JOURNAL OF BACTERIOLOGY, Mar. 1996, p. 1680–1690
0021-9193/96/$04.00+0

Vol. 178, No. 6

Identification and Characterization of the *eps* (Exopolysaccharide) Gene Cluster from *Streptococcus thermophilus* Sfi6

FRANCESCA STINGELE,* JEAN-RICHARD NEESER, AND BEAT MOLLET

Nestlé Research Center, Nestec Ltd., Vers-chez-les-Blanc, 1000 Lausanne 26, Switzerland

Received 11 October 1995/Accepted 1 December 1995

We report the identification and characterization of the *eps* gene cluster of *Streptococcus thermophilus* Sfi6 required for exopolysaccharide (EPS) synthesis. This report is the first genetic work concerning EPS production in a food microorganism. The EPS secreted by this strain consists of the following tetrasaccharide repeating unit: →3)-β-D-Gal*p*-(1→3)-[α-D-Gal*p*-(1→6)]-β-D-Glc*p*-(1→3)-α-D-Gal*p*NAc-(1→. The genetic locus was identified by Tn*916* mutagenesis in combination with a plate assay to identify Eps mutants. Sequence analysis of the gene region, which was obtained from subclones of a genomic library of Sfi6, revealed a 15.25-kb region encoding 15 open reading frames. EPS expression in the non-EPS-producing heterologous host, *Lactococcus lactis* MG1363, showed that within the 15.25-kb region, a region with a size of 14.52 kb encoding the 13 genes *epsA* to *epsM* was capable of directing EPS synthesis and secretion in this host. Homology searches of the predicted proteins in the Swiss-Prot database revealed high homology (40 to 68% identity) for *epsA*, *B*, *C*, *D*, and *E* and the genes involved in capsule synthesis in *Streptococcus pneumoniae* and *Streptococcus agalactiae*. Moderate to low homology (37 to 18% identity) was detected for *epsB*, *D*, *F*, and *H* and the genes involved in capsule synthesis in *Staphylococcus aureus* for *epsC*, *D*, and *E* and the genes involved in exopolysaccharide I (EPSI) synthesis in *Rhizobium meliloti* for *epsC* to *epsJ* and the genes involved in lipopolysaccharide synthesis in members of the *Enterobacteriaceae*, and finally for *epsK* and *lipB* of *Neisseria meningitidis*. Genes (*epsJ*, *epsL*, and *epsM*) for which the predicted proteins showed little or no homology with proteins in the Swiss-Prot database were shown to be involved in EPS synthesis by single-crossover gene disruption experiments.

Lactic acid bacteria (LAB) are capable of producing a wide array of extracellular polysaccharides which are secreted into the medium. Besides the commonly known homopolysaccharides of the dextran type produced, for example, by *Leuconostoc mesenteroides*, LAB produce a very heterogeneous group of extracellular heteropolysaccharides referred to as exopolysaccharides (EPS) (6, 7). Fermentation of milk by LAB producing EPS generally leads to a culture with a ropy or viscous texture, even though only small amounts of EPS are being produced (60 to 400 mg/liter) (6, 7). The texture can be extremely slimy, as with Scandinavian fermented milk products ("taette," "viili," or "longfil"), to moderately creamy, as with yogurt (6, 7, 43, 52). In addition to improving texture, EPS has been shown to reduce syneresis (whey separation) in yogurt, and some reports have claimed advantageous biological properties, such as antitumor activity (6, 7, 53).

Interest in EPS has increased over the last decade in the food industry, and several EPS structures have been elucidated. These include those of the EPS from *Streptococcus thermophilus* (12), *Lactobacillus delbrueckii* subsp. *bulgaricus* (25), *Lactobacillus helveticus* TY1-2 and 766 (60, 77), *Lactococcus lactis* subsp. *cremoris* H414 and SBT 0495 (24, 50), and *Lactobacillus sake* (61). They show few common features, which raises the question about the relation between these EPS structures and the texturizing properties they confer. It would be of interest to understand this relation and to have the means to produce EPS tailored to more desirable properties. Since EPS are produced in situ by food microorganisms, they have considerable potential to become better alternatives for the currently used thickeners and stabilizers.

The instability of EPS production strains at the genetic level, as well as the instability of the ropy texture itself, is a well-known problem in the dairy industry (6, 7). Even at optimum growth temperatures, ropiness can be lost after numerous transfers and prolonged periods of incubation (6, 7, 43). Therefore, ropy strains have to be periodically reselected from cultures to conserve the EPS production characteristics for industrial strains. In mesophilic LAB (e.g., *L. lactis* subsp. *cremoris* and *Lactococcus lactis* subsp. *lactis*), EPS production is generally associated with a plasmid, and genetic instability can be explained by the loss of that plasmid (43, 52, 73–75). Numerous mucoidity plasmids have been identified (52, 72–75). von Wright and Tynkkynen (75) and Vedamuthu and Neville (73) were able to transfer the mucoidity plasmid from *L. lactis* subsp. *cremoris* ARH87 and MS, respectively, to nonmucoid *L. lactis* strains, thereby making them mucoid. On the other hand, the EPS production in thermophilic LAB has not been associated with the presence of plasmids (6, 7, 74). The genes required for EPS synthesis seem therefore to be chromosomally located, and the genetic instability could be due to mobile genetic elements or to a generalized genomic instability, including deletions and rearrangements. Both phenomenons were observed for *L. delbrueckii* subsp. *bulgaricus* and *S. thermophilus* (22, 26, 44, 63).

Despite scientific interest and economical importance, no genetic study of EPS production in LAB has been reported so far. Here we present the first genetic study of EPS production with *S. thermophilus*. This microorganism is used together with *L. delbrueckii* subsp. *bulgaricus* as a starter organism in yogurt fermentation. The smooth, creamy texture of yogurt is determined by the ability of at least one of these organisms to produce EPS. However, during fermentations of short dura-

* Corresponding author. Mailing address: Nestlé Research Center, Nestec Ltd., P.O. Box 44, Vers-chez-les-Blanc, 1000 Lausanne 26, Switzerland. Phone: 41-21-7858923. Fax: 41-21-7858925. Electronic mail address: francesca.stingele@chlsnr.nestrd.ch.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
S. thermophilus strains		
ST11	Yogurt production strain, nonropy	Nestlé strain collection
YS4	Yogurt production strain, nonropy	Nestlé strain collection
Sfi6	Yogurt production strain, ropy	Nestlé strain collection
Sfi6-1	Same as Sfi6, but streptomycin resistant	This study
E. coli strains		
XL1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB laqI*^{q}ZΔ*M15* Tn*10* (Tet^r)]	Stratagene
XLOLR	Δ(*mcrA*)*183* Δ(*mcrBC-hsdSMR-mrr*)*173 endA1 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB laqI*^{q}ZΔ*M15* Tn*10* (Tet^r)]	Stratagene
E. faecalis JH2-2(pAM180)	Carries Tn*916* on pAM180	21
L. lactis MG1363	Plasmid free (Lac$^-$ Prt$^-$)	20
Plasmids		
pCI182	Carries *tetM* from Tn*919*	30
pSA3	Temperature-sensitive shuttle vector	10
pJIM2279	Gram-positive cloning vector	58
pCMV	*E. coli* cloning vector inserted in λ-ZAP-Express	Stratagene
pFS14	pCMV with residues 9962–15250 from *eps* gene cluster	This study
pFS15	pCMV with residues 1–3807 from *eps* gene cluster	This study
pFS26	pCMV with residues 8636–11901 from *eps* gene cluster	This study
pFS30	pCMV with residues 2349–8639 from *eps* gene cluster	This study
pFS33	pCMV with residues 2008–7877 from *eps* gene cluster	This study
pFS49	pCMV with residues 5047–9125 from *eps* gene cluster	This study
pFS50	pCMV with residues 1443–8324 from *eps* gene cluster	This study
pFS65	pCMV with residues 6645–11901 from *eps* gene cluster	This study
pFS73	pCMV with residues 9122–14137 from *eps* gene cluster	This study
pFS80	pCMV with residues 4197–7877 from *eps* gene cluster	This study
pFS86	pCMV with residues 1443–4200 from *eps* gene cluster	This study
pFS101	pJIM2279 containing 14.5-kb *Sac*I-*Bam*HI fragment from *eps* gene cluster	This study

tion, we have shown that if EPS-producing strains of both species are used, the EPS produced by *S. thermophilus* is predominant (unpublished observation). We previously presented two approaches to chromosomal mutagenesis in *S. thermophilus* (69). One approach consisted of disrupting genes by homologous integration of plasmids containing randomly cloned genomic fragments (45, 69). The other approach involved transposon mutagenesis with the conjugative transposon Tn*916*. Tn*917* and IS*946*, which were reported as active in *L. lactis* (32, 62), were also tested but could not be successfully used in *S. thermophilus*. In this paper, we present the *S. thermophilus eps* gene cluster which was identified by Tn*916* mutagenesis and a plate assay to differentiate ropy and nonropy colonies.

MATERIALS AND METHODS

Bacterial strains, media, and plasmids. The plasmids and bacterial strains used in this study are listed in Table 1. *Escherichia coli* strains were routinely cultured in Luria-Bertani broth, *S. thermophilus* strains were cultured in HJL broth (3% tryptone, 1% yeast extract, 1% lactose, 0.2% beef extract, 0.5% KH_2PO_4 [pH 6.5]) or M17 broth supplemented with 1% lactose (LM17), *L. lactis* MG1363 was cultured in M17 broth supplemented with 1% glucose (GM17), and *Enterococcus faecalis* was cultured in M17. When required, antibiotics were added at the following concentrations: ampicillin, 100 μg/ml; kanamycin, 50 μg/ml; chloramphenicol, 30 μg/ml; tetracycline, 2.5 μg/ml for *S. thermophilus* and 10 μg/ml for *E. faecalis*; streptomycin, 2,000 μg/ml; and erythromycin, 5 μg/ml for *L. lactis* and *S. thermophilus* and 150 μg/ml for *E. coli*. For solid media, Bacto Agar was added at a final concentration of 1.5%. Ruthenium red milk plates consisted of 0.5% yeast extract, 10% skim milk powder, 1% sucrose, 1.5% agar, and 0.08 g of ruthenium red (Fluka) per liter (69); additionally, this agar had to be enriched with amino acids and glucose (1%) for growth of *L. lactis* MG1363.

