J. Korean Soc. Microbiol., Vol. 35, No. 4, 2000

Overexpression, Purification and Truncation Analysis of RmlC Protein of *Mycobacterium tuberculosis*

Jong-Seok Lee[3], Tae-Yoon Lee*, Jae-Ho Park[1], Jong-Sun Kim[2], Tae-Jin Lee[3], Jai-Youl Lee[3] and Sung-Kwang Kim

Department of Microbiology, College of Medicine, Yeungnam University, 317-1 Daemyung-5-dong, Nam-gu, Taegu 705-035, Korea; [1]Department of Internal Medicine, Keimyung University School of Medicine, 194 Dongsan-dong, Chung-gu, Taegu 700-712, Korea; [2]Department of Microbiology, College of Medicine, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-752, Korea; [3]Department of Microbiology, College of Natural Sciences, Kyungpook National University, 1370 Sankyuk-dong, Puk-gu, Taegu 702-701, Korea

dTDP-rhamnose provides L-rhamnose to the bridge-like structure between mycolyl arabinogalactan and peptidoglycan of the mycobacterial cell wall. dTDP-rhamnose is composed of glucose-1-phosphate and dTTP by four enzymes encoded by *rmlA-D*. To determine the region(s) of RmlC protein essential for its dTDP-4-keto-6-deoxyglucose epimerase activity, we overexpressed both whole (202 amino acids) and three different truncated (N-terminal 106 or 150 or C-terminal 97 amino acids) RmlC proteins of *Mycobacterium tuberculosis*. The RmlC enzyme activity in the soluble lysates of Δ*rmlC E. coli* strain SΦ874 (DE3 PlysS) expressing the wild type or truncated *rmlC* genes was initially analyzed by three sequential reactions from dTDP-glucose to dTDP-rhamnose in the presence of purified RmlB and RmlD. All three soluble lysates containing the truncated RmlC proteins showed no enzyme activity, while that containing the wild type RmlC was active. This wild type RmlC was then overexpressed and purified. The incubation of the purified RmlC enzyme so obtained with dTDP-4-keto-6-deoxyglucose resulted in the conversion of dTDP-4-keto-rhamnose. The results show that the truncated regions of the RmlC protein are important for the RmlC enzyme activity in *M. tuberculosis*.

Key Words: RmlC, *M. tuberculosis*, Overexpression, HPLC

INTRODUCTION

The two major components of the mycobacterial cell wall are mycolyl arabinogalactan and peptidoglycan (4). These two structural units are linked by a small bridge-like structure [L-rhamnose (rha)*p*-(1→3)-D-N-acetyl glucosamine-(1→phosphate)]. Thus it has been suggested that this bridge structure could be an important drug target (12). One of the components of the bridge structure, L-rha, is provided from a deoxythymidine diphosphate (dTDP)-rha. Thus the inhibition of dTDP-rha biosynthesis is likely to induce cell death of mycobacteria. dTDP-rha is also involved in O antigen biosynthesis of various enteric bacteria. Genetic study of O antigen biosynthesis revealed that four genes were in-

Received for Publication : July 13, 2000, Accepted for Publication : August 9, 2000

* To whom correspondence should be addressed: Tel: 82-53-620-4362, Fax: 82-53-653-6628, E-mail: doxr7p@medical.yeungnam.ac.kr

volved in the biosynthesis of dTDP-rha from glucose-1-phosphate and dTTP (1,6,15,17,20,23, 25). These genes were originally named *rfbA-D*, but recently they have been renamed *rmlA-D* (16). We have previously reported the nucleotide sequence of *rmlB* and *rmlC* of a Korean *Mycobacterium tuberculosis* isolate (9). Since then the whole genomic sequence of *M. tuberculosis* H37Rv has been determined (2), and it was revealed to have all 4 *rml* genes (11). According to the genomic sequence data, *rmlA* and *rmlB* genes have more than 2 copies whereas *rmlC* and *rmlD* have only one copy in the *M. tuberculosis* genome, which indicates that *rmlC* and *rmlD* genes are better drug targets. We have previously shown, using an *Escherichia coli rmlC* mutant, that the partially purified RmlC of *M. tuberculosis* H37Rv has dTDP-4-keto-6-deoxyglucose epimerase activity (19). However, the RmlC protein purified in the previous study had 15 additional amino acids on the N-terminus (19).

This study was carried out to resolve these purification problems and to identify the important region(s) in the RmlC protein, which are responsible for its enzyme activity. We overexpressed and purified the RmlC protein of *M. tuberculosis* H37Rv without any tag or fusion, and we also expressed three different truncated mutant RmlC proteins and determined their enzyme activity. The purified wild type RmlC was active, while all three truncated RmlC proteins had no enzyme activity. Our results suggest that the truncated regions of these RmlC proteins contain the amino acids responsible for its enzyme activity.

