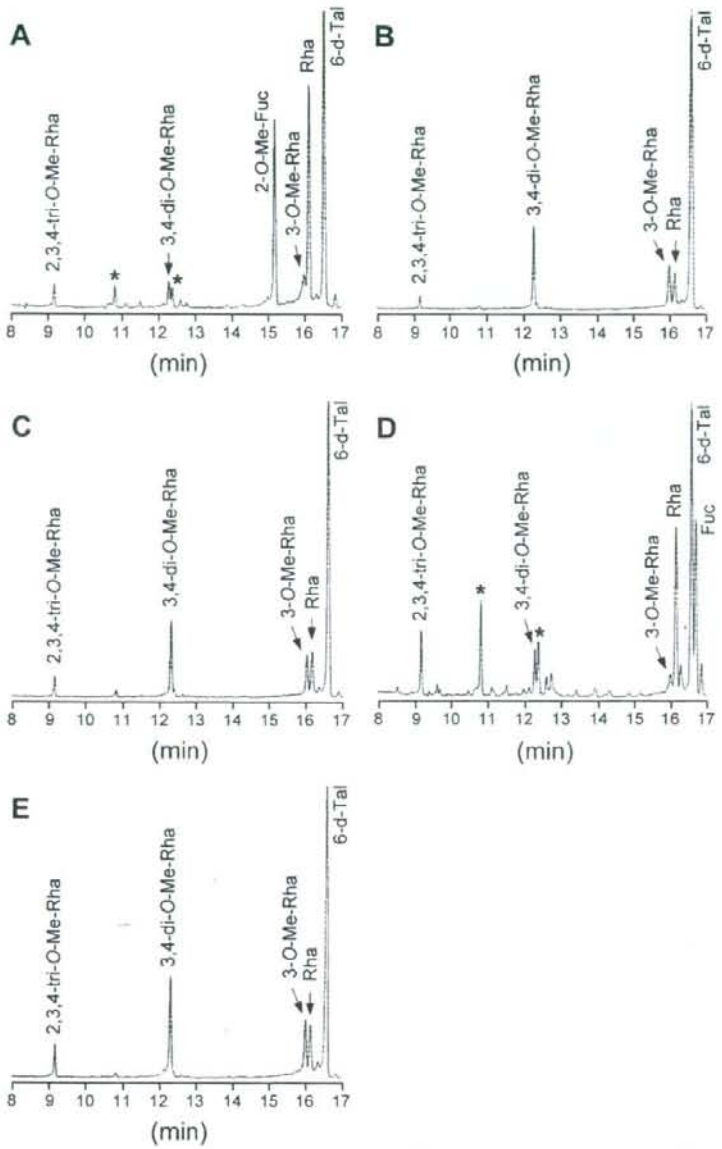


FIG. 3. GC-MS analyses of alditol acetates of sugars released from crude GPLs. GPLs were extracted from rtfA-int/pMVΔgtfC (A), rtfA-int/pMVΔmdhtA (B), rtfA-int/pMVΔmerA (C), rtfA-int/pMVΔmtfF (D), and rtfA-int/pMVΔgtfD (E). Alditol acetate derivatives were prepared from the total lipid fraction after mild alkaline hydrolysis. Asterisks indicate noncarbohydrates.

Fig. 1B, and examined the sugar moieties of the alkaline-stable GPL extracts from each recombinant strain by GC-MS analysis (Fig. 3). The results show that the profile of rtfA-int/pMVΔgtfC was the same as that of rtfA-int/pMVgtfCD, indicating that the *gtfC* gene does not participate in the formation of 2-*O*-Me-Fuc (Fig. 2B, graph b, and 3A). In rtfA-int/pMVΔmtfF, Fuc was detected instead of 2-*O*-Me-Fuc, demonstrating that the *mtfF* gene encodes fucosyl 2-*O*-methyltransferase (Fig. 3D). On the other hand, no Fuc derivatives were detected in the recombinant strains rtfA-int/pMVΔmdhtA, rtfA-int/pMVΔmerA, and rtfA-int/pMVΔgtfD (Fig. 3B, C, and E). These results indicated that the three genes *mdhtA*, *merA*, and *gtfD* are all indispensable for the formation of the Fuc residue in serovar 2 GPL.

Functional analysis of the *gtfD* gene. Although the *gtfD* gene was predicted to encode a type of glycosyltransferase based on

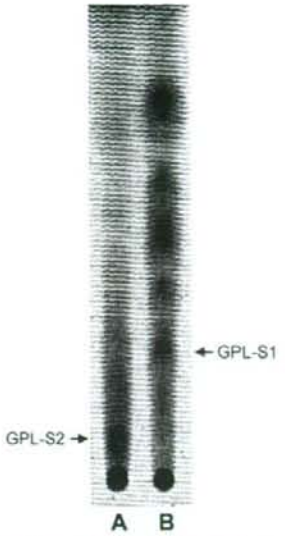


FIG. 4. TLC analyses of crude GPL extracts from rtfA-mdhtA-merA-int/pMVgtfD (lane A) and rtfA-mdhtA-merA-int/pMV261 (lane B). The total lipid fraction after mild alkaline hydrolysis was spotted onto plates and developed in $CHCl_3$-CH_3OH (9:1 [vol/vol]).

