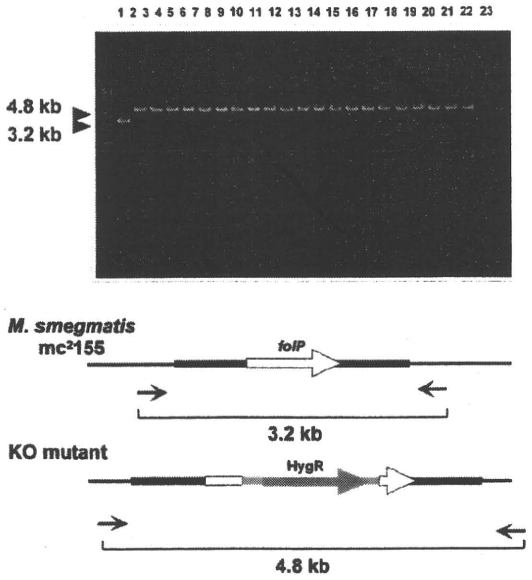

FIG. 2. PCR analysis to confirm the disruption of *folP*. Black arrows represent primers MSFPF and MSFPR for the PCR amplification. Lane 1, *M. smegmatis* mc²155; lanes 2 to 22, *M. smegmatis* strains carrying the *M. leprae folP1* without mutation and *folP1* with mutations 48-1,48-2, 48-3, 48-5, 48-6, 53-1, 53-2, 53-3, 53-4, 53-5, 53-6, 54-1, 54-2, 54-3, 55-1, 55-2, 55-3, 55-4, 55-5, and 55-6, respectively; lane 23, negative control. KO, knockout.

in amino acids from threonine to serine (T53S) encoded by different nucleotide sequences (53-3 and 53-6) were more susceptible to dapsone than strains with the wild-type *folP1* sequence. MIC values for strains with mutations at codon 48 or 54 were comparable to MICs for strains with the wild-type sequence. MIC values of dapsone for the recombinant *M. smegmatis* strains are listed in Table 3. Using a multicopy plasmid may affect the expression levels of the *M. leprae folP1* and MIC values. Therefore, we tested all the mutations using

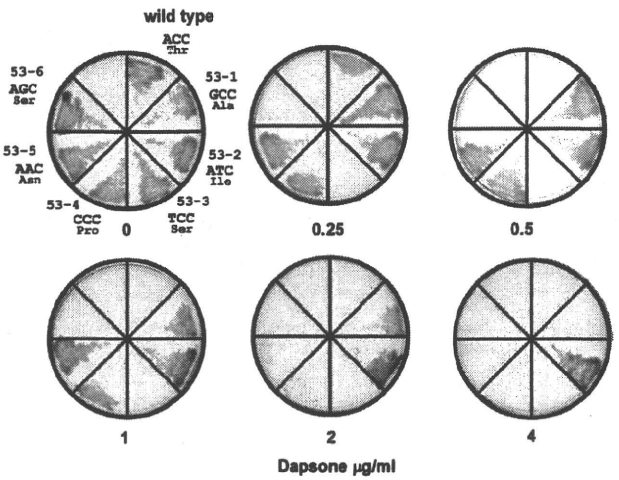

FIG. 3. Dapsone susceptibility of recombinant *M. smegmatis*. Results for *M. smegmatis* strains with point mutations at codon 53 of the *M. leprae folP1* are shown. Dapsone concentration is depicted below each plate.

TABLE 3. Dapsone susceptibility of the recombinant *M. smegmatis* strains

Strain or mutation	Dapsone MIC (μg/ml)	Reference of footpad test
Wild type	0.5	
48-1 (Val → Ile)	0.5	
48-2 (Val → Leu)	0.5	
48-3 (Val → Phe)	1.0	
48-4 (Val → Asp)	—[a]	
48-5 (Val → Ala)	1.0	
48-6 (Va → Gly)	1.0	
53-1 (Thr → Ala)	4.0	6
53-2 (Thr → Ile)	8.0	6, 10, 16
53-3 (Thr → Ser)	0.25	
53-4 (Thr → Pro)	2.0	
53-5 (Thr → Asn)	2.0	
53-6 (Thr → Ser)	0.25	
54-1 (Arg → Arg)	0.5	
54-2 (Arg → Gly)	1.0	
54-3 (Arg → Trp)	0.5	
55-1 (Pro → Thr)	1.0	
55-2 (Pro → Ala)	2.0	
55-3 (Pro → Ser)	2.0	
55-4 (Pro → His)	2.0	
55-5 (Pro → Arg)	8.0	6, 16
55-6 (Pro → Leu)	4.0	6, 10

[a] Isolation of an *folP*-disrupted *M. smegmatis* strain carrying the *M. leprae folP1* with mutation 48-4 was unsuccessful.

pNN301, a single-copy integrative vector, instead of pMV261 and obtained MIC values identical to those obtained with pMV261, suggesting that the expression levels did not influence the MIC values.

DISCUSSION

We first attempted using *E. coli* C600 Δ*folP*::Km^r transformants to determine the MIC of dapsone, but susceptibility of the recombinant *E. coli* strains to dapsone was not stable even in Mueller-Hinton medium. Subsequently, we tried to isolate a *folP*-deficient *M. smegmatis* strain by allelic exchange, given the closer association of *M. smegmatis* to *M. leprae* than *E. coli*. The selection held great promise as total-sequence comparison of *M. leprae* DHPS with *M. smegmatis* DHPS indicated 83% identity, whereas the identity between *M. leprae* DHPS and *E. coli* DHPS is only 41%, indicating the higher potential of *M. smegmatis* as a host for measuring MIC values of dapsone for *M. leprae* DHPS. However, isolation of *folP*-deficient *M. smegmatis* was unsuccessful. In *E. coli*, DHPS is not essential for bacterial growth when the cells are cultured with thymidine (7), but DHPS activity may be essential for the growth of *M. smegmatis* as it could not be replaced by any of the supplemented culture media tested. Hence, we then attempted to disrupt the *folP* gene on the *M. smegmatis* chromosome after introducing the *M. leprae folP1* gene into the cell to compensate for DHPS activity.

Comparison of the DHPS structures of *E. coli*, *Staphylococcus aureus*, and *Mycobacterium tuberculosis* has suggested that Ser53 and Pro55 of the *M. tuberculosis* DHPS, which correspond to Thr53 and Pro55 in *M. leprae*, may be the major sites of interaction with PABA, dapsone, and sulfonamides (1, 2, 8). In the present study, all mutations that cause amino acid sub-

stitutions at codon 55 resulted in dapsone resistance. Mutations at codon 53 also gave rise to dapsone resistance except for the T53S substitution, which resulted in less resistance to dapsone than the wild-type sequence (Fig. 3). The results for mutation 53-1, 53-2, 55-5, and 55-6 are consistent with the mouse footpad dapsone susceptibility testing of the *M. leprae* clinical isolates (6, 10, 16). Mutations at codon 48 or 54 showed comparable levels of susceptibility to dapsone as the wild-type sequence using dapsone susceptibility testing, but the MIC values for mutations 48-3, 48-5, 48-6, and 54-2 were slightly higher than the MIC for the wild-type sequence. Mutation 48-5 for V48A, which has been detected in our clinical samples (unpublished data), might give rise to low-level resistance to dapsone in *M. leprae*. This level of resistance should be very carefully examined by comparison with the results of footpad testing and clinical data. These data will help the molecular diagnosis of dapsone-resistant *M. leprae* with the goal of avoiding the wrong choice of drugs for chemotherapy.

Although these results should always be initially confirmed by clinical susceptibility testing as well, we believe that the method established in this study should have great utility in further attempts to determine the mutations responsible for giving rise to the dapsone resistance of *M. leprae*. The advantage of this method lies in the ability to functionally replace an essential gene of fast-growing mycobacteria with the *M. leprae* counterpart. The method may also be applicable to analysis of the rifampin resistance and quinolone resistance of *M. leprae*.

ACKNOWLEDGMENTS

E. coli C600 Δ*folP*::Km^r was kindly given by G. Swedberg (Uppsala University, Uppsala, Sweden). pYUB854 and phAE87 were kindly given by W. R. Jacobs, Jr. (Albert Einstein College of Medicine, New York, NY).

This work was supported by grants from the Ministry of Health, Labor and Welfare (Emerging and Re-Emerging Infectious Diseases) and the Ohyama Health Foundation.

REFERENCES

1. **Achari, A., et al.** 1997. Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. Nat. Struct. Biol. **4**:490–497.
2. **Baca, A. M., R. Sirawaraporn, S. Turley, W. Sirawaraporn, and W. G. Hol.** 2000. Crystal structure of *Mycobacterium tuberculosis* 7,8-dihydropteroate synthase in complex with pterin monophosphate: new insight into the enzymatic mechanism and sulfa-drug action. J. Mol. Biol. **302**:1193–1212.
3. **Bardarov, S., et al.** 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. **Baumstark, B. R., L. L. Spremulli, U. L. RajBhandary, and G. M. Brown.** 1977. Initiation of protein synthesis without formylation in a mutant of Escherichia coli that grows in the absence of tetrahydrofolate. J. Bacteriol. **129**:457–471.
5. **Brown, G. M.** 1962. The biosynthesis of folic acid. II. Inhibition by sulfonamides. J. Biol. Chem. **237**:536–540.
6. **Cambau, E., L. Carthagena, A. Chauffour, B. Ji, and V. Jarlier.** 2006. Dihydropteroate synthase mutations in the *folP1* gene predict dapsone resistance in relapsed cases of leprosy. Clin. Infect. Dis. **42**:238–241.
7. **Fermer, C., and G. Swedberg.** 1997. Adaptation to sulfonamide resistance in *Neisseria meningitidis* may have required compensatory changes to retain enzyme function: kinetic analysis of dihydropteroate synthases from *N. meningitidis* expressed in a knockout mutant of *Escherichia coli*. J. Bacteriol. **179**:831–837.
8. **Hampele, I. C., et al.** 1997. Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. J. Mol. Biol. **268**:21–30.
9. **Kai, M., et al.** 1999. Diaminodiphenylsulfone resistance of *Mycobacterium leprae* due to mutations in the dihydropteroate synthase gene. FEMS Microbiol. Lett. **177**:231–235.
10. **Maeda, S., et al.** 2001. Multidrug resistant *Mycobacterium leprae* from patients with leprosy. Antimicrob. Agents Chemother. **45**:3635–3639.
11. **Rees, R. J.** 1964. Mycobacterial disease in man and animals. Studies on leprosy bacilli in man and animals. Proc. R. Soc Med. **57**:482–483.
12. **Richey, D. P., and G. M. Brown.** 1969. The biosynthesis of folic acid. IX. Purification and properties of the enzymes required for the formation of dihydropteroic acid. J. Biol. Chem. **244**:1582–1592.
13. **Stover, C. K., et al.** 1991. New use of BCG for recombinant vaccines. Nature **351**:456–460.
14. **Swedberg, G., S. Ringertz, and O. Skold.** 1998. Sulfonamide resistance in *Streptococcus pyogenes* is associated with differences in the amino acid sequence of its chromosomal dihydropteroate synthase. Antimicrob. Agents Chemother. **42**:1062–1067.
15. **Vedantam, G., G. G. Guay, N. E. Austria, S. Z. Doktor, and B. P. Nichols.** 1998. Characterization of mutations contributing to sulfathiazole resistance in *Escherichia coli*. Antimicrob. Agents Chemother. **42**:88–93.
16. **Williams, D. L., L. Spring, E. Harris, P. Roche, and T. P. Gillis.** 2000. Dihydropteroate synthase of *Mycobacterium leprae* and dapsone resistance. Antimicrob. Agents Chemother. **44**:1530–1537.

MAJOR ARTICLE

Analysis of Drug-Resistant Strains of *Mycobacterium leprae* in an Endemic Area of Vietnam

Masanori Kai,[1] Nhu Ha Nguyen Phuc,[2] Hoang An Nguyen,[2] Thi Hoang Bich Diu Pham,[2] Khanh Hoa Nguyen,[2] Yuji Miyamoto,[1] Yumi Maeda,[1] Yasuo Fukutomi,[1] Noboru Nakata,[1] Masanori Matsuoka,[1] Masahiko Makino,[1] and Thanh Tan Nguyen[2]

[1]Department of Mycobacteriology, Leprosy Research Center, National Institute of Infectious Diseases, Higashimurayama, Tokyo, Japan; and [2]Quyhoa National Leprosy and Dermato-Venereology Hospital, Quynhon City, Binhdinh, Vietnam

(See brief report by Ramien and Wong, e133–e135.)

Background. Multidrug therapy has effectively reduced the number of leprosy cases in the world. However, the rate of reduction has decelerated over the years, giving early detection of *Mycobacterium leprae* and epidemiological study of relapse renewed relevance in attempts to eliminate the disease.

Methods. A molecular epidemiological survey for drug-resistant *M. leprae* was conducted in the central and highland regions of Vietnam. A total of 423 samples taken from patients, including 83 patients with new cases, 321 patients receiving treatment, and 19 patients with relapse, were studied for detection of *M. leprae* with mutations relating to drug resistance by sequencing the drug resistance determining region of the *folP1*, *rpoB*, and *gyrA* genes, which are responsible for dapsone, rifampicin, and ofloxacin resistance, respectively.

Results. Nineteen mutations were found in the *folP1* gene samples, and no mutations relating to drug resistance were found in either the *rpoB* or *gyrA* genes. Samples from patients with relapse showed *folP1* mutation rates as high as 57%, and the mutation rates in samples from new and recent cases were <10%. Patients with relapse who had histories of treatment with dapsone monotherapy showed high mutation rates (78%), compared with patients with relapse who had previously only received multidrug therapy (33%).

