

A Cluster of Genes Involved in Polysaccharide Biosynthesis from *Enterococcus faecalis* OG1RF

YI XU,¹ BARBARA E. MURRAY,^{2,3,4} AND GEORGE M. WEINSTOCK^{1,2,3,4*}

Department of Biochemistry and Molecular Biology,¹ Department of Microbiology and Molecular Genetics,² Division of Infectious Diseases, Department of Medicine,³ and Center for the Study of Emerging and Re-Emerging Pathogens,⁴ University of Texas Medical School, Houston, Texas 77030

Received 21 April 1998/Returned for modification 28 May 1998/Accepted 17 June 1998

Our previous work identified a cosmid clone containing a 43-kb insert from *Enterococcus faecalis* OG1RF that produced a nonprotein antigen in *Escherichia coli*. In the present work, we studied this clone in detail. Periodate treatment of lysates of the clone confirmed that the antigen was carbohydrate in nature. Analysis of DNA sequences and transposon insertion mutants suggested that the insert contained a multicistronic gene cluster. Database comparison showed that the cluster contained genes similar to genes involved in the biosynthesis of dTDP-rhamnose, glycosyltransferases, and ABC transporters involved in the export of sugar polymers from both gram-positive and gram-negative organisms. Insertions in several genes within the cluster abolished the immunoreactivity of the clone. This is the first report on a gene cluster of *E. faecalis* involved in the biosynthesis of an antigenic polysaccharide.

Many studies and reviews have discussed the important roles of polysaccharides in bacterial physiology and pathogenesis (2, 15, 46). As components of bacterial cell walls, capsules, lipopolysaccharides of gram-negative organisms, and accessory polymers of gram-positive cell walls, they not only offer structural rigidity for maintaining the shape of cells but also affect surface permeability, charge, and hydrophobicity and consequently affect the way bacteria interact with the environment. One important aspect of polysaccharides relates to their immunological and biological properties in tissues of the host, such as inhibition of phagocytosis and stimulation of cytokine production (17). These properties are critical in determining the outcome of the intricate interactions between bacteria and the host immune system. Many polysaccharides are also the immunodeterminants of serotype-specific antigens, including the O-antigen side chains of lipopolysaccharide and group and type antigens of some streptococci. Thus, it is clear that knowledge of cell surface architecture is important in understanding bacterial behavior inside the host, as well as in designing new therapeutics and serodiagnostic tools.

Biosynthesis of polysaccharides generally involves synthesis of sugar precursors in the cytoplasm, formation and polymerization of repeating units, and export to the cell surface. The formation of repeating units usually is initiated by the transfer of a sugar precursor onto a lipid carrier, such as undecaprenol-phosphate, to which subsequent sugar residues can be added. The sugar residue attached to the lipid carrier is often the first sugar in the repeating units (1, 44, 56), although there are exceptions as in *Escherichia coli* O8 and O9 serotype-specific O antigens in which the sugar residues attached to the lipid carrier do not appear in the final polysaccharides (41). Genes for the biosynthesis of polysaccharides are generally arranged in clusters of one or several transcriptional units (13, 32, 44, 56).

Enterococci (formerly group D streptococci) are among the leading causes of nosocomial infections in the United States, with the majority of clinical isolates being classified as *Entero-*

coccus faecalis (34). Besides their low-level intrinsic resistance to many antibiotics, enterococci have also developed high-level resistance to other antibiotics, including vancomycin and β -lactams (34). Very little information is available on cell surface-associated polysaccharides of enterococci. Early studies on the type-specific antigens in enterococci indicated that they were polysaccharides in the cell walls (6, 19, 45, 47). Type 1 antigen seems to be the most frequently identified type among *E. faecalis* isolates, at least in some studies (45, 47). Crude cell wall extracts from *E. faecalis* type 1 strains were found to contain rhamnose, glucose, galactose, *N*-acetylglucosamine (GlcNAc), and *N*-acetylgalactosamine (6). Pazar et al. demonstrated that enzymatic activities to synthesize thymidine diphosphate glucose (TDP-Glc) and to convert TDP-Glc to thymidine diphosphate rhamnose (TDP-Rha) were present in the cell extracts of *E. faecalis* N (40). TDP-Glc and TDP-Rha are the precursors from which glucose and rhamnose can be transferred to synthesize polysaccharides. An immunogenic glycan was also found in the cell walls of *E. faecalis* N (38, 39). The glycan was shown to be a diheteroglycan of glucose and galactose, but no rhamnose was present. No genetic information on polysaccharide biosynthesis in enterococci is available.

In a previous study, we reported a cosmid clone carrying a DNA fragment from *E. faecalis* OG1RF that produced a nonprotein antigen in *E. coli* (57). We hypothesized that the antigen could be a polysaccharide of OG1RF. To test this hypothesis, we have carried out further studies including DNA sequence analysis and characterization of transposon mutants. The results showed that the DNA fragment contains a gene cluster with homology to genes important in the assembly and export of polysaccharide from various organisms. Transposon insertions in three open reading frames (ORFs) within the cluster caused the loss of antigen expression in *E. coli*.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. Strains and plasmids used in this study are listed in Table 1. *E. coli* cells were grown in Luria-Bertani (LB) broth or on LB agar with appropriate antibiotics overnight at 37°C. Enterococci were grown in brain heart infusion (BHI) broth or on BHI agar (Difco) overnight at 37°C for routine purposes. Antibiotics were used at the following concentrations: chloramphenicol at 25 μ g/ml, kanamycin at 50 μ g/ml, and ampicillin at 10

* Corresponding author. Mailing address: Department of Microbiology and Molecular Genetics, University of Texas Medical School, 6431 Fannin St., Houston, TX 77030. Phone: (713) 500-6083. Fax: (713) 500-5499. E-mail: georgew@utmmg.med.uth.tmc.edu.

TABLE 1. Strains and plasmids used in this study

Strain, plasmid, or cosmid	Characteristics	Reference and/or source
Strains		
<i>E. coli</i>		
DH5 α	F ⁻ 80 <i>dlacZ</i> Δ (<i>lacZYA-argF</i>) <i>U169 endA1 recA1 hsdR17</i> (<i>r_K⁻ m_K⁺</i>) <i>deoR thi-1 supE44</i> λ ⁻ <i>gyrA96 relA1</i>	C. Whitfield (36)
S ϕ 874	F ⁻ <i>trp lacZ strA thi upp ΔsbcB-rfb</i>	P. Rick (42)
AB1133	K-12; <i>thr-1 leuB6</i> (Δ <i>gpt-proA</i>)66 <i>hisG4 argE3 thi-1 rfbD1 lacY ara-14 galK2 xyl-5 mtl-1 mgl-51 rpsL31 kdgK51 supE44</i>	P. Rick (42)
21548	AB1133; <i>rfe::Tn10</i> Tc ^r	Formerly called BO4B6I (57)
LO6E6II	<i>E. coli</i> library clone of OG1RF	This study
TX5159	DH5 α (pTEX5159)	
TX5169	S ϕ 874(pTEX5159)	
<i>E. faecalis</i>		
OG1RF		35
TX52	Clinical isolate from a patient with endocarditis	57
Plasmids and cosmids		
pBeloBAC11	Cosmid vector, F' replicon, Cm ^r	H. Shizuya (57)
pBluescript SK (-)	High-copy-number cloning vector derived from pUC19, Ap ^r	Stratagene
pTEX5159	pBeloBAC11 with a 43-kb insert from OG1RF which contains a gene cluster for polysaccharide biosynthesis	57
pTEX5159A-U	pTEX5159::Tn7 Cm ^r Km ^r	This study
pTEX5175	3,195-bp <i>EcoRI-HindIII</i> fragment of pTEX5159 subcloned in pBluescript SK (-), Ap ^r	This study

μ g/ml. Isopropylthio- β -D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl- β -galactoside (X-Gal) were used at 0.5 mM and 80 μ g/ml, respectively.

Patient sera were collected from patients with *E. faecalis* endocarditis infections and had high titers against *E. faecalis* strains (57).

In vitro transposition. Cosmid DNA from clone TX5159 (formerly BO4B6I) (57) was prepared by the alkaline-sodium dodecyl sulfate (SDS) method and purified in a CsCl-ethidium bromide gradient as described by Sambrook et al. (43). In vitro transposition using a Tn7-based system (3, 4, 48) was performed by Matthew C. Biery and Nancy L. Craig, Department of Molecular Biology and Genetics, Howard Hughes Medical Institute, Johns Hopkins University.

DNA manipulations and transformation of *E. coli*. DNA preparation, purification, restriction digestion, agarose gel electrophoresis, and ligation were performed by using standard methods (43) unless otherwise stated. Routine preparation of *E. coli* competent cells and transformation of DNA into *E. coli* were performed by the one-step procedure (9). Transformation of cosmid DNA that had been subjected to in vitro transposition was carried out by using Subcloning Efficiency DH5 α competent cells (GIBCO BRL, Life Technologies, Gaithersburg, Md.); transformants were randomly picked and stored in 96-well microtiter dishes in LB broth containing 15% glycerol.

Proteinase K, periodate treatment, SDS-polyacrylamide gel electrophoresis (PAGE), and Western blot analysis. Proteinase K treatment of the cell lysates of *E. coli* clones was based on a method described previously (57), with slight modifications. Briefly, 20 ml of an overnight culture of *E. coli* clones was centrifuged and resuspended in 700 μ l of phosphate-buffer saline. The suspension was split into two aliquots; 50 μ l of proteinase K (50 mg/ml; Fisher Scientific, Fair Lawn, N.J.) was added to the first aliquot, and 4 μ l of 100 mM phenylmethylsulfonyl fluoride and 50 μ l of H₂O were added to the second. The aliquots were incubated at 50°C overnight and stored at -20°C. Treatment by periodate followed part of the procedure for the ECL Glycoprotein Detection Module (Amersham Life Science, Little Chalfont, Engalnd). Twenty-five microliters of each of the above aliquots (proteinase K treated and untreated) were mixed with 25

μ l of 200 mM acetate buffer (pH 5.5.), then 25 μ l of 30 mM sodium metaperiodate (freshly prepared) was added, and the mixture was incubated at room temperature for 1 h in the dark. The reaction was stopped by addition of 25 μ l of 20 mM sodium metabisulfite and stored at -20°C in the dark. For controls, 25 μ l of 200 mM acetate buffer instead of 30 mM sodium metaperiodate was used.