a homology search of its deduced amino acid sequences, it is not clear whether the product of *gtfD* functions as the glycosyltransferase that transfers the Fuc residue via 1→3 linkage to the Rha residue of serovar 1 GPL to form the oligosaccharide part of serovar 2 GPL. To elucidate the function of *gtfD*, we constructed a recombinant strain by introducing a chromosomal integrating vector expressing *rtfA*, *mdhtA*, and *merA* (pYMrtfA-mdhtA-merA-int). We then characterized the product formed when *gtfD* was expressed solely by the plasmid vector (pMVgtfD). TLC analysis of the recombinant strains showed that the *gtfD*-expressing strain (rtfA-mdhtA-merA-int/pMVgtfD) caused the appearance of a new spot, termed GPL-S2, in connection with the disappearance of GPL-S1 (Fig. 4, lane A), while the vector control (rtfA-mdhtA-merA-int/pMV261) did not produce GPL-S2 (Fig. 4, lane B). To determine the structure of sugar moieties of GPL-S2, perdeuteriomethylation was performed on purified GPL-S2, and derived alditol acetates were analyzed by GC-MS. The GC-MS profile yielded four peaks corresponding to 6-d-Tal, Rha, Fuc, and 2,3,4-tri-*O*-Me-Rha (data not shown). The characteristic spectra of Fuc, Rha, and 6-d-Tal are shown in Fig. 5. The spectrum of Fuc gave fragment ions at *m/z* of 121, 134, and 168, which represent the presence of deuteriomethyl groups at positions C-2, C-3, and C-4 (Fig. 5A). In contrast, the detection of fragment ions at *m/z* of 121, 134, 193, and 240 from Rha indicated that a deuteriomethyl group was introduced at positions C-2 and C-4 of Rha, whose C-3 position was acetylated (Fig. 5B). Additionally, positions C-3 and C-4 of 6-d-Tal were found to be deuteriomethylated, with the detection of fragment ions at *m/z* of 134, 181, and 193 (Fig. 5C). These observations demonstrated that position C-1 of Fuc is linked to position C-3 of Rha but not position C-2 of 6-d-Tal, because it was previously determined that position C-1 of Rha is linked to position C-2 of 6-d-Tal in the oligosaccharide of serovar 1 GPL through the catalytic reaction of RtfA (14). Furthermore, we compared the molecular mass of GPL-S2 to that of GPL-S1 purified from the vector control strain by MALDI-TOF (mass spectrometry). The main pseudomolecular ion $[M + Na]^+$ from both compounds revealed that the difference between GPL-S2 (*m/z*, 1,479.9) and GPL-S1 (*m/z*, 1,333.8) was 146 mass units, suggesting that a Fuc residue was further added to the GPL-S1 (data not shown). Accordingly, the structure of GPL-S2 was determined to have Fuc-(1→3)-Rha-(1→2)-6-d-Tal at D-*allo*-Thr and 2,3,4-tri-*O*-Me-Rha at D-alaninol (Fig. 6), demonstrating that *gtfD* encodes the glycosyltransferase that transfers a Fuc residue via 1→3 linkage to the Rha residue of serovar 1 GPL.

DISCUSSION

The gene cluster involved in synthesizing ssGPLs has been cloned from *M. avium* strains. In this cluster, the functions of several genes responsible for the biosynthesis of serovar 1 GPL, such as *rtfA*, *gtfA*, *gtfB*, *mtfB*, *mtfC*, and *mtfD*, have been elucidated (14, 16, 19). On the other hand, the genes associated with the conversion of serovar 1 GPL to other serotypes, including serovar 2 GPL, have not been identified. In this study, we focused on the five genes assumed to encode the enzymes associated with fucosylation in the biosynthesis of serovar 2 GPL and experimentally showed that GtfD is responsible for the transfer of the Fuc residue to the Rha residue of serovar 1 GPL (Fig. 6). Gene deletion experiments revealed that *mdhtA* and *merA* also contribute to the formation of the Fuc residue in serovar 2 GPL. The deduced amino acid sequences of *mdhtA* and *merA* showed high levels of similarity to GDP-D-mannose-4,6-dehydratase and GDP-6-deoxy-4-keto-D-mannose-3,5-epimerase-4-reductase, respectively. These are enzymes involved in the synthesis of L-Fuc from D-mannose and are highly conserved among other bacteria (1, 26). For mycobacteria, there are no homologues of *mdhtA* and *merA* in the genome databases for *M. bovis*, *M. leprae*, and *M. smegmatis*. However, for *M. tuberculosis*, the deduced amino acid sequences of Rv1511 and Rv1512 show 89 and 84% homology to those of *mdhtA* and *merA*, respectively. This observation is supported by the fact that several strains of *M. tuberculosis* produce the Fuc-containing phenolic glycolipid, whereas *M. bovis* and *M. leprae* lack the Fuc residue, and other carbohydrate components having the Fuc residue have not been reported from the above-mentioned three species. Thus, it is strongly suggested that *mdhtA* and *merA* encode synthetases involved in the conversion of D-mannose to L-Fuc that can be transferred by GtfD to form the Fuc residue of serovar 2 GPL (Fig. 6). Before performing the functional analyses of these genes, we speculated that the glycosyltransferase involved in the fucosylation was encoded by *gtfC*, but not *gtfD*, from the observations that *gtfC* includes a putative glycosyltransferase motif and its homologue in *M. tuberculosis*, Rv1514c, is adjacent to Rv1511 and Rv1512, which are predicted to be responsible for the Fuc synthesis, while a *gtfD* homologue, Rv2957, is located far from Rv1511, Rv1512, and Rv1514c. However, the deletion analysis demonstrated that *gtfC* does not contribute to the fucosylation of the GPLs. This result raises the possibility