Conclusions. Our study indicated high rates of dapsone resistance in patients with relapse, compared with patients with new and recent cases of leprosy. Moreover, it was observed that many of the patients with relapse who had dapsone-resistant mutations had histories of treatment with dapsone monotherapy.

Leprosy is a chronic infectious disease caused by infection with *Mycobacterium leprae*. The present strategy for leprosy control is based on the multidrug therapy (MDT), recommended by the World Health Organization (WHO) [1], which has successfully reduced the number of leprosy cases in the world. However, transition in the number of registered cases and new cases amounting to ~210,000 and ~250,000, respectively, has almost come to a standstill [2]. Drug-resistant strains were first found in 1964, 1976, and 1997 [3–5]. MDT was designed to prevent the emergence and spread of drug-resistant strains. However, a strain showing resistance to both dapsone and rifampicin was reported in 1993 [6], and at present, there are further reports indicating the emergence of *M. leprae* strains resistant to multiple drugs [5, 7]. At present, the rapid detection and control of such drug-resistant strains is essential in countries approaching leprosy elimination levels, such as Vietnam.

MDT has been quite successful in Vietnam, and elimination of leprosy (prevalence rate, < 1/10,000 population) was achieved on the national level in 1995 [8]. The prevalence rate per 10,000 population in 2006

Received 28 August 2010; accepted 7 December 2010.

Correspondence: Masanori Kai, MD, Department of Mycobacteriology, Leprosy Research Center, National Institute of Infectious Diseases, 4-2-1 Aoba-cho, Higashimurayama, Tokyo 189-0002, Japan (mkai@nih.go.jp).

Clinical Infectious Diseases 2011;52(5):e127–e132

1058-4838/2011/525-0001$37.00
DOI: 10.1093/cid/ciq217

was .07 [8, 9]. However, the majority of patients with leprosy are found in the central and highland regions of Vietnam [10], consisting of 11 provinces, including 4 provinces in the highland region and 7 provinces in the delta region. In 2005, the number of patients with leprosy was 236, spread through 4 provinces of the highland region; the prevalence rate of newly detected cases was 3.5 cases/10,000 population, although the overall prevalence rate was .25 cases/100,000 population on the national level. The rate of newly detected cases in the 7 delta region provinces was 1.38 cases/10,000 population [8, 9]. These cases not only present the danger of being possible infectious sources for leprosy but also harbor the risk of developing into relapse cases. However, little is known regarding the effects of drug-resistant *M. leprae* in patients with leprosy, especially in cases of relapse.

Therefore, in the present study, molecular epidemiological studies on drug-resistant strains were conducted in 11 provinces primarily in the central and highland regions that represent the areas where leprosy is endemic in Vietnam.

MATERIALS AND METHODS

Sensitivity of Polymerase Chain Reaction

The number of bacilli isolated from nude mice footpads was counted using the method described by Shepard et al [11]. Serial 10-fold dilutions of the enumerated *M. leprae* bacilli were used for polymerase chain reaction (PCR) in our study.

Clinical Specimens

Samples (from slit-skin smears or punch biopsies) were taken from patients with leprosy after receipt of informed consent in primarily the central and highland regions of Vietnam (including 11 provinces: Danang, Quangnam, Quangngai, Binhdinh, Phuyen, Khanhhoa, Ninhthuan, Kontum, Gialai, Daknong, and Daklak), and the samples were classified as new (before starting MDT), recent (receiving MDT), and relapse cases. Relapse was defined as development of new skin lesions after completion of MDT and increase in bacterial index by >2 log units in any lesion.

The total of 423 samples included those from 83 patients with new cases, 321 patients with recent cases (receiving treatment), and 19 patients with relapse (collection period: March 2004–August 2009). Among 16 patients with relapse who had positive results of *M. leprae*–specific PCR, 9 cases were determined to be relapse after dapsone monotherapy (7–20 years), 3 as relapse after complete MDT, 2 as second relapse (the first after dapsone monotherapy and the second after MDT), and 2 as relapse after ofloxacin treatment. Samples were obtained from the skin lesions of patients (smear on blade or biopsy soaked in 1 mL of 70% ethanol at room temperature in the field, before being sent to Quyhoa National Leprosy & Dermato-Venereology Hospital laboratory).

DNA Extraction, Nested PCR, and Sequencing

M. leprae templates from both dilutions of *M. leprae* bacilli and slit-skin smears were prepared by treatment with lysis buffer at 60°C overnight, as described elsewhere [12]. Nested PCR amplification of the RLEP regions of *M. leprae* was performed under conditions described elsewhere with minor modifications, using the primers listed in Table 1 [13]. In brief, PCR amplification using special reagents (20 mM Tris-HCl [pH, 7.5], 8 mM magnesium chloride, 7.5 mM DTT, 2.5 mg BSA, 150 μM deoxynucleotides, 1.5 mM magnesium sulphate, and 2.5 units KOD-plus-Ver.2 DNA polymerase [Toyobo]) was performed using sample DNA as templates. Both first and second PCR conditions were as follows; strand separation at 94°C for 4 min, denaturing at 94°C for 40 s, annealing at 55°C for 1 min, and extension at 72°C for 20 s plus 1-s increment per cycle for 25 cycles. Products from the first PCR (0.5 uL) were used as templates in the second PCR. The nested PCR for DRDR was performed using the primer pairs listed in Table 1. Mutations were measured on the *folP1* gene for dapsone [14], the *rpoB* gene for rifampicin, and the *gyrA* gene for ofloxacin [15, 16]. Nested PCR conditions for drug resistance were different from that for RLEP-nested PCR. In brief, PCR amplification using standard reagents (10 mM Tris-HCl [pH, 8.3], 2 mM magnesium chloride, 250 μM dNTPs, and 2.5 units TaKaRa Ex Taq DNA polymerase [Takara shuzo]) was performed using sample genomic DNA as templates. The primer pairs used to amplify the specific drug-resistant genes are shown in Table 1. The reaction condition was 30 s at 94°C, 30 s at 60°C, and 1 min at 72°C for 35 cycles.

The amplicons were visualized by agarose gel electrophoresis, and DNA was recovered from the gel using Mini-Elute gel extraction kits (Qiagen). The recovered DNA molecules were sequenced using the ABI Prism BigDye Terminator Cycle Sequencing kit (Perkin-Elmer Applied Biosystems) and run on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems). The sequence data were analyzed by DNA analysis program Genetyx-MAC, version 15 (GENETYX), and were compared with those in the GenBank database.

RESULTS

PCR Sensitivity

Serial dilutions of the bacilli of 1×10^8–1×10^0 were prepared to determine PCR sensitivities. Genomic DNAs were extracted from the diluents with use of methods described under Materials and Methods [11]. The previously reported RLEP-nested PCR (named RLEP-L) was capable of detecting 1×10^2 bacilli in samples (Figure 1a) [13]. The newly designed RLEP-nested PCR, using K1 and K2 primers for the first PCR and LP1 and LP2 primers for the second PCR (named RLEP-K), is capable of detecting comparable counts of bacilli (Figure 1b), and RLEP-K

Table 1. Sequences of Primers Used in this Study

Name	Usage	Gene	Sequence, 5′→3′	Reference	Size, bp
K1	First PCR (F)	RLEP	CGTGGGTGTGAGGATAGTTGT-	Present study	268
K2	First PCR (R)	RLEP	GATCATCGATGCACTGTTCACT-	Present study	
LP1	First or second PCR (F)	RLEP	TGCATGTCATGGCCTTGAGG-	13	129
LP2	First or second PCR (R)	RLEP	CACCGATACCAGCGGCAGAA	13	
LP3	Second PCR (F)	RLEP	TGAGGTGTCGGCGTGGTC	13	99
LP4	Second PCR(R)	RLEP	CAGAAATGGTGCAAGGGA	13	
F1	Second PCR (F)	*folP1*	GCAGGTTATTGGGGTTTTGA	Present study	312
F2	First PCR(R)	*folP1*	CCACCAGACACATCGTTGAC	Present study	
F3	Second PCR (F)	*folP1*	CTTGATCCTGACGATGCTGT	Present study	245
F4	Second PCR(R)	*folP1*	ACATCGTTGACGATCCGTG	Present study	
F5	Sequencing primer (F)	*folP1*	ATCCTGACGATGCTGTCCA	Present study	–
F4	Sequencing primer (R)	*folP1*	ACATCGTTGACGATCCGTG	Present study	–
R1	First PCR (F)	*rpoB*	CAGACGCTGATCAATATCCGT	Present study	358
R2	First PCR (R)	*rpoB*	CAGCGGTCAAGTATTCGATC	Present study	
R3	Second PCR (F)	*rpoB*	CAATATCCGTCCGGTGGTC	Present study	337
R4	Second PCR (R)	*rpoB*	GTATTCGATCTCGTCGCTGA	Present study	
R5	Sequencing primer (F)	*rpoB*	ACGCTGATCAATATCCGTCC	Present study	–
R6	Sequencing primer (R)	*rpoB*	CGACAA TGAACCGATCAGAC	Present study	–
G1	First PCR (F)	*gyrA*	ACGCGATGAGTGTGATTGTGG	Present study	336
G2	First PCR (R)	*gyrA*	TCCCAAATAGCAACCTCACC	Present study	
G3	Second PCR(F)	*gyrA*	GATGGTCTCAAACCGGTACA	Present study	291
G4	Second PCR (R)	*gyrA*	CCCAAATAGCAACCTCACCA	Present study	
G3	Sequencing primer (F)	*gyrA*	GATGGTCTCAAACCGGTACA	Present study	–
G4	Sequencing primer (R)	*gyrA*	CCCAAATAGCAACCTCACCA	Present study	–

products are visualized more clearly with less smear bands. Therefore, the new RLEP-K system was used for detection in further experimentation with use of clinical samples.

Using DNAs extracted from the serial dilutions of *M. leprae*, we determined the sensitivity of the nested PCR for DRDRs. The limit of amplification by PCR was 1×10^3–1×10^4 bacilli (Figure 1 c–e).

RLEP-nested PCR for Clinical Samples

The PCR methods were applied on 423 clinical samples collected from areas of endemicity in Vietnam. First, we tested RLEP-K for detection of *M. leprae* after extraction of DNA from smear samples. Positive bands were obtained by gel electrophoresis using RLEP-K on 290 samples. The positivity rate was 69%. The patients supplying the 290 samples were divided into 3 categories: new, relapse, and recent cases. Positive rates of RLEP-K by category were 75%, 84%, and 66%, respectively (Table 2).

Mutations in Clinical Samples

Samples positive by RLEP-nested PCR were applied for mutation experiments on the DRDRs of *folP1*, *rpoB*, and the *gyrA* gene. Nineteen mutations were found in 187 *folP1* samples, but no mutations related to drug resistance were noted in 163 *rpoB* and 147 *gyrA* gene samples. The mutations detected on *folP1* were as follows: 6 cases of ACC to ATC in codon 53(threonine to isoleucine), 9 cases of CCC to CGC in codon 55 (proline to arginine), and 4 cases of CCC to CTC (proline to leucine). Two new cases, 8 relapse cases, and 9 recent cases had mutations on *folP1*. Mutation rates in the 3 categories were 6.1%, 57%, and 6.4%, respectively (Table 3).

Some missense mutations, of which the association with drug resistance is unknown, were detected in the *rpoB* gene from clinical samples. The mutations were detected in 7 patients at codons 517, 532, and 556. One patient with relapse showed a mutation from CAG (glutamine) to CAT (histidine) at codon 517. One new patient showed 2 mutations at codon 517 from CAG (glutamine) to CAT (histidine) and at codon 532 from GCG (alanine) to TCG (serine). Sequence electropherograms indicated double peaks of a second nucleotide at codon 556 in 3 patients categorized as having recent cases. One peak was G (identical to that of wild-type), and the other peak was T, which changed the amino acid from glycine (GGC) to valine (GTC; data not shown).

The Relation between Treatment and Drug-Resistant Mutations in Patients with Relapse

Patients with relapse were categorized into 4 groups, by treatment history (Table 4). Group 1 comprised those treated with dapsone monotherapy. Group 2 was treated with MDT for 24 months. Group 3 included patients who had received

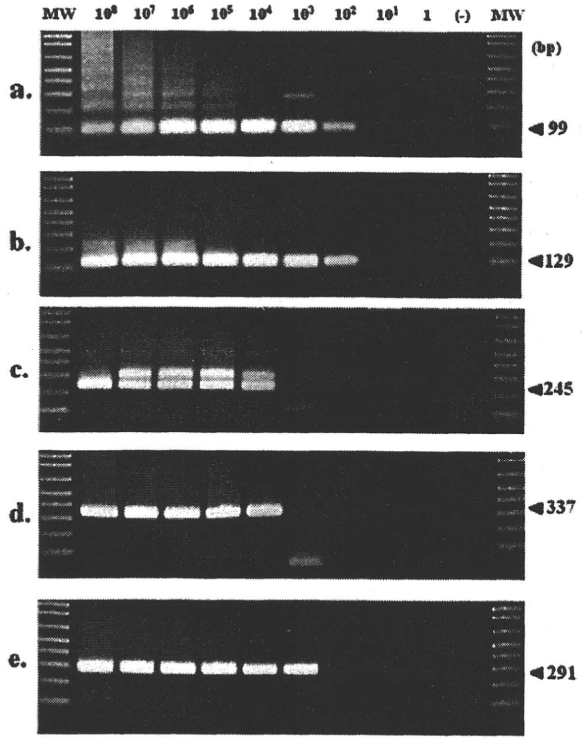

Figure 1. Sensitivity of nested polymerase chain reaction (PCR). The nested PCR products were visualized on 2 % agarose gel. A, RLEP-nested PCR (RLEP-L) using primers, LP1-LP4 (final products size, 99 bp). B, RLEP-nested PCR (RLEP-K) using primers, K1, K2, LP1, and LP2 (final products size, 129bp). C, *folP1*-nested PCR using F1-F4. D, *rpoB*-nested PCR using R1-R4. E, *gyrA*-nested PCR using G1-G4.

a diagnosis of second relapse—once after treatment with dapsone monotherapy and, subsequently, after MDT for 24 months. Group 4 was treated with ofloxacine monotherapy. Eight of the 14 patients with *folP1*-amplified relapse cases (57%) had mutations on the *folP1* gene. Seven (78%) of 9 patients with relapse who were categorized in groups 1 and 3 also had *folP1* mutations. However, 2 patients in group 4 had no mutations on any of the 3 genes.

