JOURNAL OF BACTERIOLOGY, Oct. 1995, p. 5539–5546 Vol. 177, No. 19
0021-9193/95/$04.00+0

Genetic Analysis of the dTDP-Rhamnose Biosynthesis Region of the *Escherichia coli* VW187 (O7:K1) *rfb* Gene Cluster: Identification of Functional Homologs of *rfbB* and *rfbA* in the *rff* Cluster and Correct Location of the *rffE* Gene

CRISTINA L. MAROLDA AND MIGUEL A. VALVANO*

Department of Microbiology and Immunology, University of Western Ontario, London, Ontario N6A 5C1, Canada

Received 10 April 1995/Accepted 5 August 1995

The O-repeating unit of the *Escherichia coli* O7-specific lipopolysaccharide is made of galactose, mannose, rhamnose, 4-acetamido-4,6-dideoxyglucose, and *N*-acetylglucosamine. We have recently characterized the genes involved in the biosynthesis of the sugar precursor GDP-mannose occurring in the *E. coli* O7:K1 strain VW187 (C. L. Marolda and M. A. Valvano, J. Bacteriol. 175:148–158, 1993). In the present study, we identified and sequenced the *rfbBDAC* genes encoding the enzymes for the biosynthesis of another precursor, dTDP-rhamnose. These genes are localized on the upstream end of the rfb_{EcO7} region, and they are strongly conserved compared with similar genes found in various enteric and nonenteric bacteria. Upstream of *rfbB* we identified a DNA segment containing the *rfb* promoter and a highly conserved untranslated leader sequence also present in the promoter regions of other surface polysaccharide gene clusters. Also, we have determined that *rfbB* and *rfbA* have homologs, *rffG* (o355) and *rffH* (o292), respectively, located on the *rff* cluster, which is involved in the synthesis of enterobacterial common antigen. We provide biochemical evidence that *rffG* and *rffH* encode dTDP-glucose dehydratase and glucose-1-phosphate thymidylyltransferase activities, respectively, and we also show that *rffG* complemented the *rfbB* defect in the $O7^+$ cosmid pJHCV32. We also demonstrate that *rffG* is distinct from *rffE* and map the *rffE* gene to the second gene of the *rff* cluster.

Bacterial outer membranes of most gram-negative microorganisms contain lipopolysaccharide (LPS), a complex molecule consisting of an O-specific polysaccharide chain attached to lipid A via the core oligosaccharide (36, 40, 48). The O-specific polysaccharide, made of repeating oligosaccharide units, is the most variable portion of the LPS molecule in terms of carbohydrate composition and chemical structure (23, 34).

Most of the specific enzymes involved in the biosynthesis of O polysaccharides are encoded by genes clustered in the *rfb* locus, which maps at 44 min on the *Escherichia coli* K-12 linkage map and at 42 min on the *Salmonella enterica* LT2 linkage map (40, 48). These enzymes are responsible for the biosynthesis and transfer of the nucleotide sugar precursors onto a lipid carrier to complete the O-repeating oligosaccharide units. Other *rfb* gene components encode products involved in the polymerization and assembly of the O-specific polysaccharide.

To elucidate the genetic organization at the molecular level of O-specific LPS genes in *E. coli* and to understand the mechanism of biosynthesis and assembly of O-specific polysaccharides, we are studying the *rfb* genes (rfb_{EcO7}) from strain VW187, a pathogenic isolate of *E. coli* O7:K1 (1, 26, 28, 44, 45). The O7 repeat consists of a linear backbone composed of *N*-acetylglucosamine (GlcNAc), 4-acetamido-4,6-dideoxyglucose, mannose, and galactose and a side chain of rhamnose α1→2-linked to the mannose residue (24). Genes governing the biosynthesis of the O7 precursors UDP-Gal and UDP-GlcNAc map outside the *rfb* region, whereas the genes for the synthesis of the remaining precursors are present within the *rfb* gene cluster. Recently, we showed that the *rfe* gene is also necessary for the formation of the O7 repeat in mediating the incorporation of GlcNAc onto the lipid acceptor, the initial step in the assembly of the O7 subunit (1).

We previously reported the genetic and functional characterization of two genes, $rfbM_{EcO7}$ and $rfbK_{EcO7}$, encoding GDP-mannose pyrophosphorylase and phosphomannomutase, respectively, which are located on the downstream end of the *rfb* cluster (26). Both genes function in the formation of GDP-mannose, and they are essential for the synthesis of the O7-specific repeating unit (26). In the present study we have localized the genes *rfbBDAC*, encoding the enzymes for the biosynthesis of dTDP-rhamnose, and the *rfb* promoter region present upstream of *rfbB*.

Work by other investigators has suggested, on the basis of sequence relatedness, possible relationships between the fourth (o355/*rffE*) and fifth (o292) genes in the *rff* cluster and *rfbB* and *rfbA*, respectively (25, 38, 42). The *rff* cluster maps at 85 min of the *E. coli* chromosome and contains genes encoding the biosynthesis of enterobacterial common antigen, an outer membrane glycolipid consisting of GlcNAc, mannosaminuronic acid (ManNAc), and 4-acetamido-4,6-dideoxygalactose (20). In this work, we present biochemical and genetic evidence that these *rff* genes, which we name *rffG* and *rffH*, are functional homologs of *rfbB* and *rfbA*, respectively. We also establish that *rffG* and *rffE* are different genes and define the correct map position of *rffE* in the *rff* cluster.

MATERIALS AND METHODS

Bacterial strains, plasmids, and materials. Strains and plasmids used in this study are described in Table 1. CLM5 is an *rff*::Tn*10-66* derivative of strain SØ874 which was constructed by P1 transduction (8) with a P1 lysate obtained from *E. coli* 21566. CLM9 is an o355::Km derivative of strain SØ874 constructed by P1 transduction with a phage lysate obtained from CLM8. The latter strain is a derivative of JC7623 which was obtained by transforming pCM134 (Table 1)

* Corresponding author. Phone: (519) 661-3996. Fax: (519) 661-3499. Electronic mail address: mvalvano@uwo.ca.