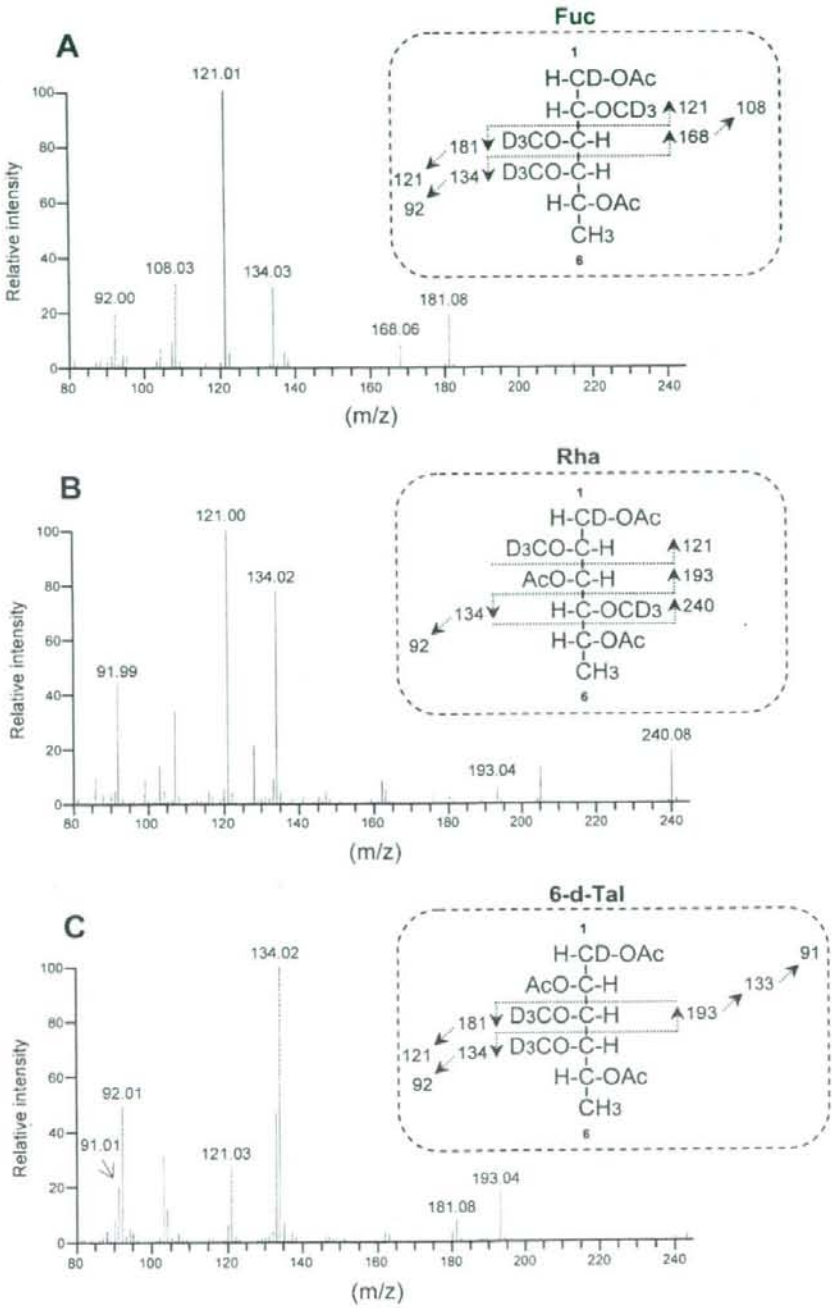


FIG. 5. GC mass spectra and fragment ion assignment of Fuc (A), Rha (B), and 6-d-Tal (C), which are derived from alditol acetates of sugars released from deuteriomethylated GPL-S2. Ac, acetate; D, deuterium.

that *gtfC* is involved in the transfer of another sugar moiety, such as glucuronic acid and Rha, followed by fucosylation, as observed for the serovar 3, 9, and 4 GPLs (9). As for *gtfD*, Rv2957 is reported to be one of the glycosyltransferase genes involved in the biosynthesis of phenolic glycolipid in *M. tuberculosis*, but its catalytic functions, such as sugar substrate and glycosidic linkage, are not clear (22). Thus, our findings implied that Rv2957 is the fucosyltransferase gene responsible for

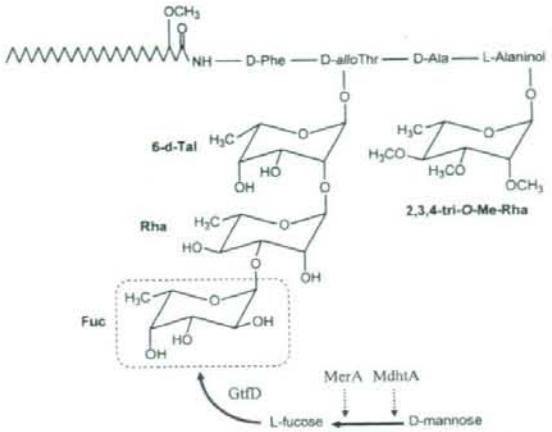


FIG. 6. Proposed structure and biosynthetic pathway of fucosylated GPL (GPL-S2).

the transfer of the Fuc residue in the phenolic glycolipid. Moreover, morphological observations showed that the surface of colony of the GPL-S2-producing strain was rougher than that of the vector control strain, suggesting that the presence of a Fuc residue in the GPL structure affected the cell surface properties (data not shown). Taking the results together, our study is the first report identifying the genes involved in the fucosylation pathway in mycobacteria and might provide a clue to understanding its role in the biosynthesis of glycolipids, including GPLs.

ACKNOWLEDGMENT

This work was supported in part by a grant in aid for research on emerging and reemerging infectious diseases from the Ministry of Health, Labor, and Welfare of Japan.

