Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the Mycobacterium tuberculosis gene encoding α-D-glucose-1-phosphate thymidylyltransferase

Yufang Ma, Jonathan A. Mills,† John T. Belisle, Vara Vissa, Mark Howell, Kelly Bowlin, Michael S. Scherman and Michael McNeil

Author for correspondence: Michael McNeil. Tel: +1 970 491 1784. Fax: +1 970 491 1815. e-mail: mmcneil@vines.colostate.edu

INTRODUCTION

The mycobacterial cell wall consists of an outer lipid layer of mycolic acids connected, via arabinogalactan polysaccharide, to an inner peptidoglycan layer. An α-L-rhamnopyranosyl residue has been shown to be a key component linking the mycolated arabinogalactan to the peptidoglycan and, therefore, the biosynthesis of L-rhamnose (Rha) in mycobacteria was investigated as the first step of developing inhibitors of its biosynthesis. Biochemical assays were used to show that dTDP-Rha was synthesized in Mycobacterium smegmatis from α-D-glucose 1-phosphate (α-D-Glc-1-P) and dTTP by the same four enzymic steps used by Escherichia coli and other bacteria. PCR primers based on consensus regions of known sequences of the first enzyme in this series, α-D-Glc-1-P thymidylyltransferase (RfbA) were used to amplify rfbA DNA from M. tuberculosis. The entire rfbA gene was then cloned and sequenced. The deduced amino acid sequence revealed a 31 362 Da putative protein product which showed similarity to RfbA proteins of other bacteria (59% identity to that found in E. coli). Sequencing of DNA flanking the rfbA gene did not reveal any of the other rfb genes required for dTDP-Rha biosynthesis. Therefore, the four Rha biosynthetic genes are not clustered in M. tuberculosis. The enzymic activity of the sequenced gene product was confirmed by transformation of E. coli with pBluescript KS(−) containing the rfbA gene from M. tuberculosis. Analysis of enzyme extracts prepared from this transformant revealed an 11-fold increase in α-D-Glc-1-P thymidylyltransferase activity.

Keywords: rhamnose, rfb, mycobacterial cell wall, drug development, thymidylyltransferase
peptidoglycan (McNeil et al., 1990), as shown in Fig. 1. Thus, it is attractive to hypothesize that inhibition of the biosynthesis of l-Rha would be lethal to the mycobacteria by making formation of the linker unit impossible. Therefore, we have begun studies on its biosynthesis in mycobacteria.

In Gram-negative bacteria, l-Rha residues are often present in LPS and are synthesized at the sugar nucleotide level, dTDP-Rha being the nucleotide product (Shibaev, 1986). dTDP-Rha then functions as the Rha donor in all bacteria studied (Shibaev, 1986) and, indeed, in a recent report it has been shown to donate the Rha residue for the linker region in mycobacteria (Mikusova et al., 1996). The pathway for dTDP-Rha biosynthesis has been studied extensively in Gram-negative bacteria (Shibaev, 1986) and dTDP-Rha has been shown to be produced from TTP and glucose 1-phosphate (Glc-1-P) as shown in Fig. 2. The genes encoding these four biosynthetic transformations (known as rfbA, rfbB, rfbC and rfbD in Gram-negative organisms) have been cloned and sequenced in many organisms (Minnikin, 1982; Stevenson et al., 1994; Rajakumar et al., 1994). They have invariably been found in a biosynthetic operon with the genes usually in the order rfbBDAC (Stevenson et al., 1994).

Here we present evidence that mycobacteria produce dTDP-Rha by the same sequence of reactions shown in Fig. 2. We also report the cloning, sequencing and expression in Escherichia coli, of the first of the

**METHODS**

**Chemicals, biochemicals and enzymes.** All chemicals, enzymes and biochemicals were purchased from Sigma unless otherwise indicated. dTDP-[U-14C]glucose (dTDP-[14C]Glc), specific activity 257 mCi mmol⁻¹ (9.5 × 10⁶ Bq mmol⁻¹), was purchased from ICN and is unfortunately no longer available. Restriction enzymes were obtained from Boehringer Mannheim and used according to the supplier’s instructions.