TABLE 1. *E. coli* strains and plasmids used in this study

Strain or plasmid	Relevant properties[a]	Source or reference
Strains		
DH5α	F$^-$ ϕ80*lacZ* M15 *endA recA hsdR*($r_K^-m_K^-$) *supE thi gyrA relA*? Δ(*lacZYA-argF*)U169	Laboratory stocks
MC1061-5	*araD* Δ(*ara-leu*) Δ*lacZ galU galK rpsL hsr*	S. Bourgeois
JC7623	*thr ara leuB* Δ(*gpt-proA*) *lacY sbcC tsx* qsr′$^-$ *glnV galK* λ$^-$ rac$^-$ *sbcB hisG rfbD recB recC rpsL kdgK xylA mtl argE3 thi*	13
SØ874	*lacZ trp* Δ(*sbcB-rfb*) *upp rel rpsL*	32
AB1133	*thr leuB*(Δ*gpt-proA*) *hisG argE thi rfbD lacY ara galK xyl mtl mgl rpsL kdgK supE*	B. Bachman
21566	AB1133 *rff*::Tn*10-66*	30
CLM4	SØ874 *recA*	26
CLM5	SØ874 *rff*::Tn*10-66*	This work
CLM8	JC7623 o355::Km	This work
CLM9	SØ874 o355::Km	This work
Plasmids		
pGEM3	Cloning and sequencing vector; Ap^r	Promega Biotec
pJHCV31	Cosmid clone containing rfb_{EcO7}; Tc^r O7$^+$	44
pJHCV32	Cosmid clone containing rfb_{EcO7}; Tc^r O7$^+$	44
pKI32	*Hin*dIII-*Nru*I fragment from λ*nfrC37* in pZ105; Ap^r	17
pTL61T	Promoterless cloning vector; Ap^r	22
pUC18	Cloning vector; Ap^r	46
pCM40	1.8-kb *Pst*I fragment from pCM10 cloned into pGEM3; Ap^r	This work
pCM48	1.7-kb *Sal*I-*Hin*dIII fragment from pCM10 cloned into pGEM3; Ap^r	This work
pCM111	5.67-kb *Eco*RI fragment from pJHCV31 cloned into pGEM3; Ap^r	This work
pCM117	1.2-kb *Eco*RI-*Hin*dIII fragment from pCM10 cloned into pTL61T; Ap^r *lac*$^+$	This work
pCM118	2.1-kb *Cla*I-*Hin*dIII fragment from pCM10 cloned into pUC18; Ap^r	This work
pCM122	2.1-kb *Cla*I-*Hin*dIII fragment from pCM10 cloned into pACYC184 (6); Cm^r	This work
pCM127	685-bp *Eco*RI-*Sph*I fragment from pCM10 cloned into pTL61T; Ap^r *lac*$^+$	This work
pCM134	Km^r gene of pUC4K (49) inserted in *Bst*EII site of pMAV35-4; Ap^r	This work
pLD1	2.3-kb *Hin*dIII fragment from pJHCV31 cloned into pGEM3; Ap^r	This work
pMAV23	4.8-kb *Sph*I fragment of pCM111; Ap^r $rfbBDAC_{EcO7}^+$	This work
pMAV30	1.74-kb *Hin*cII fragment of pKI32 in pUC18; Ap^r	This work
pMAV35-4	1.13-kb PCR fragment containing o355 cloned in pUC18; Ap^r	This work

[a] Ap, ampicillin; Tc, tetracycline; Cm, chloramphenicol; Km, kanamycin.

previously linearized by digestion with *Sca*I. Ap^s Km^r transformants suggested integration of the mutated o355 gene into the chromosome of JC7623, which was confirmed by PCR using oligonucleotide primers 19 and 20 (see below).

Chemicals and antibiotics were obtained from either Sigma Chemical Co. (St. Louis, Mo.) or Boehringer Mannheim (Dorval, Quebec, Canada). Restriction enzymes, T4 DNA ligase, *Taq*I DNA polymerase, and the nonradioactive DIG DNA Labeling and Detection Kit were purchased from Boehringer Mannheim. A T7 Sequencing Kit was purchased from Pharmacia LKB Biotechnology (Baie d'Urfe, Quebec, Canada). The DNA-directed transcription-translation system was purchased from Amersham Canada (Oakville, Ontario, Canada). Kits and enzymes were used under the conditions suggested by the suppliers.

Recombinant DNA methods. Isolation and electrophoresis of plasmid DNA were carried out as described elsewhere (28, 43). Chromosomal DNA was isolated by the method of Owen and Borman (35). Transformations were carried out either by the calcium chloride method (7) or by electroporation with a Gene Pulser apparatus (Bio-Rad Laboratories Ltd., Mississauga, Ontario, Canada).

DNA sequencing and analysis. DNA was sequenced by the dideoxy method of Sanger et al. (39) modified for use with the T7 Sequencing Kit obtained from Pharmacia LKB Biotechnology. Various overlapping fragments spanning a 4.8-kb *Pst*I-*Mun*I fragment (Fig. 1 and Table 1) were cloned into pGEM3. These recombinant plasmids were sequenced by using the T7 or SP6 promoter primers as sequencing endpoints. Nested deletions from pCM48, pCM111, and pLD1 (Table 1 and Fig. 1) obtained by the method described by Henikoff (11) were utilized to complete and confirm the DNA sequence of some regions. Specific nucleotide primers were designed to cover some gaps in the DNA sequence.

DNA sequences were analyzed with the Genetics Computer Group package, version 7 (10). The BLAST program (2) was used for searching protein and DNA sequence databases. Paired sequencing alignments and multiple alignments were done with the Genetics Computer Group programs BESTFIT and PILEUP, respectively.

PCR. To map the location of the *rff*::Tn*10-66* insertion, chromosomal DNA cleaved with various restriction endonucleases was used as a source of template for PCRs performed according to standard methods (14). The oligonucleotide primers were purchased from Procyon/Vetrogen (London, Ontario, Canada), and they were as follows: 14, 5′-ACTGCATTGCTGTTGATCC-3′, located 748 bases downstream of the initiation codon of *rffD*; 15, 5′-CATGACGAGCTTCG GACTGA-3′, located 366 bases downstream of the initiation codon of *rffH* (o292), complementary strand; 16, 5′-CCACGTCGTGCCTTCCTGAT-3′, located 103 bases upstream of the initiation codon of *rffE*; 17, 5′-CAGGTTCGAC GATATGGAT-3′, located 153 bases downstream of the initiation codon of *rffD*, complementary strand; 19, 5′-CAAAAGGAGTCTGGCGCTATGAGAA-3′, located 18 bases upstream of the initiation codon of *rffG* (o355); 20, 5′-AACCGC CCGCCAGCATAATA-3′, located 27 bases upstream of the initiation codon of *rffH* (o292), complementary strand; 21, 5′-CACGGATCAACAGCAATGCAG T-3′, located 811 bases downstream of the initiation codon of *rffD*, complementary strand; and 22, 5′-ATCCATATCGTCGAACCT-3′, located 135 bases downstream of the initiation codon of *rffD*. PCR products were concentrated with ethanol, and 5-μl aliquots were examined in 0.7% agarose gels. The cloning of PCR products in pUC18 to construct pMAV35-4 was carried out according to the procedure described by Wang et al. (47) using primers 19 and 20 with pKI32 DNA as a template.

Hybridizations. Hybridization experiments using a nonradioactive detection system were carried out to investigate the presence of the *rfbBDAC* genes in different serotypes of *E. coli*, *Shigella flexneri*, *Shigella boydii*, and *Pseudomonas aeruginosa* that contain rhamnose as part of their O side chain. A 3,125-bp *Eco*RV-*Nco*I fragment containing parts of *rfbB* and *rfbC* and whole *rfbD* and *rfbA* genes was isolated from pCM111 (Fig. 1, Rha probe). The DNA fragment was recovered by electrophoresis into a strip of DE-81 filter paper (Whatman, Inc., Clifton, N.J.) as previously described (28) and labeled by random incorporation of digoxigenin-labeled dUTP as recommended by the supplier. Southern blot hybridizations of genomic DNA cleaved with *Eco*RI were carried out at 42°C and followed by three washes performed at room temperature with 2×, 0.5×, and 0.1× 150 mM sodium chloride–15 mM sodium citrate–0.1% sodium dodecyl sulfate (SDS). DNA hybrids were detected with an anti-digoxigenin-alkaline phosphatase antibody conjugate upon reaction with 5-bromo-4-chloro-3-indolylphosphate and 4-nitroblue tetrazolium chloride.

Analysis of polypeptide gene products. The 5.6-kb *Eco*RI fragment of pCM111, isolated from an agarose gel, served as a template for the synthesis of mRNA and polypeptides with a DNA in vitro transcription-translation kit. Upon labeling with [^{35}S]methionine (ICN Pharmaceuticals, Montreal, Quebec, Canada), polypeptides were separated by SDS-polyacrylamide gel electrophoresis (PAGE) and treated with En^{3}Hance (NEN Research Products, Boston, Mass.). The dried gel was exposed to a Kodak X-Omat film at −110°C for 24 to 48 h. For the in vivo expression of polypeptides, the maxicell method was used (5). Labeled polypeptides were separated by SDS-PAGE, and the gel was processed as described above.