EPS isolation and characterization. Fermentation of *S. thermophilus* Sfi6 was performed in 10% reconstituted skim milk powder supplemented with a mixture of amino acids in quantities corresponding to those found in 1% Proteose Peptone No. 3 (Oxoid). Fermentations were carried out in a 1-liter-scale fermentor with a magnetic stirrer (60 rpm) for 24 h at 40°C and were regulated at pH 5.5 with 2 N NaOH. The EPS was extracted as follows. Proteins were removed from the spent culture with an equal volume of trichloroacetic acid (40%). After centrifugation, an equal volume of acetone was added to precipitate the EPS. The precipitated EPS was recovered by centrifugation and redissolved in water, and the solution was adjusted to pH 7.0 before dialysis against water for 24 h. Insoluble material was removed by ultracentrifugation, and the supernatant containing the EPS was freeze-dried. The EPS produced by *L. lactis* MG1363 containing pJIM2279 or pFS101 was isolated accordingly, except that 1% glucose was added to the growth medium and fermentations were conducted at 30°C. The total neutral sugar content was determined by the phenol-sulfuric acid method (15). The monosaccharide composition was determined by gas-liquid chromatography after acid hydrolysis (51). The molecular weight of the EPS was estimated by gel permeation chromatography with a Superose-6 column connected to a fast protein liquid chromatography system (Pharmacia) previously calibrated with commercial dextrans (Sigma).

DNA manipulations. All common DNA manipulations, including preparation of plasmid DNA from *E. coli*, restriction nuclease digestions, ligation, agarose gel electrophoresis, and Southern blot hybridizations of DNA, were done according to Sambrook et al. (66). Plasmid DNA was isolated from *L. lactis* as described previously (11). Restriction enzymes and T4 ligase were purchased from Boehringer, and ^{32}P-labeled probes were prepared with the random priming labeling kit from Boehringer. Conventional PCR and inverted PCR (InPCR) were performed as reported previously (65, 71). The sequences of the primers used are indicated in Table 2. Genomic DNA of *S. thermophilus* was prepared as reported previously (67). High-molecular-weight DNA for pulsed-field gel electrophoresis was obtained as previously described (63). Pulsed-field gel electrophoresis was carried out with the Rotaphor R23 system (Biometra) for 40 h with linear

TABLE 2. Primers cited in this study

Primer	Sequence
1536	GTTGCGGCCGCGATAAAGTGTGATAAGTCCAG
1537	TACGCGGCCGCACATAGAATAAGGCTTTACG
1545	ATAGCGGCCGCTTAGCTCATGTTGATGCGG
1546	CCTGCGGCCGCGCTTCCTAATTCTGTAATCG
2132	ATATGCCACCGATTTTT
2148	TTATTATGGAGGTGAAC
2151	TCTCACTCGCTATCAAA
2156	TAAGTGGCGGATTTTTA
2159	AGAAATATCAGTTTGGC
2176	CCAACCACTGCAAGGAA
2195	ACTTGAGTAATAAATGG
2198	CTCTTGAATCGTTCTGG

switching intervals from 8 to 35 s and a linear angle field from 120 to 135° on a 1% agarose gel (Bio-Rad).

Genetic transformation methods. *L. lactis* MG1363 and *E. coli* strains were transformed by electroporation (11, 13). *S. thermophilus* strains were electroporated by a modification of the method of Slos et al. (67), in which 30 ml of Belliker broth (Elliker broth [16] with 1% beef extract) supplemented with 20 mM DL-threonine was inoculated with 1% of an overnight culture grown in Belliker broth and the inoculum was grown to an optical density at 600 nm (OD_{600}) of 0.2 to 0.4. The cells were harvested, washed twice with precooled electroporation buffer (272 mM sucrose, 1 mM EDTA, 7 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid] [pH 6.5], and 15% glycerol), and resuspended in the same buffer to a final OD_{600} of 1.8. The cells were aliquoted and frozen at −80°C. For electroporations, 0.1 μg of plasmid DNA and 200 μl of thawed cells were mixed in a precooled tube and transferred to a 2-mm-gap electroporation cuvette (Bio-Rad). After a single 2.1-kV, 25-μF, 400-Ω (time constant, 6 to 8 ms pulse) was delivered with a Bio-Rad Gene Pulser, the cells were immediately resuspended in 1 ml of HJL broth. Phenotypic expression was allowed for 4 to 6 h at 42°C. Aliquots were plated on LM17 plates supplemented with the appropriate antibiotic and incubated at 42°C anaerobically overnight.

Directed genomic integration in *S. thermophilus* with temperature-sensitive vector pSA3. *S. thermophilus* cells were transformed with pSA3 containing a fragment of the *S. thermophilus* Sfi6 genome, incubated for 6 h at 37°C in HJL broth, plated on LM17 plates with erythromycin, and incubated for 24 h at 37°C. Transformants were selected and grown at 37°C to an OD_{600} of 0.2 in 2 ml of HJL broth supplemented with erythromycin. The cultures were shifted to 46°C and grown to an OD_{600} of 1.0. Dilutions of the cultures were plated on LM17 plates supplemented with erythromycin and incubated at 45°C. pSA3 integrants appeared as erythromycin-resistant colonies after 16 to 24 h of incubation.

Tn*916* mutagenesis of *S. thermophilus* Sfi6. Overnight cultures of *E. faecalis* JH2-2 and *S. thermophilus* Sfi6-1 were mixed at a donor-to-recipient ratio of 1:10, centrifuged, and resuspended in 100 μl of HJL broth. The cells were deposited on an LM17 plate, overlaid with LM17 top agar, and incubated for 20 h at 37°C. After mating, the cells were resuspended in 10 ml of HJL broth and incubated for 4 h at 42°C. Transconjugants, donors, and recipients were titrated on LM17 plates supplemented with 2.5 μg of tetracycline per ml and 2,000 μg of streptomycin per ml, on LM17 plates supplemented with 10 μg of tetracycline per ml, and on LM17 plates supplemented with 2,000 μg of streptomycin per ml, respectively.

Construction of a genomic library of *S. thermophilus* in λ-ZAP Express. The λ-ZAP Express system from Stratagene was utilized to construct a library of the *S. thermophilus* Sfi6 genome. Genomic DNA was partially *Sau*3A digested, and fragments with sizes of from 4 to 12 kb were extracted from an agarose gel. Genomic DNA (100 ng) was ligated with 500 ng of λ-ZAP DNA which was previously ligated to protect the *cos* ends from dephosphorylation, cut with *Bam*HI, and dephosphorylated. The ligation mixture was packaged with the Giga-pack gold III system (Stratagene) and plated with a lawn of *E. coli* XL1 Blue cells. Recombinant phage plaques were screened by plaque hybridization with Hybond-N membranes (Amersham) and ^{32}P-labeled probes. pCMV vectors containing the genomic inserts were excised from phages of positive plaques and transferred to *E. coli* XLOLR cells as described by the supplier.

DNA sequencing methods and sequence analysis. pCMV clones obtained by screening the genomic library were used to sequence the *eps* gene cluster. DNA sequencing was performed with the f-mol sequencing kit (Promega), with 100 ng of plasmid DNA or 50 ng of PCR product and [γ-^{33}P]ATP-labeled oligonucleotides being used as primers. The sequences of the pCMV clones were completed by primer walking. Custom-made sequencing primers were purchased from Microsynth (Balgach, Switzerland). Searches of the Swiss-Prot database for protein sequences homologous to the deduced amino acid sequences were performed with the Pearson and Lipman algorithm (54) in FASTA, and hydropathy plots were performed with the Kyte and Doolittle algorithm (35) in PEPPLOT of the Wisconsin Sequence Analysis Package, Genetics Computer Group (University of Wisconsin).

Cloning of the *eps* gene cluster in *L. lactis* MG1363. Genomic DNA (150 μg) from *S. thermophilus* Sfi6 was digested with *Bam*HI and *Sac*I and size fractionated on a 0.7% agarose electrophoresis gel, and the region between 12 and 16 kb was extracted. DNA from that fraction was ligated into vector pJIM2799 previously digested with *Sac*I and *Bam*HI and dephosphorylated. DNA (100 ng) was transformed into *L. lactis*, and transformants were analyzed by colony hybridization with probes P1 and P3 (see Fig. 4).

Nucleotide sequence accession number. The nucleotide sequence reported here has been submitted to GenBank under accession number U40830.

RESULTS

EPS-producing properties of *S. thermophilus* Sfi6. The stabilities of 20 ropy *S. thermophilus* strains of the Nestlé strain collection were examined in milk culture and on ruthenium red milk plates. Ruthenium red stains the bacterial cell wall, thus producing red colonies for nonropy strains. The production of EPS prevents this staining, and hence ropy colonies appear white on the same plates. All ropy *S. thermophilus* strains showed white colonies, but after several passages of fermentation, red colonies appeared, implying an instability of the ropy character in these strains. Strain Sfi6 presented a stable ropy phenotype in this plate assay and was therefore chosen for this study. After growth of strain Sfi6 for 200 generations, we were not able to detect a red colony in 1,000 colonies tested. Further investigation of Sfi6 EPS production revealed the following characteristics. When Sfi6 was used for fermentation until coagulation (pH 5.2), the EPS brought an important normal force of 110×10^{-3} N to the skim milk broth (43a). Under fermentation conditions with pH compensation at pH 5.5, Sfi6 yielded 175 mg of EPS per liter. This EPS had a high molecular mass of approximately 2×10^6 Da. Finally, its composition was determined by gas-liquid chromatography, which revealed a monosaccharide composition of Glc-Gal-GalNAc in molar ratios of 1:2:1. Lemoine identified the structure of this EPS as consisting of the →3)-β-D-Gal*p*-(1→3)-[α-D-Gal*p*-(1→6)]-β-D-Glc*p*-(1→3)-α-D-Gal*p*NAc-(1→ tetrasaccharide repeating unit (39). This structure had been previously reported for three other strains of *S. thermophilus* (12).

Identification of the *eps* locus. In order to inactivate and genetically tag the genes involved in EPS production, transposon mutagenesis was employed in combination with the ruthenium red plate assay. The *E. faecalis* JH2-2 donor strain carrying a copy of Tn*916* on conjugative plasmid pAM180 was used to introduce Tn*916* into *S. thermophilus* Sfi6-1. Approximately 2×10^4 tetracycline-resistant transconjugants from 20 independent mating experiments were replica plated onto ruthenium red plates, giving a frequency of 10% for the number of red cells in the total number of Tn*916* integrants. Red colonies ($n = 800$) were selected and tested for the loss of their ropy character and stability in milk culture. Only 25% (ca. 200) of the red colonies tested proved to have stably lost their ropy character, whereas the others were still ropy or reverted to the ropy phenotype after one or two transfers in milk.

