Molecular cloning and expression of a novel glycolipid sulfotransferase in *Mycobacterium tuberculosis*

Carlos A. Rivera-Marrero,¹ Jeffrey D. Ritzenthaler,¹ Sarah A. Newburn,¹ Jesse Roman¹ and Richard D. Cummings²

Author for correspondence: Carlos A. Rivera-Marrero. Tel: +1 404 321 6111 ext. 7178. Fax: +1 404 728 7750. e-mail: crivera@emory.edu

¹ Department of Medicine, Pulmonary and Critical Care Division, Atlanta VA Medical Center, Emory University School of Medicine, Room 12C 106, 1670 Clairmont Rd, Decatur, GA 30033, USA

² Department of Biochemistry and Molecular Biology, Oklahoma Center for Medical Glycobiology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

Sulfated trehalose glycolipids are among the most characteristic cell wall molecules of virulent strains of *Mycobacterium tuberculosis*. They comprise a family of trehalose-2-sulfate esters with an array of acyl fatty acids at various positions of the trehalose moiety. Although their structure has been well characterized, most of the enzymes involved in their biosynthesis, such as sulfotransferases, are unknown. It is demonstrated here by metabolic labelling with ³⁵S abundant incorporation into sulfolipids of *M. tuberculosis* strains, in comparison to *Mycobacterium avium*, *Mycobacterium bovis* BCG and *Mycobacterium smegmatis*. The most abundant sulfolipid, sulfolipid I, is present in virulent strains H37Rv and Erdman, but absent in attenuated H37Ra. Sulfotransferase assays with the donor substrate 3'-phosphoadenosine-5'-[³⁵S]phosphosulfonate and whole cell lysates of H37Ra resulted in the synthesis of four major sulfolipids (I, II, IV and VI). A search for sulfotransferase gene sequences in *M. tuberculosis* yielded gene *Rv1373*, a 981 bp gene slightly homologous (24 % identity) to eukaryotic aryl-sulfotransferases. *Rv1373* was cloned by PCR and expressed as a 39 kDa recombinant his-tagged protein. The recombinant *M. tuberculosis* aryl-sulfotransferase exhibited activity towards the cerebroside glycolipids glucosyl- and galactosylceramide. No activity was detected with sulfatide (3'-sulfated galactosylceramide), suggesting that sulfation of galactosylceramide may occur at C-3 of the galactose. Treatment of sulfated products with ceramide glycanase resulted in the release of ³⁵S-labelled material showing that sulfation was at the saccharide moiety (galactose or glucose) of the ceramide. Assays with the *M. tuberculosis* aryl-sulfotransferase and total H37Ra glycolipids showed one major product corresponding to sulfolipid IV. These results demonstrate that *Rv1373* encodes a novel glycolipid sulfotransferase with activity towards typical ceramide glycolipids and mycobacterial trehalose glycolipids.

Keywords: *Mycobacterium*, gene expression, glycolipid sulfotransferase, sulfolipid

INTRODUCTION

Mycobacterium tuberculosis is the causative agent of human tuberculosis, a worldwide health problem caus-

Abbreviations: PAPS, 3'-phosphoadenosine-5'-phosphosulfonate; SL, sulfolipid; ST, sulfotransferase; sulfatide, 3'-sulfate galactosylceramide.

The GenBank/EMBL/DBJ accession number for the sequence (gene *Rv1373*) reported in this paper is Z81011.

ing approximately 3 million deaths per year (Bloom & Murray, 1992; Dye *et al.*, 1999). During the early stages of infection, a key determinant of virulence is the ability to survive and replicate within the phagosome of host phagocytic cells, primarily the macrophage (Riley, 1996). The components of the mycobacterial cell envelope are considered prime determinants for intracellular survival and virulence. The cell wall functions as a permeability barrier that shields the organism from

environmental stress, protects from the intracellular attack mechanisms of the macrophage and provides resistance to many antibiotics (Brennan & Nikaido, 1995). Identification of the molecular determinants associated with virulence, and elucidating the mechanisms of phagolysosomal evasion and intracellular survival are currently among the most important aspects of tuberculosis research.

Sulfated glycolipids of the cell wall of *M. tuberculosis* were implicated several years ago in virulence. Early studies showed that virulent strains could be distinguished cytochemically by their staining with the cationic dye neutral red (Dubois & Middlebrook, 1948), leading to the discovery of the neutral-red-reactive sulfolipids (SLs) (Middlebrook *et al.*, 1959). Since then, several studies have established a correlation between the presence of sulfated glycolipids and the degree of virulence of *M. tuberculosis* strains (Gangadharam *et al.*, 1963; Mitchison, 1964; Goren *et al.*, 1974, 1982). It has also been proposed that sulfated glycolipids may be involved in intracellular survival of virulent *M. tuberculosis* by their interaction with phagosomes and prevention of lysosomal fusion (Goren *et al.*, 1976; Goren, 1977; D'Arcy Hart & Young, 1988; Fujiwara, 1997). More recently, it was shown that *M. tuberculosis* sulfolipid-I (SL-I) blocked the LPS and gamma interferon (INF- γ) activation of human macrophages for enhanced release of superoxide (Pabst *et al.*, 1988; Brosna *et al.*, 1991). In neutrophils however, SL-1 stimulated superoxide production and primed neutrophil responses to several metabolic agonists such as *N*-formyl methionyl-leucyl-phenylalanine (FMLP) and phorbolmyristate acetate (PMA) (Zhang *et al.*, 1988, 1991). Although these studies strongly implicate SLs as potential virulence factors, the role of SLs in virulence, immune modulation and pathogenicity of tuberculosis remains unclear.