Enzyme assays. DH5 cells containing plasmids pTL61T, pCM117, and pCM127 were incubated for 3 h, and 100-μl aliquots were processed to deter-

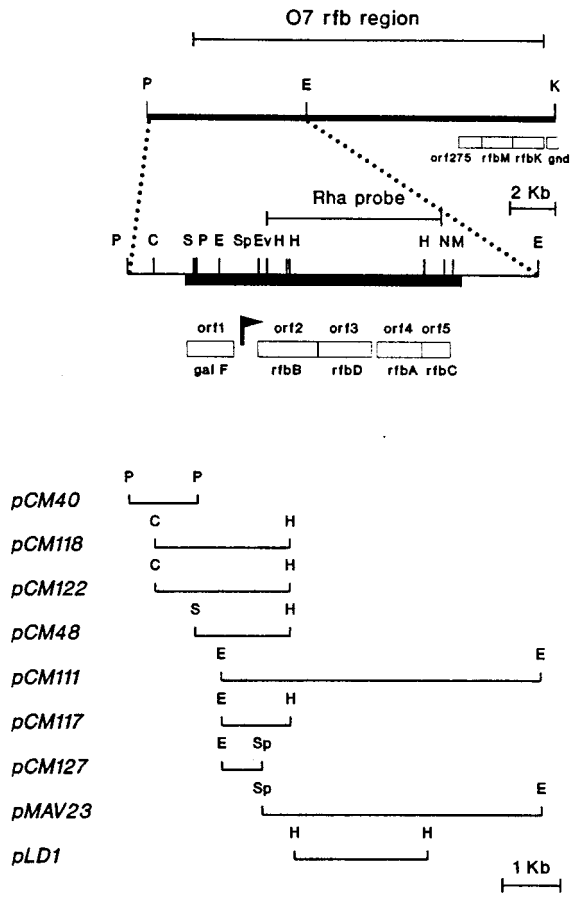

FIG. 1. Genetic analysis of the rfb_{EcO7} region. Restriction endonuclease maps of the 15.9-kb *Pst*I-*Kpn*I fragment encompassing the O7 *rfb* region and the expanded 7.25-kb *Pst*I-*Eco*RI fragment containing the dTDP-rhamnose biosynthesis genes and the *galF* gene are shown. ORF275, *rfbM*, *rfbK*, and part of *gnd* were identified previously (26). Rha probe indicates the boundaries of the probe used for hybridization experiments. The thick line underneath the 7.25-kb *Pst*I-*Eco*RI fragment indicates the 4,860-bp fragment that has been sequenced. ORF1 (*galF*), ORF2 (*rfbB*), ORF3 (*rfbD*), ORF4 (*rfbA*), and ORF5 (*rfbC*) were identified by DNA sequencing. The flag upstream of ORF2 indicates the location of the O7 *rfb* promoter region and the direction of transcription. P, *Pst*I; E, *Eco*RI; K, *Kpn*I; C, *Cla*I; S, *Sal*I; Sp, *Sph*I; Ev, *Eco*RV; H, *Hin*dIII; N, *Nco*I; M, *Mun*I.

mine β-galactosidase activity as described elsewhere (28). Glucose-1-phosphate thymidylyltransferase activity was measured by monitoring the change in concentration of dTTP and dTDP-glucose by high-pressure liquid chromatographic (HPLC) analysis exactly as described by Lindquist et al. (21). The activity of dTDP-D-glucose 4,6-dehydratase (dTDP-glucose oxidoreductase) was assayed spectrophotometrically by taking advantage of the increased light absorption at 320 nm of the reaction product dTDP-6-deoxy-D-*xylo*-hexulose in the presence of 0.1 M NaOH (29). The initial rate was expressed in enzyme units, defined as micromoles of product formed per min. The combined activities of dTDP-6-deoxy-D-*xylo*-4-hexulose 3,5-epimerase and NADPH:dTDP-6-deoxy-L-*lyxo*-4-hexulose 4-reductase resulting in the synthesis of dTDP-rhamnose were determined essentially as described by Marumo et al. (29).

LPS analysis. LPS was purified from proteinase K-treated cell envelopes and extracted with hot phenol as previously described (28). LPS extracts for analysis of lipid A-core bands were separated by electrophoresis in 10 to 20% polyacrylamide-SDS-Tricine gels (Novex, San Diego, Calif.) and stained with silver (28).

Nucleotide sequence accession number. The nucleotide sequence presented in this paper was submitted to GenBank under accession number U23775.

RESULTS

Identification of the dTDP-rhamnose biosynthesis genes. The DNA sequence of a 4,860-bp segment extending from 165 bp upstream of a *Pst*I site to the *Mun*I site, as shown in Fig. 1, revealed five open reading frames (ORFs) coding for 302 amino acids (ORF1), 384 amino acids (ORF2), 302 amino acids (ORF3), 292 amino acids (ORF4), and 180 amino acids (ORF5), all transcribed towards *gnd* (Fig. 1). No ORFs of significant length were apparent in the opposite strand of the sequenced region. ORF1 was separated from ORF2 by an untranslated DNA segment of 387 bp, and ORF3 was separated from ORF4 by an untranslated segment of 62 bp. No other extensive untranslated DNA regions were found between any of the remaining ORFs, suggesting that they are translationally coupled.

ORF1 encoded a putative polypeptide of 302 amino acids with a predicted molecular mass of 33.2 kDa and a theoretical pI of 5.5. Because of its location (Fig. 1) and its strong conservation with other *galF* genes (ranging from 92 to 97% amino acid identity [data not shown]), ORF1 was designated $galF_{EcO7}$. ORF2, ORF3, ORF4, and ORF5 encoded polypeptide products showing strong amino acid sequence conservation (ranging from 60 to 92% depending on the specific strain source used for the comparisons [data not shown]) with the products of the *rfbBDAC* genes in other microorganisms, which are known to be involved in the biosynthesis of dTDP-rhamnose, and they were designated *rfbB*, *-D*, *-A*, and *-C*, respectively. The *rfbBDAC* gene block encodes the enzymes glucose-1-phosphate thymidylyltransferase (RfbA), dTDP-D-glucose 4,6 dehydratase (RfbB), dTDP-4-keto-L-rhamnose 3,5 epimerase (RfbD), and dTDP-6-deoxy-L-*lyxo*-4-hexulose reductase (RfbC).

Analysis of the deduced primary amino acid sequences of $RfbB_{EcO7}$, $RfbC_{EcO7}$, $RfbA_{EcO7}$, and $RfbD_{EcO7}$ predicted polypeptides with molecular masses of 43.1, 32.8, 32, and 21 kDa, respectively. The expression of these polypeptides was examined by using an in vitro transcription-translation system with the isolated 5.6-kb *Eco*RI fragment of pCM111 (Fig. 1) as a DNA template. Three distinct bands with apparent molecular masses of 45, 32, and 19 kDa appeared in the autoradiograph (Fig. 2, lane A). The 45- and the 19-kDa bands were in close agreement with the predicted molecular masses of RfbB and RfbC, respectively. Since RfbD and RfbA have almost identical molecular weights, the 32-kDa band was interpreted as a doublet.

The enzyme activities encoded by the *rfbBDAC* genes were determined by using cell extracts of strain SØ874 containing pMAV23 (Fig. 1 and Table 1), a subclone from pCM111 in which the $rfbBDAC_{EcO7}$ genes were placed under the control of the *lac* promoter to ensure the overexpression of the gene products. The extracts displayed glucose-1-phosphate thymidylyltransferase and dTDP-glucose dehydratase activities, as well as the enzyme activities corresponding to RfbC and RfbD (data not shown). A low dTDP-glucose dehydratase activity was detected in a control extract prepared from strain SØ874 carrying the vector pUC18 (see below).