For SDS-PAGE and Western blot analysis, twofold serial dilutions of the samples were made in PBS phosphate-buffered saline; 15 μ l from each dilution was mixed with 15 μ l of 2 \times loading buffer (25 mM Tris Tris-Cl [pH 6.8], 9% [vol/vol] glycerol, 2.5% [vol/vol] β -mercaptoethanol, 0.0025% bromophenol blue, 4% SDS), boiled for 5 min, and loaded onto the gel. The conditions for SDS-PAGE and Western blotting were based on the method described previously (57).

Immunoblot analysis of bacterial colonies. Colonies were inoculated onto NitroBind nitrocellulose transfer membranes (Micron Separations, Inc., Westborough, Mass.) placed on LB agar plates containing proper antibiotics and were then incubated at 37°C overnight. The membranes were lifted, placed in a chloroform chamber for 15 min, and incubated in lysis buffer. Subsequent steps were as described previously (57).

PCR amplification from the ends of the transposon and purification of PCR products. NLC94 and NLC272 are primers that allow DNA synthesis to proceed outward from the ends of the Tn7 transposon. Primers GW248, GW250, GW251, GW254, GW273, GW280, and GW283 are primers hybridizing to regions in the insert of the cosmid. The sequence and location of each primer are summarized in Table 2 (see also Fig. 3).

Miniprep DNA from cosmid clones with transposon insertions was used as templates. The GeneAmp XL PCR kit (Perkin-Elmer, Branchburg, N.J.) was used to amplify DNA by using combinations of one of the NLC primers and one of the GW primers as instructed by the manufacturer. The reactions were performed in a model 9600 GeneAmp PCR system (Perkin-Elmer) with an initial denaturation at 96°C for 1 min, a regular three-step cycling reaction with 94°C denaturation for 20 s, 50°C annealing for 20 s, and 72°C extension for 7 min for

TABLE 2. Oligonucleotides used in this study

Oligonucleotide	Sequence (5'-3')	Position (5'-3')
GW248	GAAGACATTGGTTATTTGAC	8342-8323 in contig C, complement strand
GW250	TCCTACATGCTAAAACGGTC	3152-3171 in contig C
GW251	AAAATCCACAATAAAGTTAG	2989-2970 in contig C, complement strand
GW254	TGGGATGGGAAATCGTTCTG	17646-17665 in contig C
GW273	ATAAAATGAAATGTTCCACCC	501-482 bp from the end of contig B (proximal to contig C)
GW278	GGAAACGGTGGGAGTATACG	16324-16343 in contig C
GW280	CTTTAGCAGCATTTTGGAGT	1882-1901 in contig C
GW283	TTTGTITGGTTCATCTGTC	9916-9897 in contig C, complement strand
NLC94	AAAGTCCAGTATGCTTTTTTCACAGCTAAC	End of Tn7
NLC272	ATTTTCGTATTAGCTTACGACGCTACACCC	End of Tn7

16 cycles, followed by a second three-step cycling reaction using the same temperatures as in the previous one but with an increase in extension time at 72°C for 15 s per cycle for 16 cycles, and finally an incubation at 72°C for 10 min.

Agarose gel electrophoresis and ethidium bromide staining were performed to examine the PCR products. The desired products were excised from the gels and purified by using a QIAquick gel extraction kit (Qiagen Inc., Santa Clarita, Calif.). The DNA concentration of each purified product was measured in a model 7625 Microplate Fluorometer (Cambridge Technology).

Subcloning of restriction fragments into pBluescript SK (-). Subcloning was based on the method described previously (57), with slight modifications. Cosmid DNA was digested with *EcoRI*, *BamHI*, or *DraI* and purified by phenol-chloroform extraction followed by ethanol precipitation. Vector pBluescript SK (-) DNA was digested with *EcoRI*, *BamHI*, or *EcoRV* (to ligate with *DraI*) and dephosphorylated with shrimp alkaline phosphatase (United States Biochemical, Cleveland, Ohio). After ligation, the mixture was transformed into *E. coli* DH5 α and plated on LB plates with ampicillin, IPTG, and X-Gal. White colonies were picked and analyzed by restriction digestion of miniprep DNA.

DNA sequencing and sequence analysis. Three types of templates were prepared for sequencing: DNA from pBluescript SK (-) clones prepared by using a Qiagen Plasmid Mini kit (Qiagen Inc., Chatsworth, Calif.); purified PCR products; and TX5159 cosmid DNA purified in a CsCl-ethidium bromide gradient. Primers hybridizing to the T3 and T7 regions in pBluescript SK (-) and to the ends of the Tn7 transposon, as well as primers hybridizing to the ends of the sequences from the previous rounds, were used as sequencing primers. Sequencing reactions were performed by the *Taq* Dye-deoxy Terminator method and a model 377 DNA sequencing system (Applied Biosystems, Foster City, Calif.). The GelAssemble program in the Genetics Computer Group (GCG; Madison, Wis.) software package was used to assemble the nucleotide sequences. Gapped BLASTx at the National Center for Biotechnology Information (NCBI) was used to search for homologous sequences in the database. ORF Finder at NCBI was used to identify potential ORFs in the sequences. The terminator program in the GCG package was used to search for terminator-like sequences. The ProtParam program at ExPASy was used to calculate the theoretical parameters including molecular weight, isoelectric point, and hydrophobicity of each deduced amino acid sequence. Hydrophathy plots were obtained by using the ProtScale program at ExPASy according to the method of Kyte and Doolittle (24) with a window size of 9. The DAS (dense alignment surface method) (10) transmembrane segment prediction program was also used to predict transmembrane domains. Potential signal peptide cleavage sites were analyzed by using SignalP (37). Overall amino acid sequence comparison was done with the Bestfit program in the GCG package, using the default parameters. Potential ribosome-binding sites (RBSS) and promoter sequences were examined manually.

Elution of antibodies specific to the polysaccharide. Cell lysates of TX5159 were treated with proteinase K, subjected to SDS-PAGE, transferred to nitrocellulose membranes, and incubated with one of the patient sera as described above. The membranes were then incubated with 10 ml of 100 mM glycine (pH 2.5) for 30 min at room temperature. After neutralization with 1 ml of 1 M Tris (pH 8.0), the solution was transferred to a clean tube and stored at -20°C until ready to use (16).

Nucleotide sequence accession number. The nucleotide sequence reported here was submitted to GenBank under accession no. AF071085.

RESULTS

Periodate treatment of *E. coli*(pTEX5159) lysates. In our previous study, TX5159 (formerly BO4B6I) was found to react with four patient sera but not with a rabbit serum prepared against surface protein extracts from an *E. faecalis* strain isolated from a patient with endocarditis (57). The supernatant from an overnight culture of TX5159 also reacted with one of the patient sera in a dot blot (data not shown). The antigenic material in the TX5159 cell lysates was resistant to extensive proteinase K degradation, suggesting the material was not a protein. To further examine the nature of this antigen, the lysate was subjected to treatment with sodium metaperiodate, which preferentially oxidizes carbohydrates, and then subjected to Western blot analysis using one of the patient sera (Fig. 1). The immunostaining disappeared completely after periodate treatment, suggesting that the antigenic material is a carbohydrate. The material was produced in *E. coli* in sufficient amounts so that after a 1:32 dilution of the sample (equivalent to 3 to 4 μ l of the overnight culture), there was still a strong reaction on the Western blot. The appearance of the diluted material in the Western blot resembled the ladder pattern seen with repeating units of polysaccharides; 10 bands could be clearly observed. A cosmid clone previously found to encode a

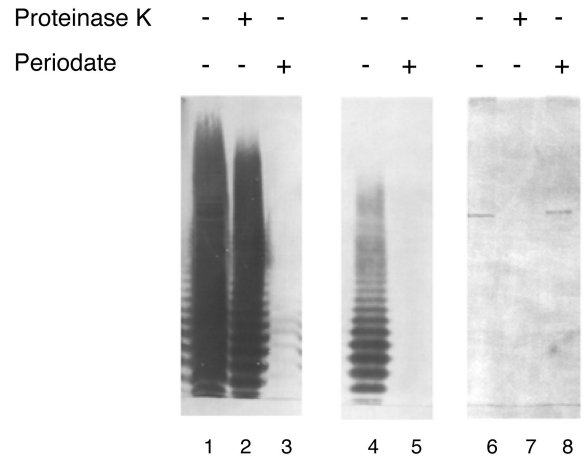


FIG. 1. Effects of proteinase K and periodate on cell lysates of TX5159. Cell lysates of an overnight culture were treated with either proteinase K or periodate. The samples were loaded for SDS-PAGE and transferred to a membrane after electrophoresis. The membrane was blotted with patient serum at 1:500 dilution in 1% skim milk. Lane 1, TX5159; lane 2, TX5159 plus proteinase K; lane 3, TX5159 plus periodate; lane 4, TX5159, diluted 32-fold; lane 5, diluted TX5159 lysates plus periodate; LO6E6II was (used as a control for the treatments); lane 6, LO6E6II lane 7, LO6E6II plus proteinase K; lane 8, LO6E6II plus periodate.

protein antigen, the autolysin of *E. faecalis* (57), was used as a control. This antigen band disappeared after proteinase K treatment but remained intact after incubation with periodate.