The major sulfated glycolipids of *M. tuberculosis* consist of a family of trehalose-2-sulfate esters with an array of acyl fatty acids (phthioceratanate, hydroxyphthioceratanate, palmitate, stearate) at various positions of the trehalose molecule (Goren, 1984). SL-1, the principal sulfatide of *M. tuberculosis*, has been identified as 2,3,6,6'-tetraacyl- α,α' -D-trehalose-2'-sulfate (Goren, 1970a, b). Although the structure of *M. tuberculosis* trehalose SLs has been well characterized, their biosynthetic pathways, including enzymes such as sulfotransferases (STs) possibly involved in their biosynthesis and regulation, are not known. The identification of genes encoding these enzymes and the generation of allelic knockout mutants is crucial for determining their role in the biology and pathogenesis of tuberculosis.

In this study we demonstrate that gene *Rv1373*, a putative aryl-ST gene in *M. tuberculosis*, encodes a novel glycolipid ST with activity towards the eukaryotic glycolipids galactosylceramide and glucosylceramide, and endogenous mycobacterial glycolipids. This is the first report of a mycobacterial glycolipid ST possibly involved in the biosynthesis of the biologically relevant trehalose SLs of the cell wall of *M. tuberculosis*.

METHODS

Bacterial strains and growth conditions. *M. tuberculosis* strains, attenuated H37Ra (ATCC 25177), virulent H37Rv (ATCC 27294), Erdman (ATCC 35801), *Mycobacterium avium* (MAC 1, ATCC 700898) colony variants smooth opaque (SO) and transparent (ST), *Mycobacterium bovis* BCG (ATCC 35734) and *Mycobacterium smegmatis* mc²155 (ATCC 700084) were purchased from American Type Culture Collection, Manassas, VA, USA. Bacteria were grown in Middlebrook 7H9 broth supplemented with ADC enrichment (Difco Laboratories), containing 2% glucose and 1% (v/v) glycerol. Mycobacteria were grown at 37 °C with slow shaking for 5–14 days. Bacterial cultures of 10 ml at 3×10^8 bacteria ml⁻¹, were centrifuged at 3000 *g* for 15 min and wet pellets processed for preparation of cell lysates and glycolipid extracts.

Metabolic labelling and glycolipid isolation. Mycobacterium species *M. avium* smooth opaque and transparent, *M. bovis* BCG, *M. smegmatis* mc²155, and *M. tuberculosis* strains H37Ra, H37Rv and Erdman were grown in 50 ml cultures of 7H9 broth containing 100 μ Ci ³⁵S (as sulfate; 3.7 MBq; 37 TBq mmol⁻¹; Amersham Pharmacia) for 5–7 days. Cultures were centrifuged at 3000 *g* for 10 min, the bacterial pellets were washed in PBS, resuspended in 1.0 ml PBS and sonicated extensively (10 times, 20 s pulses). Glycolipids were extracted by a modification of the Folch method (Morrison, 1994). Two volumes of chloroform/methanol (2:1) were added, the samples were sonicated (20 s), centrifuged (9500 *g* for 1 min), and the upper (aqueous) and lower (organic) phases saved. The upper phase contains gangliosides and neutral lipids with long carbohydrate chains and contaminating protein; the lower phase contains the neutral glycolipids with short carbohydrate chains, neutral lipids and phospholipids. The lower organic phase, containing insoluble material, was extracted sequentially in chloroform/methanol/water (4:8:3), chloroform/methanol (1:1), chloroform/water (2:1) and 100% ethanol. After extraction, the pellet of insoluble material was discarded and the extracted aqueous and organic phases were combined, dried *in vacuo* and resuspended in 50–100 μ l chloroform/methanol (2:1). Glycolipid samples (approx. 10000 c.p.m.) were separated by TLC on Silica gel 60 plates (EM Science) with chloroform/methanol/water (65:25:4), allowed to dry and exposed to X-ray film or a scanning phosphorimager. TLC plates were also stained with orcinol ferric chloride spray (Sigma) and heated at 100 °C until the brownish colour bands of glycolipids were revealed.

Gene cloning. The 0.98 kb putative aryl-ST gene homologue *Rv1373*, in *M. tuberculosis* cosmid SCY02B12 (accession no. z81011) (Philipp *et al.*, 1996; Cole *et al.*, 1998), was cloned by PCR amplification of *M. tuberculosis* genomic DNA. A set of forward (5'-CCCCGAATTCGATGAATTCAGAACACCCGAT-3') and reverse (5'-CCCCAAGCTTTCAGTTGGCCGGTTCGTATC-3') oligonucleotide primers, with *Eco*RI and *Hind*III restriction sites (underlined) incorporated at the 5' ends, were utilized for PCR amplification. PCR amplification was done in 50 μ l reaction mixtures under the action of *Taq* and *Pwo* DNA polymerases and the Expand Long Template PCR System (Roche). Amplification conditions were as follows: 1 cycle of 94 °C for 2 min; 5 cycles of 94 °C for 45 s, 65 °C for 45 s, 68 °C for 5 min; 35 cycles of 94 °C for 45 s, 55 °C for 45 s, 68 °C for 5 min; and a final extension of 68 °C for 7 min. The reaction product was separated by low-melting-point agarose electrophoresis and visualized by ethidium bromide staining. The discrete 1012 bp band was excised from the gel and isolated by spin filtration. After digestion with

Fig. 1. Sulfated lipids synthesized by *Mycobacterium* spp. (a) TLC of total lipids stained with orcinol (left panel) and autoradiography of ³⁵S-labelled lipids (right panel) isolated from *M. avium* opaque (MavO) and translucent (MavT) strains, *M. bovis* BCG, and *M. smegmatis* (M. smeg). (b) TLC of total lipids stained with orcinol (left panel) and autoradiography of ³⁵S-labelled lipids (right panel) isolated from *M. tuberculosis* H37Ra, H37Rv and Erdman strains. The major SLs detected in Erdman are designated I–VI based on their relative mobility.

