Exopolysaccharide biosynthesis in *Lactococcus lactis*; A molecular characterisation

Richard van Kranenburg

Promotor: dr. W. M. de Vos
Hoogleraar in de microbiologie

Exopolysaccharide biosynthesis in *Lactococcus lactis*; A molecular characterisation

Richard van Kranenburg

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magificus
van Wageningen Universiteit,
dr. C. M. Karssen,
in het openbaar te verdedigen
op maandag 22 november 1999
des namiddags te 13.30 uur in de Aula

Chapter 2 was printed with the kind permission of Blackwell Science Ltd

Chapters 3, 4, and 5 were printed with the kind permission of the American Society for Microbiology

ISBN 90-5808-135-4
Druk: Ponsen & Looijen BV, Wageningen

Stellingen

1. De indeling van bacteriële polysachariden in homo- en heteropolysachariden die door sommige auteurs gebruikt wordt is gekunsteld en zou vervangen moeten worden door een indeling gebaseerd op hun biosynthese route.

2. Rubens en medewerkers kunnen uit hun experimenten, waarbij ze in een assay van celextracten en radiogelabelde UDP-galactose alleen kijken naar de inbouw van radioactiviteit in de lipidefractie, niet concluderen dat CpsD galactosyltransferase activiteit heeft. De mogelijke aanwezigheid van epimeraseactiviteit maakt identificatie van de ingebouwde suikers noodzakelijk.
   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**:843-855.

3. De geconserveerde aminozuren die Wang en medewerkers aanduiden als bepalend voor de glucosyltransferase of galactosyltransferase specificiteit zijn niet terug te vinden in de overeenkomstige glycosyltransferases van gram-positieve bacteriën en zeggen waarschijnlijk meer over de onderlinge verwantschap van de desbetreffende enzymen dan over de substraatspecificiteit.
   Wang, L., D. Liu, and P. R. Reeves. 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J. Bacteriol. **178**:2598-2604.

4. Het toekennen van functies aan genen op basis van homologieën zonder daaropvolgende ondersteuning met experimentele data heeft als gevaar dat deze functies in de loop der tijd toch als feit aangenomen worden.

5. De toenemende sequentie data van polysacharide genclusters laten zien dat horizontale genoverdracht in het verleden al heeft geleid tot polysacharide engineering avant la lettre.

6. Bij de consumentenacceptatie van toepassing van genetisch gemodificeerde organismen in voedingsmiddelen in Europa spelen emoties een belangrijkere rol dan argumenten.

7. Quality assurance (QA) in een research-instelling impliceert kwaliteitsgarantie van het onderzoek, maar garandeert slechts het handelen volgens standaard werkwijzen.

8. De gewenning aan automatisering maakt dat veel werknemers niet meer kunnen functioneren als het netwerk uitgevallen is.

9. De prijzen van voetballers stijgen sneller dan die van huizen.
Vissers, W. De Volkskrant 15 juli 1999.

10. Rekeningrijden is een vorm van betaald fileparkeren.

11. Dit is een millenium-*proof*schrift.

Stellingen behorend bij het proefschrift
"Exopolysaccharide biosynthesis in *Lactococcus lactis*;
A molecular characterisation".
Richard van Kranenburg, Wageningen, 22 november 1999.

Contents

General Introduction

Chapter 1

General introduction

Many bacteria are known to produce cell-surface polysaccharides, that are involved in a wide variety of biological functions including prevention of desiccation or other environmental stresses, adherence to surfaces, and pathogenesis or symbiosis (Roberts, 1996, Whitfield and Valvano, 1993). The cell-surface polysaccharides comprise O antigens of lipopolysaccharides (LPSs), capsular polysaccharides (CPSs) or exopolysaccharides (EPSs). Both LPSs and CPSs are linked to the cell surface, while EPSs are only loosely attached or completely excreted into the environment. The O antigens of LPSs are linked to the outer cell membrane of Gram-negative bacteria via an oligosaccharide core and lipid-A and the CPSs are covalently attached to either phospholipid or lipid-A molecules in the cell membrane (Whitfield and Valvano, 1993).