MATERIALS AND METHODS

Primers and polymerase chain reaction (PCR)

The sequence of PCR primers used for amplifying whole or truncated *rmlC* genes were as follows:

P1: 5'-GGAATTCCATATGAAAGCACGC-GAACTCGACGTC-3'

P2: 5'-CGGGATCCTAGGTGCCGCGCAT-CTCCCC-3' (W-609)

P3: 5'-CGGGATCCTATTGGTCGTCGAG-CAGCACCGAC-3' (N-318)

P4: 5'-CGGGATCCTAATCTGTGGCGCA-GATGGTGTGCTC-3' (N-450)

P5: 5'-GGAATTCCATATGGACCGTAGG-ACGATCTAC-3' (C-291)

P1 and P2 were used for amplifying the whole 609 bp *rmlC* gene (W-609), P1 and P3 for the N-terminal 318 bp (N-318), P1 and P4 for N-terminal 450 bp (N-450), and P5 and P2 for C-terminal 291 bp of the *rmlC* gene (C-291). P1 and P5 contained *Eco*RI (underlined) and *Nde*I (double underlined) sites, while P2, P3, and P4 contained *Bam*HI (dotted underlined) sites. The *Nde*I sequence that contained initiation codon (**ATG**) was included in sense primers (P1 and P5). The antisense primers (P2, P3, and P4) contained complementary sequence for stop codon (**CTA**). The *Eco*RI and *Bam*HI sites in the primers were for the efficient cloning of PCR products into pUC18. Each PCR reaction contained 2.5 U of Vent polymerase (New England Biolabs, Inc., Beverly, Mass.), 0.5 μM of each primer, 200 μM of each of the 4 deoxyribonucleoside triphosphates (Boehringer Mannheim, Germany), and 0.1 μg of purified *M. tuberculosis* H37Rv DNA in the PCR buffer supplied with the polymerase. The PCR reaction was carried out in a Perkin-Elmer 480 Cycler (Perkin-Elmer Co., Norwalk, Conn.) using the following steps: one five-min denaturation at 96℃, then 35 cycles of 95℃ for one min, 55℃ for one min, and 72℃ for one min, followed by a final extension at 72℃ for 10 min.

Cloning of the PCR products and sequence confirmation

Each PCR product was separated on a 1.6% agarose gel and purified by elution using a Ge-

neClean II kit (BIO101, Inc., La Jolla, Calif.). Each purified PCR product was cloned into the *Eco*RI-*Bam*HI site of pUC18; the inserted DNA was sequenced to ensure proper insertion and PCR fidelity. Plasmids containing wild type *rmlC* gene, N-terminal 318 bp, N-terminal 451 bp, and C-terminal 291 bp were named pRMC 609, pRMC318, pRMC451, and pRMC291, respectively.

Overexpression and protein preparation

The pUC18 clone, containing truncated *rmlC* gene, was digested with *Nde*I and *Bam*HI and subcloned into overexpression vector pET23b (Novagen, Inc., Madison, WI). The RmlC proteins were expressed in a *rmlC* deletion *Escherichia coli* strain (SΦ874 [DE3 PlysS]) to avoid basal level RmlC contamination during the determination of RmlC enzyme activity. The host strain containing pRMC318, pRMC451, or pRMC291 was grown at 37℃ overnight in the presence of ampicillin (100 μg/ml) and chloramphenicol (34 μg/ml). The overnight culture was grown until the Abs_{600} reached 0.5 to 1.0, and IPTG was then added to a final concentration of 1 mM. Cells were further cultivated at 37℃ for 3 hrs. After harvesting, cells were resuspended in 2 ml of cell washing buffer (30 mM Tris [pH 7.4] and 10 mM NaCl). The cell suspension was placed on ice and one-min sonicated 10 times at one min intervals. The cell lysate was centrifuged at 8,000 x g for 20 min and the supernatant (soluble protein fraction) was analyzed by SDS-polyacrylamide gel electrophoresis (PAGE).

Analysis of the RmlC enzyme activity from the crude protein preparations

The RmlC enzyme activity was initially analyzed by performing three sequential reactions from dTDP-glucose (glc) to dTDP-rha in the presence of RmlB, RmlD, and NADPH. A total 50 μl of reaction mixture contained: 20 nmoles of TDP-glc, 10 nmoles of NADPH, 2.5 μg of RmlB, 1 μg of RmlD, 2 mM of $MgCl_2$, 2.5 mM of NAD, and 1.4 μg of crude protein preparation to be examined, all in 50 mM Hepes buffer (pH 7.6). Initially, RmlB was preincubated with 2.5 mM NAD at room temperature for 1 hr. RmlC, RmlD, NADPH, NAD, $MgCl_2$, Hepes buffer, and dTDP-glc were then added sequentially. The mixture was incubated at room temperature and Abs_{340} was measured every 10 min for 1.5 hr. A decrease of Abs_{340} of more than 0.15, in the absence of an Abs_{340} decrease in the control reaction, was regarded as positive RmlC activity.