Stable nonropy mutants ($n = 100$) were analyzed by Southern blotting of *Hin*dIII-digested genomic DNA (*Hin*dIII cuts once within the *tetM* gene of Tn*916*; see Fig. 3) and hybridization with the *tetM* gene (excised from pCI182) to map the Tn*916* integration site. A comparison of the banding patterns revealed one major hot spot into which Tn*916* integrated in 90% of the mutants analyzed (Fig. 1). The consistent two *Hin*dIII bands (8.5 and 13 kb in size) which correspond to the two Tn*916* arms (left and right of the *Hin*dIII site in *tetM*) indicate that Tn*916* (16.5 kb) had integrated into a 5-kb *Hin*dIII fragment. The fact that in independent nonropy mutants the same locus was tagged strongly suggested that this locus was involved in ropiness. The two additional bands of Tn*916* integrants in lanes 1 to 5 and 7 to 9 in Fig. 1 indicate that in these mutants, Tn*916* integrated into a second genomic locus which is different in each of them. This second copy of Tn*916* in these mutants could be eliminated by repeated transfers and growth in milk and reselection of red colonies. These red colonies eventually showed a *Hin*dIII banding pattern like that of the Tn*916* integrant in lane 10 of Fig. 1.

The Tn*916* integrant in lane 10 of Fig. 1 which carries only a single copy of Tn*916* was selected for further analysis. We tested whether upon Tn*916* excision the ropy phenotype could be reestablished. Chaparon and Scott have shown that Tn*916* excision, which spontaneously occurs after the tetracycline selection pressure is relieved, leaves the original target site basically intact (8). After growth of this Tn*916* integrant without tetracycline for 100 generations, white colonies appeared on ruthenium red plates. Milk cultures of these revertants were

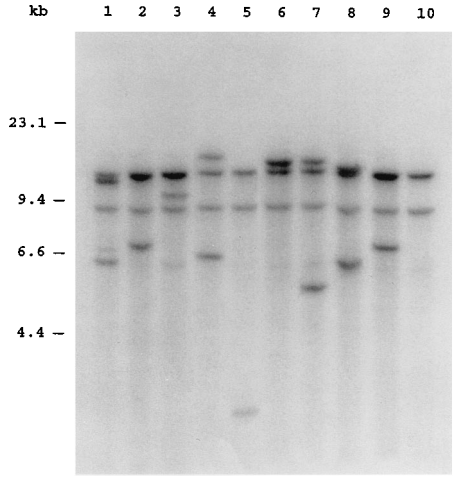


FIG. 1. Southern blot hybridization of independent Tn*916* integration mutants that have lost their ropy phenotype. Chromosomal DNA was digested with *Hin*dIII and probed with the *tetM* gene from pCI182. Lane 10 represents a single-copy Tn*916* integrant, whereas lanes 1 to 5 and 7 to 9 represent double-copy Tn*916* integrants. (In lanes 2, 5, and 9, only three bands are visible. The fourth band is at 13.5 kb and therefore cannot be seen as a distinct band.) The integrant in lane 6 carries a tandem Tn*916* integration.

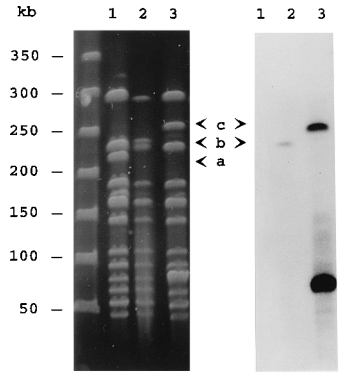


FIG. 2. Pulsed-field gel electrophoresis (left) of *Sma*I-digested genomic DNA from Tn*916* integrants and Southern hybridization (right) with the *tetM* gene from Tn*916* being used as the probe. Lanes 1, *S. thermophilus* Sfi6; lanes 2 and 3, two independent Tn*916* integration mutants, of which the latter carries a tandem integration in the *eps* locus. Shown are the 230-kb *Sma*I band (a) from *S. thermophilus* Sfi6 into which Tn*916* integrated once (b) and twice (c) as a tandem integration. The lower 70-kb *Sma*I band hybridizing in lane 3 indicates the integration of a third copy of Tn*916*, but that is irrelevant for this study.

ropy, proving that integration of Tn*916* was responsible for the loss of ropiness.

We further analyzed the growth medium of this Tn*916* integrant for the presence of EPS. Under fermentation conditions in which 175 mg of EPS per liter was produced by Sfi6, no EPS was detected for the nonropy mutant. This result proved that the Tn*916* integrant had lost the ability to produce EPS and that the ropy character in Sfi6 can be directly linked to EPS production. This new genetic locus, identified by Tn*916* mutagenesis and involved in EPS synthesis, was hereafter named the *eps* locus.

Mapping of the *eps* locus. To verify that the *eps* locus was genomically located, two Tn*916* integration mutants were selected and analyzed by pulsed-field gel electrophoresis and Southern hybridization, with the *tetM* gene of Tn*916* being used as a probe (Fig. 2). As expected, the Sfi6 wild-type strain did not show a hybridization signal (Fig. 2, lane 1). The first mutant showed a signal at 245 kb (Fig. 2, lane 2, band b), which corresponds to the 230-kb *Sma*I band (lane 1, band a) increased in size by 16.5 kb as a consequence of the Tn*916* integration. The second mutant has a tandem Tn*916* integration, and the band hybridizing is increased by 33 kb to 260 kb (Fig. 2, lane 3, band c). The third Tn*916* insertion site, which is irrelevant for the ropy phenotype, is located on the 70-kb *Sma*I band. From previous genomic mapping experiments with *S. thermophilus*, we know that this 230-kb band is genomic and constitutes 13% of the 1.8-Mb genome (63) (data not shown). The *eps* locus is therefore located on the 230-kb genomic *Sma*I fragment of *S. thermophilus* Sfi6.

Cloning and sequencing of the *eps* locus. A λ-ZAP Express genomic library of Sfi6 was constructed to recover the gene regions adjacent to the Tn*916* insertion site. Regions flanking Tn*916* were obtained by InPCR and were used as probes to identify recombinant phage plaques carrying genomic fragments of the *eps* locus. The positions and orientations of primers employed for InPCR are indicated in Fig. 3. With *Hin*dIII-digested and religated genomic DNA and primers 1536 and 1545 or primers 1546 and 1537, the right and left flanking regions could be recovered, yielding fragments with sizes of 1.0 and 4.0 kb, respectively (probes P1 and P2) (Fig. 4). From the first round of screening of 3,000 recombinant plaques, 9 positive plaques were isolated. pCMV clones excised from these positive recombinant phage plaques were named pFS14, pFS26, pFS30, pFS33, pFS49, pFS50, pFS65, pFS73, and pFS80 (Fig. 4). They were mapped by restriction enzyme analysis and PCR, with the T7 or T3 primers located at the left and right of the multiple cloning site of pCMV and custom-made primers constructed on the basis of preliminary sequence data obtained from the 1- and 4-kb InPCR fragments. pCMV clones pFS15 and pFS86 were obtained after a second round of screening with probe P3, which was obtained by PCR of clone pFS50 with custom-made primers constructed on the basis of preliminary sequencing data from clone pFS50 (Fig. 4). The exact beginnings and ends of the genomic fragments in pFS clones relative to the whole gene region are given in Table 1. Interestingly, clone pFS14 was very difficult to propagate in *E. coli* XLOLR (maximal growth to an OD_{600} of 0.3) and pCMV clones downstream of this clone could not be recovered from the recombinant phages, suggesting that a gene region incompatible with *E. coli* cell growth is located downstream of pFS73.

The genomic inserts of the pFS subclones were sequenced and revealed a 15.25-kb region encoding 15 open reading frames (ORFs) (Fig. 4). The first incompletely sequenced ORF and the last ORF, which is oriented in the direction opposite to that of the other genes, are predicted not to be involved in EPS biosynthesis. Homology values (percent identity and percent similarity) for the predicted amino acid sequences of the other ORFs and the proteins in the Swiss-Prot database are indicated in Table 3. High homology (>50% identity) was found for proteins involved in capsule synthesis in *Streptococcus pneumoniae* and *Streptococcus agalactiae*, and moderate to low homology (37 to 18% identity) was found for genes involved in EPS synthesis in *Rhizobium meliloti*, in lipopolysaccharide

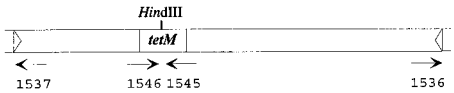


FIG. 3. Schematic representation of Tn*916* showing the unique *Hin*dIII site and the positions and orientations of primers used for InPCR to recover regions flanking Tn*916*.

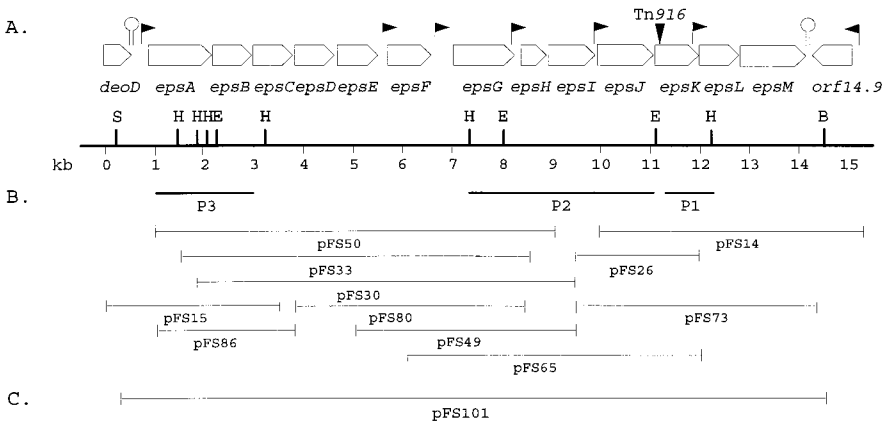


FIG. 4. (A) Physical map of the *eps* gene cluster of *S. thermophilus* Sfi6. Putative promoters and terminators are represented as flags and hairpins, respectively. The arrowhead indicates the Tn*916* insertion site of nonropy *S. thermophilus* Sfi6 Tn*916* integration mutants. Arrows represent potential ORFs, and gene designations are indicated below the arrows. Restriction enzyme abbreviations are as follows: S, *Sac*I; H, *Hin*dIII; E, *Eco*RI; and B, *Bam*HI. (B) pCMV subclones from the λ-ZAP Express genomic library of *S. thermophilus* Sfi6. P1, P2, and P3 indicate the positions of probes used for screening. Genomic inserts of subclones pFS14 to pFS86 which were recovered from positive phage plaques are represented as lines. The exact positions of genomic inserts of these subclones, relative to the sequenced gene region, are given in Table 1. (C) Genomic insert of pFS101 comprising the entire *eps* gene cluster (from *Sac*I to *Bam*HI) which was cloned into pJIM2279.