The coding sequence of *rfbC* had a G+C content of 33.6%, which is lower than the typical G+C content of *E. coli* (55%), whereas the G+C contents of *galF*, *rfbB*, *rfbD*, and *rfbA* were 51.6, 42.9, 47.3, and 43.9%, respectively.

Conservation of dTDP-rhamnose biosynthesis genes in other enteric bacteria. The strong conservation of the $rfbBDAC_{EcO7}$ genes in relation to those of *Salmonella* strains (15, 36), *E. coli* K-12 (42, 50), and *Shigella dysenteriae* type 1 (18) prompted us to investigate whether these genes are also conserved in other strains of *E. coli* and *Shigella* known to posses rhamnose in their O antigens. Upon digestion of chromosomal DNA with *Eco*RI, DNA fragments were separated in a 0.7% agarose gel and Southern blots were hybridized with the Rha probe (Fig. 1). Genomic DNA of *E. coli* K-12 strain W3110; *E. coli* B; *E. coli* serotypes O1, O2, O4, O75, and O141; *S. flexneri*; and *S.*

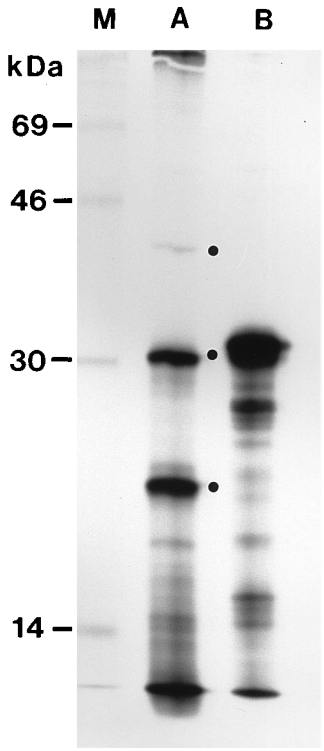

FIG. 2. Autoradiography of polypeptides expressed in an in vitro transcription-translation system. DNA templates were processed as described in Materials and Methods, and polypeptides were labeled with [^{35}S]methionine. Samples were run on an 12% SDS–PAGE gel. The following DNA templates were used: pCM111 (lane A) and pGEM3 (lane B). Dots indicate the polypeptides expressed by the DNA insert in pCM111. The positions of ^{14}C-labeled molecular mass markers (M) are as follows: bovine serum albumin, 69 kDa; ovalbumin, 46 kDa; carbonic anhydrase, 30 kDa; and lysozyme, 14 kDa.

boydii type 12 gave positive signals with this probe, and in most cases the hybridization was localized to a 5-kb *Eco*RI fragment (data not shown), suggesting that the sequence conservation extended further upstream of the *rfb* cluster. DNA purified from the *E. coli* K-12 strain CLM4 (Table 1), which carries a deletion eliminating the *rfb* cluster, did not hybridize with the probe and served as a negative control for these experiments. These results demonstrate that the dTDP-rhamnose biosynthetic genes are strongly conserved in enteric organisms. In contrast, genomic DNAs of six *P. aeruginosa* serotypes (O3, O4, O6, O10, O13, and O14) which also contain rhamnose in their O antigens did not hybridize with the probe under our experimental conditions.

Genes in the *rff* region encode RfbB and RfbA homologs. The sequences of the *rfbBDAC* region and *rfbMK* (27) permitted us to exactly position the rfb_{EcO7} cluster in the original O7$^+$ cosmid clones pJHCV31 and pJHCV32 (Fig. 3a) (44). Even though pJHCV32 lacks the DNA segment containing *galF* and part of *rfbB* (Fig. 3a), it can direct the expression of O7 polysaccharide not only in standard *E. coli* K-12 strains carrying an intact *rfbB* gene but also in strain SØ874, which has a deletion removing the rfb_{EcK12} region (44), suggesting that a gene located outside the rfb_{EcK12} region can complement the *rfbB* function. Homology searches in the databases revealed strong identities in the amino acid sequences of $RfbB_{EcO7}$ and $RfbA_{EcO7}$ with o355 and o292, respectively, which are both part of the *rff* cluster (Fig. 3b) (9, 30). The *rff* cluster is located at 85 min of the chromosomal map and contains a block of genes involved in the biosynthesis of enterobacterial common antigen (20). It has been proposed by others that o355 and o292 may be functional homologs of *rfbB* and *rfbA*, respectively (25, 38, 42).

To determine the functions of the o355 and o292 genes, we constructed plasmids pMAV35-4 and pMAV30 (Fig. 3b). pMAV35-4 was constructed by cloning into pUC18 a DNA fragment containing the o355 gene which was obtained by PCR as described in Materials and Methods. pMAV30 was obtained by cloning a 1.8-kb *Hin*cII fragment of pKI32 (17) containing part of o355 and the entire o292 gene into the *Sma*I site of pUC18 in an orientation such that a translational fusion was generated between the N-terminal segment of β-galactosidase from pUC18 and a truncated o355 polypeptide. Since o355 and o292 appear to be translationally coupled, this strategy ensures expression of the o292 polypeptide under the control of the *lac* promoter of pUC18. Cell lysates from strain SØ874 containing

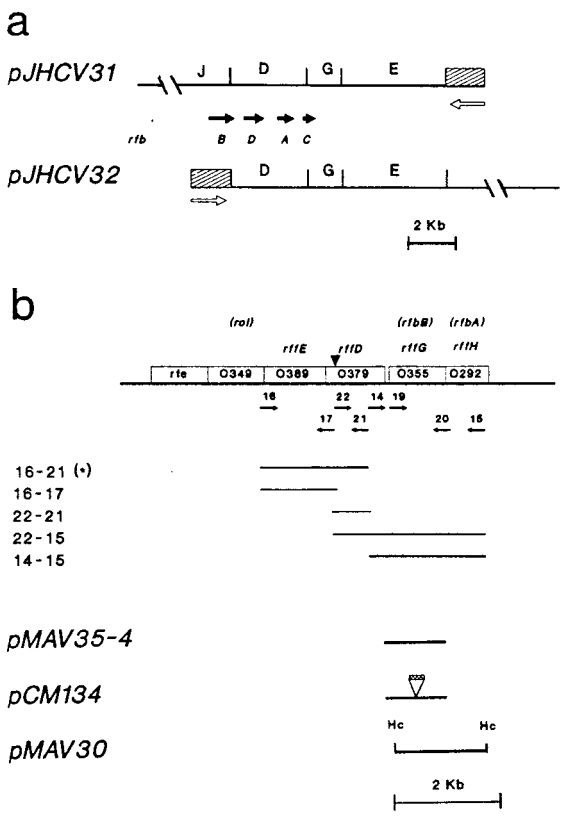

FIG. 3. Genetic analysis of the O7 *rfb* region in pJHCV31 and pJHCV32 and the *rff* cluster of *E. coli* K-12. (a) Partial map of O7$^+$ cosmid clones pJHCV31 and pJHCV32 (44). J, D, G, and E indicate some of the various *Hin*dIII fragments contained in the cosmids. Diagonal parallel bars indicate the remainder of the DNA inserts in pJHCV31 and pJHCV32. Closed arrows indicate the positions of *rfbB* (*B*), *rfbD* (*D*), *rfbA* (*A*), and *rfbC* (*C*). Hatched boxes represent the Km^r gene present in the vector pVK102. Open arrows indicate the direction of transcription of the Km^r gene. (b) Partial map of the *E. coli* K-12 *rff* cluster (*rfe*, o349, o389 [*rffE*], o379 [*rffD*], o355 [*rffG*], o292 [*rffH*]). Genes in parentheses are homologous to the corresponding *rff* genes located beneath them, but they map within or near the *rfb* cluster. The closed inverted triangle indicates the position of the *rff*::Tn*10-66* insertion in strain 21566. Small arrows indicate the positions and orientations of the primers generated for PCR. Thin lines indicate the boundaries of the DNA fragments amplified with primer pairs 16-21, 16-17, 22-21, 22-15, and 14-15 with 21566 (*rff*::Tn*10-66*) and AB1133 DNAs as templates. The asterisk indicates that this fragment was not amplified when 21566 (*rff*::Tn*10-66*) DNA was used as a template. The boundaries of DNA inserts cloned in pMAV35-4, pCM134, and pMAV30 are also indicated. The triangle with the hatched box shows the position of the Km^r gene that was inserted in pCM134. Hc, *Hin*cII.