Purification of the wild type RmlC protein

E. coli BL21 (DE3) PlysS transformed with pRMC609 was grown in 3 liters of LB media containing 50 μg/ml ampicillin, and induced with 0.5 mM IPTG at a late log phase for four hours. After harvesting bacterial cells by centrifugation, the RmlC protein was purified by using conventional chromatography techniques. The bacterial pellet was lysed in 200 ml of 20 mM Tris-HCl buffer (pH 7.5) with, 1 mM EDTA, 0.1 mM PMSF, 1 g/ml leupeptin, 0.1 mg/ml lysozyme and 1 g/ml DNase by the freeze and thaw method. Following a 20 min centrifugation at 12,000 x g, the supernatant was applied to a DEAE Sepharose CL6B column equilibrated in 20 mM Tris-HCl buffer (pH 7.5) with 2 mM EDTA. Protein was eluted with a linear gradient of 0~0.5 M NaCl in 20 mM Tris (pH 7.5). Fractions containing the RmlC protein were identified by Coomassie blue staining of SDS-PAGE gels. The fractions of RmlC protein were pooled and further purified by taking advantage of the heat stability of the protein. The RmlC protein solution was heated in a 65℃ water bath for 10 min. Following a 20 min centrifugation at 12,000 x g, the supernatant was concentrated using a centriprep10 concentrator (Amicon, Beverly, MA). Final purification was achieved by FPLC gel-filtration chromatography using a Superdex 75 column equilibrated in 20

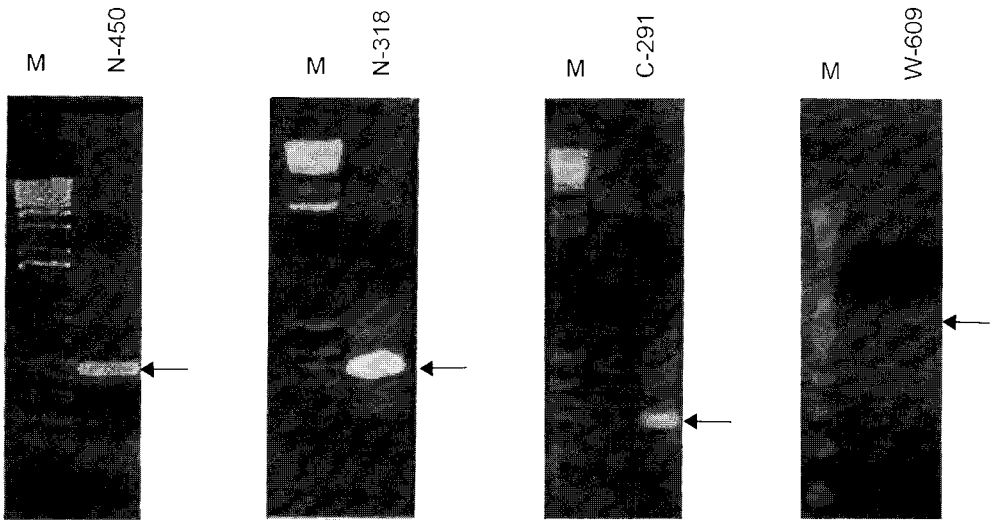

Figure 1. PCR amplification of wild type and truncated mutant *rmlC* genes. N-450, N-318, C-291 and WT-609 are PCR products containing N-terminal 450-bp and 318-bp, C-terminal 291-bp and whole 609-bp of the *rmlC* gene, respectively (arrows). M, 100-bp ladders.

