Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-keto-rhamnose by the *rmlC* gene products of *Escherichia coli* and *Mycobacterium tuberculosis*

Richard J. Stern,¹ Tae-Yoon Lee,² Tae-Jin Lee,² Wenxin Yan,¹ Michael S. Scherman,¹ Varalakshmi D. Vissa,¹ Soo-Ki Kim,³ Barry L. Wanner³ and Michael R. McNeil¹

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

¹ Department of Microbiology, Colorado State University, Fort Collins, CO 80523, USA

² Department of Microbiology, College of Medicine, Yeungnam University, 317-1 Daemyung-5-dong, Namku, Taegu 705-305, Korea

³ Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

dTDP-rhamnose is made from glucose-1-phosphate and dTTP by four enzymes encoded by *rmlA–D*. An *Escherichia coli rmlC* mutant was constructed and a crude enzyme extract prepared from it did not produce dTDP-4-keto-rhamnose, in contrast to a crude enzyme extract prepared from a wild-type *E. coli* strain where small amounts of this intermediate were found after incubation with dTDP-glucose in the absence of NADPH. These results showed that dTDP-4-keto-rhamnose, the product of RmlC, exists as a free intermediate. Further, the *Mycobacterium tuberculosis rmlC* gene was expressed and incubation of the resulting purified *M. tuberculosis* RmlC enzyme with dTDP-4-keto-6-deoxyglucose resulted in the conversion of approximately 7% of dTDP-4-keto-6-deoxyglucose to dTDP-4-keto-rhamnose. The enzyme also allowed for the incorporation of two deuterium atoms from deuterium oxide solvent into dTDP-4-keto-glucose. Thus the *rmlC* gene encodes dTDP-4-keto-6-deoxyglucose epimerase capable of epimerizing at both C-3' and C-5'; this enzyme produces free dTDP-4-keto-rhamnose but the equilibrium of the 4-keto sugar nucleotides lies strongly on the side of the *gluco* configuration.

Keywords: rhamnose, *rmlC*, mycobacterial cell wall, drug development, dTDP-4-keto-6-deoxyglucose epimerase

INTRODUCTION

A single L-rhamnosyl residue plays a key structural role in the cell wall core of mycobacteria and inhibition of its biosynthesis is likely to lead to mycobacterial cell death (McNeil *et al.*, 1990). In addition a 'rhamnolipid' in *Pseudomonas aeruginosa* is considered to be essential for virulence (Ochsner & Reiser, 1995) and rhamnose may be important for virulence in *Streptococcus suis* (Charland *et al.*, 1998). Thus it is now apparent that inhibition of the enzymes involved in the formation of the L-rhamnosyl donor, deoxythymidine diphosphate rhamnose (dTDP-rha), is an important target for drug development in several genera of bacteria. Since L-rhamnose is a common constituent of Gram-negative O antigens, the biosynthesis of dTDP-rha was studied

intensively in the late sixties and early seventies from an enzymic standpoint (Gaugler & Gabriel, 1973; Zarkowsky *et al.*, 1970). These studies lead to the conclusion that dTDP-rha is synthesized from thymidine triphosphate (TTP) and glucose-1-phosphate via four enzymes (Fig. 1). It should be noted that the keto products shown in Fig. 1 are designated in this manuscript as follows: RmlB product as dTDP-4-keto-6-deoxyglucose rather than dTDP-6-deoxy-D-xylo-4-hexulose and the RmlC product as dTDP-4-keto-rhamnose rather than dTDP-6-deoxy-L-lyxo-4-hexulose. These names are used as they indicate the relationships between sugar intermediates more clearly.

The biochemistry shown in Fig. 1 was substantiated by genetic studies of O-antigen biosynthesis in which four dTDP-rha formation enzymes were found in the O-antigen synthetic operons in many organisms (Xiang *et al.*, 1993). These genes were originally named *rfbA–D*; recently the nomenclature has been changed to *rmlA–D*

Abbreviations: dTDP-rha, deoxythymidine diphosphate rhamnose; dTDP-glc, deoxythymidine diphosphate glucose.

Fig. 1. The formation of dTDP-rha from glucose-1-phosphate and dTTP as catalysed by the four enzymes encoded by *rmlA-D*.

(Reeves *et al.*, 1996). The genes encoding the first two enzymes in the pathway, *rmlA* (α-D-glucose-1-phosphate thymidyltransferase) and *rmlB* (dTDP-D-glucose-4',6'-dehydratase) have been cloned and expressed from several different organisms (Ma *et al.*, 1997; Lindquist *et al.*, 1993; Romana *et al.*, 1991; Marumo *et al.*, 1992) and shown to catalyse the expected reactions shown in Fig. 1. However, there are not yet any reports of the last two enzymes having been purified to homogeneity or expressed from their corresponding genes. Indeed, until recently there has been some confusion regarding the assignment of these two genes to their respective enzymes (Stevenson *et al.*, 1994). The putative product of *RmlC*, dTDP-4-keto-L-rhamnose, has never been isolated as such and therefore was thought to exist only as an enzyme-bound intermediate (Melo & Glaser, 1968; Gaugler & Gabriel, 1973). This conclusion was disputed by Wahl & Grisebach (1979) who presented evidence for small amounts of dTDP-4-keto-L-rhamnose in incubations containing epimerase but lacking reductase activity as these workers isolated small amounts of dTDP-rha after chemical reduction of the 4-keto group. However, as shown by Naundorf & Klaffke (1996) and by ourselves in initial experiments, 4-keto sugars readily epimerize under the basic conditions used by Wahl & Grisebach (1979) to reduce the 4-keto group

and thus the existence of dTDP-4-keto-L-rhamnose, enzyme-bound or not, remains to be demonstrated.