*Eco*RI and *Hind*III, the gene was subcloned into the *Eco*RI/*Hind*III site of expression vectors pET-23b (Novagen) and pTRcHisC (Invitrogen). The resulting construct STpET contains the full length *M. tuberculosis* aryl-ST gene with a 6-histidine fusion at the carboxyl terminus, while the STpTRcHis contains a 6-histidine peptide and an enterokinase cleavage sequence at the amino terminal end of the protein. The *M. tuberculosis* aryl-ST gene was also subcloned out-of-frame (construct OF-STpTRcHis) to generate an inactive histagged protein for control experiments. Sequence analysis of the *M. tuberculosis* aryl-ST gene plasmid constructs was done in a 373 ABI Sequencer by the *Taq* dye-deoxy terminator method using the T7 promoter and T7 terminator primers and gene-specific primers based on published genomic DNA sequence. Sequence data were analysed using GenBank and the NCBI BLAST server (Altschul *et al.*, 1990).

Expression and purification of recombinant gene *Rv1373*.

Expression of *M. tuberculosis* gene *Rv1373* as a recombinant fusion protein in *Escherichia coli* was conducted following Novagen protocols. Plasmid constructs STpET, STpTRcHis and OF-STpTRcHis were introduced into the BL21(DE3) host *E. coli* strain and T7 RNA polymerase gene-specific transcription induced by addition of 0.8 mM IPTG for 3 h at 37 °C. After induction, cell extracts were prepared by lysis under non-denaturing conditions and separated by metal chelation affinity chromatography using His-Trap resin (Amersham Pharmacia). Histidine-bound proteins were eluted stepwise with 60 mM, 300 mM and 500 mM imidazole buffer (all in 20 mM phosphate, 0.5 M NaCl), pH 7.5. The eluted proteins were dialysed against PBS pH 7.2, concentrated in a Centriprep-10 filter (Amicon), and analysed by SDS-PAGE and Coomassie blue staining. Purified recombinant fusion proteins were stored at –70 °C until tested for ST activity.

ST assays. These were conducted with sonicated cell extracts of *M. tuberculosis* bacilli and with the purified recombinant *M. tuberculosis* aryl-ST. Assays with *M. tuberculosis* extracts consisted of 100 µg protein (containing enzyme and acceptor substrate), 1.0 µCi (37 kBq) donor substrate 3'-phosphoadenosine-5'-[³⁵S]phosphosulfonate, PAP³⁵S, (NEN Life Science Products; 96.2 GBq mmol⁻¹), in 100 µl reaction buffer (40 mM MES/NaOH pH 7.0, 0.05 M NaCl, 0.2 % Triton X-100, 5 mM EDTA). Assays with purified recombinant protein from clones

STpET (his-tag at C terminus), STpTRcHis (his-tag at N terminus) and OF-STpTRcHis (out-of-frame) contained 100 µg protein and 100 µg acceptor substrates galactosylceramide, glucosylceramide or sulfatide (Sigma). Assays with *M. tuberculosis* glycolipids as substrate contained 50–100 µg total glycolipid. Reaction mixtures were incubated for 6 h at 37 °C, then 900 µl water was added to stop the reactions and the samples were separated by reverse phase chromatography in a Sep-Pak C:18 column pre-equilibrated in water. The column was washed with 15 × 1 ml aliquots of water to remove hydrophilic molecules and eluted with 10 × 1 ml aliquots of methanol to elute the hydrophobic lipids. The radiolabelled SL products eluted were determined by scintillation counting, and by TLC and autoradiography on Silica gel 60 plates (EM Science) with chloroform/methanol/water (65:25:4). ST assays were done in duplicate and controls without enzyme or acceptor substrate were included.

Ceramide glycanase digestion. The ³⁵S-labelled glycolipid fractions (approx. 10000–20000 c.p.m.) obtained by elution of C:18 columns in methanol were dried in Eppendorf tubes, and resuspended in 35 µl water, 10 µl 250 mM phosphate buffer pH 5.0 and 1 µl (0.24 mU) ceramide glycanase (Sigma). Enzyme digestions were done for 5 h or overnight at 37 °C. After digestion, reaction mixtures were diluted in 1 ml water and separated by reverse phase chromatography. The ³⁵S-labelled digestion products were determined by liquid scintillation counting, and by TLC and autoradiography as described.

RESULTS

Differences in SL content among *Mycobacterium* species

Studies were first conducted by metabolic radiolabelling to determine potential differences in glycolipid sulfation among non-pathogenic and pathogenic mycobacterial species. Mycobacteria such as *M. avium* smooth transparent and opaque strains, BCG, and *M. smegmatis* were grown in the presence of ³⁵S and total lipids isolated by Folch extraction. Analysis by TLC and autoradiography revealed very few sulfated (³⁵S-labelled) lipids in these species (Fig. 1a). As shown, only a

couple of highly polar SLs and one less polar SL are present in *M. avium* strains and *M. smegmatis*, while BCG is almost devoid of SLs (Fig. 1a). By contrast, similar analyses with *M. tuberculosis* attenuated strain H37Ra, and virulent strains H37Rv and Erdman, revealed abundant expression of SLs. As shown in Fig. 1(b), six SLs (labelled I–VI) are present in the virulent *M. tuberculosis* strain Erdman. Interestingly, the most abundant and less polar SL (SL-I) is present in virulent strains (H37Rv, Erdman) but is absent in the attenuated strain H37Ra. The mobility of this SL is slightly lower than a non-polar abundant SL seen in *M. smegmatis*. These results demonstrate that *M. tuberculosis* strains more abundantly express several SLs than non-pathogenic mycobacterial species.