Biosynthesis of bacterial polysaccharides

Biosynthesis of bacterial cell-surface polysaccharides can occur in different ways and three pathways have been described for O-antigen production by Gram-negative bacteria. The first one involves growth of the polymer at the reducing end and can be illustrated by the biosynthesis of the O-antigen polysaccharide from *Salmonella enterica* (Fig. 1). Biosynthesis is initiated by the linking of galactose-1-phosphate from UDP-galactose to the undecaprenyl phosphate lipid carrier by the priming glycosyltransferase WbaP (RfbP) (Wang and Reeves, 1994). Subsequently, specific glycosyltransferases transfer the sugar moiety from a nucleotide sugar to the lipid-linked acceptor molecule to form a complete repeating unit (McGrath and Osborn, 1991, Wang *et al.*, 1996). A multiple membrane-spanning protein Wzx (RfbX) is thought to serve as a flippase to translocate the lipid-linked repeating units to the periplasmic side of the cytoplasmic membrane where it is polymerised at the reducing end by Wzy (Rfc), with Wzz (Rol/Cld) controlling the O antigen chain length (Liu *et al.*, 1996, McGrath and Osborn, 1991, Reeves *et al.*, 1996). Similar pathways have been proposed for assembly of the repeating unit of various other polysaccharides including xanthan gum of *Xanthomonas campestris* (Ielpi *et al.*, 1993), succinoglycan of *Rhizobium meliloti* (Reuber and Walker, 1993), and the CPS of *Streptococcus pneumoniae* serotype 14 (Kolkman *et al.*, 1997).

A second pathway seems confined to simple homopolymer chains (mannan or galactan), like those of *Escherichia coli* 09 or *Klebsiella pneumoniae* 01 O antigens that are synthesised entirely by glycosyltransferases on the cytoplasmic side of the cytoplasmic membrane without the involvement of Wzy (Clarke, 1992, Kido *et al.*, 1995). Synthesis is initiated by the WecA (Rfe) enzyme that links *N*-acetylglucosamine (GlcNAc) from the nucleotide sugar UDP-GlcNAc to the undecaprenyl phosphate lipid carrier (Whitfield, 1995). This lipid-linked GlcNAc is the primer for the assembly of the polysaccharides. Polymerisation occurs by sequential transfer of glycosyl residues to the non-reducing end, and an ATP-binding cassette

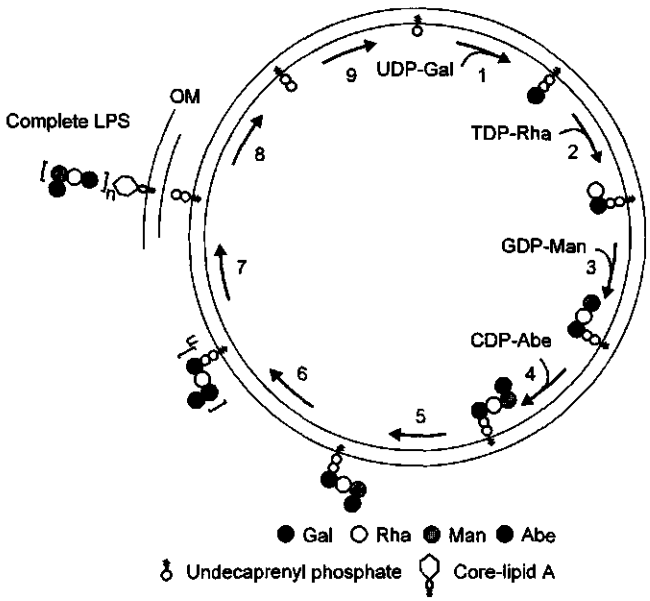


Fig. 1. Biosynthetic pathway of O-antigen of *S. enterica* group B (Liu *et al.*, 1996). Steps 1 to 4 represent the assembly of O units by transfer of galactose-phosphate, rhamnose, mannose, and abequose, respectively, onto the lipid carrier undecaprenylphosphate. These reactions are catalysed by glycosyltransferases and occur on the cytoplasmic side of the cytoplasmic membrane. Lipid-linked repeating units are translocated across the membrane by the activity of Wzx (5) and polymerisation on the periplasmic side of the cytoplasmic membrane is catalysed by Wzy (6). Finally, WaaL transfers the O-antigen chain to core-lipid A (7). The undecaprenylpyrophosphate is retranslocated over the cytoplasmic membrane (8) and undecaprenylphosphate is regenerated by phosphatase activity (9). The outer membrane (OM) is only partially depicted.

(ABC) transporter is necessary for transport of the complete O antigen across the cytoplasmic membrane (Kido *et al.*, 1995).

A third pathway was described for the assembly of the poly-*N*-acetylmannosamine (ManNAc) O antigen (factor 54) of *Salmonella enterica* serovar Borreze, consisting of disaccharide repeating units (Keensleyside, 1996). This is an Wzy-independent pathway that requires a lipid-linked GlcNAc primer, which is provided by WecA activity. First, WbbE (RfbA), a non-processive glycosyltransferase, adds a single ManNAc to this primer and subsequently, WbbF, a processive glycosyltransferase (synthase) resembling hyaluronic acid synthase (HasA) from *Streptococcus pyogenes*, is involved in polymerisation. By analogy with the HasA family of proteins, WbbF is believed to have two catalytic domains allowing the catalysis of two β-glycosidic bonds, either simultaneously or sequentially (Saxena *et al.*, 1995). There is no Wzy homologue or ABC transporter in the system, but the C-terminal part of WbbF is predicted to form a pore or channel in the membrane through which growing chain is extruded, thus combining glycosyltransferase activity and transport (Keenleyside and Whitfield, 1996). Other members of the HasA family, involved in biosynthesis of hyaluronic

acid of *S. pyogenes*, the type 3 capsule of *S. pneumoniae*, or glucosaminoglycan of *Staphylococcus epidermidis*, may also possess this combined transferase/transport function (Keenleyside and Whitfield, 1996).