To address these issues we have prepared an *Escherichia coli* mutant where the *rmlC* gene is deleted and developed procedures to reduce the 4-keto group of dTDP-4-keto-6-deoxyglucose so that no chemical epimerization takes place. This allowed us to assay for the production of dTDP-4-keto rhamnose by crude enzyme extracts made from bacteria with and without the *rmlC* gene. In addition, we cloned and overexpressed the *Mycobacterium tuberculosis rmlC* gene and purified its product. The reaction catalysed by this enzyme was then investigated with and without additional *Rml* proteins from *E. coli* and in the presence of D₂O. Taken as a whole, the studies showed that *RmlC* catalyses an equilibrium between dTDP-4-keto-6-deoxyglucose and non-enzyme-bound dTDP-4-keto-rhamnose that lies predominantly on the side of dTDP-4-keto-6-deoxyglucose.

METHODS

Plasmids. The plasmid pCANTS₇NdeI* was from A. S. Lynch, Harvard Medical School, USA; pCP15 and pCP20 were from W. Wackernagel, University of Oldenburg, Germany; pCR2.1 was from Invitrogen; pGZ119HE was from M. Russel, Rockefeller University, USA; and pPR1474 was from P. Reeves, University of Sydney, Australia. pLD55 (Metcalf *et al.*, 1996), pSK49Δ*uidA*2 (Haldimann *et al.*, 1996), and pSLF52 and pSPORT::merR (Haldimann *et al.*, 1997) were from laboratory stocks. Others are described below. Conditional replication plasmids having the *oriR*_{R6K7} origin of replication were maintained in the moderate-copy-number (*pir*⁺) host BW23473, the high-copy-number (*pir*-116) host BW23474 or other suitable hosts (Haldimann *et al.*, 1998; Metcalf *et al.*, 1994).

Preparation of *E. coli* BW24476. The strains used in this study were all derived from *E. coli* BW24476, a derivative of *E. coli* K-12. The characteristics of BW24476 are shown in Table 1. BW24476 is a descendent of the BD792-derived strain

Table 1. Key bacterial strains and plasmids

All bacteria are derivatives of *E. coli* K-12 (see Methods).

Strain or plasmid	Relevant characteristics
Strains	
BW24476	F ⁻ λ ⁻ <i>lacI</i> ^a <i>rrnB</i> _{T14} Δ <i>lacZ</i> _{WJ16} Δ <i>phoBR</i> 580 <i>hsdR</i> 514 Δ <i>araBAD</i> _{AH33} Δ <i>rhaBAD</i> _{LD78} <i>rfb</i> -50
BW24599	As BW24476, except <i>rfb</i> ⁺ (<i>rfb</i> -50 repaired)
BW24861	As BW24599, except Δ <i>rmlC</i> :: <i>kan</i> _{SK95}
BW24970	As BW24599, except Δ <i>rmlC</i> _{SK95}
BW24972	As BW24970, except <i>attλ</i> ::pTJ09
BW24973	As BW24861, except <i>attλ</i> ::pTJ09
Plasmid	
pTJ09	Conditional replication (<i>oriR</i> _{R6K7}), integration (<i>attPλ</i>) plasmid that expresses <i>rmlC</i> behind <i>P</i> _{tac}

BW23660 (Haldimann *et al.*, 1997); it was constructed in a series of phage P1 crosses using as donors: BW12070 [$\Delta(trpR\ creABCD\ arcA\ thr)$ 104::Tn5-132] (Fisher *et al.*, 1995); BW22875 ($\Delta rhaBAD_{LD78}$) (Haldimann *et al.*, 1998); BW24447 ($\Delta hsdR514$); JW383 [$metF159(Am)\ zii-510::Tn10$] (Haldimann *et al.*, 1998); MG1655 ($rpoS^+$) (Haldimann *et al.*, 1998); and ZK1001 ($cysC95::Tn10\ rpoS::kan$) (Haldimann *et al.*, 1998); (relevant markers are shown in parentheses). The $hsdR514$, $\Delta rhaBAD_{LD78}$ and $rpoS^+$ alleles were introduced in two steps. A parental strain was made Thr^- , Met^- or Cys^- by using an appropriate donor lysate and selecting tetracycline- or kanamycin-resistant transductants. Thr^+ , Met^+ or Cys^+ transductants were then selected and scored for acquisition of the respective donor phenotype and simultaneous loss of the parental antibiotic resistance. Transductions were carried out using P1kc as described elsewhere (Wanner, 1994).

Construction of a smooth (rfb^+) *E. coli* K-12 strain. The smooth *E. coli* K-12 strain BW24599 (Table 1) was constructed using allele replacement to correct the $rfb-50$ mutation, an IS5 insertion within *wbbL* [formerly called *orf-11* (Reeves *et al.*, 1996)]. To do this, the 2.7 kbp *Pst*I fragment corresponding to the *wbbL* segment in pPR1474 [from the *E. coli* K-12 strain WG1 (Liu & Reeves, 1994)] was cloned into the allele-replacement vector pLD55 (Haldimann *et al.*, 1996). The resulting plasmid pTJ08-08 was used to transform BW24476. Tetracycline-resistant transformants were selected and purified nonselectively, after which tetracycline-sensitive recombinants were selected as described previously (Metcalf *et al.*, 1996). BW24599 was shown to be rfb^+ based on its insensitivity to phage P1 and production of O antigens.