Monitoring of Mutations in Patients

One hundred seven slit-skin smear samples from 43 patients were taken with consents at different times from each patient for monitoring mutations under treatment. Table 5 shows the difference in mutation results between 5 such patients. The other 38 patients showed no mutation during monitoring. Patients A, B, and C, who had new cases, showed a similar pattern, with no mutation at first testing and mutation in codon 53 on the *folP1* gene during MDT. However, double peaks of T and C in the second base were observed on *folP1* in the 3 patients. Patients D and E, who had relapse cases and finished dapsone monotherapy 20 years earlier, had a mutation on *folP1* in 2005 and no mutation after MDT.

Table 2. Polymerase Chain Reaction Positivity in New, Relapse, and Recent Cases

Case category	No.	RLEP	*folP1*	*rpoB*	*gyrA*
New	83	62 (75%)	33	39	43
Relapse	19	16 (84%)	14	15	13
Recent	321	212 (66%)	140	109	91
Total	423	290 (69%)	187 (64%)	163 (56%)	147 (51%)

DISCUSSION

The most popular PCR method for *M. leprae* detection with high sensitivity and specificity is probably the RLEP-nested PCR method, because the RLEP regions are specific for *M. leprae*, with >28 copies dispersed through the *M. leprae* genome [17]. New primers were designed for the RLEP-nested PCR in our study. This system using the new primers was termed RLEP-K. RLEP-K products appear to be a somewhat sharper and stronger band on agarose gel electrophoresis, compared with that that of previous RLEP-nested PCR (ie, RLEP-L). The RLEP-K detected *M. leprae* in 69% of the Vietnam samples. The remaining 31% of the samples were deduced as being cases either cleared of *M. leprae* by chemotherapy or those having <100 bacilli, which was below the detection limit of RLEP-K. We also designed new primers for amplification and sequencing of DRDR in the drug-resistance related genes *folP1*, *rpoB*, and *gyrA*, which were applied in examining the Vietnam samples. The mutation rates of *folP1* in new and recent cases were 6.1% and 6.4%, respectively. In contrast, the mutation rate in relapse cases was quite high, at 57%. The result indicated a strong correlation between mutation rate and relapse. Two possible reasons were conceived regarding the high positive rate of dapsone resistance in patients with relapse: (1) reinfection by the primary drug-resistant strain (7 of 8 samples indicating relapse were collected in the province in central Vietnam, which had the highest prevalence of leprosy and high rate of relapse (data not shown) and (2) reactivation of dapsone-resistant strains capable of

Table 3. Number of Mutations on *folP1*

Case category	No. of PCR-positive cases	No. of mutations (mutation ratio)	No. of mutation in mutation types
New	33	2 (6.1%)	2 (55th: CCC-CGC)
Relapse	14	8 (57%)	2 (53rd: ACC-ATC)
			3 (55th: CCC-CGC)
			3 (55th: CCC-CTC)
Recent	140	9 (6.4%)	4 (53rd: ACC-ATC)
			4 (55th: CCC-CGC)
			1 (55th: CCC-CTC)

Table 4. Mutations Noted in RLEP-Positive Relapse Cases, by Treatment Group

Group	Past treatment	No.	Mutation on *folP1*	Mutation on *rpoB*	Mutation on *gyrA*
1	DDS	7	5	0	0
2	MDT (24 months)	3	1	1[a]	0
3	DDS plus MDT (24 months)	2	2	0	0
4	OFX	2	0	0	0
All	...	14	8	1[a]	0

Abbreviations: DDS (diaminodiphenylsulfone), dapsone monotherapy; MDT,multidrug treatment; OFX, Ofloxacine monotherapy.

[a] Unknown DR mutation

persisting after chemotherapy, discussed below. Although it is still unclear whether the relapses are caused by reinfection by *M. leprae* or by reactivation of persistent *M. leprae*, close correlation between drug resistance and relapse have been recognized in several studies [18, 19].

The proportion of samples showing mutation on the *folP1* gene related to dapsone resistance was 10.2% (19 of 187) in samples from the central and highland regions of Vietnam (Table 3). Comparison with previous reports from South Korea (19.2%) indicates lower rates of relapse in these regions of Vietnam [20].

No mutation was found in the DRDR regions of *rpoB* in all samples. Mutation frequencies of the *rpoB* gene are also very low in other reports. Regarding other areas in Southeast Asia, no cases of rifampicin resistance have been detected in the Philippines, 1 (1.9%) of 54 cases in Myanmar, and 4 (3.3%) of 121 cases in Indonesia. However, in Japan, where the prevalence of leprosy is very low, the reported rate of rifampicin resistance is very high, at 29.5% (26 of 88 cases) [21]. The long-term use of drugs outside the standard MDT regimen in Japanese leprosy cases might have been instrumental in promoting this rifampicin resistance.

As such, no mutations have been found in the DRDR of the *M. leprae rpoB* gene derived from patients with leprosy, including relapse cases in Vietnam. A possible explanation for this could be the success of leprosy control in Vietnam and efficacy of properly administered MDT in which rifampicin—with its bactericidal properties—was effective in suppressing the occurrence of drug-resistant bacilli. In contrast, dapsone (not bactericidal in itself, although capable of suppressing growth), which had previously been used as monotherapy, may have enabled bacteria surviving in the patient receiving treatment to develop mutations, giving them resistance against the drug. Although occurrence of drug-resistant *M. leprae* was kept very low after application of MDT, 7 of 9 samples with drug-resistant mutations had previously been treated by dapsone monotherapy (Table 4). Jing et al [22] reported that patients with multibacillary leprosy who were retreated with MDT after dapsone monotherapy may have lower risk of early relapse while continuing to carry the risk of late relapse. Our observations suggest the possibility that efficacy of MDT may be hampered in some patients by the presence of surviving dapsone-resistant *M. leprae* in their bodies, which could develop into late relapse. Similar observations have been reported, suspecting involvement of the effects of dapsone monotherapy in patients with relapse [23].

There was no mutation in the major sites for drug resistance on the *rpoB* gene. However, we observed mutations at 3 positions, codons 517, 532, and 556, which have not been associated with rifampicin resistance. These mutations in the *rpoB* gene are a finding calling for further clarification.

Table 5. Monitoring of 5 Patients with Multibacillary Leprosy for *folP1* Mutation

Patient	Case category	Date of sample obtainment	Sample site (method of obtainment)	*folP1* mutation
A	New	2006 April 3	Abdomen (biopsy)	None[a]
		2007 January 30	Earlobe (smear)	53rd (ACC→ATC/ACC)
		2007 January 30	Abdomen (smear)	53rd (ACC→ATC/ACC)
B	New	2005 May 31	Earlobe (smear)	None
		2006 March 24	Skin (smear)	None
		2007 November 2	Skin (smear)	53rd (ACC→ATC/ACC)
C	New	2006 July 20	Skin (smear)	None
		2007 January 30	Skin (smear)	53rd (ACC→ATC/ACC)
		2007 January 30	Skin (smear)	53rd (ACC→ATC/ACC)
D	Relapse	2005 November	Earlobe (smear)	55th (CCC→CGC)
		2007 January	Skin (smear)	None
E	Relapse	2007 Jannuary 17	Arm (smear)	None
		2007 January 30	Earlobe (smear)	55th (CCC→CGC)
		2007 January 30	Arm (smear)	None

[a] ACC ATC/ACC indicates double peaks in second base at codon 53.

To reveal the possible relation between treatment and gene mutation, some patients with leprosy were monitored for gene mutations in light of drug treatments. The results showed incidence of dapsone-resistant *M. leprae* in patients receiving MDT, suggesting that some of the patients with relapse who were previously treated with dapsone monotherapy might have persistent infections with dapsone-resistant *M. leprae*. Furthermore, samples derived from different sites of lesions in the same patient sometimes showed different results (Table 5). The results suggest that we need to know the relation between the situation of patients with leprosy and drug resistance.

Overall, our study indicated a high ratio of dapsone resistance in patients with relapse, compared with the other patients with leprosy. In contrast, an unexpected outcome of our study was that we were unable to find mutations on the *rpoB* gene in patients with relapse. Moreover, it was shown that many of the patients with relapse who had dapsone-resistant mutations had histories of treatment with dapsone monotherapy. To clarify the relationship between relapse, drug resistance, and dapsone monotherapy, it might be necessary to investigate persistence of drug-resistant *M. leprae* through large-scale surveillance.

Acknowledgments

We thank all medical officers working in the local health centers of the 11 provinces in the central and highland regions of Vietnam for their help in collecting the clinical samples.

Financial support. Thos work was supported by Ministry of Health of Vietnam; Quyhoa National Leprosy & Dermato-Venereology Hospital, Vietnam; Japan Health Sciences Foundation; and a Health Science Research Grant—Research on Emerging and Re-emerging Infectious Diseases, Ministry of Health, Labour and Welfare, Japan.

Potential conflicts of interest. All authors: no conflicts.

References

1. WHO Study Group. Chemotherapy of leprosy for control Programmes. Geneva: Tech Resp Ser, 1982: 675.
2. World Health Organization (WHO). Global leprosy situation. Wkly Epidemiol Rec **2009**; 84:333–40.
3. Rees RJ. Mycobacterial disease in man and animals. Studies on leprosy bacilli in man and animals. Proc R Soc Med **1964**;; 57:482–3.
4. Jacobson RR, Hastings RC. Rifampin-resistant leprosy. Lancet **1976**; 2:1304–5.
5. Cambau E, Perani E, Guillemin I, Jamet P, Ji B. Multidrugresistance to dapsone, rifampicin, and ofloxacin in *Mycobacterium leprae*. Lancet **1997**; 349:103–4.
6. González AB, Maestre JL, Hernández O, et al. Survey for secondary dapsone and rifampicin resistance in Cuba. Lepr Rev **1993**; 64:128–35.
7. Matsuoka M, Kashiwabara Y, Namisato MA. *Mycobacterium leprae* isolate resistant to dapsone, rifampin, ofloxacin and sparfloxacin. Int J Lepr Other Mycobact Dis **2000**; 68:452–5.
8. World Health Organization (WHO). Global leprosy situation. Wkly Epidemiol Rec **2008**; 83:217–24.
9. Bang PD, Suzuki K, Ishii N, Kang TH. Leprosy situation in Vietnam-reduced burden of stigma. Jpn J Lep **2008**; 77:29–36.
10. Quyhoa NDH. Annual report (2004–2009). http://www.quyhoandh.org.vn/quyhoandh/vn/portal/index.jsp. Accessed 29 December 2010.
11. Shepard CC. The experimental disease that follows the injection of human leprosy bacilli into foot-pads of mice. J Exp Med **1960**; 112:445–54.
12. de Wit MY, Faber WR, Krieg SR, et al. Application of a polymerase chain reaction for the detection of *Mycobacterium leprae* in skin tissues. J Clin Microbiol **1991**; 29:906–10.
13. Donoghue HD, Holton J, Spigelman M. PCR primers that can detect low levels of *Mycobacterium leprae* DNA. J Med Microbiol **2001**; 50:177–82.
14. Kai M, Matsuoka M, Nakata N, et al. Diaminodiphenylsulfone resistance of *Mycobacterium leprae* due to mutations in the dihydropteroate synthase gene. FEMS Microbiol Lett **1999**; 177:231–5.
15. Honore N, Cole ST. Molecular basis of rifampin resistance in *Mycobacterium leprae*. Antimicrob Agents Chemother **1993**; 37:414–8.
16. Maeda S, Matsuoka M, Nakata N. Multidrug resistant *Mycobacterium leprae* from patients with leprosy. Antimicrob Agents Chemother **2000**; 45:3635–9.
17. Cole ST, Eiglmeier K, Parkhill J, et al. Massive gene decay in the leprosy bacillus. Nature **2001**; 409:1007–11.
18. Kaimal S, Thappa DM. Relapse in leprosy. Indian J Lepr **2009**; 75:126–35.
19. Lopez-Roa RI, Fafutis-Morris M, Matsuoka M. A drug-resistant leprosy case detected by DNA sequence analysis from a relapsed Mexican leprosy patient. Rev Latinoam Microbiol **2006**; 48:256–9.
20. You EY, Kang TJ, Kim SK, et al. Mutations in genes related to drug resistance in *Mycobacterium leprae* isolates from leprosy patients in Korea. J Infect **2005**; 50:6–11.
21. Matsuoka M, Budiawan T, Aye KS, et al. The frequency of drug resistance mutations in *Mycobacterium leprae* isolates in untreated and relapsed leprosy patients from Myanmar, Indonesia and the Philippines. Lepr Rev **2007**; 78:343–52.
22. Jing Z, Zhang R, Zhou D, Chen J. Twenty five years follow up of MB leprosy patients retreated with a modified MDT regimen after a full course of dapsone mono-therapy. Lepr Rev **2009**; 80:170–6.
23. Chakma JK, Girdhar A, Natrajan M, et al. Two microbiological relapses in a patient with lepromatous leprosy. Lepr Rev **2008**; 79:331–4.