Glycosyltransferases involved in polysaccharide biosynthesis

Glycosyltransferases can be divided in non-processive enzymes that catalyse the conversion of a single residue to the acceptor, and processive enzymes, such as HasA, that transfer multiple sugar residues to the acceptor. There are two major catalytic mechanisms for glycosyl transfer proceeding either by retention or by overall inversion of the anomeric configuration at the reaction centre (Fig. 2). The bacterial β-glycosyltransferase activity involved in polysaccharide biosynthesis may be viewed as the reverse reaction of the glycosyl transfer reaction performed by glycosidases (Saxena *et al.*, 1995). By analogy with the polysaccharide hydrolase systems, this hypothesis predicts that the formation of a β-glycosyl linkage from an α-linked sugar nucleotide donor would involve the same type of catalytic event as that of the inverting glycoside hydrolases (Fig. 2). Hydrolysis of glycosidic

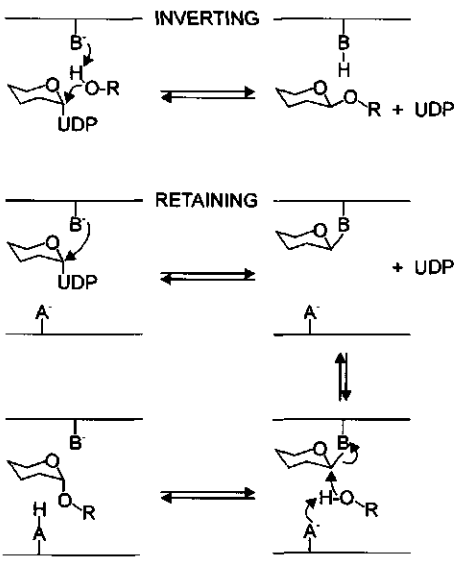


Fig. 2. The two mechanisms proposed for glycosyl transfer from nucleotide diphospho sugars (adopted from Saxena *et al.*, 1995). In the inverting mechanism, a single nucleophilic substitution at the sugar anomeric carbon leads to the formation of a β-linkage from an α-linked donor; ROH represents the acceptor, and B represents the catalytic base. The retaining mechanism involves the transient formation of a glycosyl enzyme and its subsequent addition to the acceptor. The two nucleophilic substitutions at the sugar anomeric carbon result in the formation of an α-linkage from an α-linked donor; ROH represents the acceptor, A represents the catalytic base, and B represents the nucleophile.

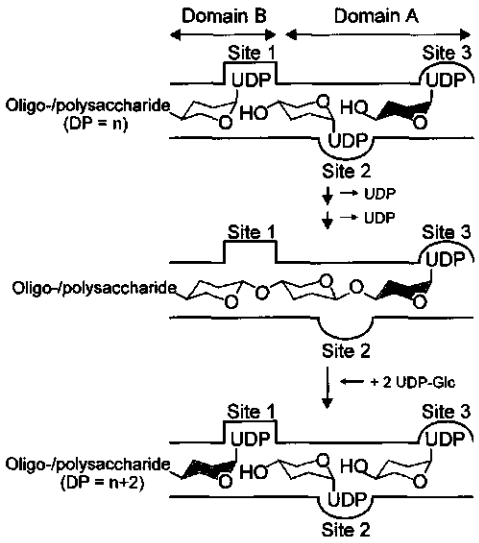


Fig. 3. Model of polymerisation by double addition showing growth of a polysaccharide chain at its reducing end using a UDP-monosaccharide as the substrate (adopted from Saxena *et al.*, 1995). **(Top)** The processive enzymes feature extended active sites able to bind three nucleotide diphospho sugars. The initiation of polymerisation does not require a primer, since it is conceivable that a UDP-monosaccharide could fill site 1. **(Middle)** Two glycosidic bonds are formed, either simultaneously or sequentially, by a mechanism resulting in the inversion of the anomeric configuration and in the release of two UDP molecules. **(Bottom)** The chain, which is elongated by two units and bears a UDP group on its reducing end (residue shaded), can move or slide by two units so as to fill site 1 with the UDP group. Sites 2 and 3 can now bind new UDP-sugars, and the double addition can proceed. DP, degree of polymerisation.

bonds by these enzymes results in a net inversion of configuration. The catalytic mechanism involves two acidic active site amino acids that act as acid-base catalysts. The two catalytic residues are located in flexible loop regions in the active site cleft, between substrate binding subsites. A model for glycosyltransferase activity is shown in Fig. 3. In this model, the function of domain A is the transfer of a glycosyl residue from a nucleotide sugar to an acceptor molecule. For non-processive enzymes, which only carry domain A, the acceptor is an intermediate in the subunit assembly and the subunit chain grows from the nonreducing end. In the case of processive enzymes that have domain B functioning along with domain A, site 3 is occupied by a nucleotide sugar as well, leading to the formation of two glycosidic linkages. The oligo- or polysaccharide chain now grows from its reducing end. The simultaneous formation of two glycosidic linkages provides a simple mechanism for the generation of the 2-fold screw axis that arises from a disaccharide repeat with two β-glycosidic bonds, without invoking a concomitant rotation of either the enzyme or the substrate (Saxena *et al.*, 1995).