REFERENCES

1. **Andrianopoulos, K., L. Wang, and P. R. Reeves.** 1998. Identification of the fucose synthetase gene in the colanic acid gene cluster of *Escherichia coli* K-12. J. Bacteriol. **180:**998–1001.
2. **Aspinall, G. O., D. Chatterjee, and P. J. Brennan.** 1995. The variable surface glycolipids of mycobacteria: structures, synthesis of epitopes, and biological properties. Adv. Carbohydr. Chem. Biochem. **51:**169–242.
3. **Bardarov, S., S. Bardarov, Jr., M. S. Pavelka, Jr., V. Sambandamurthy, M. Larsen, J. Tufariello, J. Chan, G. Hatfull, and W. R. Jacobs, Jr.** 2002. Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG, and *M. smegmatis*. Microbiology **148:**3007–3017.
4. **Belisle, J. T., K. Klaczkiewicz, P. J. Brennan, W. R. Jacobs, Jr., and J. M. Inamine.** 1993. Rough morphological variants of *Mycobacterium avium*. Characterization of genomic deletions resulting in the loss of glycopeptidolipid expression. J. Biol. Chem. **268:**10517–10523.
5. **Belisle, J. T., L. Pascopella, J. M. Inamine, P. J. Brennan, and W. R. Jacobs, Jr.** 1991. Isolation and expression of a gene cluster responsible for biosynthesis of the glycopeptidolipid antigens of *Mycobacterium avium*. J. Bacteriol. **173:**6991–6997.
6. **Billman-Jacobe, H., M. J. McConville, R. E. Haites, S. Kovacevic, and R. L. Coppel.** 1999. Identification of a peptide synthetase involved in the biosynthesis of glycopeptidolipids of *Mycobacterium smegmatis*. Mol. Microbiol. **33:**1244–1253.
7. **Bjorndal, H., C. G. Hellerqvist, B. Lindberg, and S. Svensson.** 1970. Gas-liquid chromatography and mass spectrometry in methylation analysis of polysaccharides. Angew. Chem. Int. Ed. Engl. **9:**610–619.
8. **Brennan, P. J., and H. Nikaido.** 1995. The envelope of mycobacteria. Annu. Rev. Biochem. **64:**29–63.
9. **Chatterjee, D., and K. H. Khoo.** 2001. The surface glycopeptidolipids of mycobacteria: structures and biological properties. Cell. Mol. Life Sci. **58:** 2018–2042.
10. **Ciucanu, I., and F. Kerek.** 1984. A simple and rapid method for the permethylation of carbohydrates. Carbohydr. Res. **131:**209–217.
11. **Daffe, M., and P. Draper.** 1998. The envelope layers of mycobacteria with reference to their pathogenicity. Adv. Microb. Physiol. **39:**131–203.
12. **Daffe, M., M. A. Laneelle, and G. Puzo.** 1983. Structural elucidation by field desorption and electron-impact mass spectrometry of the C-mycosides isolated from *Mycobacterium smegmatis*. Biochim. Biophys. Acta **751:**439–443.
13. **Eckstein, T. M., J. T. Belisle, and J. M. Inamine.** 2003. Proposed pathway for the biosynthesis of serovar-specific glycopeptidolipids in *Mycobacterium avium* serovar 2. Microbiology **149:**2797–2807.
14. **Eckstein, T. M., F. S. Silbaq, D. Chatterjee, N. J. Kelly, P. J. Brennan, and J. T. Belisle.** 1998. Identification and recombinant expression of a *Mycobacterium avium* rhamnosyltransferase gene (*rtfA*) involved in glycopeptidolipid biosynthesis. J. Bacteriol. **180:**5567–5573.
15. **Jeevarajah, D., J. H. Patterson, M. J. McConville, and H. Billman-Jacobe.** 2002. Modification of glycopeptidolipids by an *O*-methyltransferase of *Mycobacterium smegmatis*. Microbiology **148:**3079–3087.
16. **Jeevarajah, D., J. H. Patterson, E. Taig, T. Sargeant, M. J. McConville, and H. Billman-Jacobe.** 2004. Methylation of GPLs in *Mycobacterium smegmatis* and *Mycobacterium avium*. J. Bacteriol. **186:**6792–6799.
17. **Julander, I., S. Hoffner, B. Petrini, and L. Ostlund.** 1996. Multiple serovars of *Mycobacterium avium* complex in patients with AIDS. APMIS **104:**318–320.
18. **Mills, J. A., M. R. McNeil, J. T. Belisle, W. R. Jacobs, Jr., and P. J. Brennan.** 1994. Loci of *Mycobacterium avium ser2* gene cluster and their functions. J. Bacteriol. **176:**4803–4808.
19. **Miyamoto, Y., T. Mukai, N. Nakata, Y. Maeda, M. Kai, T. Naka, I. Yano, and M. Makino.** 2006. Identification and characterization of the genes involved in glycosylation pathways of mycobacterial glycopeptidolipid biosynthesis. J. Bacteriol. **188:**86–95.
20. **Miyamoto, Y., T. Mukai, F. Takeshita, N. Nakata, Y. Maeda, M. Kai, and M. Makino.** 2004. Aggregation of mycobacteria caused by disruption of fibronectin-attachment protein-encoding gene. FEMS Microbiol. Lett. **236:** 227–234.
21. **Patterson, J. H., M. J. McConville, R. E. Haites, R. L. Coppel, and H. Billman-Jacobe.** 2000. Identification of a methyltransferase from *Mycobacterium smegmatis* involved in glycopeptidolipid synthesis. J. Biol. Chem. **275:**24900–24906.
22. **Perez, E., P. Constant, A. Lemassu, F. Laval, M. Daffe, and C. Guilhot.** 2004. Characterization of three glycosyltransferases involved in the biosynthesis of the phenolic glycolipid antigens from the *Mycobacterium tuberculosis* complex. J. Biol. Chem. **279:**42574–42583.
23. **Recht, J., and R. Kolter.** 2001. Glycopeptidolipid acetylation affects sliding motility and biofilm formation in *Mycobacterium smegmatis*. J. Bacteriol. **183:**5718–5724.
24. **Sambandamurthy, V. K., X. Wang, B. Chen, R. G. Russell, S. Derrick, F. M. Collins, S. L. Morris, and W. R. Jacobs, Jr.** 2002. A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. Nat. Med. **8:**1171–1174.
25. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. **4:**1911–1919.
26. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J. Bacteriol. **178:**4885–4893.
27. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, S. B. Snapper, R. G. Barletta, W. R. Jacobs, Jr., and B. R. Bloom.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456–460.
28. **Sweet, L., and J. S. Schorey.** 2006. Glycopeptidolipids from *Mycobacterium avium* promote macrophage activation in a TLR2- and MyD88-dependent manner. J. Leukoc. Biol. **80:**415–423.
29. **Vergne, I., and M. Daffe.** 1998. Interaction of mycobacterial glycolipids with host cells. Front. Biosci. **3:**d865–876.
30. **Yakrus, M. A., and R. C. Good.** 1990. Geographic distribution, frequency, and specimen source of *Mycobacterium avium* complex serotypes isolated from patients with acquired immunodeficiency syndrome. J. Clin. Microbiol. **28:**926–929.

Crystallographic Investigation of the Inhibition Mode of a VIM-2 Metallo-β-lactamase from *Pseudomonas aeruginosa* by a Mercaptocarboxylate Inhibitor

Yoshihiro Yamaguchi,*[§,⊥] Wanchun Jin,[‡] Kazuyo Matsunaga,[§] Shinnji Ikemizu,[¶] Yuriko Yamagata,[¶] Jun-ichi Wachino,[⊥] Naohiro Shibata,[⊥] Yoshichika Arakawa,[⊥] and Hiromasa Kurosaki*[§,⊥]

Environmental Safety Center, Kumamoto University, 39-1 Kurokami 2-Chome, Kumamoto 860-8555, Japan, Departments of Structure-Function Physical Chemistry and Structural Biology, Graduate School of Pharmaceutical Sciences, Kumamoto University, Oe-honmachi 5-1, Kumamoto 862-0973, Japan, and Department of Bacterial Pathogenesis and Infection Control, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashi-Murayama, Tokyo 208-0011, Japan

Received August 19, 2007

The VIM-2 metallo-β-lactamase enzyme from *Pseudomonas aeruginosa* catalyzes the hydrolysis of most β-lactam antibiotics including carbapenems, and there are currently no potent inhibitors of such enzymes. We found *rac*-2-ω-phenylpropyl-3-mercaptopropionic acid, phenylC3SH, to be a potent inhibitor of VIM-2. The structure of the VIM-2−phenylC3SH complex was determined by X-ray crystallography to 2.3 Å. The structure revealed that the thiol group of phenylC3SH bridged to the two zinc(II) ions and the phenyl group interacted with Tyr67(47) on loop1 near the active site, by π−π stacking interactions. The methylene group interacted with Phe61(42) located at the bottom of loop1 through CH−π interactions. Dynamic movements were observed in Arg228(185) and Asn233(190) on loop2, compared with the native structure (PDB code: 1KO3). These results suggest that the above-mentioned four residues play important roles in the binding and recognition of inhibitors or substrates and in stabilizing a loop in the VIM-2 enzyme.