(LPS) synthesis in *Salmonella typhimurium* LT2, *E. coli*, and *Shigella flexneri*, and in capsule synthesis in *Neisseria meningitidis* and *Staphylococcus aureus*. Two distinct ORFs show homology with a regulatory protein of *Bacillus subtilis* and a family of proteins with a common acetyltransferase motif. Since the homologies of 11 of these 13 identified ORFs suggest an involvement in polysaccharide synthesis and the ORFs are organized in a gene cluster, they were designated *epsA* through *epsM*.

Putative promoters with −35 and −10 sequences similar to the *E. coli* consensus promoter sequence with 46-, 30-, 95-, 36-, 55-, 32-, and 34-bp spacings between the ends of the −10 sequence and the start codons of ORFs *epsA*, *epsF*, *epsG*, *epsH*, *epsJ*, *epsL*, and *orf14.9*, respectively, were identified and downstream of the first incompletely sequenced ORF and *epsM* putative stem-loop terminators could be identified (residues 880 to 900 and 14335 to 14346, respectively). Translation signals (Shine-Dalgarno [SD] sequences and start and stop codons) and the positions and lengths of the ORFs are indicated in Table 4. In front of all the ORFs, putative SD sequences could be identified, with 5- to 10-bp spacings between the ends of the SD sequences and the putative start codons. ORF *epsB* has the putative start codon GUG, and *orf14.9* has the start codon UUG. Overlapping reading frames could be identified only at the end of *epsK* and the beginning of *epsL* (38-bp overlaps).

The GC content of the *eps* locus is lower than the average 40% GC content of *S. thermophilus* (49). Individual values for the GC contents for *eps* genes are illustrated in Fig. 5. Within the *eps* locus, the genes could be grouped according to their GC contents, with an average of 37% for *epsA* to *epsD*, 35% for *epsE* to *epsF*, and 30% for *epsG* to *epsM*. The lowest GC content was identified in *epsK* (27%), which is the ORF where Tn*916* integration occurred. Each group of genes is preceded by a region of extremely low GC content (e.g., 18% in front of *epsA* and 22% in front of *epsG*), which supports the presence of the putative promoters identified in the DNA sequence. These groups of genes might reflect transcriptional units of the *eps* gene cluster.

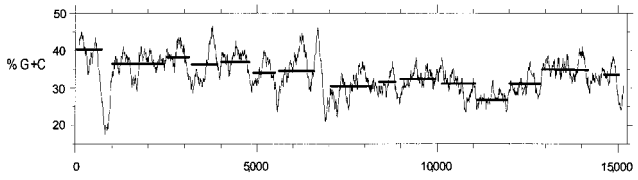


FIG. 5. Graphic representation of the G+C percentages in the *eps* region. Bars indicate the average GC contents of single genes from *deoD* to *orf14.1*. (*x* axis, DNA sequence residues [bp]; plot parameters: window, 200; shift, 4.)

Directed gene disruption of *epsJ*, *epsL*, and *epsM*. Sequence homology searches for the predicted amino acid sequences of genes located at the end of the *eps* gene cluster showed little (*epsJ* and *epsK*) or even no homology (*epsL* and *epsM*) with known proteins in the Swiss-Prot database. The involvement of *epsK* in EPS synthesis was shown by Tn*916* mutagenesis. To assess whether the other three genes were involved in EPS synthesis, we carried out gene disruption experiments. For this purpose, pSA3 integration vectors carrying internal fragments of *epsJ* (residues 10581 to 10892), *epsL* (residues 12051 to 12582), and *epsM* (residues 13167 to 13820) were constructed. *S. thermophilus* Sfi6 was transformed with each of these pSA3 integration vectors, and pSA3 integrants were selected after a temperature shift as described in Materials and Methods.

The integration of pSA3 in *epsJ* resulted in a mutated phenotype for pSA3 integrants, i.e., pSA3 integrants displayed a red phenotype on ruthenium red plates and were nonropy in milk culture, confirming the involvement of *epsJ* in EPS synthesis. Correct integration was verified by Southern hybridization (data not shown). On the other hand, integrations of pSA3 into *epsL* or *epsM* could not be achieved. Even though Sfi6 was successfully transformed with pSA3 containing the internal fragments of *epsL* or *epsM*, no pSA3 integrants could be obtained after the temperature shift. Since genomic integration of pSA3 is a well-established method in our laboratory, we concluded that inactivation of *epsL* or *epsM* is detrimental to cell growth. To further investigate the role of *epsL* and *epsM*, a pSA3 integration vector carrying the end of *epsK* and the beginning of *epsL* (residues 11742 to 12291) was constructed. After transformation of this pSA3 construct and a temperature shift of a selected transformant, pSA3 integrants were obtained. Correct integration was verified by Southern hybridization with three independent integrants (data not shown). The

pSA3 integrants displayed a nonropy phenotype in milk cultures and a light pink (but not red) phenotype on ruthenium red plates. This result indicated that small amounts of EPS were produced, and yet these amounts were not enough to establish a ropy phenotype. Single-crossover integration of this pSA3 construct leaves *epsK* intact. However, it creates a polar effect on *epsL* and *epsM* and interferes with their expression. Since the pSA3 integrants are not ropy, this outcome demonstrates that *epsL* and most probably *epsM* (they are separated by only 7 bp) are involved in EPS production. The fact that the integrants are light pink and not red on indicator plates further suggests weak expression from a promoter in front of *epsL*. In fact, a putative promoter sequence upstream of *epsL* could be identified. Hence, these integrants were capable of expressing enough *epsL* and *epsM* to permit cell growth but not enough to express the ropy phenotype.

***eps* genes in nonropy strains.** Initial data concerning the absence of *eps* genes in nonropy *S. thermophilus* strains were obtained by Southern hybridization. *S. thermophilus* YS4 and ST11, which are representatives of nonropy strains, were hybridized, with the 1- and 4-kb (probes P1 and P2) InPCR fragments being used as probes. These strains revealed no hybridization signals, indicating that nonropy strains do not carry the central region of the Sfi6 *eps* gene cluster (data not shown). More detailed information concerning the Sfi6 *eps* genes in strains YS4 and ST11 was obtained by PCR experiments with the genomic DNA of these strains, with primers at the 5′ and 3′ regions of the *eps* gene cluster. The region encoding *epsA*, *epsB*, and *epsC* (and *epsD* for ST11) and the region encoding *orf14.9* at the 3′ end of the *eps* gene cluster are conserved in both nonropy strains, whereas the region in between those genes is absent (Fig. 6).

Transfer of the *eps* gene cluster to *L. lactis* MG1363. The predicted amino acid sequence of the first incompletely sequenced ORF shows significant homology (49.5% identity) with the sequence of the purine nucleoside phosphorylase (*deoD*) of *E. coli* (29). It seemed doubtful that an enzyme involved in the scavenge pathway of purine bases would be involved in EPS synthesis. Moreover, a putative terminator could be identified downstream of *deoD*, and a putative *E. coli* consensus promoter sequence could be identified upstream of *epsA*. In addition, the sequence of *epsA* shows significant homology with the predicted amino acid sequence of the first gene of the *cps* gene cluster, *cpsA*, in *S. pneumoniae* (27). Therefore, we assumed that *epsA* is also the first gene of the *S. thermophilus* Sfi6 *eps* gene cluster. At the 3′ end, the gene cluster is flanked by *orf14.9*, which is oriented in the direction opposite to that of the *eps* genes. In order to assess if the *eps* gene cluster consisting of 13 genes between *deoD* and *orf14.9* was sufficient to establish EPS production, we cloned this gene region into the nonropy heterologous host *L. lactis* MG1363. A unique *Sac*I site (position 167) was identified within *deoD*, and

FIG. 6. *eps* genes in nonropy *S. thermophilus* ST11 and YS4. The sequences of the primers used are indicated in Table 2. Solid lines indicate the presence of genes as determined by PCR with the given primers. Broken lines indicate the approximate ends of homology. Plus and minus signs indicate if the expected PCR product was obtained.

TABLE 3. Homologies of the predicted proteins of *eps* genes

eps gene	Homologous protein from other microorganisms (reference) (% identity/% similarity)					
	B. subtilis	*S. pneumoniae*	*S. agalactiae*	*S. aureus*	*R. meliloti*	*Enterobacteriaceae* spp.	*N. meningitidis*
A	LytR (37) (28.0/51.0)	CpsA (27) 51.7/70.1)	CpsA (64) (67.5/80.2)	CapC (40) (18.2/42.3)			
B		CpsB (27) (62.1/81.1)	CpsB (64) (52.6/66.5)				
C		CpsC (27) (50.4/67.5)	CpsC (64) (56.0/74.7)		ExoP (3, 4, 59) (26.8/51.3)[a]	Cld (1) (23.1/49.3)[a,b]	
D		CpsD (27) (55.8/72.1)	CpsD (64) (40.8/57.1)[d]	CapB (40) (36.5/56.5)	ExoP (3, 4, 59) (29.7/48.9)[c]		
E		CpsE (27) (43.8/65.9)[d]			ExoY (23, 48, 59) (37.3/65.9)	RfbP (34, 76) (40.6/65.5)[4,e]	
F				CapM (40) (23.2/49.8)			
G				CapH (40) (22.1/49.6)			
H	CysE (19) (31.5/54.8)		CapG (40) (33.3/53.8)	NodL (14) (34.9/53.3)		RfaK (42) (20.3/49.8)	
I						LacA (28) (30.6/54.8)	
J						RfbV (34, 41) (24.2/46.8)[e]	
K						Orf4.1 (Rfe) (38) (20.5/50.1)[f]	LipB (18) (18.4/42.6)

[a] Homology refers to the N-terminal domain of ExoP.
[b] Homology refers to Cld of *S. typhimurium* LT2. EpsC homology with Cld of other organisms: *E. coli* O111 (1) (22.4/49.3), *E. coli* O75 (2) (21.7/45.6), and *S. flexneri* (47) (20.0/41.3).
[c] Homology refers to the C-terminal domain of ExoP.
[d] Homology refers to the C-terminal domain of CpsE, CpsD, or RfbP, respectively.
[e] Homology refers to the predicted proteins coded for by genes in the *S. typhimurium* LT2 *rfb* gene cluster.
[f] *orf4.1* is located in the *S. cholerasuis* (group C1) *rfb* gene cluster.