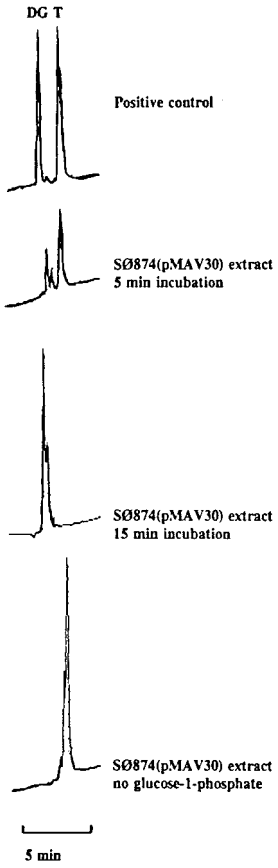

FIG. 4. Comparison of the HPLC elution profiles of dTDP-D-glucose formation at different incubation times. Glucose-1-phosphate thymidylyltransferase was assayed as described previously (21). Reaction products were separated by high-performance ion-exchange chromatography on an LC-SAX column (25 by 0.46 cm) by using a linear gradient of 50 to 400 mM potassium phosphate (pH 4.0) (21). The A_{250} was recorded. DG, dTDP-D-glucose; T, dTTP.

either pMAV30 or pUC18 were examined for glucose-1-phosphate thymidylyltransferase activity. Figure 4 shows that an extract from SØ874(pMAV30) mediates the conversion of dTTP into dTDP-glucose, but such a conversion was not detected in the absence of glucose-1-phosphate (Fig. 4) or in a cell extract prepared from SØ874 containing the vector pUC18 (data not shown). The lack of enzyme activity in the latter could be due to a low level of expression of the chromosomal o292 gene. Detection of dTDP-glucose dehydratase was carried out in cell extracts of SØ874(pMAV35-4) which displayed high levels of enzyme activity (3,774 U/mg of protein) compared with the control extract prepared from SØ874(pUC18), which yielded 83 U/mg of protein. Therefore, we provide biochemical evidence demonstrating that o355 and o292 encode dTDP-glucose dehydratase and glucose-1-phosphate thymidylyltransferase, respectively. We propose that o355 and o292 be designated *rffG* and *rffH*, respectively.

To investigate whether *rffG* is responsible for complementing the *rfbB* defect in pJHCV32, we constructed strain CLM9, a derivative of SØ874 with a Km^r gene cassette inserted into the *rffG* gene. CLM9 transformants carrying pJHCV32 were unable to express O7 LPS as determined by slide agglutination, whereas those containing pJHCV31 were slide agglutination positive. O7 LPS expression in CLM9(pJHCV32) was regained upon transformation with either pMAV35-4 containing the intact *rffG* gene or pCM111 carrying the *rfbBDAC*$_{EcO7}$ gene block. These experiments demonstrate that *rffG* can complement the function of the *rfbB* gene, and they also explain why pJHCV32 can express O7 LPS in strain SØ874.

***rffG* (o355) and *rffE* are different genes.** In a previous report, o355 was considered to be the *rffE* gene (9) encoding the UDP-GlcNAc-2-epimerase involved in the synthesis of Man-NAc, another precursor of enterobacterial common antigen (30, 37). The *rff*::Tn*10-66* insertion, which has been shown to be associated with the absence of epimerase activity (30), was introduced into SØ874 by P1 transduction, and the resulting strain, CLM5, was transformed with pJHCV32. CLM5(pJHCV32) was able to express O7-positive agglutination, suggesting that the Tn*10-66* insertion is not located in *rffG* (o355). Since we have also shown in this work that *rffG* (o355) encodes a dTDP-glucose dehydratase and that the inactivation of this gene leads to a lack of O7 LPS expression in SØ874 cells carrying pJHCV32, we have demonstrated that *rffG* (o355) and *rffE* are not the same gene. An examination of the amino acid sequence homologies of ORFs located between *rfe* and *rffG* (Fig. 3b) revealed that o349 has a strong sequence conservation with Rol proteins, which are involved in the regulation of O-polysaccharide length (3, 4, 31), whereas o389 displays 63% amino acid identity with a plasmid-encoded UDP-GlcNAc-2-epimerase (16). On the basis of this information we propose that o389 is the *rffE* gene (Fig. 3b). This is consistent with recent findings by Robertson et al. (38) showing that *rffE* is located upstream of *rffG* (o355). o379 (Fig. 3b), which corresponds to the *rffD* gene (9), has strong amino acid conservation with dehydrogenases (data not shown) as expected, taking into account the fact that the *rffD* product is the UDP-ManNAc dehydrogenase (30).

Strain 21566 has a structural defect in both *rffE* and *rffD* genes. Other investigators reported that *E. coli* 21566 carrying the *rff*::Tn*10-66* insertion is deficient in both UDP-ManNAc dehydrogenase (the product of the *rffD* gene) and UDP-GlcNAc-2-epimerase (the product of the *rffE* gene) enzyme activities (30). We determined the position of the *rff*::Tn*10-66* insertion within the *rff* cluster by a PCR strategy involving the use of various sets of primers spanning the region between o389 and o292 (Fig. 3b). DNA fragments were amplified from chromosomal DNA of *E. coli* 21566 (*rff*::Tn*10-66*) and its isogenic parent AB1133. The use of primer pairs 22-15 and 14-15 (Fig. 3b) yielded DNA fragments of 2.5 and 1.9 kb, respectively, with both chromosomal DNA templates (Fig. 5, lanes C, D, I, and J), indicating that none of the fragments amplified with each of these primers contained the Tn*10-66* insertion. In contrast, the primer pair 16-21 failed to produce a DNA fragment amplified with the chromosomal template of strain 21566, whereas a fragment of the expected molecular mass (2 kb) was identified with chromosomal DNA from AB1133 (Fig. 5, lanes G and H). The lack of amplification with 21566 DNA and primers 16 and 21 is consistent with a large DNA insertion, presumably that of the 9-kb Tn*10* element, not detected under our PCR conditions, which are not suitable to amplify fragments of DNA larger than 4 to 5 kb. We concluded that the Tn*10-66* insertion has to be located within the 2-kb fragment amplified by primers 16 and 21. To localize more precisely the Tn*10-66* insertion, primer pairs 16-17 and 22-21 were used (Fig. 3b). In both cases PCR products were obtained with 21566 DNA (Fig. 5, lanes A, B, E, and F). Taking into account the fact that primers 17 and 22 are complementary, we conclude that the Tn*10-66* insertion must be located in the region of the 19 bp spanned by these primers which is part of *rffD*. We also noticed that the extension product obtained with 21566 DNA and primers 16 and 17 was approximately 120 bp longer than that obtained with AB1133 DNA (Fig. 5, lanes A and B), suggesting a small DNA insertion within this region, most of which spans the *rffE* gene.