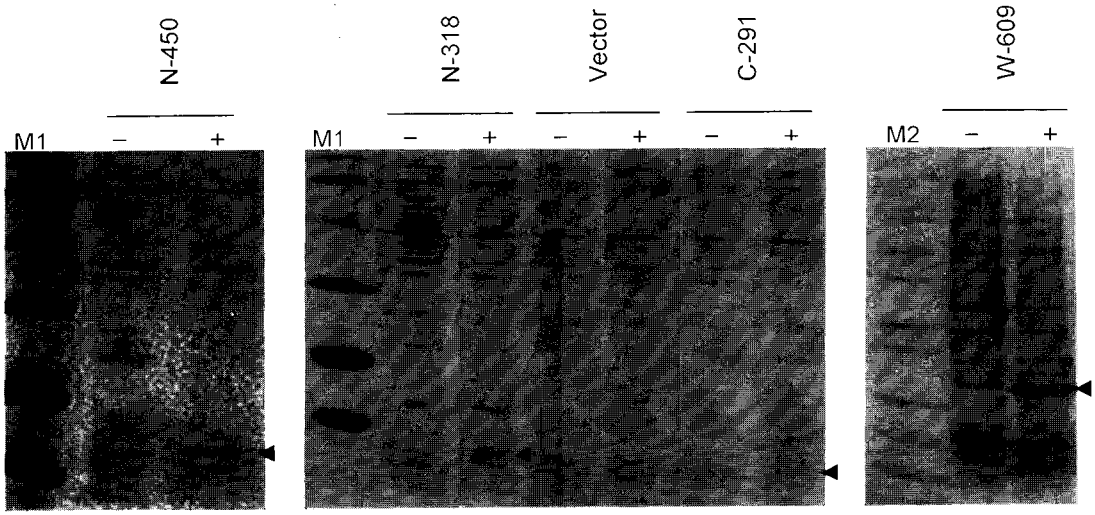

Figure 2. Protein expression of wild type and truncated mutant RmlC. Overexpression plasmid pET23b (Vector) was used for the expression of N-terminal 450-bp (N-450) and 318-bp (N-318), C-terminal 291-bp (C-291) and whole 609-bp (W-609) of the *rmlC* gene. Minus or plus indicates the absence or presence of IPTG. Arrowheads indicate the RmlC proteins expressed. M1, 97.4, 66.0, 45.0, 31.0, 21.5, and 14.5 kDa; M2, 200, 116.3, 97.4, 66.3, 55.4, 36.5, 31.0, 21.5, 14.4, and 6.0 kDa from top to bottom.

mM Tris-HCl buffer (pH 7.5) with 0.1 M NaCl and 0.02% sodium azide. Fractions containing the RmlC protein were identified by Coomassie blue staining of SDS-PAGE gels. The fractions of RmlC protein were pooled and concentrated using a centricon10 (Millipore, Bedford, MA).

HPLC analysis of purified RmlC enzyme activity

The assay mixture (50 μl) contained 2 nmol

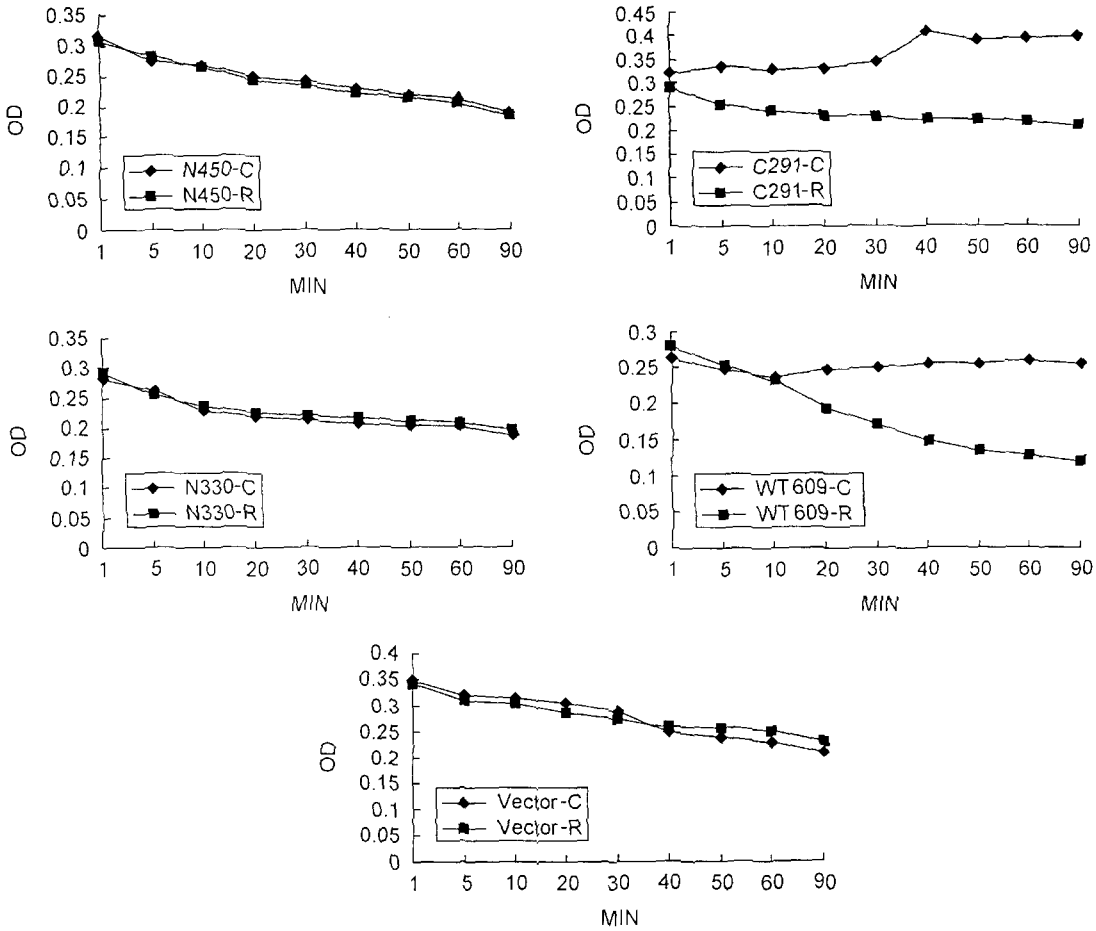

Figure 3. *Analysis of RmlC enzyme activity. C (such as N450-C) did not contain RmlC protein, while R* (such as N450-R) contained RmlC in the reaction. More than a 0.15 decrease in Abs_{340} during the R reaction was regarded as positive RmlC activity.

dTDP-glc, 6 nmol NADPH, 50 μg crude soluble protein prepared from *E. coli* BW24970 (a Δ*rmlC* strain), 2 μg RmlC protein, and 1 mM $MgCl_2$, all in 50 mM Hepes buffer (pH 7.6). The reaction was incubated for 1 hr at 37℃, which was followed by the addition of 67 μl of ethanol. Denatured protein was removed by centrifugation at 14,000 x g for 10 min and supernatant was injected onto a Dionex PA-100 HPLC column. The column was eluted with 75 mM KH_2PO_4 (isocitric) and detected by measurement of Abs_{254} (10,13,14).