JOURNAL OF BACTERIOLOGY, Nov. 2010, p. 5700–5708
0021-9193/10/$12.00 doi:10.1128/JB.00554-10

Vol. 192, No. 21

Novel Rhamnosyltransferase Involved in Biosynthesis of Serovar 4-Specific Glycopeptidolipid from *Mycobacterium avium* Complex▽

Yuji Miyamoto,[1]* Tetsu Mukai,[1] Takashi Naka,[2,3] Nagatoshi Fujiwara,[2] Yumi Maeda,[1] Masanori Kai,[1] Seiko Mizuno,[2,4] Ikuya Yano,[5] and Masahiko Makino[1]

Department of Mycobacteriology, Leprosy Research Center, National Institute of Infectious Diseases, 4-2-1 Aobacho, Higashimurayama, Tokyo 189-0002, Japan[1]; Department of Bacteriology, Osaka City University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan[2]; MBR Co., Ltd., 7-7-15 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan[3]; Faculty of Human Development, Soai University, 4-4-1 Nanko-naka, Suminoe-ku, Osaka 559-0033, Japan[4]; and Japan BCG Laboratory, 3-1-5 Matsuyama, Kiyose, Tokyo 204-0022, Japan[5]

Received 14 May 2010/Accepted 24 August 2010

Glycopeptidolipids (GPLs) are one of the major glycolipid components present on the surface of *Mycobacterium avium* complex (MAC) that belong to opportunistic pathogens distributed in the natural environment. The serovars of MAC, up to around 30 types, are defined by the variable oligosaccharide portions of the GPLs. Epidemiological studies show that serovar 4 is the most prevalent type, and the prognosis of pulmonary disease caused by serovar 4 is significantly worse than that caused by other serovars. However, little is known about the biosynthesis of serovar 4-specific GPL, particularly the formation of the oligosaccharide portion that determines the properties of serovar 4. To investigate the biosynthesis of serovar 4-specific GPL, we focused on one segment that included functionally unknown genes in the GPL biosynthetic gene cluster of a serovar 4 strain. In this segment, a putative hemolytic protein gene, *hlpA*, and its downstream gene were found to be responsible for the formation of the 4-*O*-methyl-rhamnose residue, which is unique to serovar 4-specific GPL. Moreover, functional characterization of the *hlpA* gene revealed that it encodes a rhamnosyltransferase that transfers a rhamnose residue via 1→4 linkage to a fucose residue of serovar 2-specific GPL, which is a key pathway leading to the synthesis of oligosaccharide of serovar 4-specific GPL. These findings may provide clues to understanding the biological role of serovar 4-specific GPL in MAC pathogenicity and may also provide new insights into glycosyltransferase, which generates structural and functional diversity of GPLs.

The genus *Mycobacterium* has a unique feature in the cell envelope that contains a multilayered structure consisting of peptidoglycan, mycolyl-arabinogalactan complex, and surface glycolipids (8, 12). It is known that these components play a role in protection from environmental stresses, such as antimicrobial agents and host immune responses (8, 12). Some of them are recognized as pathogenic factors related to mycobacterial diseases, such as tuberculosis and leprosy (8, 12). In case of nontuberculous mycobacteria that are widely distributed in the natural environment as opportunistic pathogens, glycopeptidolipids (GPLs) are abundantly present on the cell envelope as surface glycolipids (34). GPLs have a core structure in which a fatty acyl-tetrapeptide is glycosylated with 6-deoxy-talose (6-d-Tal) and *O*-methyl-rhamnose (*O*-Me-Rha) (2, 5, 13). This structure is common to all types of GPLs, and GPLs with this structure that have not undergone further glycosylation are termed non-serovar-specific GPLs (nsGPLs) (2, 5, 13). Structural diversity generated by further glycosylations, such as rhamnosylation, fucosylation, and glucosylation, is observed for the oligosaccharide portion linked to the 6-d-Tal residue of nsGPLs from *Mycobacterium avium* complex (MAC), a member of the nontuberculous mycobacteria consisting of two species, *M. avium* and *M. intracellulare* (2, 5, 34). Consequently, these nsGPLs with varied oligosaccharides lead to the formation of the serovar-specific GPLs (ssGPLs) that define around 30 types of MAC serovars (10).

The properties of MAC serovars are known to be notably different from each other and also to be closely associated with the pathogenicity of MAC (3, 6, 18, 30, 31, 32). Various epidemiological studies indicate that serovar 4 is the most prevalent type and is also one of the serovars frequently isolated from AIDS patients (1, 20, 33, 36). Additionally, pulmonary MAC disease caused by serovar 4 is shown to exhibit a poorer prognosis than that caused by other serovars (23). With respect to host immune responses to MAC infection, serovar 4-specific GPL is reported to have characteristic features that are in contrast to those of other ssGPLs (21, 30). Structurally, serovar 4-specific GPL contains a unique oligosaccharide in which the oligosaccharide of serovar 2-specific GPL is further glycosylated with 4-*O*-methyl-rhamnose (4-*O*-Me-Rha) residue through a 1→4 linkage (Table 1) (24). Therefore, it is thought that the presence of 4-*O*-Me-Rha and its linkage position are important in exhibiting the specificity of biological activities. The biosynthesis of the oligosaccharide portion in several ssGPLs is currently being clarified (15, 16, 17, 25, 26), while that of serovar 4-specific GPL is still unresolved. In this study, we have focused on the genomic region predicted to be associated with GPL biosynthesis in the serovar 4 strain and explored the key genes responsible

* Corresponding author. Mailing address: Department of Mycobacteriology, 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: yujim@nih.go.jp.
▽ Published ahead of print on 3 September 2010.

TABLE 1. Oligosaccharide structures of serovar 2- and 4-specific GPLs

Serovar	Oligosaccharide	Reference
2	2,3-di-*O*-Me-α-L-Fuc-(1→3)-α-L-Rha-(1→2)-L-6-d-Tal	9
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	24

for the formation of 4-*O*-Me-Rha that might determine the specific properties of MAC serovar 4.

MATERIALS AND METHODS

Bacterial strains, culture conditions, and DNA manipulation. Table 2 indicates the bacterial strains and vectors used in this study. MAC strains were grown in Middlebrook 7H9 broth (Difco) with 0.05% Tween 80 supplemented with 10% Middlebrook ADC enrichment (BBL). For GPL production, *Mycobacterium smegmatis* strains were cultured in 2× YT broth (16 g tryptone, 10 g yeast extract, and 5 g NaCl per liter) with 0.2% Tween 80. DNA manipulation of *M. smegmatis* strains was conducted as previously described (27). PCR amplification was done by two-step PCR using TaKaRa LA *Taq* with GC buffer, with the following program: denaturation at 98°C for 20 s and annealing-extension at 68°C for an appropriate time depending on the length of the targeted region. *Escherichia coli* strain DH5α was used for the routine manipulation and propagation of plasmid DNA. 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 in this study are listed in Table 3.

Construction of expression vectors. For generation of the serovar 2-specific GPL (GPL-S2)-producing strain, the vector possessing *rtfA*, *mdhtA*, *merA*, and *gtfD* genes was constructed. The *rtfA* gene was amplified from genomic DNA of *M. avium* strain JATA51-01 using primers RTFA-S and RTFA-A. The *mdhtA*, *merA*, and *gtfD* genes were amplified as one operon from the previously constructed vector pMVΔmtfF using primers MDHTA-S2 and GTFD-A2 (26). After construction of pMV261a, in which an AflII site was introduced into pMV261, the above two PCR products were digested with each restriction enzyme and cloned into the BamHI-PstI and PstI-AflII sites of pMV261a to give pMV-rtfA-mdhtA-merA-gtfD (Table 2). The fragments for construction of expression vectors were amplified from genomic DNA of MAC serovar 4 strain (ATCC 35767) using the following primers: HLPA-S and HLPA-A for *hlpA*, HLPA-S and ORF2-A for *hlpA-orf2*, and ORF3-S and ORF5-A for *orf3-orf4-orf5*. These PCR products were digested with each restriction enzyme and cloned into the EcoRI-ClaI, EcoRI-HindIII, and PstI-EcoRI sites of pYM301a to give pYM-hlpA, pYM-hlpA-orf2, and pYM-orf3-orf4-orf5, respectively (Table 2).

Isolation and purification of GPLs. To isolate whole-lipid extracts, harvested bacterial cells were mixed with $CHCl_3/CH_3OH$ (2:1 [vol/vol[) for several hours at room temperature. The extracts in organic phase were separated by adding water and evaporated until dry. To remove the lipid components except for GPLs, the whole-lipid extracts were subjected to mild alkaline hydrolysis to prepare the crude GPLs as previously described (27, 28). For analytical thin-layer chromatography (TLC), crude GPLs on silica gel 60 plates (Merck) were developed with $CHCl_3/CH_3OH/H_2O$ (30:8:1 [vol/vol/vol]), followed by spraying with 10% H_2SO_4 and charring. Purified GPLs were prepared from crude GPLs by preparative TLC on the same plates and extracted from the bands corresponding to each GPL. To determine the linkage position of sugar moieties, perdeuteriomethylation was performed for purified GPLs as previously described (7, 11, 15).

GC-MS and MALDI-TOF MS analysis. Purified and perdeuteriomethylated GPLs were hydrolyzed in 2 M trifluoroacetic acid (2 h, 120°C), and the released sugars were reduced with $NaBD_4$ and then acetylated with pyridine/acetic anhydride (1:1 [vol/vol]) at room temperature overnight. The resulting alditol acetates were analyzed by gas chromatography-mass spectrometry (GC-MS) with a GCMS-QP2010 (Shimadzu) equipped with a SP-2380 column (Supelco) using helium gas. The temperature program was from 52 to 172°C with an increase in temperature of 40°C/min, 172 to 223°C at 3°C/min, and then 223 to 270°C at 40°C/min. To determine the total mass of the purified GPLs, matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectra were acquired with an Ultraflex II (Bruker Daltonics). Samples were dissolved in chloroform-methanol (2:1 [vol/vol]) at a concentration of 1 mg/ml, 1 μl was applied directly to the sample plate, and then 1 μl of 10 mg/ml 2,5-dihydroxybenzoic acid in chloroform/methanol (1:1 [vol/vol]) was added as a matrix. The purified GPL was analyzed in the reflectron mode with an accelerating voltage operating in a positive mode of 20 kV (17).

Nucleotide sequence accession number. The 6.8-kb genomic region amplified from the MAC serovar 4 strain (ATCC 35767) by using primers GTFB-S1 and MDHTA-A2 has been deposited in the DDBJ nucleotide sequence database under accession no. AB550236.

RESULTS

Previously, the A5 strain, one of the MAC serovar 4 strains, was reported to contain a genomic region similar to the GPL biosynthetic gene cluster identified in other serovars (22). However, to date, there are no studies clarifying the biosynthetic pathways involved in the formation of 4-*O*-Me-Rha, which is unique to serovar 4-specific GPL. To explore this glycosylation pathway, we focused on one segment interposed with the *gtfB* and *mdhtA* genes whose organization was shown to vary in strains of other serovars (14, 22). In this study, using another serovar 4 strain, ATCC 35767, whose genomic information is unknown, we designed various primers for PCR amplification of a focused segment based on the sequences from other serovar strains. After the testing of primer pairs, a

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	TaKaRa
M. smegmatis mc^{2}155	Expression host	29
M. intracellulare ATCC 35767	MAC serovar 4 strain	35
M. avium JATA51-01	Source of the *rtfA* gene	26
Vectors		
pYM301a	Site-specific integrating mycobacterial vector carrying an *hsp60* promoter cassette	25
pMV261a	*E. coli-Mycobacterium* shuttle vector carrying an *hsp60* promoter cassette with an AflII site	This study
pMVΔmtfF	Source of *mdhtA*, *merA*, and *gtfD* genes	26
pMV-rtfA-mdhtA-merA-gtfD	pMV261a carrying *rtfA*, *mdhtA*, *merA*, and *gtfD* genes	This study
pYM-hlpA	pYM301a carrying the *hlpA* gene	This study
pYM-hlpA-orf2	pYM301a carrying the *hlpA* gene and ORF2	This study
pYM-orf3-orf4-orf5	pYM301a carrying ORF3, ORF4, and ORF5	This study

TABLE 3. Oligonucleotide primers used in this study

Primer	Sequence[a]	Restriction site
RTFA-S	5′-CGGGATCCCATGAAATTTGCTGTGGCAAG-3′	BamHI
RTFA-A	5′-AACTGCAGCTCAGCGACTTCGCTGCGCTTC-3′	PstI
MDHTA-S2	5′-GCTCTAGACTGCAGAAAAACCAACTTCTACTGCCTGACCTG-3′	PstI
GTFD-A2	5′-GGAATTCTTAAGTCTACGGTTCTGCGCTTCGTTCTTTG-3′	AflII
HLPA-S	5′-GGAATTCGTGACAACGACGCCACCAGT-3′	EcoRI
HLPA-A	5′-CCATCGATACTACGCTGCCGCGCTAGGCG-3′	ClaI
ORF2-A	5′-CCCAAGCTTCTCAGACTCTAACGTACAGTTC-3′	HindIII
ORF3-S	5′-CACCTGCAGAAATGACCGCCACAACCAGGGC-3′	PstI
ORF5-A	5′-GCAGAATTCCTACGGCGCCAATTCGATGAG-3′	EcoRI
GTFB-S1	5′-GGAACTCCTGCACCTTGGGGCCGT-3′	
MDHTA-A2	5′-GGTGCGGGTCAACGTAGAGGTG-3′	

[a] Underlining indicates the restriction site.