Cosmid pTEX5159 was transformed into *E. coli* S ϕ 874, which lacks the entire *rfb* region for O-antigen biosynthesis. The cell lysates of TX5159 was subjected to proteinase K treatment and Western blotting to examine whether any *E. coli* *rfb* genes were involved in generating the immunoreactive ladder pattern. The pattern of TX5159 was the same pattern as that of TX5159 (data not shown).

DNA sequencing of pTEX5159 and sequence analysis of contig C. DNA sequencing resulted in four contigs (A to D) along the insert of pTEX5159 (Fig. 2). The total number of bases sequenced is 94,434, and the total unique sequence is 34,170 bases long. The average coverage of the four contigs is 2.76 \times , and the coverage of contig C, is 3 \times , slightly higher than for the other contigs. About 17.5% of the total unique sequence are single pass, most at the ends of each contig. A restriction map of the 43-kb insert of TX5159 was generated by using the information from DNA sequencing and restriction enzyme digestion (RED) analysis of transposon insertion clones made by in vitro transposition (described below) (Fig. 2). Based on the restriction map, gaps between contigs B and C and between contigs C and D should be less than 500 bp, while a gap of about 4.5 kb is present between contigs A and B. Database comparison of the sequences from the four contigs showed that contig C contains ORFs with similarity to genes involved in polysaccharide biosynthesis.

Contig C contains 18,817 bases with an overall G+C content of 36.7%. Fifteen ORFs (*orfde1* to *orfde15*), all of which had the same transcriptional orientation (Table 3; Fig. 3), were identified by ORF Finder at NCBI. A partial coding region was found in the 3' flanking region of *orfde15* and was oriented for transcription in the same orientation as the others. This was designated *orfde16*. Manual examination of the region between *orfde5* and *orfde6* revealed that there could be an ORF of 585 bp in the same transcriptional orientation as the other ORFs in the contig, with TTG as the start codon. This was designated

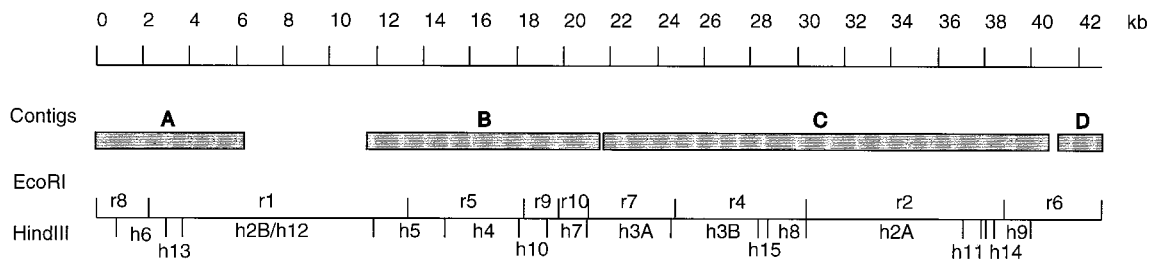


FIG. 2. Restriction map of the insert in pTEX5159. The top line shows the scale. The shaded boxes represent contigs A to D. Below the boxes, the horizontal line represents the insert, the short vertical lines above the horizontal line represent the *EcoRI* sites, and the vertical lines below show the *HindIII* sites. The *EcoRI* fragments were designated r1 to r10, with r1 being the largest in size and r10 the smallest, while the *HindIII* fragments were designated h1 to h15. Fragment h2A had the same mobility in agarose gel electrophoresis as h2B; the same was true of h3A and h3B.

orfde5_6. There was a large (1,011-bp) region between *orfde10* and *orfde11*. Numerous stop codons were found in this intergenic region, and database searches did not find any sequence similarities, suggesting this may be noncoding region.

Putative RBSs were found upstream of each coding region within 11 bp of the start codon except for *orfde3* (21 bp) and *orfde5_6* (39 bp) (Table 2). However, no obvious start sites were found 5' to the ATG at nucleotide position 2718 of *orfde3* and the TTG at 5577 of *orfde5_6*. Therefore, since the start sites for *orfde3* and *orfde5_6* were not certain, the coding region from nucleotide positions 2718 to 3710 identified by ORF Finder was used for *orfde3* and the coding region from nucleotide positions 5577 to 6161 identified by manual examination was used for *orfde5_6* in our analysis.

Putative promoter-like (-35 and -10 like) sequences were found upstream of *orfde2*, *orfde3*, *orfde4*, *orfde5_6*, *orfde6*, *orfde11*, and *orfde16*. However, since enterococcal DNA has high A+T content, the apparent promoter-like sequences may not be real promoters, and functional analysis such as primer extension or mutagenesis will be needed to verify the true promoter regions. The 3' end of *orfde4* overlapped the 5' end of *orfde5* for 74 bp; thus, these two ORFs could be transcriptionally and translationally coupled. The intergenic regions from *orfde6* to *orfde10* are 12, 24, 35, and 58 bp, respectively; these

could form a transcriptional unit. The intergenic regions from *orfde11* to *orfde13* are 12 and 13 bp, respectively. The 3' end of *orfde13* overlapped the 5' end of *orfde14* for 11 bp; therefore, *orfde11* to *orfde14* could be a transcriptional unit. There is a 221-bp intergenic region between *orfde14* and *orfde15*, but no promoter-like sequences were identified. Analysis of terminator-like sequences using the terminator program in the package GCG revealed many candidate sequences, including some within genes, and so the significance of these is not clear.

Database similarities of the genes in contig C and characteristics of the deduced gene products. BLASTx was performed to search for sequence similarities in the databases (Table 4). *Orfde1* showed strong similarity with a putative methionine aminopeptidase A (*map*) from *Synechocystis* sp. (21) and various other organisms. *Orfde2* showed some similarity to the serum resistance locus BrkB protein of *Bordetella pertussis* (11), which is required for the resistance of *B. pertussis* to complement-dependent killing by normal human serum and contains domains with homology to some transporters. Hydrophobicity analysis indicated that the *orfde2* gene product is a hydrophobic protein with four to six potential transmembrane regions. A putative signal peptide cleavage site was found between amino acids 50 and 51 (LTA-VG).

Orfde3 showed similarities to the undecaprenyl-phosphate

TABLE 3. Positions of ORFs in contig C and features of the predicted gene products

ORF	Putative RBS and start site ^a	Potential coding region		Predicted gene product ^b			
		Location in nucleotide sequence	% G+C	Length (amino acids)	MW (10 ³)	pI	GRAVY
<i>orfde1</i>	TTT <u>AGGAGAGATAGTATGATT</u>	778–1314	40.6	178	19.8	5.10	-1.292
<i>orfde2</i>	AATAAGGAGAGTGTGTGATGAAG	1558–2466	35.9	302	34.0	9.82	6.762
<i>orfde3</i>	ACCAGGAGAGCGACGAATCAATACCAAAAATATGCTT	2718–3710	36.3	330	36.3	9.85	9.497
<i>orfde4</i>	AGT <u>AAGGAGAAA</u> TGAGCATGCAA	3831–4670	34.2	279	32.3	5.65	-2.620
<i>orfde5</i>	AATGGGGAAAATTTTTAACAATGGGG	4597–5445	37.2	282	32.7	5.63	-1.770
<i>orfde5_6</i>	AGTGAAGAAAGATGAACATAATTATTCAGATGATGGTTCTACTGATCTACCTGGAA	5577–6161	39.0	194	22.2	9.78	-1.789
<i>orfde6</i>	TTTAAGGAGATTGTTTCATGAAA	6283–7149	38.5	288	32.1	4.82	-1.719
<i>orfde7</i>	AATAGGAGCGAGGAAACGTGAAA	7162–7734	38.2	190	21.3	5.30	-2.679
<i>orfde8</i>	AGTAGGAGAGTTTCAATTATGAAA	7759–8787	38.3	342	38.4	4.99	-5.047
<i>orfde9</i>	TGTGAGATAACAGTCGAAAGCGCATGGCA	8823–9722	38.3	299	33.8	4.68	-3.632
<i>orfde10</i>	AGTGGAAAGGTGTTTCTTTTATGAAG	9781–10506	37.7	211	26.9	6.23	0.357
<i>orfde11</i>	TACGGAAGGAATAAGTATATGTTT	11518–12312	32.6	264	30.7	9.48	6.549
<i>orfde12</i>	TATAAGGAGAGAGAACGACATGTCA	12325–13005	37.7	226	25.4	5.18	-1.195
<i>orfde13</i>	TATGGAGATACGATGGAC	13019–13543	36.0	174	20.2	4.89	-3.937
<i>orfde14</i>	AATGGAAAGGATACGACAAAAGATGAAG	13533–14714	36.5	393	45.1	4.56	-3.176
<i>orfde15</i>	TTTGGAAAGGCGATCAGGTTTATGGTT	14936–16543	39.1	536	61.0	5.39	-4.153
<i>orfde16</i>	TATAAGGAGAATGCACATGAAT	16700–18817 (3'partial)					

^a Putative RBSs are underlined, and start codons are in boldface.

^b Molecular weight (MW), isoelectric point (pI), and grand average hydropathicity (GRAVY) were calculated by using ProtParam at ExPASy. GRAVY is based on the Kyte-Doolittle formulation (24); the more negative the value is, the more hydrophilic the protein is.