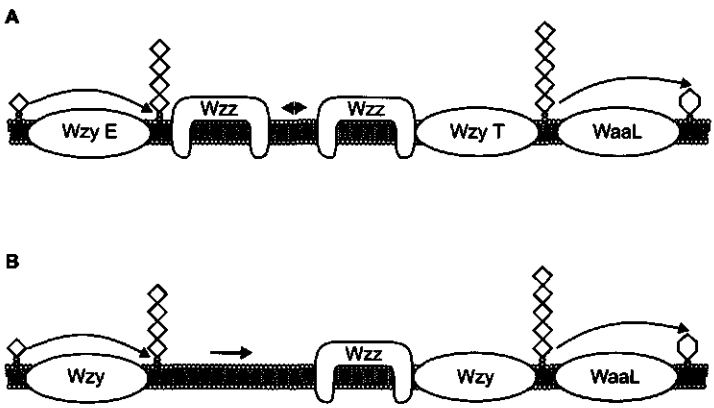


Fig. 4. Two models for the role of Wzz in O-antigen chain-length determination. (A) Wzy has two states, one of which favours polymerisation (E-state) and the other (T-state) favours transfer of the polysaccharide to WaaL that links it to core lipid A. Wzz is involved in the timing mechanism that transfers Wzy from E- to T-state (Bastin *et al.*, 1993). (B) Wzz facilitates the interaction between Wzy, WaaL, and lipid-linked polysaccharide. The Wzz-dependent ratio Wzy:Wzz determines chain-length (Morona *et al.*, 1995).

Polymerisation and export processes

Although enzymes implied in polymerisation, chain-length determination and export of the cell-surface polysaccharides have been described for several organisms, the underlying mechanisms are still poorly understood. As described above, the first pathway for O-antigen biosynthesis involves a flippase (Wzx), polymerase (Wzy), and a regulator of O-antigen chain-length (Wzz). Wzx is a hydrophobic protein with 12 potential transmembrane domains. A *Salmonella wzx* mutant strain accumulates lipid-linked O-units at the cytoplasmic side of the cytoplasmic membrane, implying that *wzx* encodes a flippase (Liu *et al.*, 1996). Wzy is an integral membrane protein located in the cytoplasmic membrane, with 12 transmembrane segments and two large periplasmic loops and is involved in polymerisation of O-repeat units into long-chain O antigen (Daniels *et al.*, 1998, Morona *et al.*, 1994, Reeves *et al.*, 1996). It has been speculated that apart from a polymerase, Wzy might also be a permease using the electrochemical gradient to drive O-antigen polymerisation or acting as a pump to retranslocate the lipid carrier to the cytoplasmic side of the membrane after the O-antigen repeating unit has been transferred (Daniels *et al.*, 1998). Furthermore, Wzy has an impact on O-antigen chain length by the Wzy/Wzz ratio, as overproduction of *Shigella flexneri* Wzy results in an unregulated O-chain length which can be modulated by introducing *wzz* on a low-copy plasmid (Daniels *et al.*, 1998). Wzz proteins have two highly conserved potential transmembrane domains in the N- and C-terminal regions and are located in the cytoplasmic membrane with the central domain exposed to the periplasm (Morona *et al.*, 1995, Whitfield *et al.*, 1997). As it has been demonstrated that single or double amino acid changes in Wzz can have an effect on the chain length of O antigens, the heterogeneity of O-antigen chain length might be the result of amino acid sequence variation of the Wzz protein (Franco *et al.*,