**Bacterial strains and plasmids.** The bacterial strains and plasmids used in this study are described in Table 1. M. smegmatis ATCC 8348 and M. tuberculosis H37Rv were grown on glycerol-alanine-salts medium (Takayama et al., 1973). E. coli DH5α and S8874 were grown on LB medium (Life Technologies). When necessary, antibiotics were used at

---

**Fig. 1.** Structure of the mycobacterial cell wall emphasizing the position and role of the l-Rha-containing linker disaccharide.

**Fig. 2.** The biosynthetic pathway for formation of dTDP-Rha as found in Gram-negative bacteria and, as shown here, in mycobacteria. RfbA, α-β-Glc-1-P thymidylyltransferase; RfbB, dTDP-Glc 4,6-dehydratase; RfbC, dTDP-6-deoxy-4-ketoglucone epimerase; RfbD, dTDP-Rha synthase.

**Mycobacterium tuberculosis** Rha biosynthetic genes, namely the rfbA gene which encodes α-β-Glc-1-P thymidylyltransferase (RfbA). Sequence data in the vicinity of rfbA show that, in contrast to Gram-negative organisms, the rfbA gene is not clustered with the remaining Rha genes.
counted for radioactivity using liquid scintillation. A control phosphate, 1 mM; crude M.
flow rate of 1 ml min⁻¹. Seventy 1 ml fractions were then
were omitted.

4
To assay for RfbA activity,
reaction was run in the same fashion except the PP, and NaF
NH₄H₂PO₄ from 5 to 185 mM over a period of 75 min with a
final concentrations of reagents as follows

MgCl₂, 0.1 mM; MgCl₂,
0-5

0.1 mM; MgCl₂,

5

M. smegmatis

NH₄H₂PO₄, 5 mM; MgCl₂,

0.2 mM PMSF and 1 mM DTT. The

Cell-free assays using
M. smegmatis. M. smegmatis ATCC
8548 was prepared as a cell-free extract by sonicating the cells
at 4 °C with a 4710 series Cole Parmer Ultrasonic homogenizer in a buffer solution containing 50 mM HEPES, pH 7.6, 10 mM MgCl₂, 0.5 mM MnCl₂, 0.2 mM PMSF and 1 mM DTT. The cell homogenate was then centrifuged at 12000 g for 10 min
and the protein concentration of the resulting supernatant was measured using the Bio-Rad Protein Assay.

To assay for RfbA activity, 0-5 μCi (1.85 x 10⁴ Bq) dTDP-[¹⁴C]Glc was incubated in a total volume of 110 μl, with the final concentrations of reagents as follows: cold dTDP-Glc, 0.1 mM; MgCl₂, 5 mM; NaF, 5 mM; tetrasodium pyrophosphate, 1 mM; crude M. smegmatis protein extract, 0.5 mg ml⁻¹. The sample was incubated for 45 min at 30 °C and then injected onto an HPLC system consisting of a 4 x 250 mm Partisil 10-SAX anion exchange HPLC column (Whatmann). The sample was eluted with a gradient of NH₄H₂PO₄ from 5 to 185 mM over a period of 75 min with a flow rate of 1 ml min⁻¹. Seventy 1 ml fractions were then counted for radioactivity using liquid scintillation. A control reaction was run in the same fashion except the PP, and NaF were omitted.

The assay for RfbB activity was performed as described by Wang & Gabriel (1969) except for the use of 50 mM HEPES, pH 7.6, instead of Tris buffer.