6.8-kb fragment was amplified with primers GTFB-S1 and MDHTA-A2 (Fig. 1). Nucleotide sequences of the amplified fragments were similar to that of the GPL biosynthetic gene cluster from the A5 strain (94% identity in nucleotide sequences) (GenBank accession no. AY130970.1). This segment contains five complete open reading frame (ORF) genes (Fig. 1): the ORF1 gene, similar to a putative hemolytic protein gene (*hlpA*) previously found in the GPL biosynthetic gene cluster of the serovar 2 strain (69% identity in amino acid sequences) (GenBank accession no. AF125999.1) (14); the ORF2 gene, an undefined gene showing low similarity to some *O*-methyltransferases; and the ORF3, ORF4, and ORF5 genes, with amino acid sequences almost identical to those of three proteins, including GtfTB, which were previously identified as biosynthetic enzymes for serovar 8-specific GPL (GenBank accession no. AB437139.1) (25).

Prior to functional analysis of each ORF, it was necessary to prepare a strain producing the substrate for the enzymes participating in the biosynthesis of serovar 4-specific GPL. Since serovar 4-specific GPL has a structure in which the terminal Fuc residue of serovar 2-specific GPL is further glycosylated with 4-*O*-Me-Rha, we created a recombinant *M. smegmatis* strain (termed MS-S2) by introducing the plasmid vector pMV-rtfA-mdhtA-merA-gtfD possessing *M. avium rtfA*, *mdhtA*, *merA*, and *gtfD* genes, which were previously shown to convert nsGPLs to serovar 2-specific GPL with a terminal Fuc residue (termed GPL-S2) (26). For five ORFs, we first examined the function of the ORF1 (termed *hlpA*) and its downstream ORF2 gene by TLC analysis of recombinant strains, because these have not been functionally defined and it is difficult to predict the role of each gene. In comparison with the profile of the control strain (MS-S2/pYM301a) (Fig. 2, lane A), the new products (GPL-S4) were observed for the strain with the *hlpA* gene introduced (MS-S2/pYM-hlpA) (Fig. 2, lane B). Moreover, when the expression vector covering both

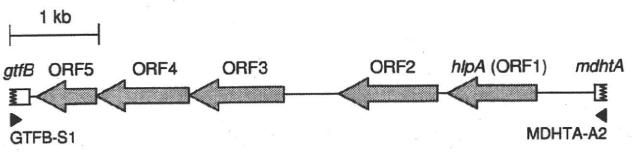

FIG. 1. Organization of the 6.8-kb genomic segment isolated from MAC serovar 4 strain (ATCC 35767). Filled triangles indicate the primers used for PCR amplification.

hlpA and ORF2 was introduced into MS-S2 (MS-S2/pYM-hlpA-orf2), another new product (GPL-S4M) appeared (Fig. 2, lane C). These observations indicated that GPL-S2 was converted to structurally different compounds by the expression of *hlpA* and that the compounds generated by *hlpA* were further modified by ORF2. As for the ORF3, ORF4, and ORF5 genes, which show a high similarity to the biosynthetic genes for serovar 8-specific GPL, we further generated a strain having three ORFs (MS-S2/pYM-orf3-orf4-orf5) and examined the GPL production by TLC analysis (Fig. 2, lane D). The results indicated the appearance of known product GPL-S8, previously shown to have a sugar residue of serovar 8-specific GPL, with no GPL-S4 and GPL-S4M (25), demonstrating that the enzymes encoded by three ORFs might act on the serovar 1-specific GPL which was produced as a precursor of GPL-S2 and subsequently yielded GPL-S8.

Because the compounds produced by *hlpA* and ORF2 were structurally unidentified, we performed a GC-MS analysis of the products GPL-S2, GPL-S4, and GPL-S4M, which were purified from recombinant strain MS-S2/pYM301a, MS-S2/pYM-hlpA, and MS-S2/pYM-hlpA-orf2, respectively. Although two spots were seen for each product, this might be due to a different methylation pattern for the fatty acid portion, which is often observed with GPL biosynthesis of *M. smegmatis* and does not affect oligosaccharide structure (19, 25). In GC-MS profiles of GPL-S2 and GPL-S4, the classes of the detected sugar residues, Fuc, 6-d-Tal, Rha, and 2,3,4-tri-*O*-Me-Rha, were found to be identical to each other (Fig. 3A and B). However, it was observed that the intensity of the Rha residue in GPL-S4 was higher than that of the other sugars, while in GPL-S2, the intensity of the Rha residue was lower than that of Fuc, indicating that the proportion of Rha content in GPL-S4 was relatively large compared to that in GPL-S2. These results implied that the *hlpA* gene mediates the transfer of an additional Rha residue to GPL-S2. In contrast, the profiles of GPL-S4M showed the presence of 4-*O*-Me-Rha that is specifically observed for serovar 4-specific GPL (Fig. 3C), demonstrating that ORF2 encodes a rhamnosyl 4-*O*-methyltransferase and that both genes are responsible for the formation of the unique sugar residue of serovar 4-specific GPL. Furthermore, we confirmed the molecular masses of products GPL-S2, GPL-S4, and GPL-S4M by MALDI-TOF MS analysis (Fig. 4). Each product contained two main pseudomolecular ions ($[M + Na]^+$) with 14 mass unit differences, indicating the

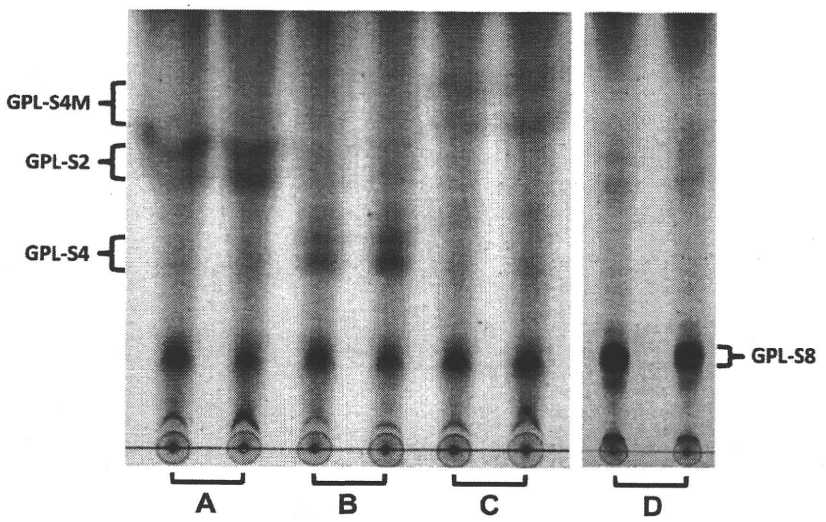

FIG. 2. TLC analysis of crude GPL extracts from recombinant *M. smegmatis* strains MS-S2/pYM301a (A), MS-S2/pYM-hlpA (B), MS-S2/pYM-hlpA-orf2 (C), and MS-S2/pYM-orf3-orf4-orf5 (D). GPL extracts were prepared from the total lipid fraction after a mild alkaline hydrolysis step. Each recombinant strain was tested by two samples derived from independent colonies. Samples were spotted and developed in $CHCl_3$-CH_3OH-H_2O (30:8:1 [vol/vol/vol]).

presence or absence of the methyl group in the fatty acid portion as described above. Thus, the results revealed that the mass unit difference between GPL-S2 (*m/z* 1,465.80, 1,479.82) and GPL-S4 (*m/z* 1,611.84, 1,625.85) was 146 and that between GPL-S2 and GPL-S4M (*m/z* 1,625.89, 1,639.90) was 160, demonstrating that the Rha and 4-*O*-Me-Rha residues were further added to the GPL-S2 to yield GPL-S4 and GPL-S4M, respectively.

The results from TLC, GC-MS, and MALDI-TOF MS analyses strongly suggested that *hlpA* and ORF2 are involved in the formation of 4-*O*-Me Rha. However, it is not clear whether the *hlpA* gene product functions as a glycosyltransferase that transfers a Rha via 1→4 linkage to a Fuc residue, which is observed only for serovar 4-specific GPL. To elucidate the function of *hlpA*, we determined the linkage of sugar moieties of GPL-S4 produced by recombinant strain MS-S2/pYM-hlpA (Fig. 2, lane B). The purified GPL-S4 was subjected to perdeuteriomethylation followed by GC-MS and gave four peaks corresponding to Fuc, 6-d-Tal, Rha, and 2,3,4-tri-*O*-Me-Rha (data not shown). The spectra of Rha and 6-d-Tal demonstrated that the linkage position between these two sugar residues is commonly observed in the oligosaccharide of all ssGPLs, and position C-3 of Rha is linked to the next one, which is consistent with the data previously reported (Fig. 5B and C) (25, 26). In addition, the detection of fragment ions at *m/z* 121, 168, and 206 in spectra of Fuc indicated that its positions C-2 and C-3 were deuteriomethylated (Fig. 5D), meaning that position C-1 of Fuc is linked to position C-3 of Rha and position C-4 of Fuc is linked to the next one. These observations were supported by the fact that GPL-S4 was structurally based on the oligosaccharide of serovar 2-specific GPL. The peak of 2,3,4-tri-*O*-Me-Rha was found to include mixed fragment ions (Fig. 5A). A group of fragment ions corresponding to the spectra of 2,3,4-tri-*O*-Me-Rha linked to alaninol of tetrapeptide was observed. The remaining fragment ions at *m/z* 121, 134, 168, and 181 indicate the presence of deuteriomethyl groups at positions C-2, C-3, and C-4 of the other Rha that is linked at the terminus of oligosaccharide in GPL-S4. These results indicate that position C-1 of terminal Rha is linked to position C-4 of Fuc. Accordingly, the oligosaccharide structures of GPL-S4 were determined to have Rha-(1→4)-Fuc-(1→3)-Rha-(1→2)-6-d-Tal at D-*allo*-Thr, demonstrating that *hlpA* encodes a rhamnosyltransferase that transfers a Rha residue via 1→4 linkage to a Fuc residue of serovar 2-specific GPL (Fig. 6).

DISCUSSION

It is known that serovar 4 is the most prevalent type, and serovar 4-specific GPL, particularly its oligosaccharide portion, plays a role in exhibiting the specific properties that belong to pathogenic factors. However, to date, the biosynthesis of its oligosaccharide portion has not been clarified. In this study, structural determination of three recombinant products, GPL-S2, GPL-S4, and GPL-S4M, revealed that *hlpA* and its downstream gene (ORF2) in the GPL biosynthetic gene cluster are involved in the formation of 4-*O*-methyl Rha, which is unique to serovar 4-specific GPL (Fig. 6). Previously, it was reported that the GPL biosynthetic gene cluster of MAC serovar 2 strains contained one gene whose amino acid sequences are similar to that of *hlpA* with 69% identity (14). This has been regarded as the gene not associated with GPL biosynthesis, because its amino acid sequences are similar to those of hemolytic proteins distributed in some species of bacteria (4). Thus, as shown in Fig. 6, it was surprising that *hlpA* from serovar 4 was found to encode a rhamnosyltransferase that plays a critical role in the pathway leading from serovar 2-specific GPL to serovar 4-specific GPL. For mycobacterium species, a BLAST analysis of HlpA revealed that its homologues are seen only in MAC serovar 2 and not in other species, including *Mycobacterium tuberculosis*. When we tested the

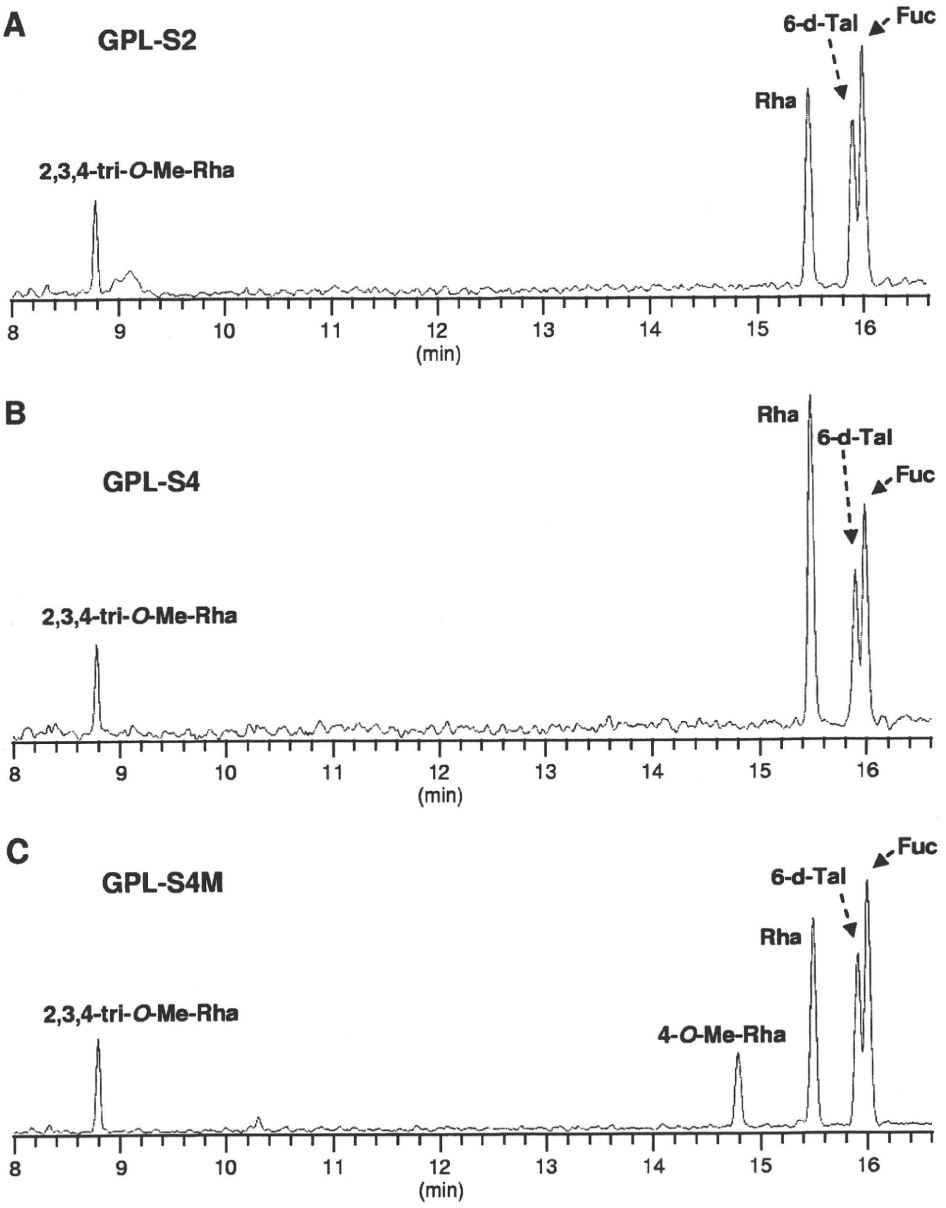

FIG. 3. GC-MS of alditol acetate derivatives from GPL-S2 (A), GPL-S4 (B), and GPL-S4M (C), which were purified from GPL extracts of recombinant *M. smegmatis* strains MS-S2/pYM301a, MS-S2/pYM-hlpA, and MS-S2/pYM-hlpA-orf2, respectively.