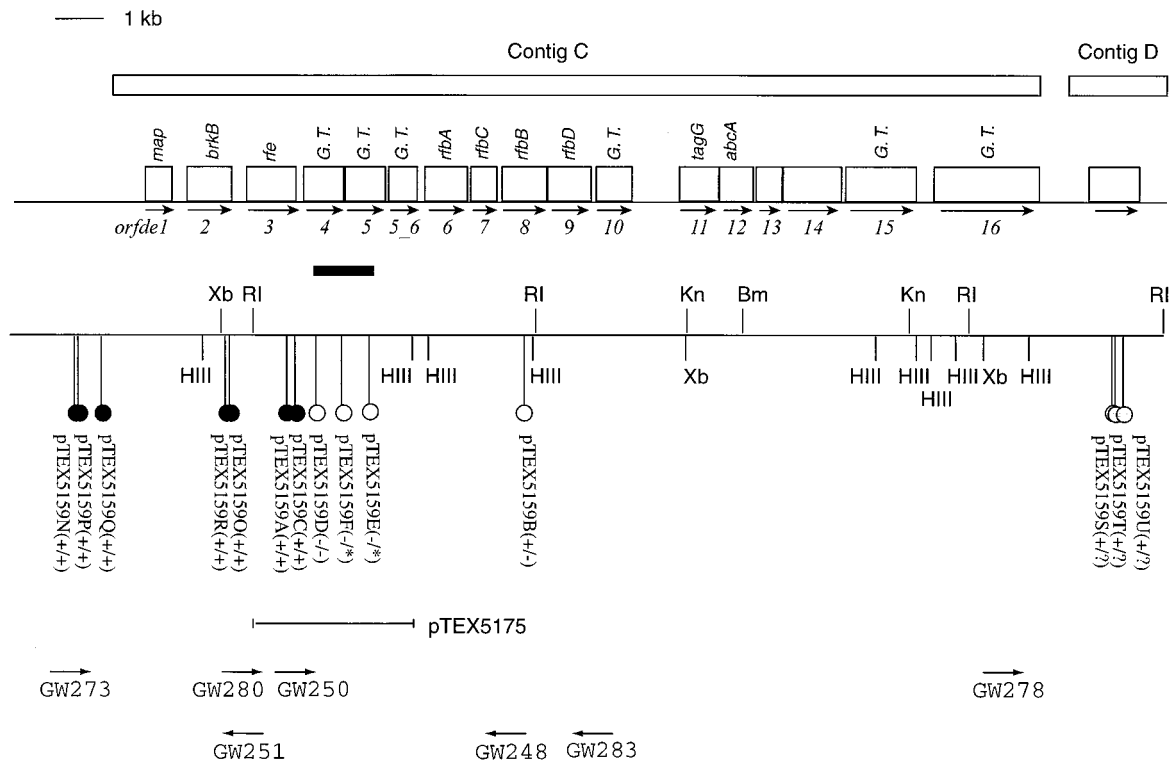


FIG. 3. The polysaccharide biosynthesis gene cluster and flanking regions. The ORFs are shown as open boxes, the direction of the arrows underneath the boxes indicates the direction of transcription, and the best-gapped BLASTx hits of some of the ORFs are indicated above the boxes. The six ORFs that showed some similarities to glycosyltransferases are labeled G. T. Tn7 transposon insertions are shown as lollipops. The name of each cosmid carrying the corresponding insertion is given underneath followed by pluses and/or minuses in parentheses to indicate the immunoreactivity. The sign before the slash indicates the immunoreactivity in DH5 α , and the sign after the slash indicates the immunoreactivity in S ϕ 874. Open-circle lollipops represent transposon insertions that rendered the clones immunonegative, filled-circle lollipops represent those that did not affect the immunoreactivity of the clones, and shaded-circle lollipops in contig D are those that had undergone DNA rearrangements in S ϕ 874. The black bar below *orfde4* and *orfde5* indicates the region where the initial 10 immunonegative clones were mapped. Three of these 10 clones, pTEX5159D to pTEX5159F, are shown as open-circle lollipops below the black bar. Primers used in characterizing transposon insertions are shown at the bottom, with the direction of arrows indicating the direction of DNA synthesis. Primers are not drawn in scale. Bm, *Bam*HI; HIII, *Hind*III; Kn, *Kpn*I; RI, *Eco*RI; Xb, *Xba*I.

α -N-acetylglucosaminyltransferase encoded by the *rfe* genes from various organisms, including *E. coli* and *Bacillus subtilis*, and a lipophilic protein which affects bacterial lysis rate and methicillin resistance level (*lrm*) of *Staphylococcus aureus* (29). A potential signal peptide cleavage site was found between amino acids 26 and 27 of Orfde3 (IIP-QD). Hydropathy plot indicated that Orfde3 is highly hydrophobic with 10 to 11 potential transmembrane domains, similar to the *E. coli* Rfe protein (31). This *E. coli* enzyme catalyzes the transfer of GlcNAc onto undecaprenyl phosphate to synthesize GlcNAc-pyrophosphorylundecaprenol, the first lipid-linked intermediate in the biosynthesis of enterobacterial common antigens (31), and was shown to be required for the initial step in the formation of repeating units of several O antigens (1, 41).

BLASTx results showed that Orfde6 to Orfde9 had extremely strong sequence homology with gene products glucose-1-phosphate thymidyltransferase, dTDP-4-dehydrorhamnose 3,5-epimerase, dTDP-glucose-4,6-dehydratase, and dTDP-4-keto-L-rhamnose reductase, respectively, of various bacteria (Table 4). These four enzymes catalyze the biosynthesis of dTDP-rhamnose from glucose-1-phosphate. The overall amino acid sequences of Orfde6 to Orfde9 were compared with the corresponding translated sequences from two gram-positive and three gram-negative organisms: *cps19fLMNO* from the *Streptococcus pneumoniae* type 19F capsular polysaccharide gene cluster (13, 32), *rmlACB* and *rmlD* of the *Streptococcus mutans* dTDP-rhamnose biosynthesis pathway (51, 52), *rfbBDAC* from

the *E. coli* O7 *rfb* gene cluster (30), *rmlBDAC* from the *Salmonella typhimurium* LT2 *rfb* gene cluster (20), and *rfbBDAC* from the *Shigella flexneri* type 2a rhamnose biosynthesis operon (28) (Fig. 4). Orfde6 showed high similarities with the glucose-1-phosphate thymidyltransferase from all five organisms, being from 70 to 78% identical. Orfde8 showed 80 and 81% identity to the *cps19fN* and *rmlC* gene products from the two gram-positive organisms, while its identity to the gram-negative proteins was lower (46 to 48%) but at the same level as the identity between the streptococcal and gram-negative proteins (47 to 50%). This was also observed with Orfde9, which showed higher identities to the two streptococcal proteins (62 and 61%) than to the gram-negative ones (37 to 40%). However, Orfde7 showed stronger identities to the three gram-negative proteins (53 to 55%) than to the gram-positive ones (both 34%). Orfde6 to Orfde9 are generally hydrophilic, and the hydropathy plots of Orfde6 to Orfde9 exhibited remarkable similarities to those of the corresponding sequences from these other organisms (data not shown).

Orfde11 showed strong similarity with the integral membrane protein component of ATP-binding cassette (ABC) transport systems involved in sugar polymer export. The highest similarity was with the teichoic acid translocation permease TagG or *B. subtilis* (25). Hydrophobicity analysis showed that the *orfde11* gene product is highly hydrophobic with five to six transmembrane domains, consistent with being an integral membrane protein. Orfde12 was similar to the ATP-binding

TABLE 4. Database homologies of the ORFs in contig C

ORF	BLASTx hit	Accession no.	Organism	Expect value	Reference(s)
<i>orfde1</i>	Methionine aminopeptidase A (<i>map</i>)	P53579	<i>Synechocystis</i> sp.	6e-59	21
<i>orfde2</i>	Serum resistance locus <i>brkB</i>	D64004	<i>Synechocystis</i> sp.	2e-07	
	Serum resistance locus <i>brkB</i>	I40328	<i>Bordetella pertussis</i>	2e-04	11
<i>orfde3</i>	Putative undecaprenyl-phosphate <i>N</i> -acetylglucosaminyltransferase (<i>tagO</i>)	gn11PIDe11 68896	<i>Bacillus subtilis</i>	1e-57	Unpublished data
	Lipophilic protein which affects bacterial lysis rate and methicillin resistance level (<i>llm</i>)	D21131	<i>Staphylococcus aureus</i>	4e-57	29
	Undecaprenyl-phosphate <i>N</i> -acetylglucosaminyltransferase (<i>rfe</i>)	S30678	<i>Escherichia coli</i>	4e-19	31
<i>orfde4</i>	EpsX	X81320	<i>Acinetobacter calcoaceticus</i>	1e-25	Unpublished data
	Hypothetical 31.1-kDa protein in <i>gnd-rfc</i> intergenic region (<i>orf264</i>)	P36667	<i>E. coli</i>	4e-12	27, 59
<i>orfde5</i>	RfbC protein	S28579	<i>Yersinia enterocolitica</i>	0.14	
	dTDP-rhamnosyltransferase (<i>rfbF</i>)	P37782	<i>Shigella flexneri</i>	0.14	33
<i>orfde5_6</i>	Unnamed protein product	AB002668	<i>Actinobacillus actinomycetemcomitans</i>	5e-18	60
	MigA	U70729	<i>Pseudomonas aeruginosa</i>	5e-11	54
	PssF	AF028810	<i>Rhizobium leguminosarum</i> bv. <i>viciae</i>	3e-09	18
<i>orfde6</i>	Glucose-1-phosphate thymidyltransferase (<i>cps19fL</i>)	U09239	<i>Streptococcus pneumoniae</i>	e-131	32
<i>orfde7</i>	dTDP-4-dehydrorhamnose 3,5-epimerase (<i>rfbC</i>)	P26394	<i>Salmonella enterica</i>	1e-44	20
<i>orfde8</i>	dTDP-glucose-4,6-dehydratase (<i>rmlB</i>)	D78182	<i>Streptococcus mutans</i>	e-150	52
<i>orfde9</i>	32.3-kDa <i>cps19fO</i> gene product	U09239	<i>S. pneumoniae</i>	2e-97	32
	dTDP-4-keto-L-rhamnose reductase (<i>rmlD</i>)	AB000631	<i>S. mutans</i>	1e-94	51
<i>orfde10</i>	Unnamed protein product	AB002668	<i>A. actinomycetemcomitans</i>	1e-30	60
	ss-1,3- <i>N</i> -Acetylglucosaminyltransferase	X85787	<i>S. pneumoniae</i>	0.001	22, 23
<i>orfde11</i>	Teichoic acid translocation permease protein TagG (<i>tagG</i>)	P42953	<i>B. subtilis</i>	2e-48	25
	O-antigen export system permease protein RfbA (<i>rfbA</i>)	Q50862	<i>Myxococcus xanthus</i>	3e-23	14
<i>orfde12</i>	ATP-binding cassette membrane transporter AbcA (<i>abcA</i>)	Q07698	<i>Aeromonas salmonicida</i>	5e-58	8
	O-antigen export system ATP-binding protein RfbB (<i>rfbB</i>)	Q50803	<i>M. xanthus</i>	8e-56	14
	Teichoic acid translocation ATP-binding protein TagH (<i>tagH</i>)	P42954	<i>B. subtilis</i>	1e-51	25
<i>orfde13</i>	No good match				
<i>orfde14</i>	No good match				
<i>orfde15</i>	Hypothetical 102.8-kDa protein Y4gI	P55465	<i>Rhizobium</i> sp. strain NGR 234	4e-68	12
	O-antigen biosynthesis protein RfbC (<i>rfbC</i>)	U36795	<i>M. xanthus</i>	5e-52	14
	EpsI	U40830	<i>Streptococcus thermophilus</i>	6e-14	50
	EspG	U93364	<i>Lactococcus lactis cremoris</i>	2e-12	53
	β -1,4-Galactosyltransferase	X85787	<i>S. pneumoniae</i>	1e-11	22, 23
<i>orfde16</i>	Hypothetical 102.8-kDa protein Y4gI O-antigen biosynthesis protein RfbC (<i>rfbC</i>)	P55465 Q50864	<i>Rhizobium</i> sp. strain NGR234 <i>M. xanthus</i>	e-128 6e-16	12 14