RESULTS AND DISCUSSION

Amplification, cloning, and sequence confirmation of whole or truncated *rmlC* genes

We amplified whole and three different truncated *rmlC* genes of *M. tuberculosis* that contained 609-bp, N-terminal 318-bp or 450-bp or C-terminal 291-bp (Fig. 1). After cloning into pUC18, each clone was confirmed to contain the correct insert by sequencing analysis (data not shown).

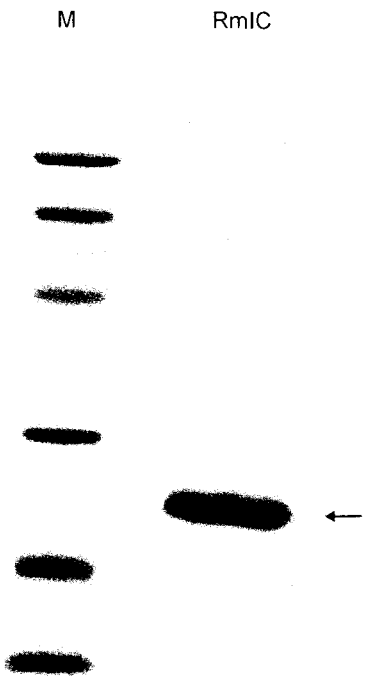

Figure 4. Purification of the wild type RmlC protein (arrow). M, same as M1 in Fig. 2.

Overexpression of wild type or truncated RmlC proteins

Wild type or each truncated *rmlC* gene was overexpressed by subcloning to overexpression vector pET23b. The wild type *rmlC* plasmid was expressed as a 22.3 kDa protein, while mutant *rmlC* plasmids containing N-terminal 318-bp or 450-bp or C-terminal 291-bp expressed 11 kDa, 16.5 kDa, or 10 kDa proteins as expected from the nucleotide sequencing data (Fig. 2). The amount of truncated RmlC proteins expressed in the SΦ874 (DE3, Δ*rmlC*) strain was lower than the wild type RmlC expressed in BL21 (DE3 PlysS, *rmlC*$^+$). This might be due to difference in expression efficiency of these two strains.

Analysis of RmlC enzyme activities

Soluble protein fractions were prepared from the Δ*rmlC* cells expressing plasmids containing whole or truncated *rmlC* genes, and the enzyme activity of each protein preparation was analyzed as described above. The reaction containing the wild type RmlC protein fraction showed a decrease of Abs_{340} as reaction time progressed. The maximum decrease observed was about 0.16 after 90 min of reaction, while the control reaction, which did not contain the RmlC protein, showed only a 0.01 decrease of Abs_{340} at the same reaction time (Fig. 3). This demonstrated that the soluble lysates of the cells expressing the wild type *rmlC* gene contained RmlC enzyme activity making dTDP-rha from dTDP-glc in addition to the two other enzymes RmlB and RmlD. This also proved that the enzyme analysis method for RmlC worked well. However, all three reactions containing truncated RmlC protein preparations showed no remarkable differences when compared with the control reactions (Fig. 3), which suggested that the truncated RmlC proteins did not contain RmlC enzyme activity.

Purification of RmlC and HPLC analysis of the enzyme activity

Since the truncated RmlC proteins did not contain enzyme activity, only the wild type RmlC was purified (Fig. 4). As shown above, the RmlC expression level in BL21 (DE3 PlysS, *rmlC*$^+$) was higher than in SΦ874 (DE3, Δ*rmlC*), the wild type RmlC expressed in BL21 (DE3 PlysS) strain. RmlC activity was further examined by HPLC by detecting the dTDP-rha made from dTDP-glc. Purified RmlC produced dTDP-rha, while the reaction containing only the protein preparations of *E. coli* BW24970 (*rmlB*$^+$, *rmlD*$^+$, and Δ*rmlC*) did not (Fig. 5). The latter reaction produced a small dTDP-rha peak, which might have been due to the contamination of RmlC protein by BL21 (DE3 PlysS) during the purification step. Taken together with the enzyme analysis result by spectrometry, this result confirmed that the purified RmlC protein was active.