The assay for the combination of RfbB, RfbC and RfbD enzymes was done in a 100 μl volume and the final concentrations of reagents were as follows: dTDP-[¹⁴C]Glc, 0.5 μCi (1.85 x 10⁴ Bq); cold dTDP-Glc, 0.5 mM; NADPH, 10 mM; NAD⁺, 5 mM; M. smegmatis extract prepared as described above, 7.5 mg protein ml⁻¹. The tube was incubated for 45 min at 30 °C and was made 50% with respect to ethanol resulting in a soluble crude sugar nucleotide fraction. After centrifugation at 12000 g for 5 min, the clear supernatant was dried, hydrolysed with 2 M trifluoroacetic acid for 1 h at 120 °C and dried. The hydrolysate was then injected onto a Dionex HPLC System with a Dionex Carbopak PA1 column (4 x 250 mm) with isocratic elution with 10 mM NaOH. Thirty 1 ml fractions were then counted for radioactivity in a liquid scintillation system.

DNA techniques. M. tuberculosis DNA was prepared using the protocol described for M. avium (Belisle et al., 1991). DNA was isolated from agarose gels with a GeneClean II kit (Bio101). Purified plasmids were prepared using a QIAGEN plasmid kit according to the manufacturer’s instructions.

PCR and sequencing of PCR products. Amino acid sequence alignment of known bacterial α-D-Glc-1-P thymidylyltransferases [Shigella flexneri (Rajakumar et al., 1994), E. coli (Stevenson et al., 1994), Yersinia enterocolitica (Zhang et al., 1993) and Streptomyces griseus (Distler et al., 1987)] revealed highly conserved sequences. Two degenerated primers were designed based on the conserved TRLYPF and AVTGLYFY stretches shown in Figs 6 and 7, and the M. tuberculosis codon usage preference (Hatfull & Sarkis, 1993). The sense primer was 5’ AC(GC)CG(CG)GT(CG)TACCC(GC)ATCAC 3’ and the antisense primer was 5’ GTAGAAGTA(GC)AG(AGCT)-CC(GC)GT(CG)AG(GC)AC(GC) GC 3’. PCR was performed in a total volume of 50 μl using a Thermal Cycler (Perkin Elmer). The reaction mixture contained 1:25 units AmpliTaq DNA polymerase (Perkin Elmer), buffer (10 mM Tris/HCl, pH 8.3, 50 mM KCl, 2.5 M MgCl₂), 0.2 mM each deoxynucleotide triphosphate (Gibco BRL), 10 pmol each primer and 0.4–10 μg M. tuberculosis H37Rv genomic DNA. The PCR conditions consisted of 1 cycle of denaturation (94 °C, 5 min) followed by 35 cycles of amplification consisting of denaturation (94 °C, 1 min), annealing (65 °C, 1 min) and primer extension (72 °C, 1 min). After the final cycle there was a 7 min incubation at 72 °C. A PCR product of correct size (495 bp) was cloned into the pBlueT7 T vector (Stratagene) resulting in plasmid pYM1.

Table 1. Properties and sources of relevant bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Species/strain</th>
<th>Relevant properties</th>
<th>Source*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacteria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. smegmatis 8548</td>
<td>Fast growing; non-pathogenic</td>
<td>ATCC</td>
</tr>
<tr>
<td>M. tuberculosis H37Rv</td>
<td></td>
<td>ATCC</td>
</tr>
<tr>
<td>E. coli DH5α</td>
<td>F+ deor recA hsdR17 (r¢,m¢) sup44¢ thi-1 gyrA96 recA1</td>
<td>Gibco BRL</td>
</tr>
<tr>
<td>E. coli S874</td>
<td>Δ(sbcB–rfl)</td>
<td>CGSC (Bachmann, 1987)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBlueT7 T</td>
<td>PCR product cloning vector</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pBluescript KS(−)</td>
<td>Sequencing and expression vector</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pGCl</td>
<td>Vector for M. tuberculosis cosmid library</td>
<td>This study</td>
</tr>
<tr>
<td>pYM1</td>
<td>M. tuberculosis H37Rv 495 bp rfbA PCR product in pT7Blue T vector</td>
<td>This study</td>
</tr>
<tr>
<td>pYM1c</td>
<td>30–40 kb of M. tuberculosis DNA containing rfbA in pGCl1 vector</td>
<td>This study</td>
</tr>
<tr>
<td>pYM2</td>
<td>5'-7 kb of rfbA containing Xbol fragment from pYM1c in pBluescript KS(−)</td>
<td>This study</td>
</tr>
<tr>
<td>pYM3</td>
<td>80 kb of rfbA containing EcoRV fragment from pYM1c in pBluescript KS(−)</td>
<td>This study</td>
</tr>
<tr>
<td>pYM4</td>
<td>5'-2 kb of rfbA containing PstI–EcoRI fragment from pYM1c in pBluescript KS(−)</td>
<td>This study</td>
</tr>
<tr>
<td>pYM2d1–d9</td>
<td>Deletion plasmids from pYM2</td>
<td>This study</td>
</tr>
</tbody>
</table>