TABLE 4. Translation signals and positions of start and stop codons of identified ORFs

ORF	Translation signal[a]	Space between SD sequence and start codon (bp)	Beginning and end of ORFs (in bp)	Size of gene product (amino acids)
epsA	tttAGGAGcaatttat**ATG**agt	8	1000–2454	484
epsB	gatGGAGGaaaaa**TAAGTG**att	8	2455–3186	243
epsC	atttaT**AG**GAGatatt**ATG**aat	5	3195–3887	230
epsD	aatt**TA**AGGAGaagaa**ATG**cct	5	3897–4646	249
epsE	aagtGGAGGgaatgag**ATG**tca	7	4699–5382	227
epsF	GGAaaaaATagtaacg**ATG**aat	7	5546–6505	319
epsG	tggaGAGGaaaataat**ATG**aaa	8	7073–8191	372
epsH	gaaagAGGAGGcataa**ATG**ctg	5	8384–8863	159
epsI	agcGtGGctAattaaa**ATG**tat	6	8869–9843	324
epsJ	ataAAGGAagcaacac**ATG**gta	8	9933–11015	360
epsK	gtttagaGAGGaaata**ATG**gag	5	11040–11990	316
epsL	agAGGAAGaagtactg**ATG**aag	8	11950–12873	307
epsM	aAAGG**GTA**aaggttta**ATG**aat	11	12881–14302	473
orf14.9	AAGGAGttagaaacag**TTG**cga	10	14953–14378	191

[a] SD sequences are underlined and capitalized; stop codons from the previous ORF and start codons are in boldface and capitalized.

a unique *Bam*HI site (position 14618) was identified within *orf14.9* (Fig. 4). The 14.5-kb *Sac*I-*Bam*HI genomic fragment, including the putative promoter upstream of *epsA* and the putative terminator downstream of *epsM*, was cloned into lactococcal multicopy vector pJIM2279 as described in Materials and Methods. The resulting recombinant plasmid, pFS101, was examined by restriction enzyme analysis with *Hin*dIII and *Eco*RI single and *Bam*HI-*Sac*I double digestions and yielded the fragment sizes predicted from the sequence data. Furthermore, in 10 PCR experiments with primers in different regions over the whole *eps* gene cluster, fragments with the expected sizes were obtained. Therefore, no detectable deletions or rearrangements have occurred in the heterologous host during the cloning procedure.

To determine whether pFS101 could induce EPS production, *L. lactis* MG1363 was retransformed with pJIM2279 and pFS101 and directly plated on selective ruthenium red plates. Of 3,000 colonies with pJIM2279, all presented a red phenotype, whereas of 800 colonies with pFS101, all except 2 presented a white phenotype. The white phenotype of MG1363 (pFS101) indicated that EPS was synthesized and secreted (Fig. 7). To obtain quantitative data on this EPS production, MG1363 with pJIM2279 and MG1363 with pFS101 were fermented for 24 h at pH 5.5, and the EPS from 200 ml of spent culture were isolated. Fermentation of MG1363 with pJIM2279 yielded only traces (<1 mg/liter) of sugar-containing material, while fermentation of MG1363 with pFS101 yielded 10 mg of EPS fraction per ml. Even though the quantities produced were low, this EPS secretion proved that the *eps* gene cluster is responsible for EPS production.

DISCUSSION

The approach involving Tn*916* mutagenesis coupled with the ruthenium red plate assay led to the identification and characterization of the gene cluster required for EPS synthesis in *S. thermophilus* Sfi6. This cluster is the first *eps* gene cluster of a food microorganism to be described. The gene cluster consists of 13 ORFs named *epsA* through *epsM* which are encoded within a 14.5-kb region and are located on the 230-kb genomic *Sma*I fragment of *S. thermophilus* Sfi6. This *eps* gene cluster was able to direct the production of EPS in the non-EPS-producing heterologous host *L. lactis* MG1363.

Genetic organization of the *eps* gene cluster. Although we did not perform biochemical experiments to assess the function of these genes, homology searches for and sequence analyses of the deduced amino acid sequences revealed information about all of the genes except for *epsB*, *epsL*, and *epsM*. On the basis of the similarities, we were able to preliminarily assign functions to the genes and thereby identify gene regions involved in regulation, biosynthesis of the tetramer repeating unit, polymerization, and export (Fig. 8). Interestingly, this overall genetic organization is similar to that of the gene cluster involved in the synthesis of the capsular polysaccharide of *N. meningitidis* (18). The region involved in regulation is located at the beginning of the gene cluster, while the central region, which is involved in biosynthesis of the polysaccharide repeating units, is flanked by two regions involved in polymerization and export.

Regulation. A region potentially involved in regulation, located at the beginning of the gene cluster, and composed of *epsA* was identified. EpsA shows significant homology to CpsA of *S. pneumoniae* and LytR of *B. subtilis* (27, 37). LytR is a basic 35-kDa protein which is involved in the regulation of the autolysin (*lytABC* operon). It acts as an attenuator of the expression of both *lytABC* and *lytR* operons and might be membrane anchored at its N-terminal end (residues 11 to 35), which is homologous with the transmembrane domain of many eucaryotic glycoproteins. EpsA is also basic and contains three highly hydrophobic segments (residues 19 to 44, 50 to 68, and 74 to 96) which could function as the membrane anchor and be the equivalent of the N-terminal hydrophobic LytR segment. Therefore, we suggest that *epsA* is involved in the regulation of EPS expression.

Biosynthesis of the tetramer repeating unit. The central region (*epsE*, *F*, *G*, *H*, and *I*) of the gene cluster encodes genes of which the deduced amino acid sequences would predict an involvement in the biosynthesis of the EPS repeating unit. EpsE, F, G, and I show homology with glycosyltransferases. EpsE shows significant homology with ExoY of *R. meliloti* (23, 48, 59) and with the C-terminal domain of RfbP of *S. typhimurium* LT2 (34, 76), with CpsE of *S. pneumoniae* (27), and with CpsD of *S. agalactiae* (64). CpsD, ExoY, and RfbP have been shown to possess galactosyltransferase activity and to catalyze the first step in the synthesis of the respective repeating units, i.e., the transfer of galactose-1-phosphate from UDP-galactose to the undecaprenyl-phosphate precursor. The galactose residues in the respective polysaccharides are joined by β-linkages. The same is true for the *S. thermophilus* EPS. We therefore assume that *epsE* encodes the galactosyltransferase catalyzing the first step in the synthesis of the repeating unit.

FIG. 7. Ruthenium red plate assay showing red *L. lactis* MG1363 harboring pJIM2279 (left) and white *L. lactis* harboring pFS101 (right).

This hypothesis is strengthened by the fact that EpsE is the only one of the potential glycosyltransferases carrying an N-terminal hydrophobic domain which would be required to interact with the undecaprenyl-phosphate lipid carrier. Recently, RfbP has been shown to be a bifunctional protein with the mentioned galactosyl-1-phosphate transferase activity (C-terminal domain) and an activity which has been assigned to the N-terminal domain which seems to be responsible for the flipping of undecaprenyl-phosphate in order to carry one O-antigen repeating unit from the cytoplasmic face to the periplasmic face of the cytoplasmic membrane (76). The missing N-terminal domain of RfbP in EpsE and ExoY might be related to an internal polymerization of the EPS. In fact, the assembly of the repeating units of these EPS would not require a flipping protein if polymerization took place at the cytoplasmic face of the cell membrane, as is the case for capsule synthesis in *E. coli* and *N. meningitidis* (18, 33, 36).

EpsF in Sfi6 and CapH and CapM, which are involved in capsule synthesis in *S. aureus* (40), all share moderate homology. On the basis of homology analysis, *capM* could encode a galactosyltransferase (40). Hence, EpsF, CapH, and CapM could be members of a family of related glycosyltransferases. EpsG shares low homology with the *N*-acetylglucosaminetransferase required for core synthesis of the *Salmonella* LPS (42), but since this sugar is not present in the streptococcal EPS, it is not possible to assign the sugar specificity to EpsG. EpsI shares homology with RfbV, previously named Orf14.1, of the *S. typhimurium* LT2 *rfb* gene cluster (34, 41). Recent results demonstrated that *rfbV* encodes an abequosyltransferase (41). Even though abequose is not a component of the streptococcal EPS, *epsI* could nevertheless encode a glycosyltransferase. Assuming that *epsE*, *F*, *G*, and *I* encode glycosyltransferases, the set of glycosyltransferases required for the synthesis of the tetramer repeating unit is complete.

EpsH shows strong homology with acetyltransferases of the NodL-LacA-CysE family and carries the conserved sequence motif IGRNCWIGSQVTILKGVTIGDNSIIGAGVVV (residues 109 to 139 [conserved residues are underlined]) (14). Therefore, we suggest that it is involved in the synthesis of *N*-acetylgalactosamine. Acetyltransferases of this type have also been identified in other gene clusters involved in polysaccharide synthesis (e.g., the *cps* gene cluster of *S. aureus* [40] and the *rfb* gene cluster of *E. coli* [68, 78]).

Polymerization and export. The regions upstream (EpsC and EpsD) and downstream (EpsJ and EpsK) of this central region could be involved in the polymerization and export of the EPS. EpsC shows the common sequence motif identified for proteins involved in the chain length determination of polysaccharides. These include the N-terminal domain of ExoP of *R. meliloti* and Cld (chain length determination) or Rol (regulation of chain length) of *S. typhimurium*, *E. coli*, and *S. flexneri* (1, 2, 4, 47) (Fig. 9). Furthermore, the hydrophobicity plot of EpsC resembles those of these proteins, with two potential membrane-spanning domains, one at the N-terminal end and the other at the C-terminal end, and a hydrophilic domain in the central region. The positioning of the conserved sequence motif relative to the C-terminal transmembrane domain is conserved in these proteins as well as in EpsC (Fig. 9). Hence, EpsC could be involved in the chain length determination of the streptococcal EPS.