protein component of several sugar polymer export systems, including the O-antigen export systems of *Myxococcus xanthus* (14) and *Klebsiella pneumoniae* (7). It contains the characteristic ABC consensus sequence of the ATP-binding proteins in the transport system. It seems likely that Orfde11 and Orfde12 comprise an ABC transport system for exporting the polysaccharide.

Orfde4, Orfde5, Orfde5_6, Orfde10, Orfde15, and Orfde16 (3' partial) showed various degrees of similarity to glycosyltransferases. No signal cleavage site was identified in any of these sequences. Except for Orfde10, they were not predicted to be membrane proteins. Orfde10 was predicted to contain two transmembrane domains, one near the N terminus and the other near the C terminus. The two proteins that showed the highest similarity to Orfde4, the *epsX* gene product of *Acinetobacter calcoaceticus* and a hypothetical 31.1-kDa protein encoded by *orf264* in the *gnd-rfc* intergenic region of *E. coli* (27, 59), have no clearly documented function. However, it has been shown that when *orf264* was introduced into *E. coli* K-12 strains EMG2 and WE110, whose *orf264* genes are defective, the two strains produced an O antigen. This led to the conclusion that *orf264* was essential for the biosynthesis of O antigen

in *E. coli* K-12 strains and was thought likely to be rhamnosyltransferase (27, 59). The third-best BLASTx expect value was with a glycosyltransferase encoded by *spsQ* of *Sphingomonas* strain S88 (58). The gene *spsQ* was identified in a gene cluster essential for the synthesis of a capsular polysaccharide in *Sphingomonas* strain S88 and is thought to be a glycosyltransferase, suggesting that *orfde4* may also encode a glycosyltransferase. Orfde5 showed only very low similarity to the dTDP-rhamnosyltransferase (encoded by *rfbF*) of *Shigella flexneri* (33). Bestfit comparison of the amino acid sequences of RfbF and Orfde5 showed a 121-amino-acid stretch with 34% similarity and 24% identity. It has been noted that the sequence similarities among glycosyltransferases are generally low (26), and thus the homology between RfbF and Orfde5 may be significant. Both Orfde5_6 and Orfde10 were found to be similar to several putative glycosyltransferases and proteins of unknown function. Orfde15 and Orfde16 shared some sequence similarity between themselves and were homologous to a similar set of proteins, most of which were glycosyltransferases. The two best matches were the hypothetical 102.8-kDa protein Y4gI of

	Orfde6	Cps19fL	RmlAsmu	RfbAeco	RmlAsen	RfbAsfl
Orfde6	100 (288)	78 (288)	78 (288)	72 (285)	71 (284)	70 (284)
	Orfde7	Cps19fM	RmlCsmu	RfbCeco	RmlCsen	RfbCsfl
Orfde7	100 (190)	34 (171)	34 (171)	53 (181)	55 (177)	53 (160)
	Orfde8	Cps19fN	RmlBsmu	RfbBeco	RmlBsen	RfbBsfl
Orfde8	100 (342)	80 (338)	81 (338)	48 (343)	47 (343)	46 (343)
	Orfde9	Cps19fO	RmlDsmu	RfbDeco	RmlDsen	RfbDsfl
Orfde9	100 (299)	62 (276)	61 (276)	37 (294)	40 (227)	38 (234)

FIG. 4. Comparison of amino acid sequences of dTDP-rhamnose biosynthesis genes from various bacteria. Percentage identity was obtained by using Bestfit; each number in parentheses indicates the length of the segment that gave the percentage identity in amino acids. Default parameters were used for comparisons. Orfde6 to Orfde9 are from *E. faecalis* OG1RF. Cps19fL to -O, *Streptococcus pneumoniae* type 19F; smu, *Streptococcus mutans*; eco, *Escherichia coli*; sen, *Salmonella enterica*; sfl, *Shigella flexneri*. Cps19fL, RmlAsmu, RfbAeco, RmlAsen, and RfbAsf1, glucose-1-phosphate thymidyltransferases; Cps19fM, RmlCsmu, RfbCeco, and RfbCsfl, dTDP-4-dehydrorhamnose 3,5-epimerases; Cps19fN, RmlBsmu, RfbBeco, RmlBsen, and RfbBsfl, dTDP-glucose-4,6-dehydratases; Cps19fO, RmlDsmu, RfbDeco, RmlDsen, and RfbDsfl, dTDP-4-keto-l-rhamnose reductases.

Rhizobium sp. strain NGR234 (12) and the O-antigen biosynthesis protein RfbC of *M. xanthus* (14).

Orfde13 and Orfde14 showed no significant similarity to any sequences in the databases.

Sequence analysis of contig D. Contig D has 2,943 bases. One ORF and 1,029 bp was identified in the middle of contig D and did not reveal any database similarities. The end of contig D was from vector pBeloBAC11. The rest of the sequence in the contig did not show any similarities to sequences in the database.

Immunoscreening of TX5159 with transposon insertions. We screened 282 transposon insertion clones, generated by in vitro transposition of a Tn7 derivative, by using a patient serum and subjected immunonegative clones to a second screen. Ten clones, DH5 α (pTEX5159D) to DH5 α (pTEX5159M), were found to be immunonegative. Five of the 10 clones had single insertions in *Eco*RI fragment r4, while the other five had double insertions, with one in r4 and the other in either r1 or r5. Clones with single insertions in only r1 or r5 were still immunopositive, suggesting that the insertions in r4 were responsible for the loss of the immunoreactivity of these clones. Cosmid DNA from three clones, DH5 α (pTEX5159D), DH5 α (pTEX5159E), and DH5 α (pTEX5159F), were transformed into S ϕ 874, and the transformants were tested by colony immunoblotting. S ϕ 874 (pTEX5159D) was negative, while transformants with pTEX5159E and pTEX5159F were not able to regrow after the initial appearance on the primary transformation plates.

Characterization of transposon insertion mutants. To determine the sites of the insertions, cosmid DNA from the 10 clones that were immunonegative was prepared and subjected to PCR amplification (Table 5). All insertions in *Eco*RI fragment r4 were localized to a small region, shown as a black bar in Fig. 3. PCR products amplified by using primer pairs GW248-NLC94 and GW250-NLC272 from three immunonegative clones, DH5 α (pTEX5159D), DH5 α (pTEX5159E), and DH5 α (pTEX5159F), were gel purified and subjected to DNA sequencing. The sequences showed that the insertions were in *orfde4* and *orfde5*, which encode potential glycosyltransferases (Fig. 3). The insertion in DH5 α (pTEX5159D) was at nucleotide position 4191 in

contig C, 361 bp downstream of the start site of *orfde4* (equivalent to amino acid 119 from the N terminus of Orfde4); the insertion in DH5 α (pTEX5159E) was at 5238, in the middle of *orfde5* (equivalent to amino acid 214 from the N terminus of

TABLE 5. Characterization of transposon mutants

Cosmid	Site of insertion ^a	Immuno-reactivity ^b	
		DH5 α	S ϕ 874
pTEX5159G	<i>orfde4</i>	—	—
pTEX5159D	<i>orfde4</i>	—	—
pTEX5159H		—	—
pTEX5159I	<i>orfde5</i>	—	—
pTEX5159E	<i>orfde5</i>	—	DNG ^c
pTEX5159J	<i>orfde5</i>	—	—
pTEX5159K	<i>orfde5</i>	—	—
pTEX5159F	<i>orfde5</i>	—	DNG
pTEX5159L	<i>orfde4/5</i>	—	—
pTEX5159M		—	—
pTEX5159A	<i>orfde3</i>	+	+
pTEX5159B	<i>orfde8</i>	+	—
pTEX5159C	<i>orfde3-orfde4</i>	+	+
pTEX5159N ^c	r7, 0.7 kb upstream of contig C	+	+
pTEX5159O ^c	<i>orfde2</i>	+	+
pTEX5159Q ^c	r7, 0.2 kb upstream of contig C	+	+
pTEX5159P ^c	r7, 0.7 kb upstream of contig C	+	+
pTEX5159R ^c	<i>orfde2</i>	+	+
pTEX5159S ^c	r6, 1.8 kb downstream of contig C	+	— ^d
pTEX5159T ^c	r6, 1.8 kb downstream of contig C	+	— ^d
pTEX5159U ^c	r6, 1.9 kb downstream of contig C	+	— ^d

^a ORF where the transposon had inserted. Sites of insertion of those that were not sequenced were estimated based on their PCR product sizes in relation to the ones that were sequenced.