1998, Klee *et al.*, 1997). Two models for Wzz activity have been proposed (Fig. 4). In the first model, Wzz is considered to act in concert with a Wzy enzyme which can exist in two functional states (Bastin *et al.*, 1993). The 'E-state' favours further polymerisation of a lipid-linked polymer, whereas the 'T-state' favours its transfer to WaaL, the lipid A-core ligase, and, upon ligation, polymerisation ends. Modality is established by a Wzz-mediated 'timing' mechanism in which Wzy moves from the E-state to the T-state after a given time period, allowing addition of a consistent number of repeating units in the polymerisation phase. Another model suggests that Wzz might act as a molecular chaperone facilitating the interaction of WaaL with Wzy and lipid-linked O-antigen chains. Specific modality would result from a given ratio of Wzy:WaaL in the Wzz-dependent complex (Morona *et al.*, 1995). Wzz homologues are found in many other gene clusters responsible for cell-surface biosynthesis, but their involvement in chain-length determination is rarely supported by experimental evidence (Whitfield *et al.*, 1997). *R. meliloti* ExoP influences succinoglycan chain length. Its N-terminus is homologous to Wzz and it has an additional C-terminal domain with an ATP binding domain. Both the C-terminal domain and a proline-rich motif ($RX_4PX_2PX_4SPKX_9IXGXMXGXG$) close to the second transmembrane helix in the N-terminal domain are involved in its activity (Becker *et al.*, 1995, Becker *et al.*, 1998). The ExoP-like proteins, containing an ATP-binding domain, are found in other CPS and EPS biosynthesis systems and are named Wzc (Reeves *et al.*, 1996). Their activity is expected to more complex than Wzz that has no ATP-binding domain (Whitfield and Roberts, 1999).

The two other pathways of O-antigen synthesis (see above) do not involve a separate polymerase. In the ABC-transporter dependent pathway, modality could be established by selection for molecules in a given size range by the transport components (Whitfield *et al.*, 1997). The ratio of the ABC-transporter and the polysaccharide synthesis enzymes is important for the modality, as overproduction of the ABC-transporter results in a decrease in O-antigen chain-length (Bronner *et al.*, 1994). If glycosyltransferases act in an efficiently coordinated complex, and transport and polymerisation are essentially continuous processes, the complexity of glycosyltransferase substrate- and acceptor-binding specificities might determine modality (Whitfield *et al.*, 1997). An alternative model is that one transferase in the system operates more slowly than others, making the synthetic process discontinuous. At the slowest point in the cycle of glycosyltransferase activities, export by the transporter may be favoured over addition of the next residue (Kido *et al.*, 1995). For *S. enterica* O:54, polymerisation and export are believed to be performed by a single enzyme, WbbF (Keenleyside and Whitfield, 1996). The chain-length of the O:54 antigen is quite broad and resembles a non-modal unregulated pattern (Whitfield *et al.*, 1997).

Genetics of polysaccharide biosynthesis

Many studies describing gene clusters involved in bacterial polysaccharide biosynthesis have appeared over the last years, in particular those directing O-antigen and CPS synthesis

in Gram-negative bacteria. In general, the genes involved in cell-surface polysaccharide biosynthesis are clustered, which allows coordinate regulation (Roberts, 1996). Although they may include one or more operons, a conserved organisation has been observed for several of these gene clusters, especially those involved in CPS biosynthesis. In all these cases glycosyltransferase genes that direct repeating unit synthesis are flanked by genes involved in polymerisation and/or transport (Roberts, 1996). In some cases genes directing the biosynthesis of specific nucleotide sugars are located in these gene clusters as well. This conserved organisation allows allelic exchange of glycosyltransferases resulting in capsule switching, which may be an important virulence mechanism of encapsulated bacterial pathogens. *Neisseria meningitidis* is able to switch from serogroup B (α(2→8)-linked polysialic acid) to serogroup C (α(2→9)-linked polysialic acid), probably after horizontal DNA transfer *in vivo* resulting in the exchange of the polysialyltransferase (Swartley *et al.*, 1997). Likewise, recombination within or around the *cps* locus involved in CPS synthesis, could have resulted in the generation of serotype 19F variants of *S. pneumoniae* (Coffey *et al.*, 1998). Serotype switching has also been proposed for O antigens. The *S. enterica* serogroup D2 is proposed to be the product of intraspecific recombination from a D1 strain with a E1 strain resulting in the exchange of the D1 *wzy* (*rfc*) polymerase gene and *wbaO* (*rfbO*) β(1→4) mannose transferase gene for the E1 *wzy* polymerase gene and *wbaU* (*rfbU*) α(1→4) mannose transferase gene. This exchange could have been mediated by the *Hinc* repeat (H-rpt) resembling an insertion sequence (Xiang *et al.*, 1994). Similarly, the *E. coli* O9a serotype is believed to be generated by recombination with the *Klebsiella* O3 O-antigen gene cluster (Sugiyama *et al.*, 1998).