Construction of $\Delta rmlC$ *E. coli* K-12 strains. A $\Delta rmlC$ mutant of BW24599 was similarly constructed by using pSK95, a derivative of pLD55 having an approximately 3 kbp insert with the structure *Not*I-*rmlA*-*Hind*III-FRT-*kan*-FRT-*Kpn*I-*wzx*-*Bam*HI. The insert was assembled by joining upstream *rmlA* and downstream *wzx* [formerly called *rfbX* (Reeves *et al.*, 1996)] sequences to a kanamycin resistance gene cassette that is flanked by FRT sites for the FLP recombinase. The *Hind*III-FRT-*kan*-FRT-*Kpn*I fragment originated from pCP15 (Cherepanov & Wackernagel, 1995). The *rmlA* and *wzx* fragments were generated using PCR with *Taq* DNA polymerase and *E. coli* BW24476 chromosomal DNA as template. PCR primers for *rmlA* were GCAAAGCTTGCG-GCCGCAT-ATGAAATGCGTAAAGGT and GCAAAG-CTTAGATCT-CATTAATTTGAATCCTTCGTC; and for *wzx* were GCAGGTACCTCCGGACAT-ATGAATACGA-ATAAATTATC and GCAGAGCTCGGATCCTT-AGCAA-ACCAGTAATTTATTATTC. Sequences preceding a hyphen correspond to extensions with a restriction site(s) that was used in subsequent cloning steps. The PCR fragments were cloned into pCR2.1 and sequenced as described elsewhere (Haldimann *et al.*, 1998). An approximately 0.5 kbp *Mfe*I-*Xmn*I fragment was deleted from the 3' end of *wzx* in the construction of pSK95. A transformant of BW20767 (Metcalf *et al.*, 1996) carrying pSK95 was mated with BW24599. Tetracycline-resistant exconjugants were selected and purified nonselectively, after which tetracycline-sensitive recombinants were selected as described previously (Metcalf *et al.*, 1996). The kanamycin-resistant and P1-sensitive recombinant BW24861 (Table 1) was shown to contain the $\Delta rmlC::kan_{SK95}$ chromosomal mutation and to no longer react with O16-specific antibody. The kanamycin resistance gene was eliminated by using the FLP plasmid pCP20 as described elsewhere (Cherepanov & Wackernagel, 1995) in order to make the kanamycin-sensitive $\Delta rmlC_{SK95}$ mutant BW24970 (Table 1).

Construction of an $rmlC^+$ integration plasmid. The plasmid pTJ09 a conditional replication, integration (*attP λ*) plasmid that expresses *rmlC* behind P_{tac} was constructed to confirm that the properties of strains BW24861 and BW24970 were due to lack of *rmlC*. The *rmlC* gene was PCR-amplified with *Nde*I and *Bam*HI sites using GCAGAATTCAT-ATGAATGTGATTAGAACTGAAATTGA and GCAGGATCCTT-ATGCAATTAATTTTAATCTGATAAGC as primers. The PCR-amplified *rmlC* gene was subcloned as an *Nde*I-*Bam*HI fragment into pSPORT::*merR* and then as an *Eco*RI-*Nhe*I fragment into pSLF52 (Haldimann *et al.*, 1997). The latter has the 1.5 kbp *Xho*I-*lacI^R*- P_{tac} -*Bam*HI fragment of pGZ119HE (Russel, 1994) joined to the *Bam*HI-*attP λ* -ori R_{6K} -chloramphenicol resistance *Sall* fragment of pCANTS γ *Nde*I* (Haldimann *et al.*, 1996). Integrants of BW24861 and BW24970 containing a single copy of pTJ09 at *att λ* were made using pINT-ts and verified by PCR as described elsewhere (Haldimann *et al.*, 1996).

Cloning, expression and purification of the protein product from *M. tuberculosis* *rmlC*. The sequence of the *rmlC* gene (Rv 3465) in *M. tuberculosis* strain H37Rv was obtained directly from the genome sequence (Cole *et al.*, 1998) and can be accessed from <http://www.sanger.ac.uk/Projects/M-tuberculosis/>. The putative dTDP-4-keto-6-deoxyglucose 3',5'-epimerase (RmlC) proteins of *M. tuberculosis* and *E. coli* K-12 are about 38% identical at the amino acid sequence level. The DNA sequence of *rmlC* in *M. tuberculosis* H37Rv was found to be very similar but not identical to that of a Korean *M. tuberculosis* clinical isolate (Lee *et al.*, 1997). The *rmlC* gene was then cloned from *M. tuberculosis* H37Rv using the Ligation Independent Cloning (LIC) system (Stratagene). Thus PCR was conducted using the 5' primer GACGACGACAAG-ATGAAAGCAGCGAAGCTCGAC (coding sequence to the right of the hyphen) and the 3' antisense primer GGAAC-AAGACCCGT-CTAGGTGCCGCGCATCTCC with *M. tuberculosis* H37Rv genomic DNA. The resulting product was cloned into the LIC plasmid and the inserted DNA was sequenced to ensure proper insertion and PCR fidelity. The gene was then expressed, and the resulting protein purified on a calmodulin-binding column and treated with enterokinase according to the manufacturer's instructions. Both matrix-assisted MS and N-terminal sequencing (City of Hope) showed that the enterokinase cleaved 15 amino acids upstream of the desired cleavage site, resulting in 15 additional amino acids on the N-terminus. The resulting polypeptide nevertheless showed the desired enzyme activity and thus was suitable for the purposes of this study.

Preparation of crude enzyme extracts. Bacteria were harvested and sonicated in 50 mM HEPES (pH 7.6) with 0.1 mM PMSF and ~1 mM DTT. The sonicates were then centrifuged at 16000 g for 5–7 min to remove debris.