* ATCC, American Type Culture Collection, Rockville, MD, USA; CGSC, E. coli Genetic Stock Center, Yale University, New Haven, Connecticut, USA.
The cloned PCR product was sequenced using the Sequenase Version 2.0 DNA Sequencing Kit (USB). The BLAST algorithm was used to search protein databases for similarity (Altschul et al., 1990).

**DNA probe preparation and DNA hybridization.** A 546 bp XbaI–EcoRI restriction fragment (containing the entire rfbA PCR product and a small amount of DNA on each end from the pBlueT7 T vector) from pYM1 was prepared and a digoxigenin non-radioactive probe made from it using a DNA labelling kit, Genius 2 (Boehringer Mannheim). For Southern blotting, 1 µg M. tuberculosis H37Rv genomic DNA was digested with 10 units BamHI, CiaI, EcoRI, EcoRV, PstI and Xhol, respectively, and separated by electrophoresis through a 0.7% agarose gel. The gel was treated for 10 min with 0.25 M HCl and then for 30 min with 0.6 M NaCl/0.4 M NaOH. DNA was transferred to a Maximum Strength Nytran Plus membrane (Schleicher & Schuell) and immobilized by UV cross-linking. The membrane was pre-hybridized at 42 °C for 2 h in pre-hybridization buffer containing 5 × SSC (1 × SSC is 0.15 M sodium chloride, 0.015 M sodium citrate), 0.1% N-lauroylsarcosine, 0.02% SDS, 1% blocking reagent (Boehringer Mannheim) and 50% deionized formamide. Following this treatment, the digoxigenin-labelled 546 bp XbaI–EcoRI rfbA probe was denatured, added to the membrane and incubated at 42 °C overnight. The membrane was washed three times at room temperature with 2 × SSC/0.1% SDS and then twice at 65 °C with 0.5 × SSC/0.1% SDS. Colorimetric detection was carried out using a nucleic acid detection kit, Genius 3 (Boehringer Mannheim).

**Construction and screening of a cosmid library.** A cosmid library of M. tuberculosis genomic DNA was constructed in pGCl, an E. coli–Mycobacterium shuttle cosmid. For this study, pGCl-based cosmid clones were used merely as a M. tuberculosis H37Rv DNA library source and the shuttle properties of the cosmid were not utilized. Vector pGCl was constructed by ligation of a 2044 bp BamHI–PstI fragment containing locS from pYUB18 (Belisle et al., 1991) with pEP2 (Radford & Hodgson, 1991) digested with BamHI–PstI. The M. tuberculosis cosmid library was constructed with 30–40 kb fragments generated by a partial Ssna3A digestion of M. tuberculosis DNA followed by ligation to BamHI-digested pGCl. The ligation mixture was packaged in vitro with Gigapack II XL Packaging Extracts (Stratagene) and transduced into DH5α cells ( Gibco BRL). To locate cosmids containing rfbA, E. coli recombinant colonies were lifted onto Optitran membranes (Schleicher & Schuell), lysed and then hybridized with the 546 bp XbaI–EcoRI probe (Maniatis et al., 1982).