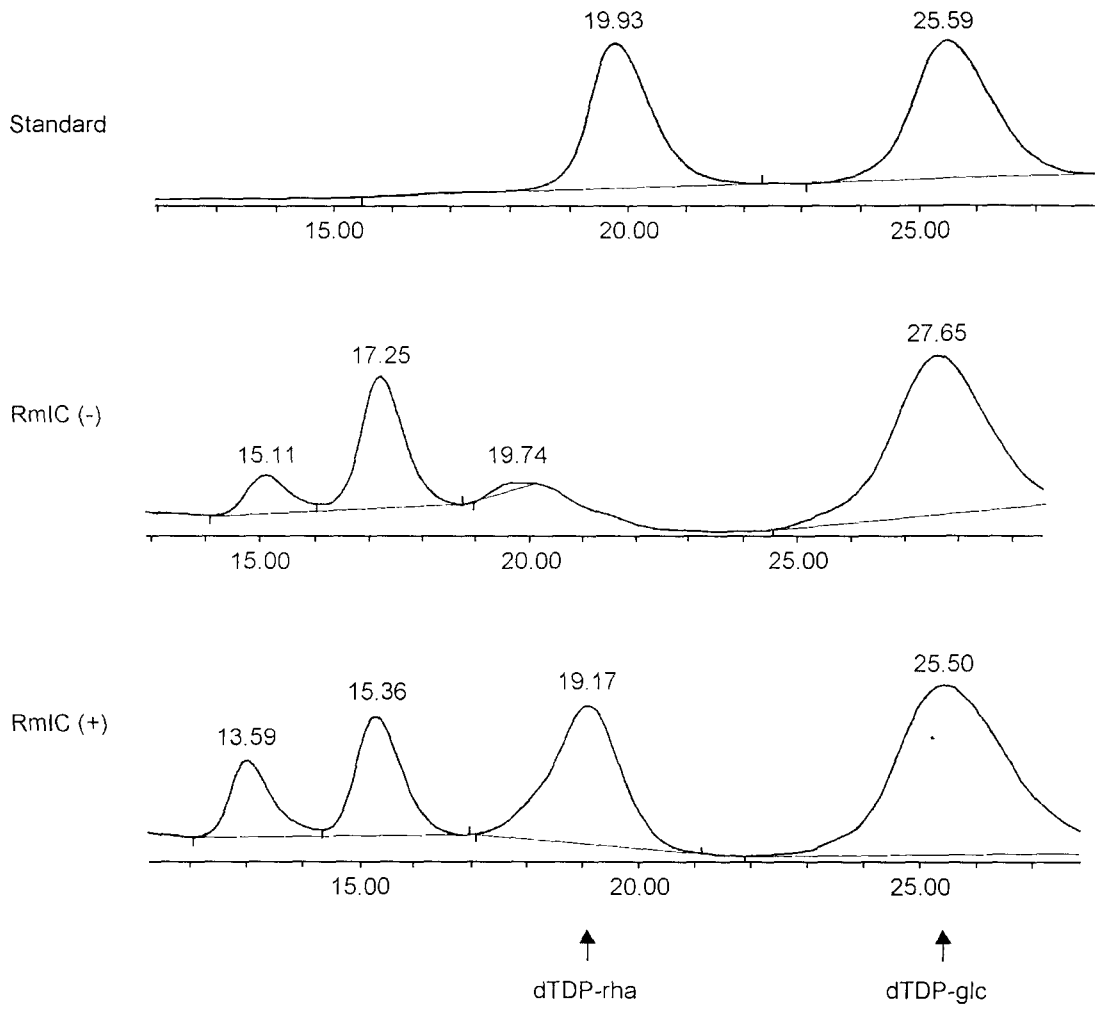

Figure 5. HPLC analysis for the production of dTDP-rha by enzyme extracts of Δ*rmlC. E. coli* BW24970 with no addition of [RmlC (-)] or after the addition of [RmlC (+)] of purified *M. tuberculosis* RmlC. In both instances dTDP-glc and NADPH were added to the enzyme extracts and after incubation for 30 min, the resulting sugar nucleotides were analyzed by HPLC.

Highly conserved amino acids are scattered along the whole RmlC sequence

Our results showed that all three truncations in the RmlC protein affected the enzyme activity. To explain this, we compared the amino acid sequence of RmlC of *M. tuberculosis* with those of seven other bacterial species (21), including *M. leprae* (5), *Serratia marcescens* (18), *Leptospira borgpetersenii* (8), *E. coli* (24), *Salmonella typhimurium* (7), *Bacteroides fragilis* (3), and *Streptococcus mutans* (22). As expected, the amino acid sequence of RmlC proteins was very well conserved among different bacterial species. There were at least 4 well-conserved domains and many identical or similar amino acids are present outside these conserved domains. The putative domains were as follows; domain I was from 16th P to 28th E, domain II from 47th Q to 63rd F, domain III from 70th Q to 132nd Y, and domain IV from 146th I to 171st D (Fig. 6). The 10 kDa truncated RmlC (C-291) did not contain N-terminal 106 amino acids (domains I and II and part of domain III)