Preparation of dTDP-rha and dTDP-4-keto-6-deoxyglucose. dTDP-rha was prepared by incubation of 5 μ mol deoxythymidine diphosphate glucose (dTDP-glc) (Sigma) with 23 mg crude *E. coli* B protein (from crude enzyme extracts) and 5 μ mol NADPH in 2 ml of a buffer consisting of 50 mM MOPS (pH 7.8), 3 mM EDTA and 1 mM DTT. After incubation for 30 min at 37 °C, an additional 5 μ mol NADPH was added and the incubation continued for 30 min. Again, 5 μ mol NADPH was added and the incubation continued for a final 30 min. Three millilitres of ethanol were added, the insoluble proteins removed by centrifugation, the volume reduced to approximately 1 ml and the dTDP-rha purified by HPLC on a Dionex Magnum 9 PA-1 column. The HPLC flow rate was 3 ml min⁻¹ and after a five min elution with 5 mM

$\text{NH}_4\text{H}_2\text{PO}_4$, a gradient from 5–50 mM $\text{NH}_4\text{H}_2\text{PO}_4$ over 40 min was applied. The OD_{260} -positive fractions were analysed for dTDP-rha using an analytical HPLC column as described below; those containing dTDP-rha were combined and desalted on a G-10 column in water. The final product was concentrated to a convenient volume; the identity of the sugar was confirmed to be rhamnose by GC-MS analysis of the resulting alditol acetates (Daffe *et al.*, 1990).

dTDP-4-keto-6-deoxyglucose was prepared by an adaptation of the procedure of Marumo *et al.* (1992). Thus, the reaction mixture (400 μl) containing 1 μmol dTDP-glc and 4 mg crude soluble protein from *E. coli* BW24970 in 50 mM HEPES buffer (pH 7.6) was incubated for 1 h at 37 °C. To assay for the completeness of the reaction, a 25 μl aliquot was withdrawn, added to 775 μl 0.1 M NaOH and reincubated at 37 °C for 20 min. The absorption was measured at 320 nm ($\epsilon_{320} = 4600 \text{ M}^{-1} \text{ cm}^{-1}$). The reaction was found to essentially go to completion. Then 400 μl ethanol was added and denatured protein was removed by centrifugation at 14000 g for 10 min. The ethanol was removed by evaporation and the crude dTDP-4-keto-6-deoxyglucose reaction applied to a G-10 (2.5 $\times$ 120 cm) column at 4 °C. Fractions (5 ml) were collected and analysed for salt by measurement of conductivity and for dTDP-4-keto-6-deoxyglucose by measurement of A_{262} . The UV-absorbing fractions were pooled, concentrated to a convenient volume and frozen.

HPLC assay for formation of dTDP-rha. The assay mixture (50 μl) contained 2 nmol TDP-glc, 6 nmol NADPH, enzyme [50 μg crude (10000 g supernatant) protein when testing *E. coli* strains; 50 μg crude soluble (10000 g supernatant) protein prepared from *E. coli* BW24970 and 2 μg *M. tuberculosis* purified RmlC when testing for RmlC complementation], 1 mM MgCl_2 , all in 50 mM HEPES buffer (pH 7.6). The reaction was incubated for 1 h at 37 °C, followed by addition of 67 μl ethanol. Denatured protein was removed by centrifugation at 14000 g for 5–10 min and supernatant was injected on a Dionex PA-100 HPLC column. The column was eluted with 75 mM KH_2PO_4 (isocratic) and detected by measurement of A_{254} . Standards of dTDP-rha and dTDP-glc were shown to be separated in this system.

Analysis of enzymically produced TDP-4-keto-6-deoxyhexoses by GC-MS. The assay mixture (35 μl) contained 140 nmol TDP-glc, 330 μg *E. coli* soluble protein (either from strain BW24970 or BW24599 depending on the experiment) in 50 mM HEPES buffer (pH 7.6). The reaction was incubated for 1 h at 37 °C and then 50 μl ethanol were added, followed by 2 mg sodium borodeuteride (to reduce the 4-keto group). After 2 h, the reduction was terminated with a few drops of glacial acetic acid, 20 μl 10% acetic acid in methanol was added and the sample dried. The sample was then hydrolysed with 2 M trifluoroacetic acid at 120 °C for 1 h, reduced with sodium borodeuteride (10 mg ml^{-1} in 1 M ammonium hydroxide), the borate removed by drying the sample after the addition of methanol (this forms volatile methyl borates), and the sample acetylated as described by York *et al.* (1986). GC-MS was performed on a BPX70 (SGE) column using a Hewlett Packard GC-MS system as described by Daffe *et al.* (1990), except that the selected ion monitoring mode was used to selectively monitor m/z 172 and 232, and in the deuterium incorporation experiments m/z 217, 218, 231 and 233.

Assay for conversion of dTDP-4-keto-6-deoxyglucose to dTDP-4-keto-rhamnose by purified *M. tuberculosis* RmlC protein. dTDP-4-keto-6-deoxyglucose (5 nmol) was incubated either with *M. tuberculosis* RmlC (7 μg) or without RmlC (control) in HEPES buffer (pH 7.6) for 3.5 h at 37 °C. The

resulting 4-keto sugar nucleotides were then derivatized and analysed by GC-MS as described above.

Assay for incorporation of deuterium from solvent into dTDP-4-keto-6-deoxyglucose. The assay mixture (18 μl) contained 1 nmol dTDP-4-keto-6-deoxyglucose, 1.26 μg *M. tuberculosis* RmlC protein in 50 mM HEPES buffers (pH 7.6). Due to $[\text{H}]_2\text{O}$ in the enzyme the resulting concentration of $[\text{H}]_2\text{O}$ was about 70%. The reaction was incubated for 30 min at 37 °C and then stopped by the addition of ethanol (50 μl). The resulting deuterated 4-keto sugar nucleotides were then reduced by the addition of $\text{NaB}[\text{H}]_4$ (not $\text{NaB}[\text{H}]_4$ as in other experiments) and then derivatized and analysed by GC-MS as described above, except that the second reduction was also with $\text{NaB}[\text{H}]_4$ rather than $\text{NaB}[\text{H}]_4$.