For EPSs the gene clusters involved in xanthan and succinoglycan synthesis are best documented. Xanthan is produced in high amounts by the phytopathogenic bacterium *X. campestris* and is widely used in food industry (see Becker *et al.*, 1998 for a review). Its synthesis is controlled by the 16-kb *gumBCDEFGHIJKLM* gene cluster. The encoded GumD is the priming glucosyltransferase, GumM, GumH, GumK, and GumI are the glycosyltransferases involved in the subsequent steps of repeating unit synthesis, GumF and GumG are mannosyl acetyltransferases, GumL is the mannosyl pyruvylase, and GumB, GumC, and GumE are involved in polymerisation and export (see Becker *et al.*, 1998 for a review). *R. meliloti* harbours a 1.5-Mb megaplasmid that contains the gene clusters involved in succinoglycan (EPS I) and galactoglucan (EPS II) biosynthesis. Both polysaccharides are involved in processes leading to nitrogen-fixing symbiosis of the bacterium and alfalfa plants (Leigh and Walker, 1994). A 27-kb *exo/exs* gene cluster directing succinoglycan biosynthesis contains 19 *exo* and 2 *exs* genes organised in 10 operons (Glucksmann *et al.*, 1993, Becker *et al.*, 1995) and is separated by 200 kb from the 32-kb *exp* gene cluster directing galactoglucan biosynthesis containing 25 genes organised in 5 operons *expA*, *expC*, *expG*, *expD* and *expE* (Becker *et al.*, 1997). The *exo* genes encode enzymes involved in repeating unit synthesis and polymerisation and export as well as those required for UDP-glucose and UDP-galactose generation (Leigh and Walker, 1994). The *exp* genes encode

enzymes for dTDP-glucose and dTDP-rhamnose synthesis, repeating unit synthesis, export, and regulation of *exp* gene expression (Becker *et al.*, 1997).

Regulation of polysaccharide biosynthesis

Several bacteria can adapt to their environment by modulating their capacity to synthesise cell-surface polysaccharides. It appears that environmental factors may affect gene expression via two component regulatory systems, but other regulatory factors have been described as well.

One of the best documented systems for regulation of polysaccharide synthesis is that of the *E. coli* colanic acid, a mucoid CPS which improves the survival of the bacterium in various hostile environments. The *cpsABCDE* genes directing colanic acid biosynthesis are positively regulated by RcsA, RcsB, and RcsF, negatively regulated by the ATP-dependent Lon protease, and probably both positively and negatively regulated by RcsC (Fig. 5). The availability of the positive regulator RcsA is normally limited, as it is rapidly degraded by the Lon protease. RcsB and RcsC are the response-regulator and sensor of a two-component, environmentally responsive, regulatory system, while RscA is thought to interact with RcsB to induce transcription of the *cps* genes (Gottesman and Stout, 1991). Furthermore, RcsA can activate its own expression and a putative RcsA binding site has been found in the *rcsA* and *cps* promoter regions (Ebel and Trempy, 1999). RcsF seems to have an accessory role in activation which could be stabilisation of RcsA (Kelly and Georgopoulos, 1997). Recently, a new factor DjlA was described which together with the DnaK and GrpE chaperones can positively regulate the RcsB/C two component system (Kelly and Georgopoulos, 1997). It has not been established whether DjlA interacts with RcsC, RcsB, or both. Stress responses influence *cps* expression. Osmotic shock induces *cps* transcription in the presence of RcsB and RcsC (Sledjeski and Gottesman, 1996), and *djlA* transcription is induced or stabilised upon cold shock (Kelly and Georgopoulos, 1997). The *rcs* system might be common to many

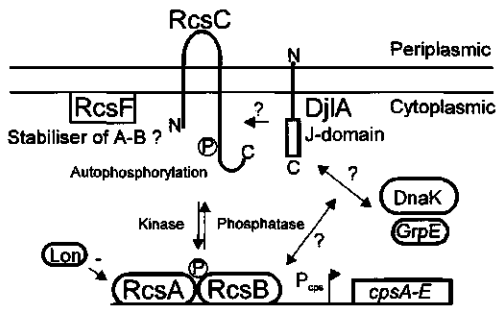


Fig. 5. Model of Rcs B/C two-component signalling system and the role(s) of DjlA, DnaK, and GrpE (adopted from Kelly and Georgopoulos, 1997).

other bacteria and Rcs homologues have been described in group 1 capsule-producing *E. coli* strains, *Klebsiella pneumoniae*, *Erwinia amylovora*, and *Salmonella typhi* (Whitfield and Roberts, 1999). In *E. amylovora*, RcsA and RcsB were shown to interact with the promoter of the *ams* operon controlling the biosynthesis of the EPS amylovoran and disruption of *rcsB* resulted in mutants which were deficient in amylovoran synthesis (Bereswill *et al.*, 1997, Kelm *et al.*, 1997). The RcsA/RcsB recognition motif of the promoter of the *ams* operon has been characterized and is also found in the promoter of the exopolysaccharide biosynthetic operon of *Pantoea stewartii* (formerly *Erwinia stewartii*) (Wehland *et al.*, 1999). In addition to *rcsA*/*rcsB*, another positive regulator of the *E. amylovora ams* operon, *rcsV*, has been identified that can suppress an *rcsA* mutation (Aldridge *et al.*, 1998). In *S. typhi* the Vi antigen genes directing CPS synthesis are controlled by two two-component regulatory systems, OmpR-EnvZ and RcsB-RcsC, that respond to osmolarity and by the positive regulatior TviA, which interacts with RcsB to promote optimal transcription of the Vi synthesis genes (Arricau *et al.*, 1998, Pickard *et al.*, 1994).