Southern analysis of five positive cosmid clones digested with a variety of restriction enzymes and probed with the 546 bp XbaI–EcoRI rfbA DNA probe yielded blots identical to those prepared with genomic DNA. Therefore, one of the five cosmid clones, pYM1, was used to subclone smaller fragments containing rfbA. Thus, 5.7 kb Xhol, 8 kb EcoRV and 5.2 kb EcoRI–PstI fragments were prepared from pYM1 and individually subcloned into pBlueScript II KS(−) (Stratagene) to produce pYM2, pYM3 and pYM4, respectively.

**DNA sequencing and analysis.** Sequences of double-stranded plasmid DNA were determined using the Taq FS Ready Reaction Kit (Applied Biosystems) on a GeneAmp PCR system 9600 (Perkin Elmer) and run on an ABI 377 PRISM Sequencer (Applied Biosystems) at the Macromolecular Resource Facility at Colorado State University, USA. To sequence rfbA, plasmids pYM2, pYM3, pYM4 and pYM1 were all used as sources of template DNA. Initial sequencing was done using the PCR primers as primers. Additional primers were then synthesized as needed using newly obtained sequence data. Both strands of the rfbA DNA were fully sequenced and these two sequences were in complete agreement. Comparisons of sequences with the database were done using the BLAST program (Altschul et al., 1990).

To sequence the 5′ and 3′ flanking regions of rfbA, the nested deletion sequencing method was used. Thus, the pYM2 plasmid was digested with SacI and Xhol. The resulting linear DNA was deleted using the Erase-a-Base System (Promega). The fragments resulting from nine different digestion times were re-circularized to form plasmids pYM2d1–9. These cosmids were sequenced as above using a primer based on the M13(−20) site present on the pBlueScript II KS(−) vector. pYM2d1 was deleted such that sequencing commenced approximately 2 kb downstream from the 3′ end of rfbA, pYM2d9 was situated so that sequencing began approximately 2 kb upstream from the 5′ end of rfbA and the sequencing commenced on the remaining plasmids in between these two extremes. For these experiments only a single strand of DNA was sequenced and the sequences were then subjected to BLAST searches. The sequence of rfbA itself was found in pYM2d5 and pYM2d6 but no sequence corresponding to any other rfb genes was obtained.

**Expression of α-α-Glc-1-P thymidylyltransferase in E. coli and measurement of its activity.** E. coli S874 was transformed with pYM2 and with pBlueScript II KS(−) only, as a control. Single colonies of each transformant were grown in 10 ml LB broth with ampicillin. The resulting cultures were transferred into 1 litre of the same medium and grown to an OD₆₆₀ of 0.6 after which IPTG was added to a final concentration of 1 mM. Cells were harvested after an additional 3 h of incubation. One half gram of each of the transformed cell preparations was resuspended in 1 ml breaking buffer (0.1 mM PMSF, 1 mM DTT, 50 mM HEPES, pH 7.4), sonicated with a 4710 series Cole Palmer ultrasonics homogenizer and centrifuged for 20 min at 16000 g. The protein concentration of the supernatants containing the crude enzyme was determined using the Bio-Rad Protein Assay.

The crude enzymes prepared from the two E. coli strains were then assayed for α-α-Glc-1-P thymidylyltransferase activity. The assay was based on the formation of Glc-1-P as described for the M. smegmatis enzyme; however, the Glc-1-P was detected by a linked enzyme assay which produced NADPH rather than by HPLC. The enzyme assays were done in a volume of 1 ml containing 184 mM HEPES, pH 7.4, 4 mM tetrathiocarbonyl pyrophosphate, 5 mM MgCl₂, 2 mM NADP, 1 mM TDP–Glc (or 1 mM UDP–Glc to determine α-α-Glc-1-P uridylyltransferase activity), 0.02 mM Glc 1,6-diphosphate, 0.2 units phosphoglucomutase, 1 unit Glc-6-P dehydrogenase and 148 µg crude enzyme. The rate of formation of NADPH was determined spectrophotometrically at 340 nm. A parallel reaction was done in the absence of PP₃.