Introduction

β-Lactamases catalyze the hydrolysis of the C−N bond of the β-lactam ring in β-lactam antibiotics, rendering them inactive.[1,2] In the family of β-lactamases, four molecular classes, A−D, were defined on the basis of their amino acid sequence homologies.[3−5] Classes A, C, and D are serine β-lactamases containing a serine residue at the active site, whereas class B metallo-β-lactamases (MBLs[a]) contain one or two zinc(II) ions at the active site. Moreover, MBLs are classified into three subclasses:[6,7] subclass B1 including BcII,[8] an IMP family (IMP-1 to IMP-18),[9] CcrA,[10] and a VIM family (VIM-1 to VIM-11a and -11b);[9] subclass B2 including CphA[11] and ImiS;[12] and subclass B3 including L1[13] and THIN-B.[14] Of all the MBLs, the VIM family and the IMP family are emerging as a worldwide source of acquired carbapenem resistance among Gram-negative bacteria. In variants of VIMs, VIM-2 was first detected from *Pseudomonas aeruginosa* in France in 1996[15] and is now one of the most widespread enzymes in well-known MBLs. The VIM-2 enzyme shows a 24–31% identity with other subclass B1 MBLs such as BlaB, CcrA, and IMP-1.[15]

The bla_{VIM-2} gene is located in a mobile plasmid-encoded gene cassette inserted into the variable region of an integron structure so that horizontal spread of resistance is possible. In fact, bla_{VIM-2} genes have been detected mostly in the Mediterranean countries of Europe, in the Far East regions, including Japan, and at present, in American regions, including the U.S., whereas bla_{IMP-1} genes were first detected mainly in Japan and subsequently isolated in Asia, Europe, and South America.[9] Moreover, in addition to the high resistance against β-lactam antibiotics, there are no clinically available inhibitors for the VIM-2 enzyme including other MBLs at present. Thus, the development of inhibitors is an important subject. The active site in an MBL is made up of the zinc(II) ion(s), the zinc(II) ion binding residues, and two loops. In the two loops, the first loop (loop1) is composed of two β-sheets and a turn constitutes a β-sheet flap, whereas the second loop (loop2) is approximately positioned in the opposite site of loop1 centered about the zinc(II) ion binding site. It is thought that these loops are responsible for substrate recognition, binding, and catalysis in MBLs.[16−22]

To develop potent and common inhibitors, it is necessary to focus on the structure and dynamics of both the zinc(II) ion binding site and loops 1 and 2 in each MBL and its complex with a lead compound.

So far, X-ray crystal structures of MBLs complexed with various inhibitors have been reported with the exception of the VIM-2 enzyme: CcrA−4-morpholineethanesulfonic acid,[23] CcrA−biphenyltetrazoles,[24] IMP-1−mercaptocarboxylate derivative,[22] IMP-1−2,3-disubstituted succinic acid derivatives,[25] IMP-1−irreversible thiol compound with a good leaving group,[26] IMP-1−dansylC4SH,[27] BlaB−D-captopril,[28] and FEZ−D-captopril.[29]

In VIM-2, in 2001, García-Sáez et al. deposited two X-ray crystal structures of VIM-2 in the reduced (denoted as the native VIM-2 enzyme) and oxidized forms of Cys221(178) from *P. aerginosa* into the Protein Data Bank (PDB codes 1KO2 and 1KO3). As far as we know, however, detailed information on the structure of the VIM-2 enzyme complexed with an inhibitor is not available.

* To whom correspondence should be addressed. For Y.Y.: phone, +81-96-342-3238; fax, +81-96-342-3237; e-mail, yyamagu@gpo.kumamoto-u.ac.jp. For H.K.: phone and fax, +81-96-371-4314; e-mail, ayasaya@gpo.kumamoto-u.ac.jp.

§ Environmental Safety Center, Kumamoto University.

‡ These authors contributed equally to work.

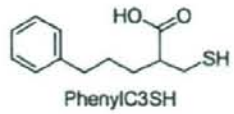
¶ Department of Structure-Function Physical Chemistry, Kumamoto University.

‖ Department of Structural Biology, Kumamoto University.

⊥ National Institute of Infectious Diseases.

[a] Abbreviations: BBL, class B β-lactamase; MAD, multiwavelength anomalous dispersion; MBLs, metallo-β-lactamases; PEG MME5000, polyethylene glycol monomethyl ether 5000; phenylC3SH, 2-ω-phenylpropyl-3-mercaptopropionic acid. In this paper, the amino acid residues of metallo-β-lactamase are designated with the BBL number and the amino acid sequence number by putting the latter in parentheses.

10.1021/jm701031n CCC: $37.00
Published on Web 12/06/2007



PhenylC3SH

Figure 1. Structure of *rac*-2-*α*-phenylpropyl-3-mercaptopropionic acid, phenylC3SH.

We are currently in the process of synthesizing several inhibitors and screening them for inhibitory activity against the VIM-2 and IMP-1 enzymes for the purpose of developing inhibitors for MBLs. In this process, the mercaptocarboxylate inhibitor, PhenylC3SH (Figure 1), was found to be a potent inhibitor of the VIM-2 enzyme with a K_i value of 220 nM, whereas it is less active for the IMP-1 enzyme (K_i = 1660 nM).[30]

For the purpose of understanding structure–activity relationships of the VIM-2 enzyme with respect to phenylC3SH, we carried out X-ray crystallography of the VIM-2 enzyme complexed with phenylC3SH.

Here, we report on the X-ray crystal structure of the VIM-2 enzyme from *P. aeruginosa* complexed with phenylC3SH determined at a resolution of 2.3 Å.