EpsD carries a nucleotide-binding motif very similar to the one previously identified in the C-terminal portion of ExoP of *R. meliloti* (3, 4). The A site (GEGKS [residues 46 to 50]) and the B site (VIID [residues 149 to 152]) of a nucleotide-binding motif (17) could be identified in EpsD. The nucleotide-binding motif identified in the C-terminal domain of ExoP is thought to fulfill a signaling rather than a transport function (i.e., as an ABC transporter), since ExoP mutants devoid of the nucleotide-binding motif still export exopolysaccharide I (EPSI). However, the quantities of EPSI produced are smaller and the EPSI has a lower molecular weight, indicating a less efficient export and an altered chain length regulation (4). EpsD could therefore also be involved in chain length determination and additionally in the transport of EPS. It is conceivable that *epsC* and *epsD*, which are located next to each other, were originally

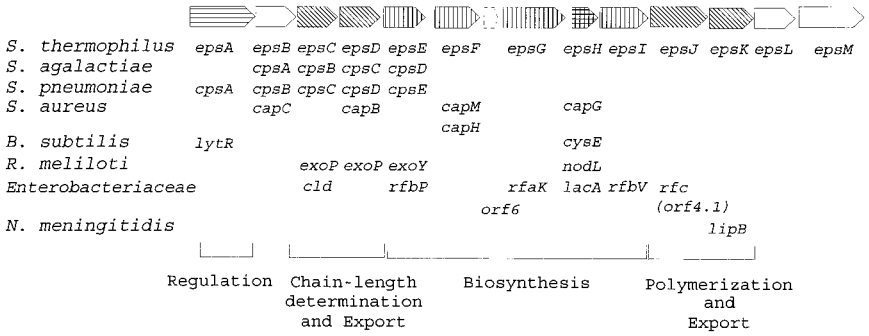


FIG. 8. Homologies of *eps* genes with genes of other microorganisms. The genes are grouped according to their function on the basis of these homologies. The ORF outlined with broken line is truncated. Horizontal hatching, gene involved in regulation; diagonal hatching, genes involved in polymerization and export; vertical hatching, glycosyltransferases; cross hatching, acetyltransferase.

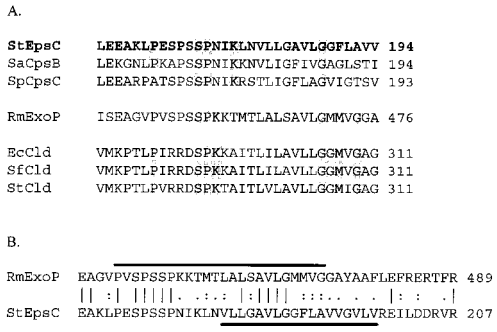


FIG. 9. (A) Conserved sequence motifs of proteins involved in chain length regulation of polysaccharides. Conserved residues are shaded. Abbreviations: StEpsC, EpsC of *S. thermophilus*; SaCpsB, CpsB of *S. agalactiae*; SpCpsC, CpsC of *S. pneumoniae*; RmExoP, ExoP of *R. meliloti*; EcCld, Cld of *E. coli*; SfCld, Cld of *S. flexneri*; StCld, Cld of *S. typhimurium*. (B) Position of the conserved sequence motif (bar above alignment) relative to that of the second potentially membrane-spanning domain (bar below alignment) in the alignment of EpsC of *S. thermophilus* and ExoP of *R. meliloti.*

one gene and were separated by a mutation or that *exoP* has evolved from a gene fusion, since EpsC and EpsD share homology with the N-terminal domain and C-terminal domain of ExoP, respectively. Moreover, the homology of EpsD with the C-terminal domain of ExoP starts at a methionine (ATG in *exoP* and the ATG start codon in *epsD*) preceded by an SD-like sequence (GGAG [6 bp before ATG]) in *exoP.*

EpsJ shows low homology with the putative O-antigen polymerase coded for by *orf4.1* of *Salmonella cholerasuis* (group C1) (38). Furthermore, it shows the characteristic hydrophobicity plot, with 11 hydrophobic peaks similar to those of the O-antigen polymerases identified so far (5, 9, 46). Nothing is known about the polymerization of EPS repeating units in LAB, but if polymerization of the tetrasaccharide repeating units can be compared to that of O-antigen repeating units, EpsJ would have the right predicted structural features to fulfill such an EPS-polymerase activity. In analogy with the model proposed by Bastin et al. (1) and Morona et al. (47), EpsJ, EpsC, and EpsD could form an enzyme complex responsible for polymerization: EpsJ would catalyze the polymerization reaction itself, and EpsC and EpsD would determine the chain length through an interaction with EpsJ and possibly initiate export. Hydrolysis of the nucleotide bound to EpsD could be signaled to EpsC in order to halt polymerization by EpsJ.

EpsK shows low homology with the *N. meningitidis* LipB which is responsible for phospholipid modification of the capsular polysaccharide before it is exported (18). EpsK could act similarly to LipB and promote the translocation by rendering the polymerized EPS more hydrophobic through a lipid modification. This would be one further argument supporting the hypothesis that as in capsule synthesis in *E. coli* and *N. meningitidis* (18, 33, 36), assembly of EPS repeating units in *S. thermophilus* takes place at the inner face of the cytoplasmic membrane.

No homologies were found for EpsL and EpsM, but indirect evidence indicated that gene disruption mutants in either of the corresponding genes are not viable. The fact that nonropy *S. thermophilus* strains do not possess *epsL* and *epsM* indicates that these genes are not required for cell growth. The lethal effect could be explained with the assumption that these genes are involved in the later stages (e.g., export and release) of EPS synthesis. A disruption would lead to an accumulation of lipid-linked EPS intermediates which would deplete the pool of lipid carriers required for cell growth and become toxic for *S. thermophilus*. In addition, the ORFs of *epsK* and *epsL* are overlapping, which might assure the coordinate expression of *epsK*, *epsL*, and *epsM.*

Interaction between EPS synthesis and primary metabolism. Interestingly, no ORF in the gene cluster, except the one of *epsH*, that showed homology with a gene involved in the synthesis of nucleotide sugar precursors was identified. Since all the components of the EPS (glucose, galactose, and *N*-acetylgalactosamine) are also sugars involved in housekeeping functions, the nucleotide sugars from primary metabolism may be utilized for the synthesis of the tetramer repeating unit. An involvement of UDPglucose 4-epimerase (*galE*) in EPS synthesis has been proposed by Poolman et al. (56). Even though enzyme activities for the Leloir pathway, UDPglucose 4-epimerase and UDPglucose-hexose-1-phosphate uridylyltransferase, have been detected in *S. thermophilus*, high levels of galactokinase cannot be expressed in this organism (31). This fact leads to an apparent Gal^- phenotype which can be converted to a Gal^+ phenotype under appropriate selective conditions (70). In consequence, only the glucose part of the lactose is metabolized, and galactose is used in a galactose-lactose antiport system (LacS) to import lactose without the requirement of additional energy (55). Therefore, it seems unlikely that UDPglucose 4-epimerase is involved in energy metabolism, but it could be involved in other pathways, such as that for the synthesis of EPS precursors. We envision the construction of *galE* mutants to verify this assumption and to test whether acceptable levels of labeled UDP-galactose or UDP-glucose can be incorporated in vivo into the EPS in order to initiate biochemical studies of EPS synthesis.

Biodiversity and evolutionary origin of *eps* gene clusters. It is worth noting that the *epsA*, *B*, *C*, *D*, and *E* genes of *S. thermophilus*, *cpsA*, *B*, *C*, and *D* genes of *S. agalactiae*, and *cpsA*, *B*, *C*, *D*, and *E* genes of *S. pneumoniae* are organized in the same manner and share very high homology (Fig. 8). The gene corresponding to *epsA* in *S. thermophilus* and *cpsA* in *S. pneumoniae* might still be identified upstream of the current *cps* genes of *S. agalactiae*. The high conservation of these genes indicates that their gene products are involved in enzymatic reactions that are independent of the specific sugar structures, since the EPS of *S. thermophilus* and capsular polysaccharides of *S. agalactiae* and *S. pneumoniae* share few common features. These three *Streptococcus* species are closely related (49), suggesting a common origin for the *eps* and *cps* genes. Interestingly, a similar conservation could be identified within nonropy *S. thermophilus* strains. The genes *epsA*, *epsB*, and *epsC* (and *epsD* for ST11) are present in nonropy strains, suggesting that these strains might have undergone spontaneous genomic deletions of a once-acquired ancestral *eps* gene cluster. Genomic instability, including deletions, has previously been observed for *S. thermophilus* (63).

The predicted amino acid sequences of genes in the central part of the *eps* gene cluster (*epsF* to *epsJ*) show mainly homology with those of proteins involved in the O-antigen and core synthesis of LPS of members of the *Enterobacteriaceae*. Interestingly, between *epsF* and *epsG*, a stretch of 240 bp of which the deduced amino acid sequence showed 60.0% identity with that of the C terminus of the protein coded for by *orf6* in the *rfb* cluster of *E. coli* K-12 was identified (68, 78) (Fig. 8). The N terminus of Orf6 shows similarity to flavin adenine dinucleotide- and/or NAD-containing proteins and is thought to be an oxidoreductase converting UDP-Gal*p* into UDP-Gal*f*. Thirty base pairs upstream of this 240-bp stretch, a start codon (AUG) and a well-positioned SD sequence (6 bp before the start codon) were identified, but they were out of frame with

the section of which the predicted amino acid sequence shows homology with that of Orf6. Sequencing errors or a possible mutation during the cloning process were excluded by sequence analysis of independent clones (i.e., pFS30 and pFS80). It seems as if these are the remnants of a gene that was transferred with other *rfb* genes and that might have once encoded an oxidoreductase in *S. thermophilus*. This observation and the homologies mentioned above support the hypothesis that the *eps* genes located in the central region of the gene cluster have the same evolutionary origin as the *rfb* genes of members of the *Enterobacteriaceae.*

The same conclusion could be drawn from the GC content of the *eps* genes. Genes *epsA* to *epsE* have an average GC content of 37%, which is close to the average streptococcal GC content of 40% (49), whereas genes *epsG* to *epsL* have an average GC content of 30%, which is close to the average GC content in the central region of the *rfb* gene cluster of *S. typhimurium* LT2 (34). In an analogy with *Salmonella* O-antigen variation (57), the highly conserved part of the *eps* gene cluster (*epsA* to *epsE*) might have provided the basis for intergenic recombinations which led to the immense variation of EPS in LAB. The central region of the *eps* gene cluster would be the target for exchange in case of a recombination event. This is the region where the putative glycosyltransferases are located and where genetic diversity among *eps* gene clusters encoding different EPS structures would be expected. Recombinations among *eps* gene clusters in ropy LAB might occur during coculture of strains in dairy fermentations.

There has been much speculation about the biological need of EPS in LAB. These organisms live in a rich, nutritive habitat in which the selective advantage to produce EPS is not evident. Moreover, EPS production is very unstable and frequently lost upon cultivation. EPS synthesis might therefore be a trait which was carried over in evolution from organisms for which the polysaccharides provided a selective advantage (e.g., capsules are virulence factors responsible for colonization of the host [18, 27, 40, 64] or EPSI of *R. meliloti* promotes infection of alfalfa root nodules [3, 4, 23, 48, 59]) and was merely kept in the gene pool of LAB by selection for ropy strains in the dairy industry. Since LAB are a heterogeneous group of microorganisms that found a common habitat after a secondary transfer into milk, the genetics of EPS gene clusters might help to shed more light onto the different evolutionary origins of LAB.