M. tuberculosis cell lysates contain ST activity

We attempted to assay the putative ST activity in whole lysates of *M. tuberculosis* using the donor substrate PAP³⁵S. After 6 h incubation at 37 °C, the potential ³⁵S-labelled SLs were isolated by reverse-phase chromatography. The elution with methanol of ³⁵S-labelled material from the C:18 column in assays with cell free lysates of H37Ra is shown in Fig. 2(a). The peaks of unbound (not shown) and bound material obtained from the C:18 column were then analysed by TLC. The C:18 unbound material, presumably protein, is highly polar and does not migrate in the TLC (Fig. 2b). The C:18 bound material, presumably lipid, exhibited four major bands corresponding to SLs I, II, IV and VI, obtained from ³⁵S-metabolic labelling of *M. tuberculosis* (Fig. 1b). The most abundant species obtained were SLs III and VI. Interestingly, the relatively non-polar SL-I, present in virulent strains but not in H37Ra (Fig. 1b), appears to be sulfated *in vitro* in ST assays with H37Ra extracts. Assays with cell lysates from H37Rv and Erdman showed consistently lower activity levels in comparison to H37Ra (data not shown). These results demonstrate that *M. tuberculosis* cell lysates contain glycolipid STs that can utilize PAPS as the donor substrate and transfer sulfate to various endogenous lipid acceptor substrates *in vitro*.

A putative aryl-ST gene homologue in *M. tuberculosis*

A gapped BLAST search (Altschul *et al.*, 1990) for ST gene sequences in mycobacteria yielded gene *Rv1373* in cosmid SCY02B12 (accession no. z81011), a 981 bp putative ST gene in *M. tuberculosis* with slight similarity to eukaryotic aryl-STs (Philipp *et al.*, 1996; Cole *et al.*, 1998). The amino acid identities between hypothetical protein Rv1373 and a human hydroxysteroid ST (accession no. U92315) are shown in Fig. 3. In a stretch of 292 residues there is 24% identity, 39% positive residues and 19% gaps. The highly conserved signature sequences of eukaryotic STs, believed to be involved in PAPS binding and transfer of sulfate, are shown in Fig. 3. A similar degree of sequence identity is seen by BLAST analysis with many eukaryotic STs, among them: a murine hydroxysteroid ST (accession no. AF026072), a

Fig. 2. ST activity in H37Ra cell lysates. (a) Separation of ST products by C:18 (Sep-Pak) reverse-phase chromatography showing elution of ³⁵S-labelled glycolipids in methanol. (b) TLC autoradiogram of the unbound and bound fractions showing four bands of ³⁵S-labelled SL products, identified as SLs I, III, IV and VI, based on their relative mobility.

guinea pig oestrogen ST (accession no. U09552), a rat aryl-ST (accession no. X52883), a rat tyrosine-ester ST (accession no. U32372), and a human oestrogen ST (accession no. U08098). No sequence identities were observed, however, with any prokaryotic STs. Gene *Rv1373* probably encodes a soluble protein since no transmembrane domains or signal peptide sequences are present, based on PredictProtein prediction analyses (<http://cubic.bioc.columbia.edu>) (Rost, 1996).

Gene *Rv1373* was cloned by PCR of genomic DNA, subcloned into expression vectors pET23b and pTRcHis, and expressed in *E. coli* as a recombinant His-fusion protein. Construct STpET contains the full length *Rv1373* gene with a 6-His fusion at the carboxyl terminus, while the STpTRcHis contains a 6-His peptide and an enterokinase cleavage sequence at the amino terminus. While the predicted size for the native protein

Rv1373	MNSEHPMTDRVVYR	14
hST	MDGPAEPQIPGLWDITYEDDISEISQKLPGEYFRYKGVPPVGLY	44
Rv1373	SLMADNLRWDALQRLDGIISAPSKSGLTWTQRLVSLVDFDGPD--LPGLSTVSPNL	71
hST	SL + +L + +RD DI I KSG TW ++ L++ +G P+ +PW	104
	<u>Region I</u>	
Rv1373	DQTIRPIEEVVATLDAQQHRRFIKTHPLD---GLVLDDRVSYICVGRDPRDAAVSMLVQ	128
hST	+ + + +L Q R + +H P+ + I +GR+PRD VS+ +	159
Rv1373	SANMNEDRMRIHEAVVPFHERIAPPFAELGHARSPTTEFRDWMEGPNQPPPGIGFTHLK	188
hST	S + + + +IA + G +P + RD++G Q G F H+K	194
Rv1373	GIGTLANILHQLGTVVVR-RHLPNVALFHYADYQADLAGELLRPARVLGIAATRDRARDL	247
hST	G + + + +W+R + N Y + Q DL G + R LG ++ +	240
Rv1373	AQYATLDAMR---SRASEIAPNTTDGIWHSDERFFRRGGSGDWQQFFTEAE	295
hST	++T AM+ S + +P+ D H F R+G GDW+ FT A+	289
	<u>Region IV</u>	
Rv1373	HLRYYYHRINQLAPPDLLAWAHEGRRGYDPAN	326
hST	SEAFDRAYRKQMRGMPTPPWDEDPEEDGSPDPEPSPEPEPKSPLEPNTSLEREPRNSSP	365
	SPSPGQASETPHPRPS	365

Fig. 3. Comparison of *M. tuberculosis* gene *Rv1373* with human hydroxysteroid ST gene *SULT2B1b* (U92315) showing 72/292 (24%) identities, 118/292 (39%) positive (+) residues, and 58/292 (19%) gaps (-). ST signature sequences, regions I and IV are shown underlined. Results were obtained by gapped BLAST analysis (Altschul *et al.*, 1990).

Fig. 4. Expression of gene *Rv1373* as a recombinant His-fusion protein in *E. coli*. Coomassie blue stained SDS-PAGE of cell lysates from STpET clone before and after induction with 0.8 mM IPTG and protein fractions isolated by nickel affinity chromatography are shown. The purified 39 kDa recombinant enzyme was eluted in 300 mM imidazole. Molecular weight markers, including carbonic anhydrase (36.9 kDa), are shown.

is 37 kDa, the size for the His-fusion protein expressed by construct STpET is 39 kDa and 42.5 kDa by STpTRcHis. Both recombinant proteins were expressed in *E. coli* BL21 cells after 3 h induction with 0.8 mM IPTG and isolated by nickel affinity chromatography. As shown in Fig. 4, induction of STpET transformants

results in the expression of a 39 kDa protein. The his-tagged protein was separated by nickel column chromatography and eluted with 300 mM imidazole (Fig. 4). Similar results were obtained for the purification of the 42.5 kDa his-tagged protein from STpTRcHis and OF-STpTRcHis transformants (data not shown).