**RESULTS**

**Demonstration of RfbA, B, C and D activities in mycobacteria**

We have previously shown that the cell wall core of M. tuberculosis and M. smegmatis contain the same linker disaccharide (Fig. 1) and that the arabinosyl and galactosyl residues of their respective AGs are arranged identically as determined by ¹H and ¹³C NMR analyses (Daffe et al., 1993). Therefore, it is highly likely that the
Ketoglucose, yields an UV-active compound after treatment with base. Incubation of varying amounts of crude enzyme prepared from *M. smegmatis* with dTDP-Glc followed by base treatment gave $A_{318}$ values that were fourfold greater than those of controls with no dTDP-Glc in the incubation. At a protein concentration of 70 μg ml$^{-1}$ the *M. smegmatis* enzymes converted 97 nmol dTDP-Glc to dTDP-6-deoxy-4-ketoglucose min$^{-1}$ (mg protein)$^{-1}$.

**Fig. 3.** Inorganic PP$_2$-dependent conversion of dTDP-$[^{14}C]$Glc to $\alpha$-d-Glc-1-P by $\alpha$-d-Glc-1-P thymidylyltransferase present in enzyme extracts of *M. smegmatis*. HPLC profiles of reaction products from dTDP-$[^{14}C]$Glc and enzyme extract without (a) and with (b) added PP$_2$. Authentic $\alpha$-Glc-1-P eluted in 25 min under the conditions used; the heterogeneity seen around 25 min in (b) probably results from further conversion of Glc-1-P into other sugar phosphates.

RfbB activity was confirmed and evidence for the activities of dTDP-6-deoxy-4-ketoglucose epimerase (RfbC) and dTDP-Rha synthase (RfbD) obtained by the conversion of dTDP-Glc to a Rha-yielding product (presumably dTDP-Rha). Thus dTDP-$[^{14}C]$Glc was treated with a *M. smegmatis* enzyme extract, the resultant sugar nucleotides hydrolysed and the free sugars analysed by Dionex anion exchange chromatography (Fig. 4). The results demonstrated the enzymic conversion of the Glc unit to Rha with a small amount of Glc being converted to 6-deoxytalose (6-dTal) (see Discussion).

**Amplification, cloning and sequencing of rfbA DNA from *M. tuberculosis* **

The $\alpha$-d-Glc-1-P thymidylyltransferase amino acid sequences from *Sh. flexneri* (Rajakumar et al., 1994), *E. coli* (Stevenson et al., 1994), *Y. enterocolitica* (Zhang et al., 1993) and *S. griseus* (Distler et al., 1987) were aligned. PCR primers were designed based on sequence similarities, mycobacterial codon bias and melting temperatures (see Fig. 6 for primer positions). Using *M. tuberculosis* H37Rv DNA as a template, a PCR product of the correct size was amplified. The PCR product was cloned into pT7Blue T vector (see Table 1), resulting in pYM1, and a partial sequence determined. A BLAST search (Altschul et al., 1990) showed significant homology with rfbB from several species. Analyses by Southern blotting of genomic *M. tuberculosis* DNA cut with various restriction enzymes using a 546 bp probe containing rfbB sequence generated from pYM1 all yielded single bands (Fig. 5), suggesting that only a single copy of the gene is present in the *M. tuberculosis* genome.

To isolate the entire rfbA gene, the 546 bp rfbA probe was used to screen the *M. tuberculosis* pGCl cosmid library by colony blot hybridization. A cosmid, pYM1, obtained from a positive colony was subsequently analysed by Southern blot analysis with the same probe. The hybridization pattern was identical to that shown in Fig. 5 for genomic DNA as expected. Thus, three
subclones, pYM2, pYM3 and pYM4, were prepared from pYMc1 by ligation of its 5.7 kb XhoI, 8.0 kb EcoRV and 5.2 PstI-EcoRI restriction fragments, respectively, into pBluescript II KS(−) (Table 1).

The complete sequence of rfbA was then obtained using plasmids pYMcl, pYM2, pYM3 and pYM4 as templates for double-stranded automated sequencing. The sequence is presented in Fig. 6. The G + C content was found to be 61% and a putative ribosome binding site (RBS), located at the PCR primers and stop codon (*) are indicated.