RESULTS

Construction of *E. coli* ΔrmlC mutants

In *E. coli* K-12, the *rmlC* gene lies within the *rfb* gene cluster that has been well characterized by Reeves and co-workers and is required for O-antigen biosynthesis (Stevenson *et al.*, 1994). We constructed an *E. coli* K-12 ΔrmlC mutant to aid in the determination of the function of this gene and in the characterization of the *M. tuberculosis* RmlC protein. As all known laboratory strains of *E. coli* K-12 are rough (Rfb^-) due to synthesis of an incomplete lipopolysaccharide structure lacking the O antigen (Liu & Reeves, 1994), this required that

Fig. 2. The scheme used to derivatize dTDP-4-ketosugars for GC-MS analysis as illustrated for dTDP-4-keto-6-deoxyglucose (a) and for dTDP-4-keto-rhamnose (b). The resulting alditol acetates are numbered according to their elution order (see Fig. 3). D, deuterium; Ac, acetate.

Fig. 3. Analysis of the dTDP-4-keto sugars after incubation of dTDP-glc in crude enzyme extracts prepared from *E. coli* with [BW24599 (trace A)] and without [BW24970 (trace B)] a functional *rmlC* gene. In both cases, the resultant dTDP-4-keto sugars were analysed after the derivatization shown in Fig. 2 by selected ion GC-MS in which m/z 172 was monitored. Ions with m/z 172 come from a complex loss of 206 atomic mass units from the molecular ion in which hydrogen or deuterium at C-4 (and C-1) is retained. Contaminating 6-deoxyhexoses not deuterated at C-4 yield ions with m/z 171. Peak 1, 1,2,3,4,5-penta-O-acetyl 1,4-dideuterio-rhamnitol; peak 2, 1,2,3,4,5-penta-O-acetyl 1,4-dideuterio-6-deoxytalitol; peak 3, 1,2,3,4,5-penta-O-acetyl 1,4-dideuterio-fucitol; peak 4, 1,2,3,4,5-penta-O-acetyl 1,4-dideuterio-6-deoxyglucitol. Compounds 1 and 2 come from dTDP-4-keto-rhamnose; compounds 3 and 4 result from dTDP-4-keto-6-deoxyglucose (see Fig. 2). In the case of trace A, a substantial amount of the m/z 172 comprising peak 1 is due to the ^{13}C isotope of m/z 171. Ions with m/z 171 are not found in any other peaks (data not shown) and in the case of peak 1 are from dTDP-rha (non-4-keto compound), which were produced due to small amounts of NADPH that was not removed from the protein preparation of *E. coli* BW24599 even by dialysis. The calculated value of m/z 172 due to dTDP-4-keto-rhamnose after subtraction of the isotope peak due to dTDP-rha is shown by the dashed line.

Fig. 4. SDS-PAGE analysis of purified *M. tuberculosis* RmlC. Lane 1, molecular mass standards; lane 2, RmlC.

Fig. 5. HPLC analysis for the production of dTDP-4-keto-rhamnose by enzymes present in crude enzyme extracts of *rmlC* *E. coli* BW24970 with no additions (trace A) or after addition of purified *M. tuberculosis* RmlC (traces B–D). In both instances dTDP-glc and NADPH were added to the extracts and after incubation for varying lengths of time the resulting sugar nucleotides analysed by HPLC. Times of incubations were: A (control), 60 min; B, 10 min; C, 30 min; D, 60 min.

we first make a smooth (*Rfb*⁺) *E. coli* K-12 strain. The rough phenotype results from one of two independent mutations in the *rfb* gene cluster, *rfb-50* or *rfb-51*. The *E. coli* K-12 strain EMG2 (Liu & Reeves, 1994), like the recently sequenced strain MG1655 (Blattner *et al.*, 1997), has the *rfb-50* mutation, an IS5 insertion in the rhamnosyl transferase (*wbbL*) gene. The *E. coli* strains used in this study are descendants of the *E. coli* K-12 strain BD792 (Haldimann *et al.*, 1997). Both BD792 and MG1655 are immediate and unmutagenized descendants of W1485, which is in turn a descendant of EMG2 (Bachmann, 1996; B. J. Bachmann, personal communication). We were therefore able to repair the *rfb* mutation by using a standard allele replacement method and a DNA fragment from the *E. coli* K-12 *rfb-51* strain WG1 (Liu & Reeves, 1994). The resulting smooth *E. coli* K-12 strain BW24599 was shown to produce O antigen by a colony blot using a polyclonal antibody preparation

Fig. 6. Formation of dTDP-4-keto-rhamnose by purified *M. tuberculosis* RmlC. The enzyme (trace A), or buffer control (trace B) was incubated with dTDP-4-keto-6-deoxyglucose and the resultant 4-keto-sugar nucleotides derivatized as in Fig. 2 and analysed by selected ion mass spectrometry as in Fig. 3. The numbered peaks correspond to those identified in Fig. 3.

known to react against the K-12 O antigen (Stevenson *et al.*, 1994). We then constructed two $\Delta rmlC$ derivatives of BW24599: strain BW24861, with a kanamycin resistance gene associated with its $\Delta rmlC$ mutation, and strain BW24970, with an unmarked $\Delta rmlC$ mutation. We showed that both could be complemented with a plasmid that expresses *rmlC*⁺ behind *P*_{tac} in a single copy at the phage λ attachment site (Table 1). In the course of this study, we also discovered that rough *E. coli* K-12 strains are sensitive and smooth ones insensitive to the phage P1. This phenotype provided a simple means to test various recombinants. The *rfb*⁺ *E. coli* K-12 strains also proved to be extremely difficult to transform. On several occasions, no or very few transformants were obtained using standard transformation or electroporation techniques (data not shown). These phenotypes may be useful in new studies of genes for polysaccharide biosynthesis. As the *rfb*⁺ strains acted normally as recipients in conjugation, particular mutants were constructed using conjugative plasmids.