The *M. tuberculosis* aryl-ST gene encodes a glycolipid ST

The purified 39 kDa and 42.5 kDa fusion proteins were first tested for PAPS ST activity towards known substrates of eukaryotic glycolipid STs such as sulfatide (3'-sulfate galactosylceramide), type I galactosylceramide, type II galactosylceramide and glucosylceramide (cerebroside). Assays contained the recombinant protein, PAP³⁵S, the glycolipid acceptor, and the reaction products were separated by reverse-phase chromatography. The unbound hydrophilic molecules were eluted in water and the bound hydrophobic lipids eluted in methanol. The radiolabelled SLs eluted were determined by scintillation counting. The recombinant *M. tuberculosis* Aryl-ST (STpET) shows ST activity towards the eukaryotic glycolipids (Fig. 5a). The highest peak of labelled product was obtained with glucosylceramide, followed by type I and type II galactosylceramides. The efficiency in the transfer of ³⁵S from PAP³⁵S to these substrates was 3.5 %, 2.4 % and 1.5 %, respectively. As expected, little activity was obtained with already sulfated galactosylceramide (sulfatide), which showed less than 0.1 % efficacy. Since assays with the 42.5 kDa (His-tag at N terminus) purified recombinant protein from STpTRcHis transformants showed similar results (data not shown) to those of the 39 kDa STpET transformants, all subsequent analyses were done with the 39 kDa (His at C terminus) protein. No activity was obtained with *E. coli* cells transformed with plasmid

Fig. 5. *M. tuberculosis* gene *Rv1373* encodes a glycolipid ST. (a) Separation by reverse-phase chromatography of ST products obtained in PAP³⁵S assays with the purified 39 kDa protein and the glycolipid substrates sulfatide (□), galactosylceramide type I (◆) and type II (○), and glucosylceramide (▲). (b) TLC separation and staining with orcinol of type I GalCer, type II GalCer and GluCer (left panel). The right panel is a TLC autoradiogram of the ³⁵S-labelled glycolipid products obtained in PAP³⁵S assays. Sulfation of glycolipids results in increased polarity and slower migration on silica plates.

without insert and the out-of-frame gene construct OF-STpTRcHis, (data not shown).

The products of ST reactions were then analysed by TLC and autoradiography as shown in Fig. 5(b). The left panel shows orcinol staining of glycolipid substrates GalCer Type I, GalCer Type II and GluCer before incubation with the enzyme. The right panel shows the autoradiogram of the ³⁵S-labelled reaction products (6000 c.p.m. per sample) after incubation with the recombinant *M. tuberculosis* aryl-ST. As expected, sulfation of the various glycolipids results in increased polarity and slower migration in the TLC. These data demonstrate that the recombinant *M. tuberculosis* aryl-ST protein encodes a functional ST capable of acting on several glycolipid acceptors.

To confirm that the sulfate residue was added to the saccharide moiety of the glycolipid acceptors, the ³⁵S-labelled glycolipids (10000–20000 c.p.m.) were treated with ceramidase and the digestion products separated by reverse-phase chromatography. Treatment of ³⁵S-galactosylceramide products (GalCer Type I and II) with ceramidase resulted in the release of ³⁵S-labelled material (~ 20–30% of total radioactivity) that eluted with the aqueous phase (unbound) instead of the lipid phase (bound) (Fig. 6a). Analysis of the aqueous (unbound) phase and the lipid (bound) phases by TLC and autoradiography revealed that the ceramidase-

digested aqueous material migrated close to the origin as free ³⁵S-galactose, while the lipid phase consisted of undigested ³⁵S-GalCer Type I and II products (Fig. 6b). A second digestion of the ceramidase-resistant material resulted in the release of more ³⁵S-galactose (20–30%) in the aqueous phase, suggesting that the ceramidase utilized has relatively low activity toward mono-saccharide glycosphingolipids such as galactosyl- and glucosylceramide. Nevertheless, these results demonstrate that the activity of the *M. tuberculosis* ST is towards the sugar moiety (galactose or glucose) of typical eukaryotic glycolipid substrates.

We then tested the *M. tuberculosis* ST for activity towards endogenous mycobacterial glycolipids. Assays were performed with total glycolipids from H37Ra bacilli prepared by Folch extraction (50–100 µg) and the purified recombinant *M. tuberculosis* ST (100 µg). H37Ra glycolipids were used because they are less sulfated than those from H37Rv and Erdman strains (as shown in Fig. 1). Separation of the reaction products by reverse-phase chromatography yielded a peak of ³⁵S-labelled lipid corresponding to approximately 30% (20000 c.p.m.) of the total radioactivity applied. Fig. 7 shows the ³⁵S-labelled reaction products isolated by reverse-phase chromatography and analysed by TLC and autoradiography. As shown, one major band (indicated by an arrow) of ³⁵S-labelled glycolipid was

Fig. 6. Ceramide glycanase digestion of sulfated products. (a) Separation by reverse-phase chromatography of ^{35}S -GalCer type I (Δ) and type II ($\circ$) sulfated products after digestion with ceramide glycanase. (b) Autoradiogram of TLC plate containing the aqueous (unbound) and lipid (bound) phases obtained after digestion of GalCer type I and II. The radiolabelled material in the aqueous phase migrates near the origin and consists of free ^{35}S -galactose, while material in the lipid phase consists of undigested ^{35}S -labelled glycolipids.