Expression of rfbA

E. coli strain 39874, a strain where the entire rfb region is deleted, was chosen to express the M. tuberculosis rfbA gene. It was recognized that this strain would express low levels of \(\alpha\)-D-Glc-1-P thymidylyltransferase activity due to the presence of a functional homologue of rfbA (Marolda & Valvano, 1995), but it was expected that this background level of activity would still be considerably less than that of a wild-type E. coli. Thus, E. coli strain 39874 was transformed with pYM2 and, as a control, with pBluescript KS(−) vector only. The transformants were grown to mid-exponential phase, treated with IPTG, harvested, disrupted by sonication, centrifuged and enzyme assays for Glc-1-P thymidylyltransferase (RfbA) and Glc-1-P uridylyltransferase were conducted. As expected, the control bacteria showed some Glc-1-P thymidylyltransferase activity (Table 2). However, bacteria containing the XhoI subclone showed an approximately 11-fold elevation of Glc-1-P thymidylyltransferase activity per mg protein.

**DISCUSSION**

The results of the biochemical studies strongly suggest that dTDP-Rha is formed by the same pathway in mycobacteria as it is in other bacteria (Fig. 2). In addition, the biosynthesis of 6-dTal was also shown to occur in M. smegmatis (Fig. 4) by use of RfbA-C followed by reduction of the keto sugar (Fig. 2) by the
dTDP-rhamnose biosynthesis in mycobacteria

Fig. 7. Comparison of the deduced amino acid sequence for α-D-Glc-1-P thymidylyltransferase (RfbA) from *M. tuberculosis* with the amino acid sequence of RfbA from three Gram-negative and one Gram-positive bacteria. The shaded areas are those where the sequence from *M. tuberculosis* is identical to one or more of the other sequences. The conserved putative activator binding site (Thorson et al., 1994; Parsons & Preiss, 1978) at K-24 (*M. tuberculosis*) is underlined as is the conserved postulated α-Glc-1-P binding site (Thorson et al., 1994) at K-159 (*M. tuberculosis*).

Table 2. Cell-free assays for α-D-Glc-1-P thymidylyltransferase and α-D-Glc-1-P uridylyltransferase activities in crude extracts of *E. coli* strain Sø874 transformed with pYM2 or with the pBluescript KS(−) vector only

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Enzyme</th>
<th>Activity [nmol min⁻¹ (mg protein)⁻¹]</th>
</tr>
</thead>
<tbody>
<tr>
<td>pYM2</td>
<td>α-D-Glc-1-P thymidylyltransferase</td>
<td>20.0</td>
</tr>
<tr>
<td>pYM2</td>
<td>α-D-Glc-1-P uridylyltransferase</td>
<td>74.9</td>
</tr>
<tr>
<td>pBluescript II KS(−)</td>
<td>α-D-Glc-1-P thymidylyltransferase</td>
<td>1.8</td>
</tr>
<tr>
<td>pBluescript II KS(−)</td>
<td>α-D-Glc-1-P uridylyltransferase</td>
<td>56.9</td>
</tr>
</tbody>
</table>

dTDP-6-dTal synthase forming the 1-talo configuration (Gaugler & Gabriel, 1973). *M. smegmatis* utilizes 6-dTal as a component of glycopeptidolipids produced by *M. smegmatis* (McNeil et al., 1989). It should be noted that since *M. tuberculosis* is not known to produce any 6-dTal-containing molecules, dTDP-6-dTal synthase is unlikely to be present in *M. tuberculosis*.