^b Tested by using colonies of each clone in immunoblot analyses with patient serum.

^c Not been subjected to PCR amplification. The insertion site was determined by RED analysis.

^d The clone was transformed into S ϕ 874, resulted in large deletions in the insert, and was immunonegative.

^e S ϕ 874 transformants of pTEX5159E and pTEX5159F did not grow (DNG) when streaked from the primary transformation plates.

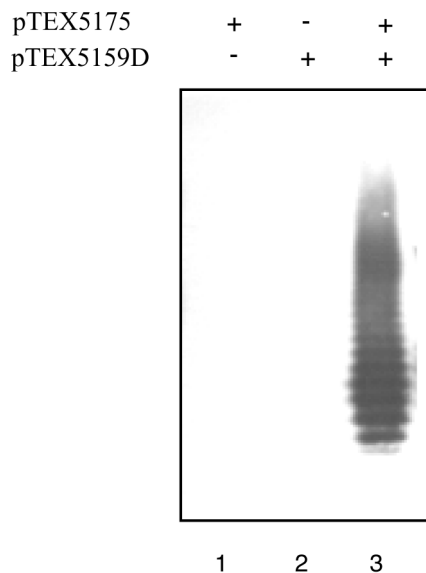


FIG. 5. Complementation of DH5 α (pTEX5159D) with pTEX5175. Samples were treated with proteinase K and analyzed as described in the text. Lane 1, DH5 α (pTEX5175); lane 2, DH5 α (pTEX5159D); lane 3, DH5 α (pTEX5159D + pTEX5175).

Orfde5); and the insertion in DH5 α (pTEX5159F) was at 4701, 105 bp downstream of the start of *orfde5* (equivalent to amino acid 35 from the N terminus of Orfde5). The sites of insertion in the other seven clones were estimated based on their PCR product sizes in relation to those from the three clones that had been sequenced. It appeared that all were in *orfde4* or *orfde5*. The insertion in DH5 α (pTEX5159L) could be either at the end of *orfde4* or at the beginning of *orfde5*.

To prove the *orfde4-5* region was responsible for the mutant phenotype, a complementation test was performed. An *EcoRI-HindIII* fragment which contains the 3' half of *orfde3*, all of *orfde4* and *orfde5*, and almost all of *orfde5_6*, was cloned in pBluescript SK (-), and the resulting construct was designated pTEX5175. The plasmid was transformed into the immunonegative insertion mutant DH5 α (pTEX5159D) as well as into DH5 α . Cell lysates of the transformants were treated with proteinase K and subjected to Western blot analysis. Plasmid pTEX5175 restored the immunoreactivity of DH5 α (pTEX5159D), while DH5 α (pTEX5175) was immunonegative (Fig. 5).

Clones with insertions in r7 or r6, at the ends of contig C, were found through RED analysis of random clones (Table 5). Five clones were found to have insertions in r7; PCR results showed that insertion in pTEX5159N, pTEX5159Q, and pTEX5159P were upstream of contig C, while insertions in pTEX5159O and pTEX5159R were in *orfde2*. Three clones, pTEX5159S, pTEX5159T, and pTEX5159U, had insertions 1.8 to 1.9 kb downstream of contig C. These clones were all immunopositive in colony blots in DH5 α . When transformed into S ϕ 874, the five clones with insertions in r7 remained positive. However, S ϕ 874 transformants of pTEX5159S, pTEX5159T, and pTEX5159U seemed to have undergone deletions or rearrangements in S ϕ 874, based on RED analysis, and were immunonegative.

Complementation by *E. coli* O-antigen biosynthesis genes. After RED analysis of random clones, three clones, DH5 α (pTEX5159A), DH5 α (pTEX5159B), and DH5 α (pTEX5159C) were found to have insertions in r4 but still reacted in colony immunoblots. PCR products amplified by using primer pairs GW283-NLC94 and GW280-NC272 from these three clones were subjected to DNA sequencing (Table 5). The insertion in

DH5 α (pTEX5159C) was at nucleotide position 3728 in contig C, which is in the region between *orfde3* and *orfde4* (Fig. 3). This finding suggested that *orfde4* was transcribed independently.

The insertion in DH5 α (pTEX5159B) was at nucleotide position 8388 in the middle of *orfde8* (equivalent to amino acid 210 from the N terminus of Orfde8), which is highly similar to the gene for dTDP-glucose-4,6-dehydratase, the second enzyme in the dTDP-rhamnose biosynthesis pathway of various bacteria. Rhamnose is a major component in the O antigens of gram-negative bacteria and in capsules of some gram-positive bacteria. Thus, it was surprising to find that DH5 α (pTEX5159B) was still immunopositive. Since most *E. coli* K-12 strains should carry genes for dTDP-rhamnose synthesis (27, 59), however, it was possible that DH5 α (pTEX5159B) was complemented by the *E. coli* genes. To test this, cosmid DNA from DH5 α (pTEX5159B) was transformed into *E. coli* S ϕ 874, which lacks the entire *rfb* region. The transformant was found to be immunonegative in colony blots (Table 5), indicating that complementation of the insertion mutation occurred in DH5 α but not in S ϕ 874.

DH5 α (pTEX5159A) had an insertion at nucleotide position 3525, which is 187 bp upstream of the 3' end of *orfde3* (equivalent to 61 amino acids from the C terminus of Orfde3). Orfde3 is similar to the undecaprenyl-phosphate α -N-acetylglucosaminyltransferase that catalyzes the transfer of GlcNAc to the lipid carrier, the first step in the synthesis of the repeating units of several O antigens (1, 41). Since the *E. coli rfe* gene is outside the *rfb* region, both DH5 α and S ϕ 874 possess this gene. To test the possibility that this clone was being complemented by the *rfe* gene of *E. coli*, cosmids pTEX5159A and pTEX5159 were transformed into an *E. coli rfe* strain, 21548, as well as into the parent strain AB1133 (*rfe*⁺). The cell lysates of the transformants were subjected to proteinase K treatment and Western blot analysis as described above. All four clones showed the characteristic ladder pattern as with TX5159 (data not shown).

Detection of the expression of this polysaccharide in *E. faecalis*. Antibody was eluted from nitrocellulose membranes following Western blotting of proteinase K-treated extracts from TX5159 with a patient serum. The eluted antibody reacted with TX5159 but not with the cloning host DH5 α . The eluted antibody was used to examine *E. faecalis* OG1RF and TX52 (57) under different growth conditions, including exposure to horse and rabbit sera, at different temperatures, and in minimal growth media. Several clinical from patients with *E. faecalis* endocarditis were also examined after growth in BHI broth. These clinical strains include isolates from patients from whom the tested sera were collected. No positive reaction was detected (data not shown).

DISCUSSION

In a previous study, we described a cosmid clone of *E. faecalis* (TX5159) that produced proteinase K-resistant antigenic material in *E. coli* (7). In this report, we provide evidence that the antigenic material was carbohydrate in nature and that its production did not involve the *E. coli rfb* genes. Furthermore, the antigen appears as a ladder after PAGE, suggesting a structure of repeating units, such as those of polysaccharides. Since the clone contains a number of genes with sequence similarity to polysaccharide biosynthesis functions, we conclude that genes for the biosynthesis of an enterococcal antigenic polysaccharide had been cloned and expressed in *E. coli*.

Early studies indicated that antigenic polysaccharides could be detected in the cell walls of enterococci and, moreover, that

many of these were rhamnose-containing polysaccharides (6, 19, 45, 47). However, no studies to our knowledge suggested the presence of capsules in enterococci. In our previous report, TX5159 was found to react with four patient sera but not with a rabbit serum prepared against surface protein extracts from an *E. faecalis* strain grown in BHI (57). The rabbit serum was made from rabbits immunized with zwittergent surface extracts of *E. faecalis*, and it is possible that the polysaccharide was not present. In this study, we failed to detect the antigen in *E. faecalis* OG1RF and TX52 (57) under different growth conditions or in several *E. faecalis* clinical isolates. This raised the possibility that the antigen is due to *E. coli* and not enterococcus. PCR amplification of OG1RF genomic DNA by using intragenic primers to *orfde2* to *orfde16* resulted in PCR products of the expected sizes (data not shown). Southern blot analysis using pTEX5159 DNA as a probe against genomic DNA of OG1RF, TX52, and DH5 α indicated that the probe hybridized to OG1RF and TX52 but not to DH5 α (data not shown), suggesting that they are enterococcal genes. Therefore, in this model, the patients would also be making antibody against the *E. coli* polysaccharide, and the enterococcal genes would be complementing *E. coli* genes that were defective in the laboratory strains. We note that in our screens for immunonegative transposon mutants in DH5 α , where all the mutants carried insertions in *orfde4* and *orfde5*, *orfde4* shows some sequence similarity to *orf264* in the *gnd-rfc* intergenic region of *E. coli* (27, 59). In most of the *E. coli* K-12 strains, there is an IS5 insertion in *orfde264*, which is responsible for the loss of O-antigen production in these K-12 strains. This finding raised the possibility that *orfde4* complemented the defective *orf264* of DH5 α and thus restored its ability to make an O antigen. This possibility can be ruled out because of the following observations. Firstly, unlike DH5 α , which still maintains most of the *rfb* genes, S ϕ 874 lacks the entire *rfb* region for O-antigen biosynthesis. However, TX5169 had the same reaction pattern with the patient serum as TX5159. Second, plasmid pTEX 5175, which contains the complete sequence of *orfde4*, *orfde5*, and their flanking regions, alone did not make DH5 α immunoreactive. Thus, the production of the antigenic polysaccharide required the presence of pTEX5159 but not the *E. coli rfb* genes.