Fig. 7. TLC autoradiogram of the SL products obtained in ST assays with the recombinant *M. tuberculosis* aryl-ST and total H37Ra glycolipids (left lane). A major band with a relative mobility similar to SL-IV (arrow) and minor bands corresponding to SLs III, V and VI are shown. The left lane shows the typical profile of total ^{35}S -labelled lipids obtained in H37Rv bacilli after ^{35}S metabolic labelling.

obtained with a relative mobility similar to SL-IV, a SL typically seen in *M. tuberculosis* strains after metabolic labelling with ^{35}S . Other minor bands possibly corresponding to glycolipids III, V and VI were also detected. For comparison, ^{35}S -labelled total lipids from H37Rv bacilli are shown in the left lane.

DISCUSSION

Sulfated glycolipids of *M. tuberculosis* consist of trehalose-2-sulfate esters that are associated with the outermost layer of the cell wall. They are considered among the most characteristic components of virulent mycobacteria (Goren, 1984). Several studies established a correlation between the presence of neutral-red-reactive SLs and the degree of virulence in guinea pigs (Goren *et al.*, 1974, 1982). We demonstrated by metabolic labelling with ^{35}S that *M. tuberculosis* strains synthesize abundant SLs (4–6 bands by TLC analysis) in comparison to non-pathogenic mycobacterial species such as *M. avium*, *M. bovis* BCG and *M. smegmatis*. In addition, we found that virulent laboratory strains of *M. tuberculosis*, H37Rv and Erdman, synthesize an abundant and relatively non-polar SL (shown as SL-I) that is not present in the attenuated strain H37Ra. Although the molecular structure of these SLs is not addressed in the present study, it is presumed that these abundant SLs correspond to those previously characterized as the SL-I–IV series (Goren, 1970a, b, 1984). SL-I, 2,3,6,6'-tetraacyl trehalose-2'-sulfate, has been shown to be the most abundant SL (Goren, 1984) and a good marker to differentiate between *M. tuberculosis* strains and other mycobacterial species (Luquin *et al.*, 1992).

Since the structure of *M. tuberculosis* trehalose SLs is well characterized, but most of their biosynthetic enzymes are not known, we tested for the presence of SL STs in cell free lysates of *M. tuberculosis*. ST assays utilizing PAP ^{35}S as the donor substrate and cell lysates of H37Ra, containing enzyme and acceptor lipid substrates, showed ST activity towards four endogenous lipids. One of the sulfated products is the abundant and relatively non-polar SL-I present in virulent strains but absent in H37Ra after metabolic labelling. This suggests

that sulfation of this particular glycolipid may be regulated *in vivo* at either the enzyme or the substrate level. We also found that assays with cell lysates of H37Rv and Erdman typically showed lower activity than those with H37Ra. We attributed these differences to the availability of unsulfated glycolipid substrates in the cell lysates of H37Ra in comparison to H37Rv and Erdman. Nevertheless, the fact that soluble ST activity was identified in crude cell lysates prompted us to search for ST gene sequences in *M. tuberculosis*.

The *M. tuberculosis* gene *Rv1373* showed slight similarity (24%) to eukaryotic aryl-STs but it contains the highly conserved signature sequences regions I and IV, which are involved in PAPS binding and transfer of sulfate (Weinshilboum *et al.*, 1997). Gene *Rv1373* showed no homology to a *Klebsiella* aryl-ST (Baek *et al.*, 1996) and is the only sequence in the *M. tuberculosis* genome with similarity to eukaryotic aryl-STs (Cole *et al.*, 1998). Other putative ST genes in *M. tuberculosis*, such as genes *Rv2392* (*cysH*), *Rv3117–3118* (*cysA3–sseC2*), and *Rv0815c–0814c* (*cysA2–sseC*), show similarity to prokaryotic cysteine thio-STs (Cole *et al.*, 1998) and are possibly involved in the assimilatory pathway of cysteine biosynthesis (unpublished data). Interestingly, a BLAST analysis of *Rv1373* against the *M. bovis* virulent strain AF2122/97 sequence database from the Sanger Centre (www.sanger.ac.uk) showed a gene homologue in Contig 281 with 99% identity. However, this sequence shows nucleotide changes such as an extra C at position 455 and a C to T change (Pro to Leu) at position 692. The extra nucleotide change would result in a frameshift at amino acid 155, a premature termination after residue 264 and truncation of 31 aa at the carboxyl terminus of the protein. This truncation would possibly result in inactivation of the enzyme since this region contains the highly conserved signature sequence region IV. Since we showed that BCG is almost devoid of sulfated glycolipids, future studies are designed to determine if the lack of SLs is due to a mutated *Rv1373* gene homologue.

Gene *Rv1373* was cloned by PCR, expressed in *E. coli* as a recombinant histidine-fusion protein and isolated by affinity chromatography. The purified fusion protein was first tested for PAP³⁵S ST activity towards known substrates of eukaryotic glycolipid STs. The eukaryotic ceramide glycolipids were used since they are simple glycolipids that are commercially available in sulfated or desulfated form. We demonstrated that the recombinant *M. tuberculosis* aryl-ST protein was able to transfer sulfate to glucosylceramide and galactosylceramide (Type I and II) but not sulfatide (3'-sulfate GalCer). The inability to sulfate sulfatide suggests that sulfation of galactosylceramide may occur at carbon 3 of the galactose. In addition, treatment of sulfated products with ceramide glycanase showed that sulfation was associated with the sugar moiety (galactose or glucose) of the glycolipid. To further determine the function of the recombinant *M. tuberculosis* aryl-ST, assays were performed with total glycolipids extracted from H37Ra bacilli. The *M. tuberculosis* aryl-ST was capable of

efficiently transferring sulfate to several glycolipids in a crude mixture. Comparison of the mobility of sulfated products in the TLC to that of typical SLs in H37Rv, obtained after metabolic labelling with ³⁵S, showed that the main product corresponds to SL-IV, while the minor bands correspond to SLs III, V and VI. Interestingly, the most abundant SL (SL-I) was not produced. We showed that SL-I is not found in H37Ra after metabolic labelling (Fig. 1), but can be produced in PAPS assays with cell free lysates of H37Ra (Fig. 2). This suggests that synthesis of SL-I requires other STs or factors only present in the cell lysates. Also, it is possible that the acceptor substrate required for SL-I synthesis is present in limited quantities in H37Ra in comparison to the virulent strains H37Rv and Erdman. Taken together, these studies show that gene *Rv1373* encodes a novel glycolipid ST in *M. tuberculosis* with broad specificity and that could be involved in the biosynthesis of trehalose SLs in *M. tuberculosis*.