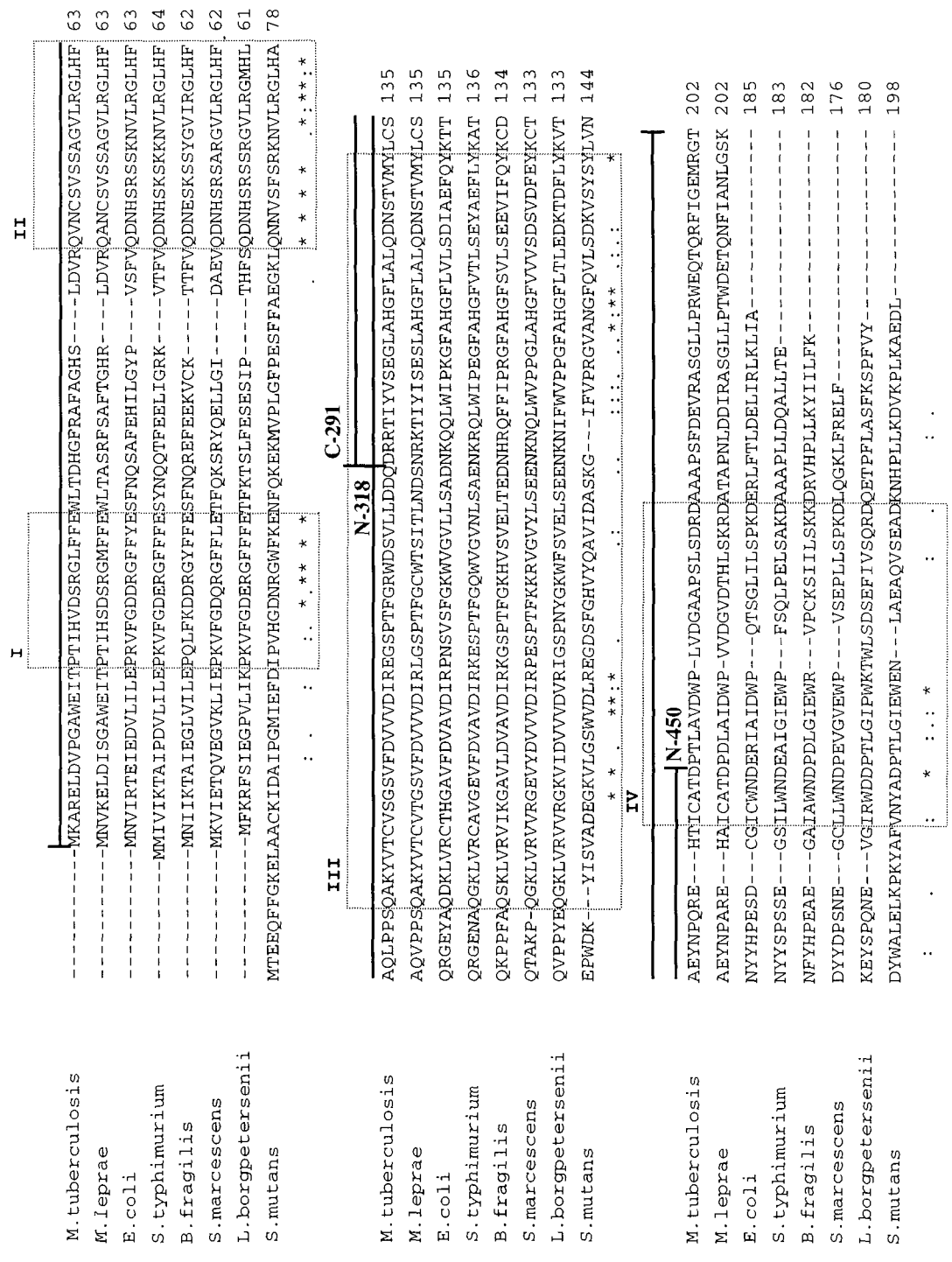

Figure 6. Multiple alignment of the RmlC protein of *M. tuberculosis* H37Rv with those of seven bacterial species. Multiple alignments were performed using CLUSTAL W (Thompson JD, et al., 1994). The alignment asterisk (*) denotes identity, a colon (:) that the position is identical in greater half of the proteins, and a period (.) that the position contains conserved amino acids in all sequences.

suggesting that these domains are important in RmlC enzyme activity. The largest truncated RmlC (N-450) contained N-terminal 150 amino acids that covered domains I to III but only part of IV, showing the importance of amino acids in the putative domain IV (such as 152nd L, 156th W, 167th L, 168th S, and 171st D) for the RmlC enzyme activity.

Since *M. tuberculosis* RmlC protein is an important drug target candidate, we undertook these truncation analyses to make the smallest active RmlC protein. Although we did not fully achieve this goal, the wild type RmlC purified in this study was small (22.3 kDa) enough to be used for 3-D structure determination by X-ray crystallography, which finally can be used for the design of RmlC enzyme inhibitor(s), an anti-tuberculosis drug candidate(s).

Acknowledgements

This study was supported by the grant from 1999 Kiturami Culture Foundation. We thank to W. Yan for her technical support and M. McNeil for providing SΦ874 (DE3) strain.