The four patient sera used were collected from endocarditis patients infected with *E. faecalis* in different regions in the United States. The fact that TX5159 reacted strongly with all four sera suggested that the polysaccharide was produced in all of these patients and was not due to one particular strain or one particular patient. It has long been known that the biosynthesis of many polysaccharides (e.g., *E. coli* group IA and group II K antigens, and colanic acid [56]) is regulated, and many regulatory strategies have been used to achieve control of expression in response to different changes in the environment. Thus, expression of the polysaccharide in *E. faecalis* may be controlled so that it is only expressed under certain conditions, such as those found in infection. The use of more sensitive detection methods with a wider range of growth conditions is under way. Preliminary results from reverse transcription-PCR using primers to *orfde4* and *orfde6* suggest that the mRNA transcripts to those two genes could be detected in RNA extracts from OG1RF in exponential phase but not in stationary phase (data not shown). The other ORFs in contigs C and D are now being tested.

Comparison of the ORFs in contig C to database sequences revealed similarities to genes involved in polysaccharide biosynthesis and export from both gram-negative and gram-positive organisms. *Orfde3* showed similarities to the undecaprenylphosphate α -N-acetylglucosaminyltransferase encoded by the

rfe genes from various organisms and Llm of *S. aureus* (29). The activity of the lipophilic protein encoded by *llm* is not known, although it has been proposed to be involved in the metabolism of cell surface components such as peptidoglycan (29). *Rfe* is required for the synthesis of the first lipid-linked intermediate in the biosynthesis of enterobacterial common antigen and several O antigens (1, 41, 42). Hydropathy plots of the enterococcal *Orfde3* and *Rfe* of *E. coli* showed remarkable similarities between the two, indicating similar overall distribution of hydrophobic residues. However, a Tn7 insertion near the 3' end of *orfde3* (pTEX5159A) did not affect the expression of the polysaccharide in *E. coli* 21548 (*rfe*), S ϕ 874 (*rfb*), AB1133 (*rfe*⁺), or DH5 α . Several possibilities may explain this result: (i) the C-terminal 60 or so amino acids may not be required for enzymatic activity, (ii) the gene is not required for the synthesis of the polysaccharide in *E. coli*, and (iii) *orfde3* codes for some other function(s) instead of the undecaprenylphosphate α -N-acetylglucosaminyltransferase and is complemented by similar functions from *E. coli*. Further mutational studies of *orfde3* may help answer some of these questions. One conclusion that can be drawn from results with pTEX5159A is that the insertion did not affect the expression of downstream genes, consistent with the result for DH5 α (pTEX5159C), which had an insertion between *orfde3* and *orfde4* and was immunopositive, as well as the observation that *orfde4* appeared to have its own promoter, based on DNA sequence.

orfde6 to *orfde9* showed high sequence similarities to genes in the dTDP-rhamnose biosynthesis pathway of various organisms. As noted above, rhamnose has been found as a component in enterococcal polysaccharides; the transposon insertion in *orfde8* (pTEX5159B) resulted in loss of immunoreactivity in S ϕ 874 (*rfb*) but not in DH5 α . This finding indicated that the *E. coli rfbBDAC* genes could provide this function and provides further support for the conclusion the *orfde6* to *orfde9* encode a dTDP-rhamnose biosynthesis pathway.

The G+C contents of *orfde6* to *orfde9* are very close to each other, from 38.2 to 38.5%, while those of the other ORFs in the cluster vary from 32.6 to 40.6%, suggesting that *orfde6* to *orfde9* may have originated from one source while the other ORFs in the cluster may have been assembled from several sources during evolution. This is consistent with observations in other bacteria such as *E. coli* and *Salmonella* spp. (20, 30, 49, 55, 59).

Several reports have shown the involvement of ABC transport systems in the translocation of sugar polymers across the cytoplasmic membrane in both gram-negative and gram-positive bacteria (14, 25, 56, 61) or the presence of ABC transporter genes within polysaccharide biosynthesis clusters (5, 61). It seems that the use of ABC transporters could be one of the common mechanisms for exporting sugar polymers. The strong sequence similarities of *orfde11* and *orfde12* to components in the ABC transport system, in particular to the ABC transporters for teichoic acids of *B. subtilis* (25) and the O antigen of *M. xanthus* (14), suggest that these two ORFs may play similar roles (14) in the export of the *E. faecalis* polysaccharide to the cell surface.

The products of six ORFs, *orfde4*, *orfde5*, *orfde5_6*, *orfde10*, *orfde15*, and *orfde16* (3' partial), showed various degrees of similarities to glycosyltransferases. One feature of these ORFs is that, except for *orfde10*, they encode hydrophilic proteins, consistent with the findings for other glycosyltransferases and the notion that the assembly of repeat units occurs on the cytoplasmic side of the membrane. The sequence homology of these ORF products to glycosyltransferases was not very strong, as has been recognized in general among glycosyltransferases. Thus, it would be difficult to propose, based on sequence com-

parison, the substrate specificity encoded by these ORFs. Further analysis, such as enzyme activity assays, is needed to elucidate their functions. Transposon insertions in *orfde4* and *orfde5* abolished the immunoreactivity in both DH5 α and S ϕ 874, suggesting that these two genes are required for the production of the polysaccharide and that there are not equivalent functions in the *E. coli* host strains. The introduction of pTEX5175 into DH5 α (pTEX5159D) restored its immunoreactivity, suggesting that its transposon insertion did not affect the expression of *orfde6* and downstream genes. This is consistent with the results from sequence analysis that *orfde4* and *orfde6* may be transcribed from their own promoters and are in two individual transcriptional units.

Analysis of Tn7 transposon insertion mutants also provided evidence that *orfde1* and *orfde2* are not essential for antigen production. The involvement of *orfde3* is not certain and requires further investigation. No Tn7 insertions in the other ORFs in contig C were identified. However, based on sequence analysis, it is likely that the other ORFs are also involved in the synthesis and export pathway. Tn7 insertions in contig D resulted in large deletions or rearrangement in the cosmid, suggesting that disruption in this region may be toxic to the host. Whether contig D is required in the biosynthesis pathway remains unclear.

In conclusion, a gene cluster of OG1RF was cloned and produced an antigenic polysaccharide in *E. coli*. Analysis of DNA sequences and transposon mutants indicated that it is a multicistronic region containing genes for dTDP-rhamnose synthesis, assembly of repeat units, and export of the polysaccharide. This is the first report of an enterococcal gene cluster involved in the biosynthesis of an antigenic polysaccharide. Further study of this cluster and the effects of such a polysaccharide on bacterial cell surface permeability, response to antimicrobial agents, and the host immune system may have important implications for research in enterococci as well as for polysaccharides in other organisms.

ACKNOWLEDGMENTS

We are grateful to Matthew C. Biery and Nancy L. Craig, Department of Molecular Biology and Genetics, Johns Hopkins University, for performing the in vitro transposition reactions, and as well as Huy Phan, University of Texas Medical School at Houston, for mapping the transposon insertion clones. We thank Chris Whitfield, Department of Microbiology, University of Guelph, Guelph, Ontario, Canada, and Paul Rick, Department of Microbiology, Uniformed Services of the Health Sciences, for providing strains and instructive advice. We also thank Miguel Valvano, Department of Microbiology and Immunology, University of Western Ontario, London, Ontario, Canada, for helpful advice.

This work was supported by the NIH grant AI 33516 to Barbara E. Murray.