Results and Discussion

In this study, we determined the crystal structure of the VIM-2 enzyme complexed with phenylC3SH in order to elucidate the detailed binding mode of the inhibitor with the enzyme, in particular, focusing on the role of the mobile loop, along with a comparison of the native VIM-2 structure (PDB code 1KO3).

Overall Structure of the VIM-2 Enzyme Complexed with PhenylC3SH. The final refined structural model contained two VIM-2−phenylC3SH molecules in an asymmetric unit, consisting of Glu30(12)−Val290(236) residue each for molecules A and B. The root-mean-squared deviation (rmsd) value between the α-carbon atoms of the two monomers was about 0.28 Å. In molecules A and B, two zinc(II) ions and one phenylC3SH molecule are located at the active sites. The final $R_{working}$ and R_{free} values were 0.212 and 0.247, respectively, and the rmsd values from ideal bond distances and angles were 0.007 Å and 2.0°, respectively (Table 1). The overall structure of the inhibitor complex adopts an $\alpha\beta\beta\alpha$ sandwich structure, where the active site containing the two zinc(II) ions is located at the bottom at the interface of the two β-sheets surrounded by two α-helices, which resemble those of other subclass B1 MBLs.[22,28,31]

In addition, a clear electron density corresponding to a phenylC3SH molecule was found in the active center (Figures 2 and 3). Although the crystals were grown in the presence of a racemic mixture of phenylC3SH, only the *S*-isomer of the inhibitor was observed in the complex.

All main chain dihedral angles were within the allowed regions of a Ramachandran plot except Asp84(64) and Ala195(158) which adopted φ, ψ angles of 73/74° (molecules A/B), 151/149°, and −151/−157°, −106/−106°, respectively. Asp84(64) is buried in the protein, whereas Ala195(158) is positioned on top of a hairpin loop formed from Tyr191(154)−Val202(165). Asp84(64) and Ala195(158) have strained main chain conformations in both the native VIM-2 enzyme and the phenylC3SH complex. In the crystal structures of BcII, CcrA, and IMP-1 metallo-β-lactamases,[22,31,32] Asp84s were found to have a common strained conformation. Therefore, the conservation of the conformation of Asp84 suggests that it is important for the folding of metallo-β-lactamases. In additional supporting evidence, tRNA maturase RNase Z (PDB code 1Y44)[33] and pre-mRNA 3′-end-processing endonuclease CPSF-73 (PDB code 2I7T)[34] in the metallo-β-lactamase family also have strained main chain conformations of the Asp residues, as has been found for Asp84 of MBLs.

The structures of molecule A in an asymmetric unit of the phenylC3SH complex and that of the native VIM-2 enzyme (PDB code 1KO3) are superimposable except for residues Lys291(237)−Asn295(241), and the rmsd for the α-carbon atoms between them was 0.56 Å. In a comparison of the deviations of the α-carbon atoms between each of the amino acid residues of the native VIM-2 enzyme and the inhibitor complex, residues that moved more than 1 Å were Val34(16) (1.4 Å), Ser35(17) (2.7 Å), Glu36(18) (2.3 Å), Ile37(19) (1.0 Å), Pro38(20) (1.2 Å), Val39(21) (1.1 Å), Ser60(41) (1.5 Å), Phe61(42) (1.8 Å), Asp62(43) (1.9 Å), Gly63(44) (2.0 Å), Ala64(45) (1.4 Å), Val66(46) (1.2 Å), and Gly232(189) (1.8 Å), respectively.

Of these residues, Val34(16), Ser35(17), Glu36(18), Ile37(19), Pro38(20), and Val39(21) were located at the N terminus and Ser60(41), Phe61(42), Asp62(43), Gly63(44) (2.3 Å), Ala64(45), and Val66(46) were located on loop1, whereas Gly232(189) was located on loop2.

Active Site of the VIM-2 Enzyme Complexed with PhenylC3SH. In the phenylC3SH complex, one of two zinc(II) ions (Zn1) was coordinated to His116(94), His118(96), and His196(159) residues with distances of 2.2/2.1 Å (in molecules A/B), 2.0/2.2 Å, and 2.2/2.2 Å, respectively.

The thiol group of the inhibitor is bridged to the two zinc(II) ions (the distances of Zn1−S and Zn2−S were 2.5/2.2 Å and 2.1/2.2 Å), respectively, forming a tetrahedral coordination around the Zn1 atom (the average angle of ligand−Zn1−ligand is 109° for molecule A and 110° for molecule B; Figure 3, Table 2). The coordination of the thiol group of an inhibitor to two zinc(II) ions is similar to those found in IMP-1 complexed with 2-[5-(1-tetrazoylmethyl)thien-3-yl]-*N*-[2-(mercaptomethyl)-4-(phenylbutyrylglycine)][22] or with dansylC4SH[27] and in BlaB complexed with D-captopril.[28]

The second zinc(II) ion (Zn2) is also tetrahedrally coordinated by Asp120(98), Cys221(178), and His263(220) residues, and the thiol group of the inhibitor with distances of 1.9/2.0, 2.2/2.2, 2.1/2.1, and 2.1/2.2 Å, respectively (the average angles of ligand−Zn2−ligand is 109° for molecules A and B, respectively; Figure 3, Table 2). The Zn1−Zn2 distance is 3.8 Å for molecule A and 3.7 Å for molecule B, and this distance is close to those found in the IMP-1 and BlaB enzymes in complex with an inhibitor containing a thiol group (the average distance of 3.6–3.7 Å).[22,27,28]

In the native VIM-2 structure, the geometry around the Zn1 atom is a distorted tetrahedral with three His residues (His116(94), His118(96), and His196(159)) and H_2O/or OH^- (this H_2O/or OH^- is thought to act as the attacking nucleophile on the carbonyl carbon atom of the β-lactam ring),[31,35,36] bridging to the Zn1 and Zn2 atoms, whereas that around the Zn2 atom is a distorted square-pyramid with the basal plane defined by a H_2O/or OH^-, Asp120(98), Cys221(178), His263(220), and Cl atom, where the Cl atom interacts weakly with the Zn2 atom with a distance of 2.9 Å. The apical position is occupied by Asp120(98). In the crystal structures of other subclass B1 MBLs (CcrA and IMP-1),[22,31] the Cl atom in the Zn2 site of the native VIM-2 enzyme is replaced by H_2O, which is thought to contribute to the catalysis of the hydrolysis of β-lactam antibiotics.[31,37] Unlike the native VIM-2 enzyme, the coordination geometry around the Zn2 atom in CcrA and IMP-1 is a distorted trigonal bipyramid.