Analysis of strains for their ability to synthesize dTDP-rha from dTDP-glc

Incubation of crude enzyme extracts of BW24599 with dTDP-glc and NADPH resulted in the formation of dTDP-rha as analysed by HPLC; in contrast incubation of extracts of strains BW24970 and BW24861 with dTDP-glc and NADPH yielded no dTDP-rha, as expected of these *rmlC* mutants (HPLC data not presented except for strain BW24970 in Fig. 5 trace A).

Incubation of enzymes present in a crude enzyme extract of *E. coli* strain BW24970 with dTDP-glc yields dTDP-4-keto-6-deoxyglucose but not dTDP-4-keto-rhamnose

To assay for the presence of 4-keto sugar nucleotides, the 4-keto group of the enzymically produced

compounds resulting from the incubation of dTDP-glc with *E. coli* BW24970 was reduced with NaB[²H]₄ in 1 M NH₄OH and the resulting sugar nucleotide diastereomers treated with acid, NaB[²H]₄ and acetic anhydride to prepare the corresponding alditol acetates (Fig. 2). These alditol acetates were analysed by GC-MS and compounds derived from both dTDP-4-keto-6-deoxyglucose and from small amounts of dTDP-4-keto-rhamnose were detected. When the conditions of the first chemical reduction were changed by quenching the enzymic reaction with ethanol and adding NaB[²H]₄ as a powder (or dissolved in ethanol) directly to the ethanol-soluble sugar nucleotides in the absence of the commonly included NH₄OH (York *et al.*, 1986), only the alditol acetates derived from dTDP-4-keto-6-deoxyglucose were produced (Fig. 3 trace A). However, when the crude enzyme extract was prepared from *E. coli* BW24599 it could be seen that a small amount of the dTDP-4-keto-6-deoxyglucose was converted to dTDP-4-keto-rhamnose as evidenced by the reduction products (Fig. 3 trace B). The straightforward interpretation of this result is that RmlC protein in the *rmlC*-containing bacteria was responsible for this conversion.

Cloning, expression and purification of *rmlC* from *M. tuberculosis*

To confirm the results reported above and as part of our programme to develop new drugs against *M. tuberculosis*, the *rmlC* gene of *M. tuberculosis* was cloned into the Ligation Independent Cloning (LIC) system of Stratagene. In this commercial system, *rmlC* is fused at the 5' end to DNA encoding a calmodulin-binding peptide with an enterokinase cleavage site directly upstream of the starting ATG of *rmlC*. After expression, the resulting protein was purified on a calmodulin column and treated with enterokinase to yield the purified polypeptide shown in Fig. 4. Inexplicably, as revealed by N-terminal sequencing, the polypeptide was cleaved 15 amino acids upstream of the enterokinase site and the protein shown in Fig. 4 thus has a 15 amino acid 'tail' at the amino terminus. However, the extra amino acids did not interfere with enzymic activity and thus when RmlC was mixed with the enzymes in a crude enzyme extract prepared from *E. coli* BW24970, dTDP-rha was readily formed from dTDP-glc as shown in Fig. 5.

Purified RmlC catalyses the conversion of dTDP-4-keto-glucose to dTDP-4-keto rhamnose

dTDP-4-keto-glucose free of dTDP-4-keto-rhamnose was prepared by incubating dTDP-glc with the enzymes in a particle-free extract prepared from *E. coli* BW24970 and purification by gel chromatography. GC-MS analysis after reduction, hydrolysis, reduction and acetylation (Fig. 6 trace A) showed the absence of dTDP-4-keto-rhamnose, although care was needed to strictly avoid basic or acidic pH during the purification. When incubated with purified dTDP-4-keto-6-deoxyglucose epimerase and analysed for 4-keto sugar nucleotides, the

Fig. 7. Incorporation of deuterium into dTDP-4-keto-6-deoxyglucose as catalysed by dTDP-4-keto-6-deoxyglucose epimerase (RmlC) from *M. tuberculosis*. After incubation in the absence (a) and presence (b) of dTDP-4-keto-glucose epimerase (RmlC) in deuterium oxide, the resultant 4-keto-sugar nucleotides were converted into alditol acetates as in Fig. 2 except that all reductions were done with $\text{NaB}[^1\text{H}]_4$ rather than $\text{NaB}[^2\text{H}]_4$. The structure of the resulting 6-deoxyglucitol pentaacetate with no deuteration and with 3,5 deuteration and the expected fragment ions are illustrated. Only the trace for the per-O-acetylated 6-deoxyglucitol derivative (peak 4, Fig. 3) is shown. No deuterium incorporation is seen in the control but incorporation at both positions 3 (*m/z* 218) and 5 (*m/z* 233) occurred in the presence of enzyme. The deuterium incorporation was not complete as seen by the presence of ions with *m/z* 217 and 231 in (b); this is in part due to the fact that $[^1\text{H}]_2\text{O}$ was present in the reaction mixture to the extent of 30%, but it is also likely that exchange between the solvent and substrate was not complete.

conversion of dTDP-4-keto-6-deoxyglucose to dTDP-4-keto-rhamnose was apparent (Fig. 6). The ratio of the peaks was the same (as analysed by GC rather than GC-MS; data not shown) for incubations of 1, 2 and 3 h, showing that the reaction had reached equilibrium and the ratio dTDP-4-keto-rhamnose to dTDP-4-keto-6-deoxyglucose was calculated to be 7:93.