REFERENCES

- 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.
- Archibald, A. R., I. C. Hancock, and C. R. Harwood. 1993. Cell wall structure, synthesis, and turnover, p. 381–410. In A. L. Sonenshein, J. A. Hock, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington, D.C.
- Bainton, R., P. Games, and N. Craig. 1991. Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. *Cell* **65**:805–816.
- Bainton, R. J., K. M. Kubo, J.-N. Feng, and N. L. Craig. 1993. Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. *Cell* **72**:931–943.
- Becker, A., S. Ruberg, H. Kuster, A. A. Roxlau, M. Keller, T. Ivashina, H.-P. Cheng, G. C. Walker, and A. Puhler. 1997. The 32-kilobase *exp* gene cluster of *Rhizobium meliloti* directing the biosynthesis of galactoglucan: genetic organization and properties of the encoded gene products. *J. Bacteriol.* **179**:1375–1384.
- Bleiweis, A. S., and R. M. Krause. 1965. The cell walls of group D streptococci. I. The immunochemistry of the type 1 carbohydrate. *J. Exp. Med.* **122**:237–249.
- Bronner, D., B. R. Clarke, and C. Whitfield. 1994. Identification of an ATP-binding cassette transporter system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumoniae* serotype O1. *Mol. Microbiol.* **14**:505–519.
- Chu, S., and T. J. Trust. 1993. An *Aeromonas salmonicida* gene which influences a-protein expression in *Escherichia coli* encodes a protein containing an ATP-binding cassette and maps beside the surface array protein gene. *J. Bacteriol.* **175**:3105–3114.
- Chung, C. T., S. Z. Niemela, and R. H. Miller. 1989. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. USA* **86**:2172–2175.
- Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane protein, the dense alignment surface method. *Protein Eng.* **10**:673–676.
- Fernandez, R. C., and A. A. Weiss. 1994. Cloning and sequencing of a *Bordetella pertussis* serum resistance locus. *Infect. Immun.* **62**:4727–4738.
- Freiberg, C., R. Fellay, A. Bairoch, W. J. Broughton, A. Rosenthal, and X. Perret. 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* **387**:394–401.
- Guidolin, A., J. K. Morona, R. Morona, D. Hansman, and J. C. Paton. 1994. Nucleotide sequence analysis of genes essential for capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 19F. *Infect. Immun.* **62**:5384–5396.
- Guo, D., M. G. Bowden, R. Pershad, and H. B. Kaplan. 1996. The *Myxococcus xanthus* *rfbABC* operon encodes an ATP-binding cassette transporter homolog required for O-antigen biosynthesis and multicellular development. *J. Bacteriol.* **178**:1631–1639.
- Hancock, I. C. 1991. Microbial cell surface architecture, p. 21–59. In N. Mozes, P. S. Handley, H. J. Busscher, and P. G. Rouxhet (ed.), *Microbial cell surface analysis: structure and physicochemical methods*. VCH Publishers, Inc., New York, N.Y.
- Harlow, E., and D. Lane. 1988. *Antibodies: a laboratory manual*, p. 498. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Henderson, B., S. Poole, and M. Wilson. 1996. Bacterial modulins: a novel class of virulence factors which cause host tissue pathology by inducing cytokine synthesis. *Microbiol. Rev.* **60**:316–341.
- Ivashina, T. V., M. R. Sadykov, A. A. Kanapin, and V. N. Ksenzenko. Unpublished data.
- Jelinkova, J., and J. Rotta. 1978. Identification and typing of enterococci. *Methods Microbiol.* **12**:199–220.
- 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.
- Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata. 1995. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. *DNA Res.* **2**:153–166.
- Kolkman, M. A., D. A. Morrison, B. A. van der Zeijst, and P. J. Nuijten. 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: identification of glycosyl transferase gene *cps14E*. *J. Bacteriol.* **178**:3736–3741.
- Kolkman, M. A. B., W. Wakarchuk, P. J. M. Nuijten, and B. A. M. van der Zeijst. 1997. Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. *Mol. Microbiol.* **26**:197–208.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105–132.
- Lazarevic, V., and D. Karamata. 1995. The *tagGH* operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids. *Mol. Microbiol.* **16**:345–355.
- Liu, D., A. M. Haase, L. Lindqvist, A. A. Lindberg, and P. R. Reeves. 1993. Glycosyl transferases of O-antigen biosynthesis in *Salmonella enterica*: identification and characterization of transferase genes of groups B, C2, and E1. *J. Bacteriol.* **175**:3408–3413.
- Liu, D., and P. R. Reeves. 1994. *Escherichia coli* K12 regains its O antigen. *Microbiology* **140**(Pt. 1):49–57.
- Macpherson, D. F., P. A. Manning, and R. Morona. 1994. Characterization of the TDP-rhamnose biosynthesis genes encoded in the *rfb* locus of *Shigella flexneri*. *Mol. Microbiol.* **11**:281–292.
- Maki, H., T. Yamaguchi, and K. Murakami. 1994. Cloning and characterization of a gene affecting the methicillin resistance level and the autolysis rate in *Staphylococcus aureus*. *J. Bacteriol.* **176**:4993–5000.
- Marolda, C. L., and M. A. Valvano. 1995. 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. *J. Bacteriol.* **177**:5539–5546.
31. Meier-Dieter, U., K. Barr, R. Starman, L. Hatch, and P. D. Rick. 1992. Nucleotide sequence of the *Escherichia coli* *rfe* gene involved in the synthesis of enterobacterial common antigen. *J. Biol. Chem.* **267**:746–753.
 32. Morona, J. K., R. Morona, and J. C. Paton. 1997. Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthesis pathway. *Mol. Microbiol.* **23**:751–763.
 33. Morona, R., M. Mavris, A. Fallarino, and P. A. Manning. 1994. Characterization of the *rfe* region of *Shigella flexneri*. *J. Bacteriol.* **176**:733–747.
 34. Murray, B. E. 1990. The life and times of the enterococcus. *Clin. Microbiol. Rev.* **3**:46–65.
 35. Murray, B. E., K. V. Singh, R. P. Ross, J. D. Heath, G. M. Dunny, and G. M. Weinstock. 1993. Generation of restriction map of *Enterococcus faecalis* OG1 and investigation of growth requirements and regions encoding biosynthetic function. *J. Bacteriol.* **175**:5216–5223.
 36. 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.* **26**:999–1001.
 37. Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**:1–6.
 38. Pazur, J. H., A. Cepure, and J. A. Kane. 1973. Glycans from streptococcal cell walls. The molecular structure of an antigenic diheteroglycan of glucose and galactose from *Streptococcus faecalis*. *J. Biol. Chem.* **248**:279–284.
 39. Pazur, J. H., and S. L. Forsberg. 1978. Determination of the sugar sequences and the glycosidic-bond arrangements of immunogenic heteroglycans. *Carbohydr. Res.* **60**:167–178.
 40. Pazur, J. H., and E. W. Shuey. 1960. The enzymatic synthesis of thymidine diphosphate glucose and its conversion to thymidine diphosphate rhamnose. *J. Biol. Chem.* **236**:1780–1785.
 41. Rick, P. D., G. L. Hubbard, and K. Barr. 1994. Role of the *rfe* gene in the synthesis of the O8 antigen in *Escherichia coli* K-12. *J. Bacteriol.* **176**:2877–2884.
 42. 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.
 43. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 44. Schnaitman, C. A., and J. D. Klena. 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol. Rev.* **57**:655–682.
 45. Sharpe, M. E., and P. M. F. Shattock. 1952. The serological typing of group D streptococci associated with outbreaks of neonatal diarrhoea. *J. Gen. Microbiol.* **6**:150–165.
 46. Shockman, G. D., and J. F. Barrett. 1983. Structure, function and assembly of cell walls of gram positive bacteria. *Annu. Rev. Microbiol.* **37**:501–527.
 47. Smyth, J., H. Matthews, M. K. Halpenny, H. Brandis, and G. Colman. 1987. Biotyping, serotyping and phage typing of *Streptococcus faecalis* isolated from dental plaque in the human mouth. *J. Med. Microbiol.* **23**:45–54.
 48. Stellwagen, A. E., and N. L. Craig. 1997. Gain-of-function mutations in TnsC, an ATP-dependent transposition protein that activates the bacterial transposon Tn7. *Genetics* **145**:573–585.
 49. 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.
 50. Stingle, F., J.-R. Neeser, and B. Mollet. 1996. Identification and characterization of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. *J. Bacteriol.* **178**:1680–1690.
 51. Tsukioka, Y., Y. Yamashita, Y. Nakano, T. Oho, and T. Koga. 1997. Identification of a fourth gene involved in dTDP-rhamnose synthesis in *Streptococcus mutans*. *J. Bacteriol.* **179**:4411–4414.
 52. Tsukioka, Y., Y. Yamashita, T. Oho, Y. Nakano, and T. Koga. 1997. Biological function of the dTDP-rhamnose synthesis pathway in *Streptococcus mutans*. *J. Bacteriol.* **179**:1126–1134.
 53. van Kranenburg, R., J. D. Marugg, I. I. van Swam, N. J. Willem, and W. M. de Vos. 1997. Molecular characterization of the plasmid-encoded *eps* gene cluster essential for exopolysaccharide biosynthesis in *Lactococcus lactis*. *Mol. Microbiol.* **24**:387–397.
 54. Wang, J., S. Lory, R. Ramphal, and S. Jin. 1996. Isolation and characterization of *Pseudomonas aeruginosa* genes inducible by respiratory mucus derived from cystic fibrosis patients. *Mol. Microbiol.* **22**:1005–1012.
 55. Wang, L., L. R. Romana, and P. R. Reeves. 1992. Molecular analysis of a *Salmonella enterica* group E1 *rfb* gene cluster: O antigen and the genetic basis of the major polymorphism. *Genetics* **130**:429–443.
 56. Whitfield, C., and M. A. Valvano. 1993. Biosynthesis and expression of cell-surface polysaccharides in gram-negative bacteria. *Adv. Microb. Physiol.* **35**:135–246.
 57. Xu, Y., L. Jiang, B. E. Murray, and G. M. Weinstock. 1997. *Enterococcus faecalis* antigens in human infections. *Infect. Immun.* **65**:4207–4215.
 58. Yamazaki, M., L. Thorne, M. Mikolajczak, R. W. Armentrout, and T. J. Pollock. 1996. Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. *J. Bacteriol.* **178**:2676–2687.
 59. Yao, Z., and M. A. Valvano. 1994. Genetic analysis of the O-specific lipopolysaccharide synthesis region (*rfb*) of *Escherichia coli* K-12 W3110: identification of genes that confer group 6 specificity to *Shigella flexneri* serotype Y and 4a. *J. Bacteriol.* **176**:4133–4143.
 60. Yoshida, Y., Y. Nakano, Y. Yamashita, and T. Koga. 1998. Identification of a genetic locus essential for serotype b-specific antigen synthesis in *Actinobacillus actinomycetemcomitans*. *Infect. Immun.* **66**:107–114.
 61. Zhang, L., A. Al-Hendy, P. Toivanen, and M. Skurnik. 1993. Genetic organization and sequence of the *rfb* gene cluster of *Yersinia enterocolitica* serotype O:3: similarities to the dTDP-L-rhamnose biosynthesis pathway of *Salmonella* and to the bacterial polysaccharide transport systems. *Mol. Microbiol.* **9**:309–321.