Table 1. Crystallographic Data Collection and Refinement Statistics for the VIM-2 Enzyme Complexed with PhenylC3SH

data collection				
data set wavelength (Å)	1.0000 (final)	1.2826 (edge)	1.2817 (peak)	1.2906 (remote)
resolution (outer shell) (Å)	50.0–2.30 (2.38–2.30)	99.0–2.53 (2.63–2.53)	99.0–2.54 (2.63–2.54)	99.0–2.55 (2.65–2.55)
cell dimensions				
a, *b*, and *c* (Å)	45.2, 90.75, 129.0			
space group	$P2_12_12_1$			
molecules/asymmetric unit	2			
completeness (outer shell) (%)	99.6 (98.4)	99.2 (96.3)	99.1 (96.7)	99.0 (94.8)
R_{merge}[a] (outer shell)	0.062 (0.267)	0.047 (0.106)	0.050 (0.097)	0.046 (0.121)
no. of observed reflns	163206	123781	122809	121255
no. of unique reflns	24156	18155	18121	17742
$\langle I/\sigma(I)\rangle$ (outer shell)	34.9 (4.45)	54.2 (25.5)	55.5 (30.1)	53.4 (23.3)
refinement statistics				
resolution (Å)	42.8–2.30			
no. of non-H atoms[b]				
protein	3392			
ligand	50			
water	212			
rmsd from ideal[c]				
bond length (Å)	0.007			
angles (deg)	2.0			
$R_{working}$[d]	0.212			
R_{free}[e]	0.247			

[a] $R_{merge} = \sum|I_j - \langle I_j\rangle|/\sum I_j$, where I_j is the observed intensity for reflection j and $\langle I_j\rangle$ is the average intensity calculated for reflection j from replicate data. [b] Per asymmetric unit. [c] rmsd: root-mean-square-deviation. [d] $R_{working} = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. [e] $R_{free} = \sum||F_o| - |F_c||/\sum|F_o|$ for 5% of the data not used at any stage of structural refinement.

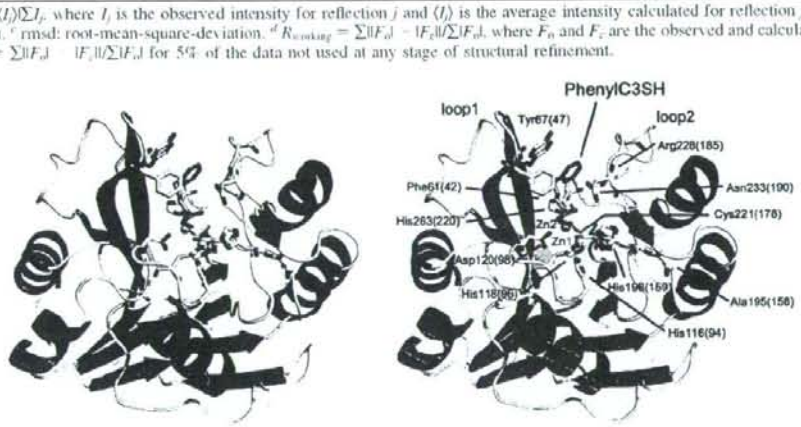


Figure 2. Overall structure of the VIM-2 enzyme from *P. aeruginosa* is complexed with phenylC3SH. Molecule A of the phenylC3SH complex is depicted, and the amino acid residues of VIM-2 are designated by a BBL number and an amino acid sequence number; the latter is in parentheses. α-Helices, β-strands, and loops are shown in red, green, and yellow, respectively. Zn(II) atoms are shown as orange spheres. Carbon, oxygen, and sulfur atoms in phenylC3SH are shown as magenta, red, and light-green sticks, respectively. The figure was prepared with PyMol software (http://pymol.sourceforgen.net/).

Comparison of Loop between the Native VIM-2 Enzyme and the PhenylC3SH Complex. The VIM-2 enzyme contains loop1 (Phe61(42)−Ala64(45)) and loop2 (Ile223(180)−Trp242(199)).

Upon inhibitor binding, significant structural changes were found in both loops. The amino acid residues on loop1 are thought to be the most important for substrate recognition and binding, catalysis, and inhibition.[16-19,22] In the phenylC3SH complex, Tyr67(47) and Phe61(42), located on loop1, moved toward the active site to facilitate the interaction with the inhibitor (Figure 4). Compared to the native VIM-2 structure, Tyr67(47) rotates by ~27° about the Cβ−Cγ bond of Tyr67(47) to provide face-to-face π−π stacking interactions between the phenyl rings of the inhibitor and Tyr67(47) with spacing planes of ~3.6 Å (Figures 3 and 4). At the same time, the phenyl ring of Phe61(42) lie ~1.3 Å closer to Zn2 (Figure 4), forming CH−π interactions with the methylene chain of the inhibitor, and these interactions may contribute to the stabilization of loop1. In addition, the active site of the phenylC3SH complex is transformed from an opened cavity into a tunnel-shaped cavity upon binding of an inhibitor (Figure 5).

In loop2 of the phenylC3SH complex, the carboxyl group (O2) of the inhibitor interacts through a hydrogen bond with the side chain ND2 of Asn233(190) (2.9/2.6 Å in molecules A/B), which is conserved in most MBLs.[36] A comparison of the structures between the native VIM-2 and the inhibitor complex showed that conformational change in Asn233(190) is accomplished by the rotation of the neighboring amino acid Gly232(189) (ψ and ψ angles are −85° and −119°, respectively, for the native VIM-2 enzyme, whereas those of the phenylC3SH complex are 73° and −143° for molecule A and 63° and −143° for molecule B, respectively; Figure 4). In addition, the binding of the inhibitor to the active site triggers a conformational change in the side chain Arg228(185) (Figure 4); the torsion angle of CB−CG−CD−NE of Arg228(185) is changed from −48° in the native VIM-2 enzyme to 172° for molecule A and −175° for molecule B, respectively, in the phenylC3SH complex. These results suggest that Phe61(42), Tyr67(47), Arg228(185), and Asn233(190) are functionally important residues that play roles in the binding and recognition of the inhibitor or substrate and in stabilizing loops 1 and 2.