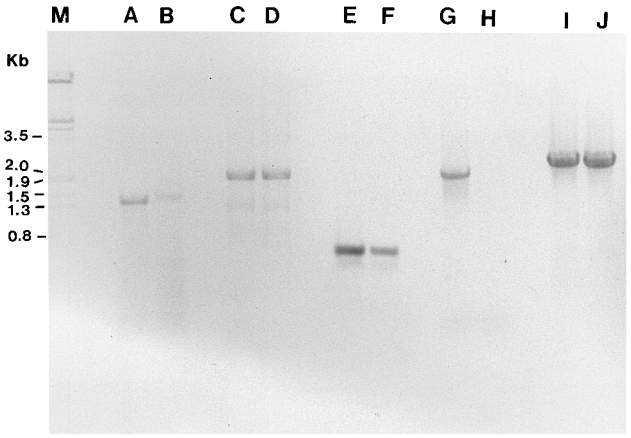

FIG. 5. DNA fragments amplified by PCR. *Eco*RI-cleaved chromosomal DNA samples were amplified by PCR using different primer pairs (Fig. 3b), and aliquots were run in a 0.7% agarose gel stained with 50 μg of ethidium bromide per ml. DNA from strain AB1133 (lanes A, C, E, G, and I) and DNA from strain 21566 (*rff*::Tn*10-66*) (lanes B, D, F, H, and J) were used as templates with the following primer pairs: 16-17 (lanes A and B), 14-15 (lanes C and D), 22-21 (lanes E and F), 16-21 (lanes G and H), and 22-15 (lanes I and J). Only the molecular masses of the relevant fragments of DNA Molecular Weight Marker III (M; Boehringer Mannheim) are indicated.

Altogether, our results strongly indicate that strain 21566 has a small insertion in the *rffE* (o389) gene and a Tn*10* insertion in *rffD*, and this explains the lack of the enzymatic activities mediated by both genes as determined by Meier-Dieter et al. (30).

Identification of a promoter region upstream of *rfbB*. The fact that the isolated *Eco*RI fragment of pCM111 was able to direct the expression of Rfb polypeptides (Fig. 2) suggested that it contains a functional promoter. This idea was verified by cloning a 1.2-kb *Eco*RI-*Hin*dIII fragment into pTL61T, a promoter cloning vector especially designed to construct transcriptional fusions to a promoterless *lac* operon (22). Cells containing the resulting plasmid, pCM117 (Fig. 1), expressed 7,681 U of β-galactosidase activity, in contrast to the 680 U (11-fold less activity) produced by cells containing the cloning vector pTL61T, confirming that a promoter region is present within the 1.2-kb fragment. A 535-bp *Sph*I-*Hin*dIII fragment of pCM117 was deleted, removing the coding region of the *rfbB* gene (Fig. 1). The resulting construct, designated pCM127, displayed 10,888 U of β-galactosidase activity. The 685-bp insert of pCM127 contains 322 bp corresponding to the carboxyl end of the *galF* gene coding region which are probably not required for promoter activity. The remaining 363 bp include the intergenetic region between *galF* and *rfbB* (Fig. 1), which could accommodate a promoter region. A computer analysis carried out with the program created by O'Neill (33) confirmed the potential of this region to harbor promoter elements and predicted possible locations for the −35 and −10 regions (Fig. 6). These regions in the rfb_{EcO7} sequence are identical to corresponding sequences in *Salmonella enterica* (15, 36) and *S. flexneri* (25) also predicted to be −10 and −35 candidate regions (Fig. 6). Moreover, the rfb_{EcO7} promoter region contains a conserved 39-bp sequence, highly rich in G's and C's and with a strong probability to form mRNA secondary structures, designated the JUMPstart element (12), which is present in a similar location in other gene clusters involved in the synthesis of surface polysaccharides in various microorganisms (Fig. 6).

DISCUSSION

In previous studies we delineated the boundaries of the O7 LPS gene cluster by deletion mapping and transposon mutagenesis (28, 44), and we also characterized in detail its downstream end, which is flanked by the *gnd* gene and contains the genes *rfbM* and *rfbK*, both involved in the biosynthesis of the O7 oligosaccharide precursor GDP-mannose (26). In the present work we examined the upstream endpoint of the rfb_{EcO7} region, which contains the *rfbBDAC* genes for the biosynthesis of another precursor, dTDP-rhamnose. The *rfbBDAC* gene polypeptides were expressed, and they were found to be functional by enzyme assays of the dTDP-rhamnose biosynthesis pathway. Upstream of *rfbB*, we identified the *galF* gene and a promoter region for the O7 *rfb* cluster.

We observed a gradual decrease in the G+C content from *rfbB* through *rfbC*, and the same occurred on the downstream end of the *rfb* cluster, where $rfbM_{EcO7}$ displays a G+C content lower than that of $rfbK_{EcO7}$ (26). Similar observations were made with *rfb* clusters of *S. enterica* serogroups (15, 36), *S. dysenteriae* (18), and *E. coli* K-12 (42, 50), in which the central *rfb* genes display a G+C content lower than that typical of the host strain. These central regions contain in general genes for the specific glycosyltransferases, the polymerase, and some of the unique carbohydrate biosynthesis pathways (36). For rfb_{EcO7}, the central region encompassed between *rfbBDAC* and *rfbMK* contains genes for the synthesis of acetamido-dideoxyglucose, a polymerase gene, a transporter gene, and genes for four glycosyltransferases (27). This finding is consistent with the idea that the central regions of the *rfb* clusters are the most variable genetically and with the suggestion that they may have originated from multiple sources with G+C content lower than that of *E. coli* or *Salmonella* spp. (36).

Hybridization experiments demonstrated that the *rfbBDAC* gene block is well conserved in other enteric strains producing

```
                          -35 ?                    -10 ?
ecrfbo7   CAGGAGTTTTCCTTGTTTCTACAGCTGTTTGGTAAGACAATTAGCGTTTGAATTTTTCGGATTTGGCACGAGG
sfrfbaj   CCAGGATTTTCCTTGTTTCTACGGCTGTTTGGTAAGACAATTAACGTTTGAATTTTTCGGATTTGGCGCGA-G
 serfbb   -AAGATTAATCCTTGTTTCCGGATGCAATTAATAAGACAATTAGCGTTTAAGTTTTAGTGAGCTTTGCCCTGC

ecrfbo7   TGGGTTACGCTCGTCACTAT----CGTGGAAATGTACG  CAGTGCTCTGGTAGCTGTTAAGCCAGGGGCGGTAGCGTG
sfrfbaj   TGGCTAACGCTCGCCAC-AT----CGTTGAGATGCATG  CAGTGCACTGGTAGCTGTTAAGCCAGGGGCGGTAGCGTG
 serfbb   TGGGCGAGGTTT------GCAACAAGTCGATATGTACG  CAGTGCACTGGTAGCTGATGAGCCAGGGGCGGTAGCGTG

ecrfbo7   CATTAATGCATCTATTAATCAAATCTGGAGCAGTCTATTTCACAGCATGCTCTCTGGCTATATGGAATAAAAAAGTG
sfrfbaj   TATTAATACCTCTATTAATCAAATCTGGAGCAGTCTATTTCACAGCATGCTCTATAGCTATATGGAATAAAAAAGGTG
 serfbb   T--GTAACGACTTGAGCAATTAATTTTTATTGGCAAATTAAATACCACATTAAATACGCCTTATGGAATAGAAAAGTG
```

FIG. 6. Comparison of the rfb_{EcO7} promoter region (ecrfbo7) with *rfb* promoter regions in *S. enterica* LT2 (serfbb) and *S. flexneri* 2a (sfrfbaj). Bold type with single underlining indicates −35 and/or −10 regions. Double underlining indicates initiation codons. Sequences within the rectangle correspond to the 39-bp JUMPstart element (12). Dashes indicate gaps in the alignment.