Environmental factors also influence the expression of the *eps* gene cluster of *Pseudomonas solanacearum* directing EPS I biosynthesis (Huang *et al.*, 1995). This gene cluster is controlled by a complex regulatory network consisting of three separate signal transduction systems: PhcA, a LysR-type transcriptional regulator, and the two-component regulatory systems VsrA/VsrD and VsrB/VsrC. PhcA and VsrA/VsrD control transcription of a sixth regulation factor *xpsR*. XpsR is required by VsrB/VsrC to activate *eps* gene expression and interconnects the three signal transduction systems (Huang *et al.*, 1995). Another factor regulating *eps* gene expression is EpsR. Depending on its phosphorylation state, EpsR reduces or induces *eps* gene expression by binding the *eps* promoter in the phosphorylated form (Chapman and Kao, 1998).

R. meliloti succinoglycan production is under control of the *exoR* and *exoS* gene products, as Tn*5*-insertions in the chromosomaly-located *exoR* and *exoS* result in an increased succinoglycan synthesis (Doherty *et al.*, 1988). The *exoS* mutant, like wild-type *Rhizobium*, synthesises less succinoglycan in the presence of ammonia, while succinoglycan-production of the *exoR* mutant is not influenced by ammonia, indicating that ExoR is involved in sensing ammonia in the medium (Doherty *et al.*, 1988). Both *exoR* and *exoS* genes have been cloned and sequenced (Cheng, 1998, Reed, 1991) and ExoR was found to be a negative regulator for the transcription of the *exo* genes (Reed, 1991). The *exoS* gene is located downstream of *chvI* and both genes may encode a two-component regulatory system with ExoS as the sensor domain and ChvI as the response regulator (Cheng, 1991). Another form of succinoglycan regulation involves *exoX* and *exoY*. ExoX may interact with ExoY, the glycosyltransferase initiating repeating unit synthesis, thereby preventing its activity (Gray *et al.*, 1990, Gray and Rolfe, 1992). Additionally, posttranscriptional regulation of succinoglycan synthesis occurs by the *mucR* gene product encoded by the chromosome and by the *exsB* gene product encoded in the *exs* operon that have a positive and negative effect, respectively (Becker *et al.*, 1995, Keller *et al.*, 1995). Production of galactoglucan by *R. meliloti* occurs at low phosphate concentrations or in strains with mutations in *expR* or *mucR*, which are located in the

chromosome. MucR represses transcription of the 5 operons of the *exp* gene cluster (directing galactoglucan synthesis), has a positive effect on the *exo* genes (directing succinoglycan synthesis) via posttranscriptional regulation (see above), and is negatively regulating its own expression (Becker *et al.*, 1997, Keller *et al.*, 1995). Furthermore, the *exp* gene cluster contains the *expG* (*mucS*) gene that may be a transcriptional activator of *exp* gene expression, as it is known to be required for the activation of at least one gene of the *expE* operon by low phosphate concentrations (Astete, 1996, Becker *et al.*, 1997).

Lactic acid bacteria

Lactic acid bacteria (LAB) are Gram-positive bacteria comprising the genera *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Streptococcus*, and *Pediococcus*. For centuries, they have been used in food fermentations for the production of *e.g.* dairy products, wine or sausages. The most important property of LAB is their rapid conversion of lactose into lactate. This results in acidification and preserves the food from spoilage. In addition, LAB provide taste and texture to fermented food products. From the LAB, the species *Lactococcus lactis* is best studied for its genetics and metabolism (Gasson and de Vos, 1994). *L. lactis* is used for the production of cheese, buttermilk and butter. Its proteolytic system consists of an extracellular proteinase, several peptide uptake systems, and intracellular peptidases and is responsible for the conversion of milk caseins into small peptides and amino acids that contribute to the determination of the flavour and texture of the final product (see Kunji *et al.*, 1996 for a review). Some lactococci are able to form diacetyl from citrate or lactose, which is an important flavour and aroma compound in products such as buttermilk, lactic butter, or cottage cheese.

Many industrially relevant characteristics of lactococci, such as lactose fermentation, protease activity, bacteriophage resistance, are encoded by plasmids that in most cases can be transferred by conjugation. Lactococcal plasmids are known to replicate via the rolling circle or theta mechanism. Rolling circle replication seems to be restricted to relatively small lactococcal plasmids with cryptic functions, while those encoding the metabolic functions all seem to replicate via theta mechanism (Khan, 1997, Seegers *et al.*, 1994). For the plasmids that can be conjugally transferred, self-transmissible conjugative plasmids, which have the ability to form effective cell-to-cell contact, and mobilisation plasmids, which are able only to prepare their DNA for transfer can be distinguished (Steele and McKay, 1989).