The fact that the Southern analysis of various restriction fragments produced from genomic *M. tuberculosis* DNA using a fragment of *rfbA* as a probe showed only single bands (Fig. 5) suggests that *M. tuberculosis* has only one copy of *rfbA*. Although the PCR approach described here was successful in amplifying the *rfbA* gene, similar approaches to amplifying the other three *rfb* genes were unsuccessful. Also, surprisingly and disappointingly, the remaining *rfb* genes were not found to be clustered with *rfbA* as in other bacteria (Minnikin, 1982; Stevenson et al., 1994; Rajakumar et al., 1994). Indeed, although we sequenced approximately 2 kb on either side of *rfbA*, no sequences with substantial homology to those found in the database were found in these areas. Confirming these results, Dr Tae-Yoon Lee (Yeungnam University, Taegu, Korea), in the context of examining RFLP patterns of various *M. tuberculosis* strains, isolated and sequenced a region of DNA which showed high homology to *rfbB* and *rfbC* but not to...
rfbA (Tae-Yoon Lee, personal communication). Several different 30–40 kb cosmids from the M. tuberculosis pGCl cosmid library containing the DNA sequences homologous to rfbB and rfbC did not hybridize with the rfbA probe (T. Y. Lee, Y. Ma & M. McNeil, unpublished results) as expected if rfbA is not linked to rfbB/rfbC.

The deduced amino acid sequence of the RfbA protein is very similar to that of other bacteria (Fig. 7). Somewhat surprisingly it is more similar to the Gram-negative bacterial proteins (e.g. 59% identity to the E. coli protein) than to the single Gram-positive α-d-Glc-1-P thymidylyltransferase protein sequence (32% identity). However, it should be noted that the S. griseus protein is not involved in Rha biosynthesis but rather catalyses the formation of dTDP-Glc as the first step in streptose biosynthesis. The M. tuberculosis α-d-Glc-1-P thymidylyltransferase contains the highly conserved lysine at residue 24 (Fig. 7) seen in most, if not all enzymes which transfer a nucleoside phosphate to Glc-1-P (Thorson et al., 1994; Parsons & Preiss, 1978) and postulated to be part of an activator binding site. It also contains the Glu-Lys-Pro (158–160) sequence which is postulated to be part of the Glc-1-P-binding region (Thorson et al., 1994) and is fully conserved, amongst α-d-Glc-1-P nucleoside transferases (Thorson et al., 1994).

Enzymic activity from the cloned RfbA M. tuberculosis protein was readily detected (Table 2) even in the absence of some E. coli background. Although the mechanism of transcription and translation of pYM2 in E. coli was not studied, we assume that the vector lacZ promoter was responsible for transcription and the putative ribosome binding site 6 bases upstream from the rfbA start codon (Fig. 6) was used for translation. E. coli transformed with pYM2 (Table 2) also showed a slight (1-3-fold) increase in α-d-Glc-1-P uridylyltransferase activity. This is not unexpected as α-d-Glc-1-P thymidylyltransferase is known to utilize UDP-Glc as a substrate (Bernstein & Robbins, 1965; Melo & Glaser, 1965).

α-d-Glc-1-P thymidylyltransferase is a potential drug target since it is likely to be an essential protein involved in mycobacterial cell wall linker synthesis and since its enzymic product, dTDP-Glc, is not found in humans. However, care must be exercised targeting this enzyme as potential inhibitors must necessarily not inhibit the very similar α-d-Glc-1-P uridylyltransferase needed in human metabolism. It thus remains important to clone the genes for the three remaining dTDP-Rha biosynthetic proteins in mycobacteria as well as the transferase gene which utilizes dTDP-Rha to transfer the Rha residue to the three positions of the linker GlcNAc residue (McNeil et al., 1990; Mikusova et al., 1996) for the purpose of drug targeting.

ACKNOWLEDGEMENTS

This work was supported by funds provided by Public Health Service grant NIAID, NIH (AI-33706). We gratefully acknowledge the gift of Pep2 from Glenn Songer, the gift of pGCl from Gerald McDonnel, the construction of the M. tuberculosis H37Rv library by David Taylor and the assay for RfbB by Wenxin Yan. The fundamental and profound contributions of Dr Patrick Brennan to this undertaking are also gratefully acknowledged.

REFERENCES


dtDP-rhamnose biosynthesis in mycobacteria


Received 24 July 1996; revised 14 October 1996; accepted 17 October 1996.