In mammals, cytosolic aryl-STs, such as the hydroxysteroid, phenol, monoamine and oestrogen-sulfating types, function in the detoxification of xenobiotic molecules and the metabolism of steroids and bile acids (Falany, 1997). The other class of STs are membrane-bound Golgi STs that are responsible for sulfation of glycosaminoglycans (GAGs), glycoproteins and glycolipids (Falany, 1997). The *M. tuberculosis* ST is a novel enzyme since it is cytosolic, shows similarity to eukaryotic cytosolic aryl-STs but not to glycolipid STs and is active towards cerebroside glycolipids. There are few examples of prokaryotic glycolipid STs. The soil proteobacterium *Rhizobium* sp. secretes specific lipo-chitoooligosaccharide signals called Nod factors that are required for infection and nodulation of legumes (Varin *et al.*, 1997). Sulfation of these glycolipids is catalysed by STs encoded by the *nodH* (Bourdineaud *et al.*, 1995; Ehrhardt *et al.*, 1995; Schultze *et al.*, 1995) and *noeE* (Hanin *et al.*, 1997) genes, which are also required for host specificity and biological activity. The NodH ST has specificity toward terminal N-acetylglucosamine residues of lipochitoooligosaccharides (Roche *et al.*, 1991), while the Node ST is fucose specific (Hanin *et al.*, 1997; Quesada-Vincens *et al.*, 1998). In *M. tuberculosis*, sulfation of cell wall glycolipids has been implicated in evasion of bacilli from the phagolysosomal compartment of the macrophage (Goren, 1977; Goren *et al.*, 1982, 1987). However, because of the lack of knowledge about the genes and biosynthetic pathways involved, their function in virulence and pathogenicity has not been defined. Recently, a gene locus encoding polyketide synthase genes (*pks2*) was shown to encode the synthase for hepta- and octamethyl-branched fatty acids in SLs. A gene knockout of *pks2* resulted in mutants of H37Rv lacking SLs (Sirakova *et al.*, 2001). This finding, and the present work showing the first glycolipid ST in *M. tuberculosis* are important steps towards defining the biological role of SLs. Other aspects of their biosynthesis such as the synthesis of trehalose, the synthesis of sulfate donor substrates (PAPS), the transfer of fatty acids to trehalose, and their translocation and assembly into the

cell wall, are equally important. Further studies are under way to fully characterize the *M. tuberculosis* glycolipid ST, determine the molecular structure of its endogenous substrates, generate gene knockout mutants, and define its role in the biology of *M. tuberculosis*.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health Grant 1 RO1 AI37937 (to J.R.), and a minority supplement and Veterans Affairs Administration Type II Merit Review Grant (to C.R.-M.). We thank Brian Mackin, Mark Burroughs, Nirupama Masse and Hilda Rivera for technical support at various stages of this project.