function of HlpA from serovar 2, it did not serve as a glycosyltransferase like HlpA from serovar 4 (data not shown). At present, the function of HlpA from serovar 2 is still unclear, because the biosynthesis of the oligosaccharide portion in serovar 2-specific GPL has been fully elucidated (14, 26). The oligosaccharide structure of serovar 2-specific GPL is basic for several ssGPLs, including serovar 4-specific GPL. In the biosynthetic gene cluster of serovar 2-specific GPL, several insertion sequence (IS) elements are observed, raising the possibility that the HlpA from serovar 2 is retained through genomic alterations that induce biosynthetic changes from other ssGPLs to serovar 2-specific GPL. Therefore, HlpA in the serovar 2 strain originally might function as a glycosyltransferase in the biosynthesis of oligosaccharides of other serovars.

Most HlpA homologues are putatively categorized as hemolytic proteins because they are similar to one protein from *Prevotella intermedia*, which is actually proved to have hemolytic activity (4). Since the amino acid sequences of HlpA show 38% identity and 54% similarity to the above protein of *P. intermedia*, we predicted that HlpA also possesses hemolytic activity as an additional function. However, none was detected when *hlpA* was expressed in *M. smegmatis* and *E. coli* by plate assay using a sheep blood agar plate (data not shown). A BLAST analysis of HlpA homologues showed that they also

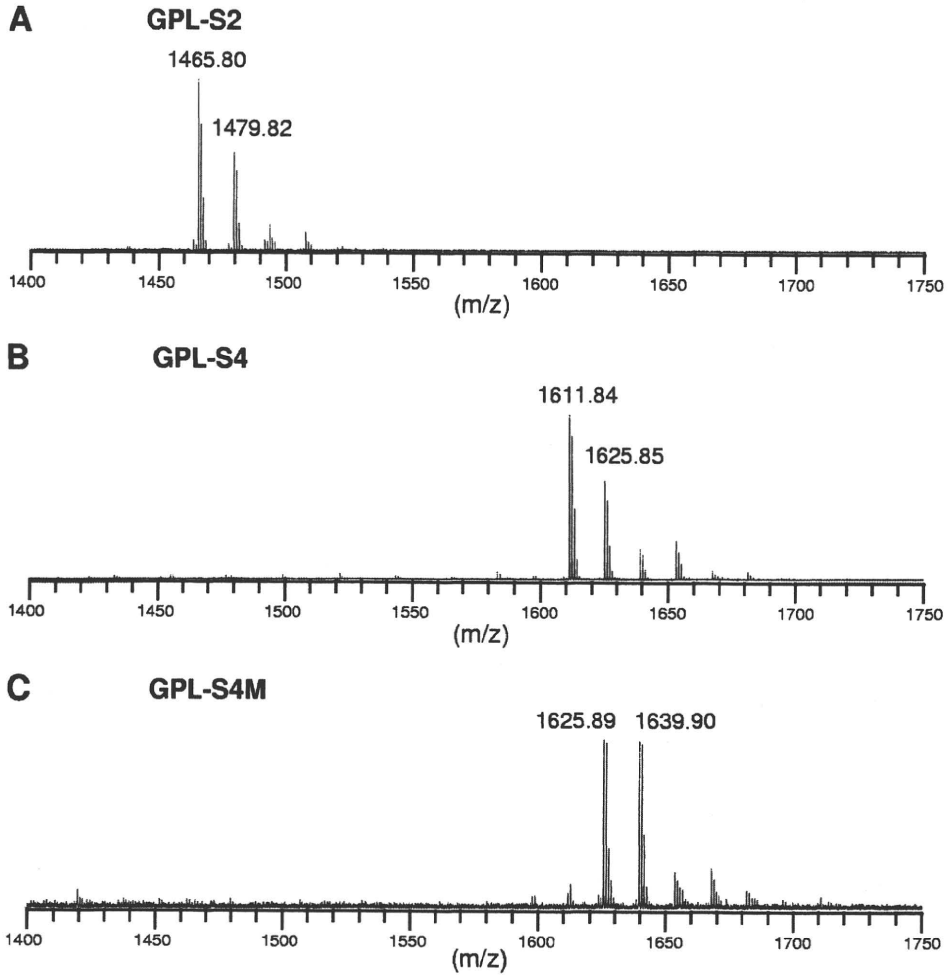

FIG. 4. MALDI-TOF MS of GPL-S2 (A), GPL-S4 (B), and GPL-S4M (C), which were purified from GPL extracts of recombinant *M. smegmatis* strains MS-S2/pYM301a, MS-S2/pYM-hlpA, and MS-S2/pYM-hlpA-orf2, respectively.

contained a partial motif of some glycosyltransferases and methyltransferases. Therefore, it is envisaged that the evolutionary ancestor of HlpA might have lost hemolytic activity in MAC or, conversely, have been altered to retain it in *P. intermedia* in the course of phylogenetic evolution between these bacterial species.

Serovar 4 strains, including ATCC 35767, have been recognized as strains producing the serovar 4-specific GPL but not the serovar 8-specific GPL (24, 35). However, as shown in Fig. 1, we found that the GPL biosynthetic gene cluster contains three known genes (the ORF3, ORF4, and ORF5 genes) previously identified as biosynthetic genes responsible for the formation of 4,6-*O*-(1-carboxyethylidene)-3-*O*-methyl glucose residue in the oligosaccharide of serovar 8-specific GPL (25). TLC analysis showed that overexpression of three ORFs potentially produces the serovar 8-specific GPL, including the 4,6-*O*-(1-carboxyethylidene)-3-*O*-methyl glucose residue (Fig. 2, lane D), demonstrating that in the serovar 4 strain, there is inefficient expression of three genes, which might be caused by genomic alterations affecting their transcription, resulting in the loss of serovar 8-specific GPL. Moreover, HlpA homologues are often found in several species of cyanobacteria but not in other bacterial groups and mycobacterium species, implying the occurrence of a certain kind of "horizontal gene transfer" between these environmental bacteria. Thus, MAC seemed to incorporate foreign genes or realign preexisting genes to modify the oligosaccharide structures of GPLs for their survival in a varied environment. In terms of sugar composition and linkage affecting the properties of ssGPLs, the functional aspects of the 4-*O*-methyl-Rha residue, which influence the interactions with the host cell, are still unclear. In addition, the sugar linkage Rha-(1→4)-Fuc is seen only in serovar 4-specific GPL and not in other ssGPLs, suggesting that it might generate unique properties that differ notably from those generated by other sugar linkages. Also, the rarity of this sugar linkage could be one of the factors that define the specificity of MAC serovar 4, which would be resolved by further studies, including the generation of an *hlpA* knockout mutant. For functional characterization of *hlpA* and ORF2, we have adopted the gene expression experiment

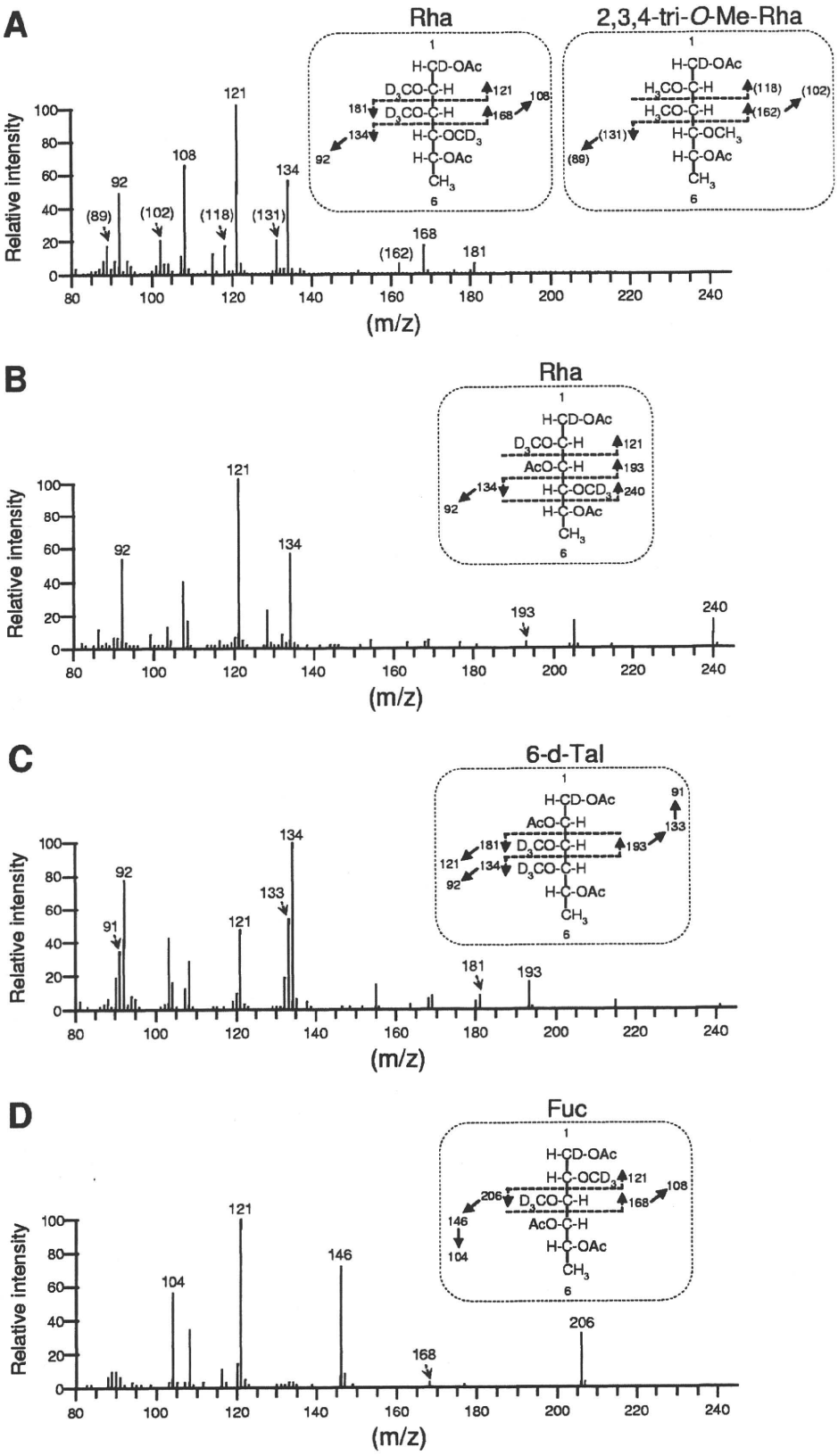

FIG. 5. GC-MS spectra and fragment ion assignments of 2,3,4-tri-*O*-Me-Rha (A), Rha (A and B), 6-*d*-Tal (C), and Fuc (D), which are derived from alditol acetates of sugars released from deuteriomethylated GPL-S4. Ac, acetate; D, deuterium.