dTDP-4-keto-6-deoxyglucose epimerase catalyses the expected deuteration of dTDP-4-keto-6-deoxyglucose from $[^2\text{H}]_2\text{O}$ solvent

Previous workers have shown that crude preparations of dTDP-4-keto-6-deoxyglucose epimerase allow for the incorporation of deuterium at carbons 3 and 5 from solvent (Melo & Glaser, 1968). We therefore checked the pure dTDP-4-keto-6-deoxyglucose epimerase from

M. tuberculosis to see if it also was capable of this exchange reaction. As shown in Fig. 7, this turned out to be the case; the single polypeptide catalysed the epimerization at both positions (Fig. 6) and the incorporation of deuterium at both positions (Fig. 7).

DISCUSSION

The experiments presented above demonstrate the validity of the prediction that the RmlC protein is a 'di' epimerase acting at both positions 5' and 3' of dTDP-4-keto-6-deoxyglucose. The earlier (Melo & Glaser, 1968; Gaugler & Gabriel, 1973) inability to detect the product, dTDP-4-keto-rhamnose, in a convincing fashion arose from the technical difficulties of analysing the unstable 4-keto sugars and from the fact that the equilibrium with dTDP-4-keto-6-deoxyglucose is such that the 4-keto rhamnose product is present only in small amounts. The equilibrium value is perhaps not unexpected given the inherent stability of equatorial phosphates (i.e. dTDP-4-keto-rhamnose) compared to axial phosphates (i.e. dTDP-4-keto-6-deoxyglucose). The fact that the first step in dTDP-rhamnose formation, the synthesis of dTDP-glucose, is essentially irreversible due to the hydrolysis of the pyrophosphate, and the fact that the equilibrium of the last reaction catalysed by RmlD (Fig. 1) lies strongly on the side of dTDP-rhamnose, results in a committed and efficient pathway in spite of the equilibrium of reactants and products produced by RmlC. It is interesting to compare the biosynthesis of dTDP-rha with that of GDP-fucose. Both 6-deoxyhexose sugar nucleotides are prepared in an analogous fashion with an oxidation/reduction at C-4' and C-6, epimerizations at C-3' and C-5', and finally reduction at C-4'. In the case of GDP-fucose the starting compound is GDP-mannose. An enzyme analogous to RmlB (Gmd) converts GDP-mannose to GDP-4-keto-6-deoxymannose (Andrianopoulos *et al.*, 1998; Stevenson *et al.*, 1996). Interestingly, however, the homologues of RmlC and RmlD occur as a single polypeptide as recently shown by Reeves and co-workers (Andrianopoulos *et al.*, 1998). The GDP-4-keto-fucose formed after epimerization also has the phosphate in the equatorial position and would be expected to be less stable than GDP-4-keto-6-deoxymannose. Combining both the reductase and the di-epimerase in a single polypeptide chain would thus make the complete conversion from GDP-4-keto-6-deoxymannose to GDP-fucose more efficient. However, clearly two polypeptides are involved in the dTDP-rha story and the epimerase can function independently of the reductase. Nevertheless, our data do not preclude the possibility that *in vivo* RmlC and RmlD (the putative dTDP-4-keto-rhamnose reductase) interact for a more efficient transformation.

Finally, it should be noted that the inability of the *rmlC* mutant *E. coli* BW24970 to synthesize dTDP-rha establishes the validity of RmlC as a drug target when inhibition of the production of rhamnosyl residues is desired. The *E. coli* strains developed in the course of these studies may be valuable in designing screens for such inhibitors.

ACKNOWLEDGEMENTS

This work was supported by funds provided by Public Health Service grants NIAID, NIH AI-33706 and U19-AI 40972 to M.R.M.; NSF 9730034-MCB to B.L.W.; and by Korea Science and Engineering Foundation, KOSEF 971-0712-397-1, to T.-Y.L. We thank A. S. Lynch, P. Reeves, M. Russel and W. Wackernagel for generously providing plasmids.