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Baek, M. C., Kim, S. K., Kim, D. H., Kim, B. K. & Choi, E. C. (1996). Cloning and sequencing of the *Klebsiella* K-36 *astA* gene, encoding an arylsulfate sulfotransferase. *Microbiol Immunol* **40**, 531–537.
- Bloom, B. R. & Murray, C. J. L. (1992). Commentary on a reemergent killer. *Science* **257**, 1055–1064.
- Bourdineaud, J. P., Bono, J. J., Ranjeva, R. & Cullimore, J. V. (1995). Enzymatic radiolabelling to a high specific activity of legume lipo-oligosaccharidic nodulation factors from *Rhizobium meliloti*. *Biochem J* **306**, 259–264.
- Brennan, P. J. & Nikaido, H. (1995). The envelope of mycobacteria. *Annu Rev Biochem* **64**, 29–63.
- Brosna, J. P., Horan, M., Rademacher, J. M., Pabst, K. M. & Pabst, M. J. (1991). Monocyte responses to sulfatide from *Mycobacterium tuberculosis* inhibition of priming for enhanced release of superoxide, associated with increased secretion of interleukin-1 and tumour necrosis factor alpha, and altered protein phosphorylation. *Infect Immun* **59**, 2542–2548.
- Cole, S. T., Brosch, R., Parkhill, J. & 39 other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.
- D'Arcy Hart, P. D. & Young, M. R. (1988). Polyanionic agents inhibit phagosome-lysosome fusion in cultured macrophages. *J Leukoc Biol* **43**, 179–182.
- Dubos, R. J. & Middlebrook, G. (1948). Cytochemical reaction of virulent tubercle bacilli. *Am Rev Tuberc Pulm Dis* **58**, 698–699.
- Dye, C., Scheele, S., Dolin, P., Pathania, V. & Raviglione, M. C. (1999). Consensus statement: global burden of tuberculosis: estimated incidence, prevalence, and mortality by country: WHO global surveillance and monitoring project. *JAMA (J Am Med Assoc)* **282**, 677–686.
- Ehrhardt, D. W., Atkinson, E. M., Faull, K. F., Freedberg, D. I., Sutherland, D. P., Armstrong, R. & Long, S. R. (1995). In vitro sulfotransferase activity of NodH, a nodulation protein of *Rhizobium meliloti* required for host-specific nodulation. *J Bacteriol* **177**, 6237–6245.
- Falany, C. N. (1997). Sulfation and sulfotransferases 3: enzymology of human cytosolic sulfotransferases. *FASEB J* **11**, 206–216.
- Fujiwara, N. (1997). Distribution of antigenic glycolipids among *Mycobacterium tuberculosis* strains and their contribution to virulence. *Kekkaku* **72**, 193–205.
- Gangadharam, P. R. J., Cohn, M. L. & Middlebrook, G. (1963). Infectivity, pathogenicity and sulpholipid fraction of some Indian and British strains of tubercle bacilli. *Tubercle* **44**, 452–455.
- Goren, M. B. (1970a). Sulfolipid I of *Mycobacterium tuberculosis*, strain H37Rv. I. Purification and properties. *Biochim Biophys Acta* **120**, 116–126.
- Goren, M. B. (1970b). Sulfolipid I of *Mycobacterium tuberculosis*, strain H37Rv. II. Structural studies. *Biochim Biophys Acta* **120**, 127–138.
- Goren, M. B. (1977). Phagocyte lysosomes: interactions with infectious agents, phagosomes and experimental perturbations in function. *Annu Rev Microbiol* **31**, 507–533.
- Goren, M. B. (1984). Biosynthesis and structures of phospholipids and sulfatides. In *The Mycobacteria: a Sourcebook*, Part A. Edited by G. P. Kubica & L. G. Wayne. New York: Marcel Dekker.
- Goren, M. B., Brokl, O. & Schaefer, W. B. (1974). Lipids of putative relevance to virulence in *Mycobacterium tuberculosis*: correlation of virulence with elaboration of sulfatides and strongly acidic lipids. *Infect Immun* **9**, 142–149.
- Goren, M. B., D'Arcy Hart, P. D., Young, M. R. & Armstrong, J. A. (1976). Prevention of phagosome-lysosome fusion in cultured macrophages by sulfatides of *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* **73**, 2510–2514.
- Goren, M. B., Grange, J. M., Aber, V. R., Allen, B. W. & Mitchison, D. A. (1982). Role of lipid content and hydrogen peroxide susceptibility in determining the guinea-pig virulence of *Mycobacterium tuberculosis*. *Br J Exp Pathol* **63**, 693–700.
- Goren, M. B., Vatter, A. E. & Fiscus, J. (1987). Polyanionic agents do not inhibit phagosome-lysosome fusion in cultured macrophages. *J Leukoc Biol* **41**, 122–129.
- Hanin, M., Jabbouri, S., Quesada-Vincens, D., Freiberg, C., Perret, X., Prome, J. C., Broughton, W. J. & Fellay, R. (1997). Sulphation of *Rhizobium* sp. NGR234 Nod factors is dependent on *noeE*, a new host-specificity gene. *Mol Microbiol* **24**, 1119–1129.
- Luquin, M., Papa, F. & David, H. L. (1992). Identification of sulpholipid I by thin-layer chromatography in the rapid identification of *Mycobacterium tuberculosis*. *Res Microbiol* **143**, 225–227.
- Middlebrook, G., Coleman, C. & Schaefer, W. B. (1959). Sulfolipid from virulent tubercle bacilli. *Proc Natl Acad Sci U S A* **45**, 1801–1804.
- Mitchison, D. A. (1964). The virulence of tubercle bacilli from patients with pulmonary tuberculosis in India and other countries. *Bull Int Union Tuberc* **35**, 287–306.
- Morrison, I. M. (1994). Glycolipids. In *Carbohydrate Analysis: a Practical Approach*, pp. 295–317. Edited by M. F. Chaplin & J. F. Kennedy. New York: Oxford University Press.
- Pabst, M. J., Gross, J. M., Brosna, J. P. & Goren, M. B. (1988). Inhibition of macrophage priming by sulfatide from *Mycobacterium tuberculosis*. *J Immunol* **140**, 634–640.
- Philipp, W. J., Poulet, S., Eiglmeier, K. & 7 other authors (1996). An integrated map of the genome of tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. *Proc Natl Acad Sci USA* **93**, 3132–3137.
- Quesada-Vincens, D., Hanin, M., Broughton, W. J. & Jabbouri, S. (1998). In vitro sulfotransferase activity of NoeE, a nodulation protein of *Rhizobium* sp. NGR234. *Mol Plant-Microbe Interact* **11**, 592–600.
- Riley, L. W. (1996). Phagocytosis of *M. tuberculosis*. In *Tuberculosis*, pp. 281–289. Edited by W. N. Room & S. Gray. Boston, MA: Little & Brown.

- Roche, P., Debelle, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Denarie, J. & Prome, J. C. (1991).** Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* **67**, 1131–1143.
- Rost, B. (1996).** PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol* **266**, 525–539.
- Schultze, M., Staehelin, C., Rohrig, H., John, M., Schmidt, J., Kondorosi, E., Schell, J. & Kondorosi, A. (1995).** *In vitro* sulfotransferase activity of *Rhizobium meliloti* NodH protein: lipochitooligosaccharide nodulation signals are sulfated after synthesis of the core structure. *Proc Natl Acad Sci USA* **92**, 2706–2709.
- Sirakova, T. D., Thirumala, A. K., Dubey, V. S., Sprecher, H. & Kolattukudy, P. E. (2001).** The *Mycobacterium tuberculosis* *pks2* gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. *J Biol Chem* **276**, 16833–16839.
- Varin, L., Marsolais, F., Richard, M. & Rouleau, M. (1997).** Sulfation and sulfotransferases. 6. Biochemistry and molecular biology of plant sulfotransferases. *FASEB J* **11**, 517–525.
- Weinshilboum, R. M., Otterness, D. M., Aksoy, I. A., Wood, T. C., Her, C. & Raftogianis, R. B. (1997).** Sulfotransferase molecular biology: cDNAs and genes. *FASEB J* **11**, 3–14.
- Zhang, L., Goren, M. B., Holzer, T. J. & Andersen, B. R. (1988).** Effect of *Mycobacterium tuberculosis*-derived sulfolipid I on human phagocytic cells. *Infect Immun* **56**, 2876–2883.
- Zhang, L., English, D. & Andersen, B. B. (1991).** Activation of human neutrophils by *Mycobacterium tuberculosis*-derived sulfolipid-I. *J Immunol* **146**, 2730–2736.
-
- Received 18 September 2001; accepted 1 November 2001.