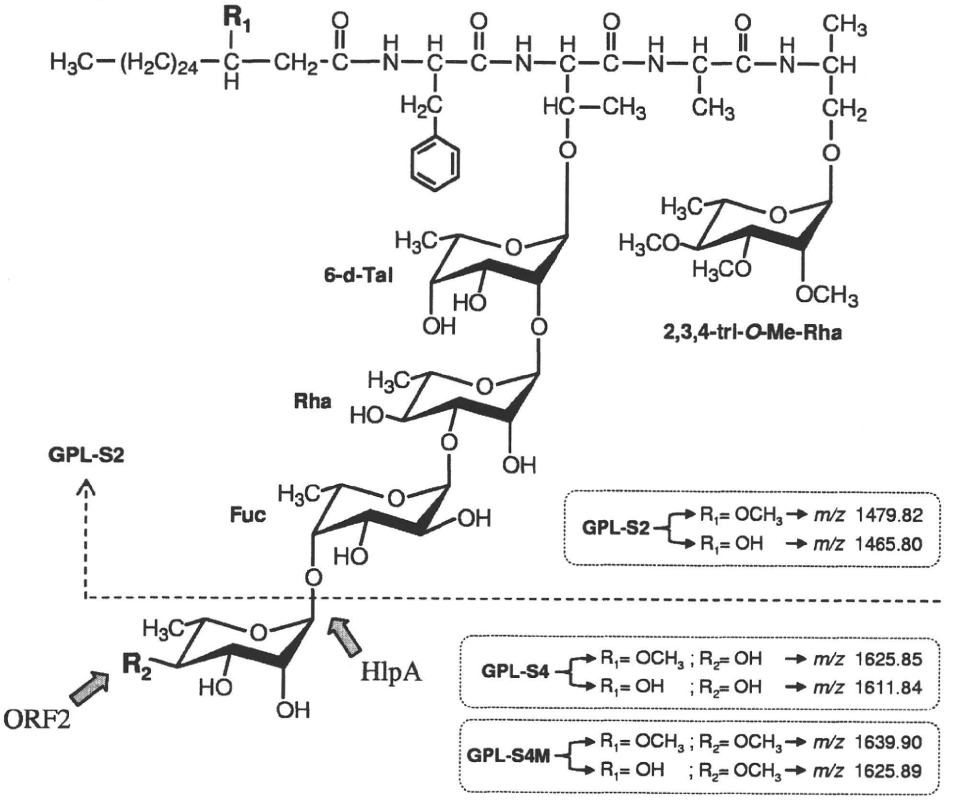

FIG. 6. Proposed structures and biosynthetic pathways for GPL-S2, GPL-S4, and GPL-S4M. Parentheses indicate structural differences between three compounds, which are deduced from MALDI-TOF MS analyses {pseudomolecular ions ($[M + Na]^+$)}.

using the *M. smegmatis* strain. Further enzymatic analyses, such as *in vitro* testing with recombinant proteins, would confirm our results. Taken together, these findings may contribute to understanding the mechanism for generation of structural and functional diversity of GPLs as well as their biological role in MAC.

ACKNOWLEDGMENTS

This research was partially supported by a Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Culture, Science and Technology of Japan and by a Research on Emerging and Re-Emerging Infectious Diseases grant from the Ministry of Health, Labor and Welfare of Japan.

REFERENCES

1. **Askgaard, D. S., S. B. Giese, S. Thybo, A. Lerche, and J. Bennedsen.** 1994. Serovars of *Mycobacterium avium* complex isolated from patients in Denmark. J. Clin. Microbiol. **32:**2880–2882.
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. **Barrow, W. W., T. L. Davis, E. L. Wright, V. Labrousse, M. Bachelet, and N. Rastogi.** 1995. Immunomodulatory spectrum of lipids associated with *Mycobacterium avium* serovar 8. Infect. Immun. **63:**126–133.
4. **Beem, J. E., W. E. Nesbitt, and K. P. Leung.** 1999. Cloning of *Prevotella intermedia* loci demonstrating multiple hemolytic domains. Oral Microbiol. Immunol. **14:**143–152.
5. **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.
6. **Birkness, K. A., W. E. Swords, P. H. Huang, E. H. White, C. S. Dezzutti, R. B. Lal, and F. D. Quinn.** 1999. Observed differences in virulence-associated phenotypes between a human clinical isolate and a veterinary isolate of *Mycobacterium avium*. Infect. Immun. **67:**4895–4901.
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. Internat. Ed. Engl. **9:**610–619.
8. **Brennan, P. J., and H. Nikaido.** 1995. The envelope of mycobacteria. Annu. Rev. Biochem. **64:**29–63.
9. **Camphausen, R. T., R. L. Jones, and P. J. Brennan.** 1986. Structure and relevance of the oligosaccharide hapten of *Mycobacterium avium* serotype 2. J. Bacteriol. **168:**660–667.
10. **Chatterjee, D., and K. H. Khoo.** 2001. The surface glycopeptidolipids of mycobacteria: structures and biological properties. Cell. Mol. Life Sci. **58:** 2018–2042.
11. **Ciucanu, I., and F. Kerek.** 1984. A simple and rapid method for the permethylation of carbohydrates. Carbohydr. Res. **131:**209–217.
12. **Daffe, M., and P. Draper.** 1998. The envelope layers of mycobacteria with reference to their pathogenicity. Adv. Microb. Physiol. **39:**131–203.
13. **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.
14. **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.
15. **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.
16. **Fujiwara, N., N. Nakata, S. Maeda, T. Naka, M. Doe, I. Yano, and K. Kobayashi.** 2007. Structural characterization of a specific glycopeptidolipid containing a novel *N-acyl*-deoxy sugar from *Mycobacterium intracellulare* serotype 7 and genetic analysis of its glycosylation pathway. J. Bacteriol. **189:**1099–1108.
17. **Fujiwara, N., N. Nakata, T. Naka, I. Yano, M. Doe, D. Chatterjee, M. McNeil, P. J. Brennan, K. Kobayashi, M. Makino, S. Matsumoto, H. Ogura, and S. Maeda.** 2008. Structural analysis and biosynthesis gene cluster of an antigenic glycopeptidolipid from *Mycobacterium intracellulare*. J. Bacteriol. **190:**3613–3621.
18. **Horgen, L., E. L. Barrow, W. W. Barrow, and N. Rastogi.** 2000. Exposure of human peripheral blood mononuclear cells to total lipids and serovar-spe-

cific glycopeptidolipids from *Mycobacterium avium* serovars 4 and 8 results in inhibition of TH1-type responses. Microb. Pathog. **29:**9–16.

19. **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.
20. **Julander, I., S. Hoffner, B. Petrini, and L. Ostlund.** 1996. Multiple serovars of *Mycobacterium avium* complex in patients with AIDS. APMIS **104:**318–320.
21. **Kano, H., T. Doi, Y. Fujita, H. Takimoto, I. Yano, and Y. Kumazawa.** 2005. Serotype-specific modulation of human monocyte functions by glycopeptidolipid (GPL) isolated from *Mycobacterium avium* complex. Biol. Pharm. Bull. **28:**335–339.
22. **Krzywinska, E., and J. S. Schorey.** 2003. Characterization of genetic differences between *Mycobacterium avium* subsp. *avium* strains of diverse virulence with a focus on the glycopeptidolipid biosynthesis cluster. Vet. Microbiol. **91:**249–264.
23. **Maekura, R., Y. Okuda, A. Hirotani, S. Kitada, T. Hiraga, K. Yoshimura, I. Yano, K. Kobayashi, and M. Ito.** 2005. Clinical and prognostic importance of serotyping *Mycobacterium avium-Mycobacterium intracellulare* complex isolates in human immunodeficiency virus-negative patients. J. Clin. Microbiol. **43:**3150–3158.
24. **McNeil, M., A. Y. Tsang, and P. J. Brennan.** 1987. Structure and antigenicity of the specific oligosaccharide hapten from the glycopeptidolipid antigen of *Mycobacterium avium* serotype 4, the dominant mycobacterium isolated from patients with acquired immune deficiency syndrome. J. Biol. Chem. **262:**2630–2635.
25. **Miyamoto, Y., T. Mukai, Y. Maeda, M. Kai, T. Naka, I. Yano, and M. Makino.** 2008. The *Mycobacterium avium* complex *gtfTB* gene encodes a glucosyltransferase required for the biosynthesis of serovar 8-specific glycopeptidolipid. J. Bacteriol. **190:**7918–7924.
26. **Miyamoto, Y., T. Mukai, Y. Maeda, N. Nakata, M. Kai, T. Naka, I. Yano, and M. Makino.** 2007. Characterization of the fucosylation pathway in the biosynthesis of glycopeptidolipids from *Mycobacterium avium* complex. J. Bacteriol. **189:**5515–5522.
27. **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.
28. **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.
29. **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.
30. **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.
31. **Sweet, L., W. Zhang, H. Torres-Fewell, A. Serianni, W. Boggess, and J. Schorey.** 2008. *Mycobacterium avium* glycopeptidolipids require specific acetylation and methylation patterns for signaling through toll-like receptor 2. J. Biol. Chem. **283:**33221–33231.
32. **Tassell, S. K., M. Pourshafie, E. L. Wright, M. G. Richmond, and W. W. Barrow.** 1992. Modified lymphocyte response to mitogens induced by the lipopeptide fragment derived from *Mycobacterium avium* serovar-specific glycopeptidolipids. Infect. Immun. **60:**706–711.
33. **Tsang, A. Y., J. C. Denner, P. J. Brennan, and J. K. McClatchy.** 1992. Clinical and epidemiological importance of typing of *Mycobacterium avium* complex isolates. J. Clin. Microbiol. **30:**479–484.
34. **Vergne, I., and M. Daffe.** 1998. Interaction of mycobacterial glycolipids with host cells. Front. Biosci. **3:**d865–876.
35. **Wayne, L. G., R. C. Good, A. Tsang, R. Butler, D. Dawson, D. Groothuis, W. Gross, J. Hawkins, J. Kilburn, M. Kubin, K. H. Schroder, V. A. Silcox, C. Smith, M. F. Thorel, C. Woodley, and M. A. Yakrus.** 1993. Serovar determination and molecular taxonomic correlation in *Mycobacterium avium*, *Mycobacterium intracellulare*, and *Mycobacterium scrofulaceum*: a cooperative study of the International Working Group on Mycobacterial Taxonomy. Int. J. Syst. Bacteriol. **43:**482–489.
36. **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.

Microbial Pathogenesis 49 (2010) 181–185

Contents lists available at ScienceDirect

Microbial Pathogenesis

journal homepage: www.elsevier.com/locate/micpath

Analysis of *Mycobacterium leprae* gene expression using DNA microarray

Takeshi Akama, Kazunari Tanigawa, Akira Kawashima, Huhehasi Wu, Norihisa Ishii, Koichi Suzuki*

Leprosy Research Center, National Institute of Infectious Diseases, 4-2-1 Aoba-cho, Higashimurayama-shi, Tokyo 189-0002, Japan

ARTICLE INFO

Article history:
Received 8 February 2010
Received in revised form
19 May 2010
Accepted 20 May 2010
Available online 27 May 2010

Keywords:
Mycobacterium leprae
Leprosy
DNA microarray
Operon

ABSTRACT

Mycobacterium leprae, the causative agent of leprosy, does not grow under *in vitro* condition, making molecular analysis of this bacterium difficult. For this reason, bacteriological information regarding *M. leprae* gene function is limited compared with other mycobacterium species. In this study, we performed DNA microarray analysis to clarify the RNA expression profile of the Thai53 strain of *M. leprae* grown in footpads of hypertensive nude rats (SHR/NCrj-*rnu*). Of 1605 *M. leprae* genes, 315 showed signal intensity twofold higher than the median. These genes include Acyl-CoA metabolic enzymes and drug metabolic enzymes, which might be related to the virulence of *M. leprae*. In addition, consecutive RNA expression profile and *in silico* analyses enabled identification of possible operons within the *M. leprae* genome. The present results will shed light on *M. leprae* gene function and further our understanding of the pathogenesis of leprosy.

1. Introduction

Mycobacterium leprae (*M. leprae*) is the causative agent of leprosy, a chronic bacterial disease of the skin and nerves that still occurs at a relatively high incidence in South and Southeast Asia, South America, and Africa [1]. *M. leprae* grows in macrophages or Schwann cells upon *in vivo* infection; however, this bacterium has not been cultured *in vitro* to date [2]. Moreover, although nine-banded armadillos or nude mice have been used as limited models of *leprosy* [3], the mechanisms of infection, parasitization, proliferation, and pathogenesis of *M. leprae* have largely been undetermined. Previous studies on the nature of *M. leprae* were limited to methods requiring no gene manipulation, such as evaluation of respiration or ATP production [4], detection of RNA, or the use of other mycobacteria to study the function of specific *M. leprae* genes [5,6].

Whole genome sequence analysis of *M. leprae* has revealed that the genome consists of 3.3 Mbp and has only 1605 genes that encode proteins, but an unexpectedly large number of pseudogenes, that is, 1115 [7]. This genome sequence information has elucidated the evolutionary process of the *M. leprae* genome by phylogeographic analysis [8] or comparative genome science [9].

In order to clarify the expression profile of the entire *M. leprae* genome, as a first step, we previously performed membrane array analysis, then designed a whole genome tiling array analysis, and showed that high levels of RNA are transcribed from pseudogenes and other non-coding regions [10,11]. Tiling array was effective in detecting RNA expression from non-coding regions and pseudogenes, as well as in comparing the expression level between genes and non-coding regions. Nevertheless, the obtained signals in the array might have been affected by possible mismatch hybridization because of the low specificity of the tiling probes. This resulted in variable signal intensities of each probe even within the same gene. To date, precise analysis of *M. leprae* gene expression has been limited and a previous report utilized *Mycobacterium tuberculosis* (*M. tuberculosis*) open reading frame (ORF) array [12]. Moreover, little information is available regarding gene structures, such as operons.

In this study, we investigated the precise gene expression profile of *M. leprae* using an ORF array on which highly specific and multiple probes were mounted. We identified several functional genes whose expression level is significantly higher than that of other genes. We also found possible operon structures based on their consecutive expression and through the use of various prediction software packages.

2. Results

2.1. Evaluation of signal intensities of M. leprae ORF array

Total RNA was extracted from 7.2×10^{10} *M. leprae* cells grown in the footpad of SHR/NCrj-*rnu* rats and treated with DNase I. The ratio of 23S/16S rRNA was 0.83, indicating that the purified RNA was of sufficiently quality to proceed with array hybridization. Samples were then reverse-transcribed, fluorescence-labeled and

* Corresponding author. Tel.: +81 42 391 8211; fax: +81 42 394 9092.
E-mail address: koichis@nih.go.jp (K. Suzuki).