REFERENCES

- Andrianopoulos, K., Wang, L. & Reeves, P. R. (1998). Identification of the fucose synthetase gene in the colanic acid gene cluster of *Escherichia coli* K-12. *J Bacteriol* **180**, 998–1001.
- Bachmann, B. J. (1996). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, pp. 2460–2488. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Blattner, F. R., Plunkett, G., III, Bloch, C. A. & 14 other authors (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474.
- Charland, N., Harel, J., Kobisch, M., Lacasse, S. & Gottschalk, M. (1998). *Streptococcus suis* serotype 2 mutants deficient in capsular expression. *Microbiology* **144**, 325–332.
- Cherepanov, P. P. & Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9–14.
- 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.
- Daffe, M., Brennan, P. J. & McNeil, M. (1990). Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by ^1H and ^{13}C -NMR analyses. *J Biol Chem* **265**, 6734–6743.
- Fisher, S. L., Jiang, W., Wanner, B. L. & Walsh, C. T. (1995). Cross-talk between the histidine protein kinase VanS and the response regulator PhoB. Characterization and identification of a VanS domain that inhibits activation of PhoB. *J Biol Chem* **270**, 23143–23149.
- Gaugler, R. W. & Gabriel, O. (1973). Biological mechanisms involved in the formation of deoxy sugars. VII. Biosynthesis of 6-deoxy-L-talose. *J Biol Chem* **248**, 6041–6049.
- Haldimann, A., Prahalad, M. K., Fisher, S. L., Kim, S. K., Walsh, C. T. & Wanner, B. L. (1996). Altered recognition mutants of the response regulator PhoB: a new genetic strategy for studying protein–protein interactions. *Proc Natl Acad Sci USA* **93**, 14361–14366.
- Haldimann, A., Fisher, S. L., Daniels, L. L., Walsh, C. T. & Wanner, B. L. (1997). Transcriptional regulation of the *Enterococcus faecium* BM4147 vancomycin resistance gene cluster by the VanS–VanR two-component regulatory system in *Escherichia coli* K-12. *J Bacteriol* **179**, 5903–5913.
- Haldimann, A., Daniels, L. L. & Wanner, B. L. (1998). Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *J Bacteriol* **180**, 1277–1286.
- Lee, T. Y., Lee, T. J., Belisle, J. T., Brennan, P. J. & Kim, S. K. (1997). A novel repeat sequence specific to *Mycobacterium tuberculosis* complex and its implications. *Tubercle Lung Dis* **78**, 13–19.
- Lindquist, L., Kaiser, R., Reeves, P. R. & Lindberg, A. A. (1993). Purification, characterization and HPLC assay of *Salmonella* glucose-1-phosphate thymidyl-transferase from the cloned *rfbA* gene. *Eur J Biochem* **211**, 763–770.
- Liu, D. & Reeves, P. R. (1994). *Escherichia coli* K12 regains its O antigen. *Microbiology* **140**, 49–57.
- Ma, Y., Mills, J. A., Belisle, J. T., Vissa, V., Howell, M., Bowlin, K., Scherman, M. S. & McNeil, M. (1997). Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the *Mycobacterium tuberculosis* gene encoding α -D-glucose-1-phosphate thymidyltransferase. *Microbiology* **143**, 937–945.
- McNeil, M., Daffe, M. & Brennan, P. J. (1990). Evidence for the nature of the link between the arabinogalactan and peptidoglycan components of mycobacterial cell walls. *J Biol Chem* **265**, 18200–18206.
- Marumo, K., Lindqvist, L., Verma, N., Weintraub, A., Reeves, P. R. & Lindberg, A. A. (1992). Enzymatic synthesis and isolation of thymidine diphosphate-6-deoxy-D-xylo-4-hexulose and thymidine diphosphate-L-rhamnose. *Eur J Biochem* **204**, 539–545.
- Melo, A. & Glaser, L. (1968). The mechanism of 6-deoxyhexose synthesis. II. Conversion of deoxythymidine diphosphate 4-keto-6-deoxy-D-glucose to deoxythymidine diphosphate L-rhamnose. *J Biol Chem* **243**, 1475–1478.
- Metcalfe, W. W., Jiang, W. & Wanner, B. L. (1994). Use of the rep technique for allele replacement to construct new *Escherichia coli* hosts for maintenance of R6K gamma origin plasmids at different copy numbers. *Gene* **138**, 1–7.
- Metcalfe, W. W., Jiang, W., Daniels, L. L., Kim, S. K., Haldimann, A. & Wanner, B. L. (1996). Conditionally replicative and conjugative plasmids carrying *lacZ* alpha for cloning, mutagenesis, and allele replacement in bacteria. *Plasmid* **35**, 1–13.
- Naundorf, A. & Klaffke, W. (1996). Substrate specificity of native dTDP-D-glucose-4,6-dehydrates: chemo-enzymatic synthesis of artificial and naturally occurring deoxy sugars. *Carbohydr Res* **285**, 141–150.
- Ochsner, U. A. & Reiser, J. (1995). Autoinducer-mediated regulation of rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* **92**, 6424–6428.
- Reeves, P. R., Hobbs, M., Valvano, M. A. & 8 other authors (1996). Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol* **4**, 495–503.
- Romana, L. K., Santiago, F. S. & Reeves, P. R. (1991). High level expression and purification of dthymidine diphospho-D-glucose 4,6-dehydratase (*rfbB*) from *Salmonella* serovar typhimurium LT2. *Biochem Biophys Res Commun* **174**, 846–852.
- Russel, M. (1994). Mutants at conserved positions in gene IV, a gene required for assembly and secretion of filamentous phages. *Mol Microbiol* **14**, 357–369.
- Stevenson, G., Neal, B., Liu, D., Hobbs, M., Packer, N. H., Batley, M., Redmond, J. W., Lindquist, L. & Reeves, P. R. (1994). Structure of the O antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster. *J Bacteriol* **176**, 4144–4156.
- Stevenson, G., Andrianopoulos, K., Hobbs, M. & Reeves, P. R. (1996). Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **178**, 4885–4893.
- Wahl, H. P. & Grisebach, H. (1979). Biosynthesis of streptomycin. *Biochim Biophys Acta* **568**, 243–252.
- Wanner, B. L. (1994). Gene expression in bacteria using *TnphoA* and *TnphoA'* elements to make a dns switch *phoA* gene, *lacZ*(op),

and *lacZ*(pr) fusions. In *Methods in Molecular Genetics*, pp. 291–310. Edited by K. W. Adolph. Orlando: Academic Press.

Xiang, S. H., Haase, A. M. & Reeves, P. R. (1993). Variation of the *rfb* gene clusters in *Salmonella enterica*. *J Bacteriol* **175**, 4877–4884.

York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T. & Albersheim, P. (1986). Isolation and characterization of plant cell walls and cell wall components. *Methods Enzymol* **118**, 3–40.

Zarkowsky, H., Lipkin, E. & Glaser, L. (1970). The mechanism of 6-deoxyhexose synthesis. V. The relation of pyridine nucleotide to the structure of the deoxythymidine diphosphate-glucose oxidoreductase. *J Biol Chem* **245**, 6599–6606.

Received 8 September 1998; revised 18 November 1998;
accepted 1 December 1998.