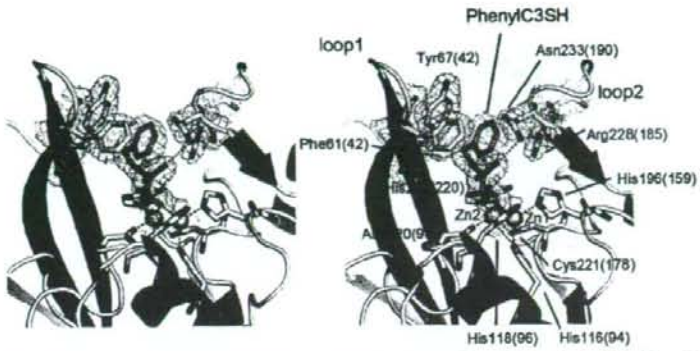


Figure 3. Stereoview of the active site of the VIM-2 enzyme complexed with phenylC3SH. Molecule A of the phenylC3SH complex is depicted, and the amino acid residues of VIM-2 are labeled with a BBL number and an amino acid sequence number; the latter is in parentheses. The electron density map (pink mesh) of phenylC3SH is shown contoured at the 1.0σ level in the $2|F_o| - |F_c|$ map. Zn(II) atoms are shown as orange spheres. Phe61(42), Tyr67(47), His116(94), His118(96), His196(159), His263(220), Asp120(98), Cys221(178), Arg228(185), and Asn233(190) residues and phenylC3SH are represented as sticks. Carbon atoms in amino acid residues are shown in gray (nitrogen, blue; oxygen, red; and sulfur, light-green), and carbon atoms in phenylC3SH are shown in magenta (oxygen, red; sulfur, light-green). The figure was prepared with PyMol software (http://pymol.sourceforgen.net/).

Table 2. Zinc(II)−Ligand Distances (Å) and Angles (deg) for the Native VIM-2 Enzyme and the PhenylC3SH Complex

		distance	
Zn(II)−ligand		native VIM-2	VIM-2−phenylC3SH complex[a]
Zn1	His116(94)NE2	2.2	2.0/2.1
	His118(96)ND1	2.1	2.0/2.2
	His196(159)NE2	2.2	2.2/2.2
	O(w)	2.1	
	S(phenylC3SH)		2.5/2.2
Zn2	Asp120(98)OD2	2.3	1.9/2.0
	Cys221(178)SG	2.3	2.2/2.2
	His263(220)NE2	2.3	2.1/2.1
	O(w)	2.5	
	S(phenylC3SH)		2.1/2.2
	Cl	2.9	
Zn1	Zn2	4.2	3.8/3.7

			angle	
ligand−Zn(II)−ligand			native VIM-2	VIM-2−phenylC3SH complex[a]
His116(94)NE2	Zn1	His118(96)ND1	103	105/105
	Zn1	His196(96)NE2	103	120/110
	Zn1	O(w)	103	
	Zn1	S(inhibitor)		116/126
His118(96)ND1	Zn1	His196(159)NE2	116	101/107
	Zn1	O(w)	119	
	Zn1	S(inhibitor)		113/112
His196(159)NE2	Zn1	O(w)	111	
	Zn1	S(inhibitor)		101/97
Asp120(98)OD2	Zn2	Cys221(178)SG	103	114/117
	Zn2	His263(220)NE2	94	96/99
	Zn2	O(w)	74	
	Zn2	S(inhibitor)		100/100
	Zn	Cl	101	
Cys221(178)SG	Zn2	His263(220)NE2	103	104/105
	Zn2	O(w)	98	
	Zn2	S(inhibitor)		124/116
	Zn2	Cl	152	
His263(220)NE2	Zn2	O(w)	157	
	Zn2	S(inhibitor)		116/119
	Zn2	Cl	103	

[a] The distances and angles for the phenylC3SH complex are quoted for molecules A and B in the asymmetric unit.

Comparison of Loops 1 and 2 between the IMP-1 and VIM-2 Enzymes. We compared the amino acid residues on loops 1 and 2 between the VIM-2 and IMP-1 enzymes (Figure 6). In the case of the IMP-1 enzyme,[22] Trp64(28) is located on the top of loop1, and this residue is thought to be important for the binding of both inhibitors and substrates. Indeed, the role of Trp64(28) in the binding of inhibitors has been demonstrated in X-ray crystal structures of the IMP-1−mercaptocarboxylate inhibitor and −dansylC4SH complexes; the indole ring of Tyr64(28) interacts with the aromatic ring of the inhibitor, and this interaction causes a dynamic movement of loop1 to cover an inhibitor into the active site.[22,27] On the other hand, Trp64(28) in the IMP-1 enzyme is replaced with Ala64(45) in the VIM-2 enzyme. Therefore, in the VIM-2 enzyme, Trp64(28) in IMP-1 is covered with the hydrophobic residues Phe61(42) and Tyr67(47) located in the root of the loop1. Phe61(42) and Tyr67(47) in the VIM-2 enzyme are replaced by Val61(25) and Val67(31), respectively, in the IMP-1 enzyme. It is thought that these residues could not participate in hydrophobic contacts with the inhibitor seen in the crystal structure of the phenylC3SH complex, reflecting the difference

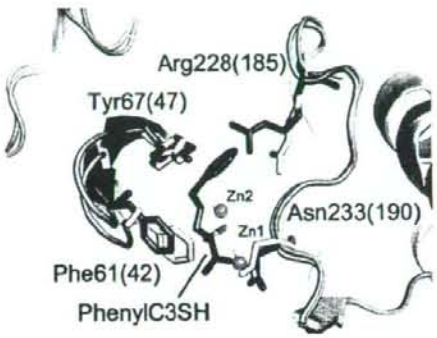


Figure 4. Conformational changes in loops 1 and 2 upon phenylC3SH binding to the VIM-2 enzyme. Superposition of native VIM-2 (green) (PDB code 1KO3) and the phenylC3SH complex (yellow). Molecule A of the phenylC3SH complex is depicted and the amino acid residues of VIM-2 are designated with a BBL number and an amino acid sequence number; the latter is in parentheses. Phe61(42), Tyr67(47), Arg228(185), and Asn233(190) residues are presented as balls and sticks. Zn(II) ions are shown as orange spheres. PhenylC3SH is presented as a stick (carbon, oxygen, and sulfur atoms are shown in magenta, red, and light-green, respectively). The figure was prepared with PyMol software (http://pymol.sourceforgen.net/).