0882-4010/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.micpath.2010.05.010

hybridized with the ORF array. In this array, 20 different probes were designed for each of the 1605 *M. leprae* genes. These probes together with the multiple control probes were randomly arranged in one block on an array, and five different blocks were used for hybridization.

After scanning the array, the signal intensity of each probe ranged from 2^0 to 2^{16} (Fig. 1). From all the probes used, 19.7% showed signal intensity twofold higher than the median (*i.e.*, 5800 or $2^{12.5}$). Average signal intensity from 20 different probes in each gene was calculated for five different blocks. Genes with average signal intensity twofold higher than the median were considered as strongly positive. As a result, 315 genes were identified as highly expressed (Table S1). The expression of those genes was confirmed by RT-PCR and real-time PCR analysis for 28 and four genes, respectively (data not shown).

2.2. Expressed genes and their functional classes

M. leprae genes are classified into six functional classes [7]. To predict the characteristics of highly expressed genes, their functional classes were compared with those of the entire genes and pseudogenes (Fig. 2). The proportion of functional classes, whose RNA was detected strongly using ORF array, was found to be similar to that of the original gene classes, although there were some differences in the proportion of the six classes between genes and pseudogenes. This evidence suggests that all the six classes of *M. leprae* genes are similarly transcribed without significant preferences. Most of the genes detected by a previous tiling array analysis [10] were also detected using ORF array (92% matched).

Table 1 shows a more detailed expression profile of the functional gene classes. The average expression level of class III (Cell process) was the highest (27.3%), and that of class I (Small molecule metabolism) was the lowest (14.5%), although there was no significant difference. Interestingly, five out of the six genes expressed in class I.A were acyl-CoA metabolic enzymes, namely, ML0132, ML0138, ML0737, ML1241, and ML2358 (*fadD2*, *fadD28*, *fadE25*, *chA12*, and *fadD26*, respectively). Among the expressed genes, 12 showed more than tenfold stronger expression than the median (Table 2). Half of these genes belonged to Functional class II (Macromolecule metabolism); notably, four genes were from class II.C (Cell envelope). A putative monooxygenase showed the highest expression followed by hypothetical protein, which was 64–77% homologous to the *PadR* transcriptional regulator family of seven *Mycobacterium* species, namely, *Mycobacterium avium* (77%), *Mycobacterium marinum* (76%), *Mycobacterium ulcerans* (76%), *Mycobacterium smegmatis* (67%), *Mycobacterium gilvan* (67%), *Mycobacterium vanbaalenii* (65%), and *Mycobacterium abscessus* (64%), but not to *M. tuberculosis*, by BLASTP search (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Two chaperons, *DnaK* and *GroEL*, were also highly expressed (third and fifth strongest, respectively).

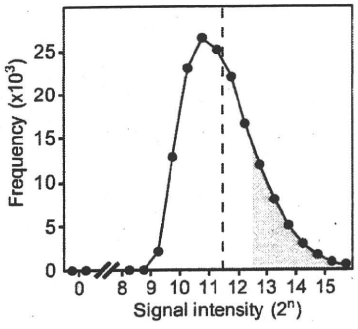

Fig. 1. Distribution of signal intensities detected using *Mycobacterium leprae* open reading frame (ORF) array. *M. leprae* ORF array was hybridized with RNA extracted from the Thai53 strain of *M. leprae*. Obtained signal intensities are shown as 2^n for each signal intensity. The median (2900) of the normalized entire signal intensities is indicated by a dashed line. The area whose signal intensity is twofold stronger than the median (*i.e.*, 5800 or $2^{12.5}$) is indicated by shading.

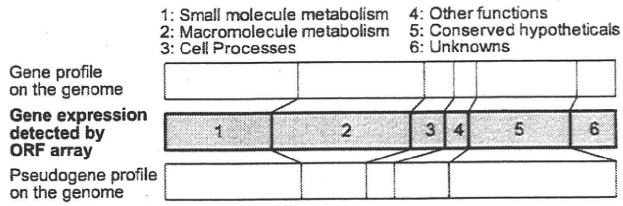

Fig. 2. Highly expressed gene within functional classes. Functional classes of highly expressed genes were compared with their original proportion in the genes and pseudogenes according to the classification by Cole et al. [7]. The significance of gene expression on each functional class was evaluated by the chi-square test ($\chi^2 = 1.0$, $p = 0.960$ for the genes and $\chi^2 = 23.4$, $p = 0.003$ for the pseudogenes).

Table 1
Functional classification of genes whose expression ratio was twofold higher than the median.

Classification	Number of genes expressed	Number of genes in the genome	Expression ratio (%)
I Small molecule metabolism			14.5
I.A Degradation	6	49	12.2
I.B Energy metabolism	18	91	19.8
I.C Central intermediary metabolism	8	30	26.7
I.D Amino acid biosynthesis	8	77	10.4
I.E Polyamine synthesis	0	0	0.0
I.F Purines, pyrimidines, nucleosides and nucleotides	11	52	21.2
I.G Biosynthesis of cofactors, prosthetic groups and carriers	6	63	9.5
I.H Lipid biosynthesis	5	35	14.3
I.I Polyketide and non-ribosomal peptide synthesis	2	13	15.4
I.J Broad regulatory functions	9	57	15.8
II Macromolecule metabolism			20.1
II.A Synthesis and modification of macromolecules	42	159	26.4
II.B Degradation of macromolecules	6	43	14.0
II.C Cell envelope	51	256	19.9
III Cell processes			27.3
III.A Transport/binding proteins	8	55	14.5
III.B Chaperones/heat shock	6	14	42.9
III.C Cell division	4	11	36.4
III.D Protein and peptide secretion	2	10	20.0
III.E Adaptations and atypical conditions	1	6	16.7
III.F Detoxification	2	6	33.3
IV Other[a]			21.5
IV.A Virulence	1	7	14.3
IV.B IS elements, repeated sequences, and phage	0	4	0.0
IV.C PE and PPE families	5	11	45.5
IV.D Antibiotic production and resistance	0	1	0.0
IV.E Bacteriocin-like proteins	0	0	0.0
IV.F Cytochrome P450 enzymes	1	1	100.0
IV.G Coenzyme F420-dependent enzymes	1	2	50.0
IV.H Miscellaneous transferases	7	35	20.0
IV.I Miscellaneous phosphatases, lyases, and hydrolases	1	15	6.7
IV.J Cyclases	0	0	0.0
IV.K Chelatases	0	0	0.0
V Conserved hypotheticals	71	360	19.7
VI Unknowns	33	141	23.4
Total	315	1604	19.6

[a] This class includes one gene that is not sub-classified.

Table 2
Genes with more than tenfold stronger expression than the median.

Fold of median	Acc. No.	Classification	Gene name
17.6	ML0065	I.B.7	Probable monooxygenase *EthA*
16.4	ML2313	VI	Hypothetical protein
15.6	ML2496	III.B	Molecular chaperone *DnaK*
15.1	ML2042	III.F	Alkyl hydroperoxide reductase
14.6	ML0317	III.B	Chaperonin *GroEL*
14.2	ML1735	I.F.3	Probable *Nrdl*-family protein
13.7	ML1394	II.A.6	Translation initiation factor *IF-3*
13.1	ML0577	III.D	Probable preprotein translocase *SecA1* subunit
12.5	ML0319	II.C.1	Possible lipoprotein *LpqE*
11.4	ML0050	II.C.2	Possible 10 kDa culture filtrate antigen homolog *EsxB* (*lhp*) (*cfp10*)
11.0	ML2135	II.C.4	Possible transmembrane protein
10.3	ML1171	II.C.4	Possible conserved integral membrane protein

2.3. Detection of expressed operons in M. leprae genome

In the course of analyzing 315 highly expressed genes, we found some sequential expressions of gene sets. We therefore mapped highly expressed genes to the entire genome structure (http://www.sanger.ac.uk/Projects/M_leprae/) and detected 19 sequential regions that have three or more genes highly expressed in the same direction (Table 3). Of these 19 regions, seven were known operons reported in *M. leprae* or *M. tuberculosis* (Table 3, Nos. 1–7). These highly expressed operons include genes related to protein synthesis of ribosome subunits ML1132-ML1134 (*rho*, *rpmE* and *prfA*), ML1957-1962 (*rpoA*, *rpsD*, *rpsK*, *rpsM*, *rpmJ* and *infA*), and ML2683-ML2685 (*prsR*, *ssb* and *rpsF*) (Table 3, Nos. 3, 4, and 7, respectively). Also, 6-kDa early secreted antigenic target (ESAT-6) and culture filtrate protein 10 (CFP-10) family genes ML0049-ML0050 (*esxA* and *esxB*) and ML2531-ML2532 (*esxH* and *esxG*), which are known virulent factors of *M. tuberculosis* (Table 3, Nos. 1 and 6, respectively), were highly expressed within operons. Others include secretion in mycobacteria (*Snm*) proteins ML0406-ML0407 (*Snm9* and *Snm10*) (Table 3, No. 2) and molecular chaperons ML2494-ML2496 (*dnaJ*, *grpE* and *dnaK*) (Table 3, No. 5).

To predict possible operon structures in the consecutively expressed gene regions that were identified, three programs were used: MicrobesOnline (http://www.microbesonline.org/operons/gnc1769.html) [13], Operon DataBase (http://operondb.cbcb.umd.edu/cgi-bin/operondb/operons.cgi) [14], and DOOR (http://csbl1.bmb.uga.edu/OperonDB/index.php) [15]. By using these programs, six out of seven known operons were successfully predicted by two or all three programs. Among 12 operon candidates identified in this study (Table 3, Nos. 8–19), 11 were predicted as operon by two or all three programs and seven by all three programs. These potential operons include ABC transporter components (ML1425-ML1427) (Table 3, No. 13) that are expected to import carbohydrates [16] and *nrd* protein genes that catalyze *de novo* synthesis of deoxyribonucleotides [17] *nrdE*, *nrdI*, and *nrdH* (ML1734-ML1736) (Table 3, No. 16).

3. Discussion

We took advantage of *M. leprae* ORF array on which gene-specific multiple probes were arranged. By using highly purified *M. leprae* RNA from SHR/NCrj-*rnu* rats, where *M. leprae* notably proliferates and causes disease resembling that of humans [18,19], it was possible to analyze the gene expression profile of the entire range of *M. leprae* genes that may cause disease resembling that seen in humans.

Highly expressed genes were detected similarly from the six functional classes without significant deflection. Among these genes, five were related to Acyl-CoA metabolic enzymes. The metabolite of these enzymes, phthiocerol dimycocerosate (DIM) lipids, composes the cell wall of only pathogenic mycobacteria and this group of lipids is thought to play a role in cell wall architecture and permeability that affect virulence [20,21]. Because all the five genes detected in *M. leprae* have their corresponding orthologues in *M. tuberculosis*, these genes may play essential roles in the virulence of both species.

Table 3
Known and possible operons detected using ORF array.

No.	Genes expressed sequentially	Refs.	Functional class	Prediction programs		
				Microbes online[a]	Operon Database[b]	DOOR[c]
Known operons						
1	ML0048, ML0049, ML0050	[10]	V, II.C.2, II.C.2	–	+	+
2	ML0405, ML0406, ML0407	[16]	V, V, V	–	+	+
3	ML1132, ML1133, ML1134	[18]	II.A.7, II.A.1, II.A.6	–	+	–
4	ML1957, ML1958, ML1959, ML1960, ML1961, ML1962	[18]	II.A.7, II.A.1, II.A.1, II.A.1, II.A.1, II.A.6	+	+	+
5	ML2494, ML2495, ML2496	[33]	III.B, III.B, III.B	+	+	+
6	ML2530, ML2531, ML2532	[17]	I.J.1, II.C.2, IV.C.1.a	–	+	+
7	ML2683, ML2684, ML2685	[18]	II.A.1, II.A.5, II.A.1	+	+	+
Novel operon candidates						
8	ML0125, ML0126, ML0127	None	IV.H, VI, VI	–	+	–
9	ML0245, ML0246, ML0247	None	II.A.1, II.C.1, III.F	–	+	+
10	ML0574, ML0575, ML0576, ML0577, ML0578	None	VI, VI, VI, III.D, I.C.1	–	+	+
11	ML0592, ML0593, ML0594	None	I.J.1, V, V	+	+	+
12	ML1180, ML1181, ML1182	None	V, V, IV.C.2	–	+	+
13	ML1425, ML1426, ML1427	None	III.A.3, III.A.3, III.A.3	+	+	+
14	ML1468, ML1469, ML1470	None	II.B.1, I.F.5, II.C.4	+	+	+
15	ML1616, ML1617, ML1618	None	II.A.2, V, II.A.1	+	+	+
16	ML1734, ML1735, ML1736	None	I.F.3, I.F.3, I.F.3	+	+	+
17	ML1905, ML1906, ML1907	None	II.A.1, II.A.7, III.D	+	+	+
18	ML2259, ML2260, ML2261	None	II.C.4, III.A.4, V	–	+	+
19	ML2706, ML2707, ML2708, ML2709	None	III.C, III.C, V, V	+	+	+

–: no prediction, +: predicted.
[a] MicrobesOnline (http://www.microbesonline.org/operons/gnc1769.html) [13].
[b] Operon DB (http://operondb.cbcb.umd.edu/cgi-bin/operondb/operons.cgi) [14].
[c] DOOR (http://csbl1.bmb.uga.edu/OperonDB/index.php) [15].