O polysaccharides with rhamnose. The absence of hybridization signals with chromosomal DNA from different *Pseudomonas* strains may be due to a low level of homology with respect to the enteric *rfbBDAC* genes. Recently, Köplin et al. (19) have characterized and sequenced the genes encoding the enzymes for the dTDP-rhamnose pathway in *Xanthomonas campestris* and found them to be in a different order, *rfbBADC*. The $rfbBDAC_{EcO7}$ genes displayed about 60% or less overall homology at the DNA sequence level with the *Xanthomonas rfbBADC* genes, although considerable similarities were observed when the corresponding protein products were compared (data not shown). Since *Pseudomonas* strains are genetically more closely related to strains of the genus *Xanthomonas* than they are to enteric bacteria, it is possible that the dTDP-rhamnose biosynthesis genes, although present in *Pseudomonas* strains, are not well detected with our rhamnose biosynthesis gene-specific probe.

Elucidation of the DNA sequence of the $rfbBDAC_{EcO7}$ genes permitted us to determine that the $O7^+$ cosmid pJHCV32 lacks most of the *rfbB* gene even though it can direct the synthesis of O7 polysaccharide in *E. coli* K-12 strains with a deletion removing the chromosomal *rfb* region. It has been suggested by other investigators that the *rff* cluster contains *rfbB* and *rfbA* gene homologs, designated o355 and o292, respectively (25, 38, 42). In this work we demonstrated that cloned o355 and o292 genes encode dTDP-glucose dehydratase and glucose-1-phosphate thymidylyltransferase activities, respectively, which are the same activities encoded by the corresponding genes *rfbB* and *rfbA*. We propose that o355 and o292 be designated *rffG* and *rffH*, respectively, and we have shown that pJHCV32 cannot support the expression of O7 LPS in strain CLM9, which has both a mutation in *rffG* and a deletion removing the *E. coli* K-12 *rfb* region. The presence of genes in the *rff* cluster which are homologous to genes in the *rfb* region is not unique. A similar duplication of functions has been reported for the GDP-mannose biosynthesis genes *rfbM* and *rfbK* with the genes *cpsB* and *cpsG* (41), respectively, which are located in the *cps* region encoding the biosynthesis of colanic acid (41).

Since we have demonstrated that *rffG* (o355) encodes dTDP-glucose dehydratase, the original assignment of o355 as *rffE* (9), the structural gene for the UDP-GlNAc-2-epimerase, is incorrect. A database search indicates that o389, located upstream of *rffD* (Fig. 3b), is conserved with a plasmid-encoded UDP-GlNAc-2-epimerase of *S. enterica* O:54 (16), and thus it is likely that o389 is *rffE*. The plasmid-encoded epimerase is the product of the $rfbC_{O:54}$ gene, and it is essential for the formation of the unusual homopolymeric ManNAc polysaccharide in this strain (16). Therefore, the gene order proposed for the upstream region of the *rff* cluster is o349, *rffE*, *rffD*, *rffG*, and *rffH* (Fig. 3b). Both *rffE* and *rffD* are involved in the biosynthesis of UDP-ManNAc, whereas *rffG* and *rffH* are likely involved in the formation of dTDP-4-acetamido-4,6-dideoxygalactose. These two precursors, together with UDP-GlcNAc, are used for the assembly of enterobacterial common antigen (37).

Meier-Dieter et al. (30) examined the biosynthesis of ManNAc in cell extracts of two *E. coli* strains with Tn*10* insertions in the *rff* region. They concluded that strain 21546 (*rff*::Tn *10-46*) possessed a defective UDP-ManNAc dehydrogenase (the product of the *rffD* gene) whereas strain 21566 (*rff*::Tn *10-66*) was defective in both UDP-GlNAc-2-epimerase (the product of the *rffE* gene) and UDP-ManNAc dehydrogenase activities, suggesting that the *rff*::Tn*10-66* insertion was located in the structural gene for the epimerase and that the lack of both epimerase and dehydrogenase activities was due to a polar effect. Our results and those of others (38) indicate that *rffE* is located upstream of *rffD*; therefore, polarity cannot be causing the double enzyme defect of strain 21566. This region in the chromosome of strain 21566 was characterized by a PCR strategy which revealed that the *rff*::Tn*10-66* insertion lies on the amino-terminal region of the *rffD* coding sequence whereas *rffE* contains a small DNA insertion. It is possible that during the Tn*10* mutagenesis two insertions took place close to one another and that further on one of the transposons was excised, leaving behind a small segment of DNA. Confirmation of this hypothesis awaits the DNA sequencing of the mutated region in strain 21566.

Also in this study, on the basis of transcriptional gene fusions to a promoterless *lac* operon, we have provided evidence supporting the existence of a promoter region located upstream of $rfbB_{EcO7}$. This region is also missing in the $O7^+$ cosmid pJHCV32 which resulted from the cloning of a partially cleaved *Hin*dIII chromosomal DNA fragment into the *Hin*dIII site of the Km^r gene of the vector pVK102 (49). Our DNA sequencing data show that because of the orientation of the insert, a translational fusion took place between the amino-terminal portion of the Km^r gene product and the carboxyl end of the RfbB protein, permitting the expression of downstream *rfb* genes under the control of the Km^r gene promoter.

ACKNOWLEDGMENTS

This work was supported by a grant from the Medical Research Council and by an NSERC-UWO Internal Research Grant.

We thank T. Viswanatha and L. Marrone for their help and advice with the HPLC experiments, Lydia Dafoe for technical assistance, and colleagues mentioned or referenced in Table 1 for providing strains and plasmids.

REFERENCES

1. **Alexander, D. C., and M. A. Valvano.** 1994. Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing *N*-acetylglucosamine. J. Bacteriol. **176:** 7079–7084.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–441.
3. **Bastin, D. A., G. Stevenson, P. K. Brown, A. Haase, and P. R. Reeves.** 1993. Repeat unit polysaccharides of bacteria: a model for polymerisation resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. Mol. Microbiol. **7:**725–734.
4. **Batchelor, R. A., G. E. Haraguchi, R. A. Hull, and S. I. Hull.** 1991. Regulation by a novel protein of the bimodal distribution of lipopolysaccharide in the outer membrane of *Escherichia coli*. J. Bacteriol. **173:**5699–5704.
5. **Boulnois, G. J., and K. N. Timmis.** 1984. Synthesis of plasmid-encoded polypeptides in maxicells, p. 204–211. *In* A. Puhler and K. N. Timmis (ed.), Advanced molecular genetics. Springer-Verlag, Berlin.
6. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable DNA cloning vehicles derived from P15A cryptic plasmid. J. Bacteriol. **134:**1141–1156.
7. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. USA **69:**2110–2114.
8. **Curtiss, R.** 1981. Gene transfer, p. 243–265. *In* P. Gerhardt, R. G. E. Murray, R. N. Costilow, E. W. Nester, W. A. Wood, N. R. Krieg, and G. B. Phillips (ed.), Manual of methods for general bacteriology. American Society for Microbiology, Washington, D.C.
9. **Daniels, D. L., G. Plunkett III, V. D. Burland, and F. R. Blattner.** 1992. Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. Science **257:**771–778.
10. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
11. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
12. **Hobbs, M., and P. R. Reeves.** 1994. The JUMPstart sequence: a 39 bp element common to several polysaccharide gene clusters. Mol. Microbiol. **12:**855–856.
13. **Horii, Z. I., and A. J. Clark.** 1973. Genetic analysis of the RecF pathway of genetic recombination in *Escherichia coli* K-12. Isolation and characterization of mutants. J. Mol. Biol. **80:**327–344.
14. **Innis, M. A., and D. H. Gelfand.** 1990. Optimization of PCR's, p. 3–12. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR

protocols. A guide to methods and applications. Academic Press, New York.
15. **Jiang, X.-M., B. Neal, F. Santiago, S. J. Lee, L. K. Romana, and P. R. Reeves.** 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar *typhimurium* (strain LT2). Mol. Microbiol. **5:**695–713.
16. **Keenleyside, W. J., and C. Whitfield.** 1995. Lateral transfer of *rfb* genes: a mobilizable ColE1-type plasmid carries the $rfb_{O:54}$ (O:54 antigen biosynthesis) gene cluster from *Salmonella enterica* serovar borreze. J. Bacteriol. **177:** 5247–5253.
17. **Kiino, D. R., R. Licudine, K. Wilt, D. H. C. Yang, and L. R. Rothman-Denes.** 1993. A cytoplasmic protein, NfrC, is required for bacteriophage N4 adsorption. J. Bacteriol. **175:**7074–7080.
18. **Klena, J., and C. Schnaitman.** 1993. Function of the *rfb* gene cluster and the *rfe* gene in the synthesis of O antigen by *Shigella dysenteriae* 1. Mol. Microbiol. **9:**393–402.
19. **Köplin, R., G. Wang, B. Hötte, U. B. Priefer, and A. Pühler.** 1993. A 3.9-kb DNA region of *Xanthomonas campestris* pv. *campestris* that is necessary for lipopolysaccharide production encodes a set of enzymes involved in the synthesis of dTDP-rhamnose. J. Bacteriol. **175:**7786–7792.
20. **Kuhn, H.-M., U. Meier, and H. Mayer.** 1988. ECA, the enterobacterial common antigen. FEMS Microbiol. Rev. **54:**195–222.
21. **Lindquist, L., R. Kaiser, P. R. Reeves, and A. A. Lindberg.** 1993. Purification, characterization and HPLC assay of *Salmonella* glucose-1-phosphate thymidyltransferase from the cloned *rfbA* gene. Eur. J. Biochem. **211:**763–770.
22. **Linn, T., and R. St. Pierre.** 1990. Improved vector system for constructing fusions that ensures independent translation of *lacZ*. J. Bacteriol. **172:**1077–1084.
23. **Luderitz, O., O. Westphal, A. M. Staub, and H. Nikaido.** 1971. Isolation and chemical and immunological characterization of bacterial lipopolysaccharides, p. 145–233. *In* G. Weinbaum, S. Kadis, and S. J. Ajl (ed.), Microbial toxins, vol. IV. Academic Press Inc., London.
24. **L'vov, V. L., A. S. Shashkov, B. A. Dimitriev, N. K. Kochtkov, B. Jann, and K. Jann.** 1984. Structural studies of the O-specific side chain of the lipopolysaccharide from *Escherichia coli* O:7. Carbohydr. Res. **126:**249–259.
25. **Macpherson, D. F., P. A. Manning, and R. Morona.** 1994. Characterization of the dTDP-rhamnose biosynthesis genes encoded in the *rfb* locus of *Shigella flexneri*. Mol. Microbiol. **11:**281–292.
26. **Marolda, C. L., and M. A. Valvano.** 1993. Identification, expression, and DNA sequence of the GDP-mannose biosynthesis genes encoded by the O7 LPS *rfb* region of strain VW187 (*Escherichia coli* O7:K1). J. Bacteriol. **175:** 148–158.
27. **Marolda, C. L., and M. A. Valvano.** Unpublished observations.
28. **Marolda, C. L., J. Welsh, L. Dafoe, and M. A. Valvano.** 1990. Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. J. Bacteriol. **172:**3590–3599.
29. **Marumo, K., L. Lindquist, N. Verma, A. Weintraub, P. R. Reeves, and A. A. Lindberg.** 1992. Enzymatic synthesis and isolation of thymidine diphosphate-6-deoxy-D-*xylo*-4-hexulose and thymidine diphosphate-L-rhamnose. Production using cloned gene products and separation by HPLC. Eur. J. Biochem. **204:**539–545.
30. **Meier-Dieter, U., R. Starman, K. Barr, H. Mayer, and P. D. Rick.** 1990. Biosynthesis of enterobacterial common antigen in *Escherichia coli*. J. Biol. Chem. **265:**13490–13497.
31. **Morona, R., L. Van den Bosch, and P. A. Manning.** 1995. Molecular, genetic, and topological characterization of O-antigen chain length regulation in *Shigella flexneri*. J. Bacteriol. **177:**1059–1068.
32. **Neuhard, J., and E. Thomassen.** 1976. Altered deoxyribonucleotide pools in P2 eductants of *Escherichia coli* K-12 due to deletion of the *dcd* gene. J. Bacteriol. **126:**999–1001.
33. **O'Neill, M. C.** 1989. Consensus methods for finding and ranking DNA binding sites—application to *Escherichia coli* promoters. J. Mol. Biol. **207:** 301–310.
34. **Orskov, I., F. Orskov, B. Jann, and K. Jann.** 1977. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriol. Rev. **41:**667–710.
35. **Owen, R. J., and P. Borman.** 1987. A rapid biochemical method for purifying high molecular weight bacterial chromosomal DNA for restriction enzyme analysis. Nucleic Acids Res. **15:**3631.
36. **Reeves, P. R.** 1994. Biosynthesis and assembly of lipopolysaccharide, p. 281–317. *In* J.-M. Ghuysen and R. Hakenbeck (ed.), Bacterial cell wall. Elsevier Science B. V., Amsterdam.
37. **Rick, P. D., H. Mayer, B. A. Neumeyer, S. Wolski, and D. Bitter-Suermann.** 1985. Biosynthesis of enterobacterial common antigen. J. Bacteriol. **162:**494–503.
38. **Robertson, B. D., M. Frosch, and J. P. M. Van Putten.** 1994. The identification of cryptic rhamnose biosynthesis genes in *Neisseria gonorrhoea* and their relationship to lipopolysaccharide biosynthesis. J. Bacteriol. **176:**6915–6920.
39. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
40. **Schnaitman, C. A., and J. D. Klena.** 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. Microbiol. Rev. **57:**655–682.
41. **Stevenson, G., S. J. Lee, L. K. Romana, and P. R. Reeves.** 1991. The *cps* gene cluster of *Salmonella* LT2 includes a second mannose pathway: sequence of two genes and relationship to genes in the *rfb* gene cluster. Mol. Gen. Genet. **227:**173–180.
42. **Stevenson, G., B. Neal, D. Liu, M. Hobbs, N. H. Packer, M. Batley, J. W. Redmond, L. Lindquist, and P. Reeves.** 1994. Structure of the O antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster. J. Bacteriol. **176:**4144–4156.
43. **Valvano, M. A., and J. H. Crosa.** 1984. Aerobactin iron transport genes commonly encoded by certain ColV plasmids occur in the chromosome of a human invasive strain of *Escherichia coli* K1. Infect. Immun. **46:**159–167.
44. **Valvano, M. A., and J. H. Crosa.** 1989. Molecular cloning and expression in *Escherichia coli* K-12 of chromosomal genes determining the O7 lipopolysaccharide antigen of a human invasive strain of *E. coli* O7:K1. Infect. Immun. **57:**937–943.
45. **Valvano, M. A., and C. L. Marolda.** 1991. Relatedness of O-specific lipopolysaccharide side chain genes from strains of *Shigella boydii* type 12 belonging to two clonal groups and from *Escherichia coli* O7:K1. Infect. Immun. **59:**3917–3923.
46. **Vieira, J., and J. Messing.** 1992. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
47. **Wang, K., B. F. Koop, and L. Hood.** 1994. A simple method using T4 DNA polymerase to clone polymerase chain reaction products. BioTechniques **17:** 236–238.
48. **Whitfield, C., and M. A. Valvano.** 1993. Biosynthesis and expression of cell-surface polysaccharides in bacteria. Adv. Microb. Physiol. **35:**135–246.
49. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
50. **Yao, Z., and M. A. Valvano.** 1994. 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.