ACKNOWLEDGMENTS

We express our thanks to P. Renault for sending us pJIM2279 prior to publication, J. Lemoine for the communication of the Sfi6 EPS structure, N. Favre for the EPS analysis of Sfi6 and the nonropy Sfi6 Tn*916* integrant, and O. Mignot and D. De Maleprade for rheological measurements of ropy *S. thermophilus* strains. We further thank D. R. Pridmore, D. De Pasquale, and A. Constable for critical reading of the manuscript.

This work was supported by the Commission of the European Communities (DG XII) within the framework of the Agro-Industrial Research Program.

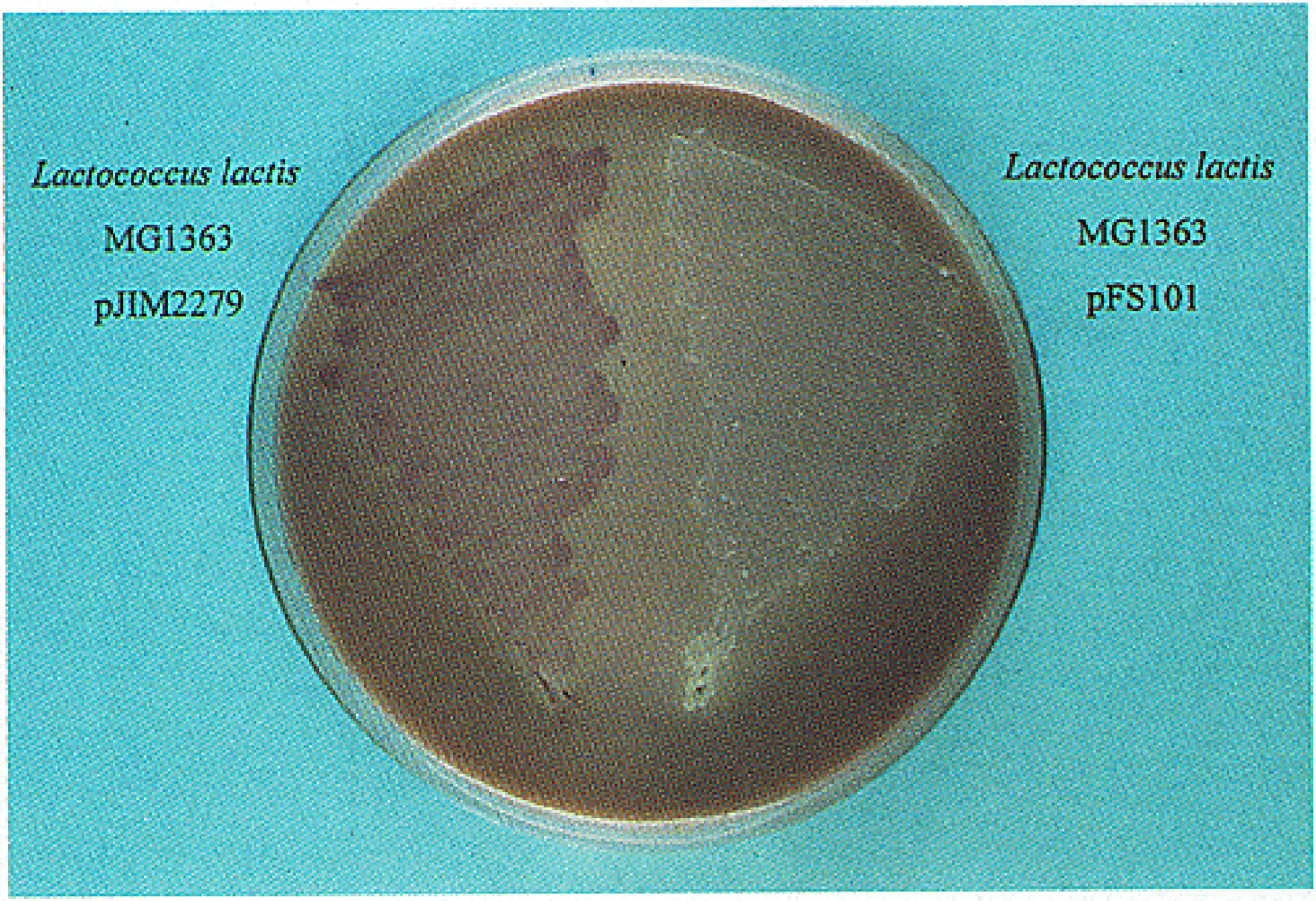
REFERENCES

1. **Bastin, D. A., G. Stevenson, P. K. Brown, A. Haase, and P. R. Reeves.** 1993. Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid synthase, with a novel mechanism for determining chain length. Mol. Microbiol. **7:**725–734.
2. **Batchelor, R. A., P. Alifano, E. Biffali, S. I. Hull, and R. A. Hull.** 1992. Nucleotide sequence of the genes regulating O-polysaccharide antigen chain length (*rol*) from *Escherichia coli* and *Salmonella typhimurium*: protein homology and functional complementation. J. Bacteriol. **174:**5228–5236.
3. **Becker, A., A. Kleickmann, M. Keller, W. Arnold, and A. Pühler.** 1993. Identification and analysis of the *Rhizobium meliloti exoAMONP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. Mol. Gen. Genet. **241:**367–379.
4. **Becker, A., K. Niehaus, and A. Pühler.** 1995. Low-molecular-weight succinoglycan is predominantly produced by *Rhizobium meliloti* strains carrying a mutated ExoP protein characterized by a periplasmic N-terminal domain and a missing C-terminal domain. Mol. Microbiol. **16:**191–203.
5. **Brown, P. K., L. K. Romana, and P. R. Reeves.** 1992. Molecular analysis of the *rfb* gene cluster of *Salmonella* serovar *muenchen* (strain M67): the genetic basis of the polymorphism between groups C2 and B. Mol. Microbiol. **6:**1385–1394.
6. **Cerning, J.** 1990. Exocellular polysaccharides produced by lactic acid bacteria. FEMS Microbiol. Rev. **87:**113–130.
7. **Cerning, J.** 1994. Polysaccharides exocellulaires produits par les bactéries lactiques, p. 309–329. *In* H. de Rossartand and F. M. Luquet (ed.), Bactéries lactiques, vol. I. Lorica, Uriage, France.
8. **Chaparon, M. G., and J. R. Scott.** 1989. Excision and insertion of the conjugative transposon Tn*916* involves a novel recombination mechanism. Cell **59:**1027–1034.
9. **Collins, L. V., and J. Hackett.** 1991. Molecular cloning, characterization, and nucleotide sequence of the *rfc* gene, which encodes an O-antigen polymerase of *Salmonella typhimurium*. J. Bacteriol. **173:**2521–2529.
10. **Dao, M. L., and J. J. Ferretti.** 1985. *Streptococcus-Escherichia coli* shuttle vector pSA3 and its use in the cloning of streptococcal genes. Appl. Environ. Microbiol. **49:**115–119.
11. **de Vos, W. M., H. de Haard, and I. Boerrigter.** 1989. Cloning and expression of the *Lactococcus lactis* subsp. *cremoris* SK11 gene encoding an extracellular serine proteinase. Gene **85:**169–176.
12. **Doco, T., J.-M. Wieruszeski, and B. Fournet.** 1990. Structure of an exocellular polysaccharide produced by *Streptococcus thermophilus*. Carbohydr. Res. **198:**313–321.
13. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *Escherichia coli* by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
14. **Downie, J. A.** 1989. The *nodL* gene from *Rhizobium leguminosarum* is homologous to acetyltransferases encoded by *lacA* and *cysE*. Mol. Microbiol. **3:** 1649–1651.
15. **Dubois, M. A., K. A. Gilles, J. K. Hamilton, P. A. Rebers, and F. Smith.** 1956. Colorimetric method for determination of sugars and related substances. Anal. Chem. **28:**350–356.
16. **Elliker, P. R., A. W. Anderson, and G. Hannesson.** 1956. An agar culture medium for lactic acid streptococci and lactobacilli. J. Dairy Sci. **89:**1611–1612.
17. **Fath, M. J., and R. Kolter.** 1993. ABC transporters: bacterial exporters. Microbiol. Rev. **57:**995–1017.
18. **Frosch, M., and A. Müller.** 1993. Phospholipid substitution of capsular polysaccharides and mechanism of capsule formation in *Neisseria meningitidis*. Mol. Microbiol. **8:**483–493.
19. **Gagnon, Y., R. Breton, H. Putzer, M. Pelchat, M. Grunbergmanago, and J. Lapointe.** 1994. Clustering and co-transcription of the *Bacillus subtilis* genes encoding the aminoacyl-tRNA synthetase specific for glutamate and for cysteine and the first enzyme for cysteine biosynthesis. J. Biol. Chem. **269:** 7473–7482.
20. **Gasson, M. J.** 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. **154:**1–9.
21. **Gawron-Burke, C., and D. B. Clewell.** 1982. A transposon in *Streptococcus faecalis* with fertility properties. Nature (London) **300:**281–283.
22. **Germond, J.-E., L. Lapierre, M. Delley, and B. Mollet.** 1995. A new mobile genetic element in *Lactobacillus delbrueckii* subsp. *bulgaricus*. Mol. Gen. Genet. **248:**407–416.
23. **Glucksmann, M. A., T. L. Reuber, and G. C. Walker.** 1993. Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: a model for succinoglycan biosynthesis. J. Bacteriol. **175:** 7045–7055.
24. **Gruter, M., B. R. Leeflang, J. Kuiper, J. P. Kamerling, and J. F. G. Vliegenthart.** 1992. Structure of the exopolysaccharide produced by *Lactococcus lactis* subspecies *cremoris* H414 grown in a defined medium or skimmed milk. Carbohydr. Res. **231:**273–291.
25. **Gruter, M., B. R. Leeflang, J. Kuiper, J. P. Kamerling, and J. F. G. Vliegenthart.** 1993. Structural characterization of the exopolysaccharide produced by *Lactobacillus delbrückii* subspecies *bulgaricus* rr grown in skimmed milk. Carbohydr. Res. **239:**209–226.
26. **Guédon, G., F. Bourgoin, M. Pebay, Y. Roussel, C. Colmin, J. M. Simonet, and B. Decaris.** 1995. Characterization and distribution of two insertion sequences, IS*1191* and iso-IS*981*, in *Streptococcus thermophilus*: does intergenic transfer of insertion sequences occur in lactic acid bacteria co-cultures? Mol. Microbiol. **16:**69–78.
27. **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.
28. **Hediger, M., D. F. Johnson, D. P. Nierlich, and I. Zabin.** 1985. DNA sequence of the lactose operon: the *lacA* gene and the transcriptional termination region. Proc. Natl. Acad. Sci. USA **82:**6414–6418.