REFERENCES

1) Bachmann BJ: Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. *In Escherichia coli* and *Salmonella:* Cellular and Molecular Biology, 2nd ed, American Society for Microbiology, Washington, DC., pp 2460-2488, 1996.
2) Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, Mclean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulton JE, Taylor K, Whitehead S, Barrell BG: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393:** 537-544, 1998.
3) Comstock LE, Coyne MJ, Tzianabos AO, Pantosti A, Onderdonk AB, Kasper DL: Analysis of a capsular polysaccharide biosynthesis locus of *Bacteroides fragilis. Infect Immun* **67:** 3525-3532, 1999.
4) Daffe M, Brennan PJ, McNeil M: 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 13C-NMR analyses. *J Biol Chem* **265:** 6734-6743, 1990.
5) Eiglmeier K, Honore N, Woods SA, Caudron B, Cole ST: Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae. Mol Microbiol* **7:** 197-206, 1993.
6) Gaugler RW, Gabriel O: Biological mechanisms involved in the formation of deoxy sugars. VII. Biosynthesis of 6-deoxy-L-talose. *J Biol Chem* **248:** 6041-6049, 1973.
7) Jiang XM, Neal B, Santiago F, Lee SJ, Romana LK, Reeves PR: Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar *typhimurium* (strain LT2). *Mol Microbiol* **5:** 695-713, 1991.
8) Kalambaheti T, Bulach DM, Rajakumar K, Adler B: Genetic organization of the lipopolysaccharide O-antigen biosynthetic locus of *Leptospira borgpetersenii* serovar Hardjobovis. *Microb Pathog* **27:** 105-117, 1999.
9) Lee TY, Lee TJ, Belisle JT, Brennan PJ, Kim SK: A novel repeat sequence specific to *Mycobacterium tuberculosis* complex and its implications. *Tuber Lung Dis* **78:** 13-19, 1997.
10) Lindquist L, Kaiser R, Reeves PR, Lindberg AA: Purification, characterization and HPLC assay of *Salmonella* glucose-1-phosphate thymidylyl-transferase from the clon-

ed *rfbA* gene. *Eur J Biochem* **211:** 763-770, 1993.

11) Ma Y, Mills JA, Belisle JT, Vissa V, Howell M, Bowlin K, Scherman MS, McNeil M: Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the *Mycobacterium tuberculosis* gene encoding alpha-D-glucose-1-phosphate thymidylyltransferase. *Microbiology* **143:** 937-945, 1997.
12) McNeil M, Daffe M, Brennan PJ: Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J Biol Chem* **265:** 18200-18206, 1990.
13) Marumo K, Lindqvist L, Verma N, Weintraub A, Reeves PR, Lindberg AA: Enzymatic synthesis and isolation of thymidine diphosphate-6-deoxy-D-xylo-4-hexulose and thymidine diphosphate-L-rhamnose. *Eur J Biochem* **204:** 539-545, 1992.
14) Melo A, Glaser L: 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, 1968.
15) Naundorf A, Klaffke W: Substrate specificity of native dTDP-D-glucose-4,6-dehydratase: chemo-enzymatic syntheses of artificial and naturally occurring deoxy sugars. *Carbohydr Res* **285:** 141-150, 1996.
16) Reeves PR, Hobbs M, Valvano MA, Skurnik M, Whitfield C, Coplin D, Kido N, Klena J, Maskell D, Raetz CR, Rick PD: Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol* **4:** 495-503, 1996.
17) Romana LK, Santiago FS, Reeves PR: 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, 1991.
18) Saigi F, Climent N, Pique N, Sanchez C, Merino S, Rubires X, Aguilar A, Tomas JM, Regue M: Genetic analysis of the *Serratia marcescens* N28b O4 antigen gene cluster. *J Bacteriol* **181:** 1883-1891, 1999.
19) Stern RJ, Lee TY, Lee TJ, Yan W, Scherman MS, Vissa VD, Kim SK, Wanner BL, McNeil MR: Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-keto-rhamnose by the *rmlC* gene products of *Escherichia coli* and *Mycobacterium tuberculosis*. *Microbiology* **145:** 663-671, 1999.
20) Stevenson G, Neal B, Liu D, Hobbs M, Packer NH, Batley M, Redmond JW, Lindquist L, Reeves P: Structure of the O antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster. *J Bacteriol* **176:** 4144-4156, 1994.
21) Thompson JD, Higgins DG, Gibson TJ, Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22:** 4673-4680, 1994.
22) Tsukioka Y, Yamashita Y, Oho T, Nakano Y, Koga T: Biological function of the dTDP-rhamnose synthesis pathway in *Streptococcus mutans*. *J Bacteriol* **179:** 1126-1134, 1997.
23) Xiang SH, Haase AM, Reeves PR: Variation of the *rfb* gene clusters in *Salmonella enterica*. *J Bacteriol* **175:** 4877-4884, 1993.
24) Yao Z, Valvano MA: Genetic analysis of the O-specific lipopolysaccharide Biosynthesis region (*rfb*) of *Escherichia coli* K-12 W3110: identification of genes that confer group 6 specificity to *Shigella flexneri* serotypes Y and 4a. *J Bacteriol* **176:** 4133-4143, 1994.
25) Zarkowsky H, Lipkin E, Glaser L: 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, 1970.