29. **Hershfield, M. S., S. Chaffee, L. Koro-Johnson, A. Mary, A. A. Smith, and S. A. Short.** 1991. Use of site-directed mutagenesis to enhance epitope shielding effect of covalent modification of proteins with polyethylene glycol. Proc. Natl. Acad. Sci. USA **88:**7185–7189.
30. **Hill, C., G. Venema, C. Daly, and G. F. Fitzgerald.** 1988. Cloning and characterization of the tetracycline resistance determinant of and several promoters from within the conjugative transposon Tn*919*. Appl. Environ. Microbiol. **54:**1230–1236.
31. **Hutkins, R., H. A. Morris, and L. L. McKay.** 1985. Galactokinase activity in *Streptococcus thermophilus*. Appl. Environ. Microbiol. **50:**777–780.
32. **Israelsen, H., and E. B. Hansen.** 1993. Insertion of transposon Tn*917* derivatives into the *Lactococcus lactis* subsp. *lactis* chromosome. Appl. Environ. Microbiol. **59:**21–26.
33. **Jann, B., and K. Jann.** 1991. Structure and biosynthesis of the capsular antigens of *Escherichia coli*. Curr. Top. Microbiol. Immunol. **150:**695–713.
34. **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.
35. **Kite, J., and R. F. Doolittle.** 1982. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **157:**105–132.
36. **Kröncke, K.-D., G. Boulnois, I. Roberts, D. Bitter-Suermann, J. R. Golecki, B. Jann, and K. Jann.** 1990. Expression of the *Escherichia coli* K5 capsular antigen: immunoelectron microscopic and biochemical studies with recombinant *E. coli*. J. Bacteriol. **172:**1085–1091.
37. **Lazaveric, V., P. Margot, B. Soldo, and D. Karamata.** 1992. Sequencing and analysis of the *Bacillus subtilis lytRABC* divergon: a regulatory unit encompassing the structural genes of the *N*-acetylmuramoyl-L-alanine amidase and its modifier. J. Gen. Microbiol. **138:**1949–1961.
38. **Lee, S. J., L. K. Romana, and P. R. Reeves.** 1992. Sequence and structural analysis or the *rfb* (O antigen) gene cluster from a group C1 *Salmonella enterica* strain. J. Gen. Microbiol. **138:**1843–1855.
39. **Lemoine, J.** Personal communication.
40. **Lin, W. S., T. Cunnen, and C. Y. Lee.** 1994. Sequence analysis and molecular characterization of genes required for the biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*. J. Bacteriol. **176:**7005–7016.
41. **Liu, D., L. Lindqvist, and P. R. Reeves.** 1995. Transferases of O-antigen biosynthesis in *Salmonella enterica*: dideoxyhexosyltransferases of group B and C2 and acetyltransferase of group C2. J. Bacteriol. **177:**4084–4088.
42. **MacLachlan, P. R., S. K. Kadam, and K. E. Sanderson.** 1991. Cloning, characterization, and DNA sequence of the *rfaLK* region for lipopolysaccharide synthesis in *Salmonella typhimurium* LT2. J. Bacteriol. **173:**7151–7163.
43. **Macura, D., and P. M. Townsley.** 1984. Scandinavian ropy milk—identification and characterization of endogenous ropy lactic streptococci and their extracellular excretion. J. Dairy Sci. **67:**735–744.

43a. **Mignot, O., and D. De Maleprade.** Unpublished data.

44. **Mollet, B., and M. Delley.** 1990. A β-galactosidase deletion mutant of *Lactobacillus bulgaricus* reverts to generate an active enzyme by internal DNA sequence duplication. Mol. Gen. Genet. **227:**17–21.
45. **Mollet, B., J. Knol, B. Poolman, O. Marciset, and M. Delley.** 1993. Direct genomic integration, gene replacement, and integrative gene expression in *Streptococcus thermophilus*. J. Bacteriol. **175:**4315–4324.
46. **Morona, R., M. Mavris, A. Fallarino, and P. Manning.** 1994. Characterization of the *rfc* region of *Shigella flexneri*. J. Bacteriol. **176:**733–747.
47. **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.
48. **Müller, P. M., M. Keller, W. M. Weng, J. Quandt, W. Arnold, and A. Pühler.** 1993. Genetic analysis of the *Rhizobium meliloti exoYFQ* operon: ExoY is homologous to sugar transferases and ExoQ represents a transmembrane protein. Mol. Plant-Microbe Interact. **6:**55–65.
49. **Mundt, J. O.** 1986. Lactic acid streptococci, p. 1065–1066. *In* P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 2. The Williams and Wilkins Co., Baltimore.
50. **Nakajiama, H., T. Hirota, T. Toba, and S. Adachi.** 1992. Structure of the extracellular polysaccharide from slime-forming *Lactococcus lactis* subsp. *cremoris* SBT 0495. Carbohydr. Res. **224:**245–253.
51. **Neeser, J.-R., and T. Schweizer.** 1984. A quantitative determination by capillary gas-liquid chromatography of neutral and amino sugars (as *O*-methyloxime acetates), and a study on hydrolytic conditions for glycoproteins and polysaccharides in order to increase sugar recoveries. Anal. Biochem. **142:**58–67.
52. **Neve, H., A. Geis, and M. Teuber.** 1988. Plasmid-encoded functions of ropy lactic acid streptococcal strains from Scandinavian fermented milk. Biochimie (Paris) **70:**437–442.
53. **Oda, M., H. Hasegawa, S. Komatsu, M. Kambe, and F. Tsuchiya.** 1983. Anti-tumor polysaccharide from *Lactobacillus* sp. Agric. Biol. Chem. **47:** 1623–1625.
54. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
55. **Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt.** 1989. Lactose transport system of *Streptococcus thermophilus*: a hybrid protein with homology to the melibiose carrier and enzyme III of phosphoenolpyruvate-dependent phosphotransferase systems. J. Bacteriol. **171:**244–253.
56. **Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt.** 1990. Carbohydrate utilization in *Streptococcus thermophilus*: characterization of the genes for aldose 1-epimerase (mutarotase) and UDPglucose 4-epimerase. J. Bacteriol. **172:**4037–4047.
57. **Reeves, P.** 1993. Evolution of *Salmonella* O-antigen variation by interspecific gene transfer on a large scale. Trends Genet. **9:**17–22.
58. **Renault, P., G. Corthier, C. Delorme, N. Goupil, and S. D. Ehrlich.** A set of vectors for gram-positive bacteria which can be switched from high to low copy number. Unpublished data.
59. **Reuber, T. L., and G. C. Walker.** 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. Cell **74:** 269–280.
60. **Robijn, G. W., J. R. Thomas, H. Haas, D. J. C. van den Berg, J. P. Kamerling, and J. F. G. Vliegenthart.** 1995. The structure of the exopolysaccharide produced by *Lactobacillus helveticus* 766. Carbohydr. Res. **276:**137–154.
61. **Robijn, G. W., D. J. C. van den Berg, H. Haas, J. P. Kamerling, and J. F. G. Vliegenthart.** 1995. Determination of the structure of the exopolysaccharide produced by *Lactobacillus sake* 0-1. Carbohydr. Res. **276:**117–136.
62. **Romero, D., and T. R. Klaenhammer.** 1992. IS*946*-mediated integration of heterologous DNA into the genome of *Lactococcus lactis* subsp. *lactis*. Appl. Environ. Microbiol. **58:**699–702.
63. **Roussel, Y., M. Pebay, G. Guedon, J. M. Simonet, and B. Decaris.** 1994. Physical and genetic map of *Streptococcus thermophilus* A054. J. Bacteriol. **176:**7413–7422.
64. **Rubens, C. E., L. M. Heggen, R. F. Haft, and M. R. Wessels.** 1993. Identification of *cpsD*, a gene essential for type III capsule expression in group B streptococci. Mol. Microbiol. **8:**843–885.
65. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:**487–491.
66. **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.
67. **Slos, P., J.-C. Bourquin, Y. Lemoine, and A. Mercenier.** 1991. Isolation and characterization of chromosomal promoters of *Streptococcus salivarius* subsp. *thermophilus*. Appl. Environ. Microbiol. **57:**1333–1339.
68. **Stevenson, G., B. Niel, D. Liu, M. Hobbs, N. H. Packer, M. Batley, J. W. Redmond, L. Lindqvist, 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.
69. **Stingele, F., and B. Mollet.** 1995. Homologous integration and transposition to identify genes involved in exopolysaccharide production in *Streptococcus thermophilus*. Dev. Biol. Stand. **85:**487–493.
70. **Thomas, T. D., and V. L. Crow.** 1984. Selection of galactose-fermenting *Streptococcus thermophilus* in lactose-limited chemostat cultures. Appl. Environ. Microbiol. **48:**186–191.
71. **Triflia, T., M. G. Peterson, and D. J. Kemp.** 1988. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res. **16:**8186.
72. **van Kranenbug, R., J. D. Marugg, and W. de Vos.** 1995. Genetics of exopolysaccharide production in *Lactococcus lactis*, poster abstr. P-77. *In* Abstracts of the Carbohydrate Bioengineering Meeting, April 23–26, 1995, Elsinore, Denmark. Sintef Unimed, Trondheim, Norway.
73. **Vedamuthu, E. R., and J. M. Neville.** 1986. Involvement of a plasmid in production of ropiness (mucoidness) in milk cultures by *Streptococcus cremoris* MS. Appl. Environ. Microbiol. **51:**677–682.
74. **Vescovo, M., G. L. Scolari, and V. Bottazzi.** 1989. Plasmid-encoded ropiness production in *Lactobacillus casei* ssp. *casei*. Biotechnol. Lett. **2:**709–712.
75. **von Wright, A., and S. Tynkkynen.** 1987. Construction of *Streptococcus lactis* subsp. *lactis* strains with a single plasmid associated with mucoid phenotype. Appl. Environ. Microbiol. **53:**1385–1386.
76. **Wang, L., and R. P. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O-antigen subunit processing. J. Bacteriol. **176:**4348–4356.
77. **Yamamoto, Y., S. Murosaki, R. Yamauchi, K. Kato, and Y. Sone.** 1994. Structural study on an exocellular polysaccharide produced by *Lactobacillus helveticus* TY1-2. Carbohydr. Res. **261:**67–78.
78. **Yao, Z., and M. A. Valvano.** 1994. Genetic analysis of the O-specific 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.


Lactococcus lactis
MG1363
pJIM2279
Lactococcus lactis
MG1363
pFS101