Exopolysaccharides from LAB

Several LAB are known to produce EPSs that can be beneficial for the structure of dairy products. It has been demonstrated that EPSs play a significant role in the rheology of stirred yoghurt (van Marle *et al.*, 1998). Furthermore, they may be used to replace polysaccharides

that are used in food industry as thickeners, stabilisers, emulsifiers, bodying agents, foam enhancers, gelling agents, or fat replacers. EPSs produced by LAB could serve as alternatives for xanthan gum produced by the phytopathogenic *X. campestris*, with the advantage that LAB are food-grade organisms that have a long history of safe use in food fermentations. The EPSs can be produced *in situ* resulting in a 'natural' product with no need for additives to improve the structure. In addition, some studies indicate EPS from LAB may be health beneficial and that consumption of fermented products containing EPS has a stimulatory effect on the immune response, an antitumoral activity, or cholesterol-lowering activity (Kitazawa *et al.*, 1991a, Kitazawa *et al.*, 1991b, Nakajima *et al.*, 1992a).

EPS-producing LAB strains have been isolated from dairy products such as Scandinavian ropy fermented milk products (Macura and Townsley, 1984, Nakajima *et al.*, 1990) and yoghurts (Cerning *et al.*, 1986, Cerning *et al.*, 1988), and from other fermented foods such as salami sausages and olives (van den Berg *et al.*, 1993). The nature of the slime material produced by these ropy LAB has been a matter of dispute. In early studies the slime extracted from ropy sour milk was found to consist of a protein-like material (Nilsson and Nilsson, 1958, Sundman, 1953). Later, Macura and Townsley reported that the slime produced by *L. lactis* strain L416, the parental strain of NIZO B40, consisted of a glycoprotein (Macura and Townsley, 1984). However, the reported protein content resembled that of the growth medium. Forsén *et al.* (Forsén *et al.*, 1989) focused on the differences in lipoteichoic acid from the cell wall of *L. lactis* strains in relation to slime production. Other studies showed that the carbohydrate parts from the carbohydrate-protein mixture of the produced slime can be purified and indeed are polysaccharides (see below). In general, these polymers contain (branched) repeating units consisting of a various number of sugar residues that are coupled via different types of linkages, and can be decorated with non-sugar groups (acetyl, phosphate, or glycerol). The molecular structures of EPSs formed by various *Streptococcus thermophilus* strains, *Lactobacillus acidophilus* LMG9433, *Lactobacillus delbrückii* subsp. *bulgaricus* rr, various *Lactobacillus helveticus* strains, *Lactobacillus paracasei* 34-1, *Lactobacillus sake* 0-1, and *L. lactis* subsp. *cremoris* strains H414, SBT 0495 and NIZO B40 have been elucidated (see Fig. 6).

Some strains produce more than one type of EPS. *Lb. delbrückii* subsp. *bulgaricus* NCFB2772 produces a high-molecular weight (HMW) EPS with a molecular mass of 1.7×10^6 and a low-molecular weight (LMW) EPS (4×10^4) that differ in types of sugar linkages (Grobben *et al.*, 1997). The production of HMW EPS is dependent on the carbohydrate source in the growth medium, while LMW EPS is produced continuously. Similarly, *L. lactis* LC330 produces HMW ($>1 \times 10^6$) and LMW (approximately 1×10^5) EPS with different sugar composition (Marshall *et al.*, 1995). Moreover, the production of LMW EPS by this strain is not influenced by growth conditions, while that of HMW EPS is. For *S. thermophilus* LY03, a HMW (1.8×10^6) and LMW (4.1×10^5) fraction of EPS with identical sugar composition can be distinguished, the ratio and production level of which is dependent on carbon/nitrogen ratio of the growth medium (Degeest *et al.*, 1998).

Little is known about the intrinsic characteristics of EPS that determine its texturising capacities. The structure of the repeating unit will affect the secondary and tertiary conformation. However, the variety of structures of EPSs found (Fig. 6) gives no clue about essential common features. Two *S. thermophilus* strains, Rs and Sts, from yoghurt starters that differ markedly in viscosifying properties, were found to produce similar amounts of EPS consisting of identical repeating units, but with different molecular masses (2.6×10^6 and 3.7×10^3, respectively), indicating that for these bacteria the size of the EPS determines the viscosity of the stirred milk cultures (Faber *et al.*, 1998).

Genetics of EPS biosynthesis by LAB

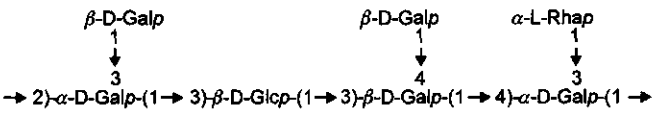
EPS production by LAB is an unstable property and this is manifested by either instability of the texture itself or instability of synthesis at the genetic level (Cerning, 1990). For several strains of *L. lactis* and for two strains of *Lactobacillus casei*, plasmids have been associated with EPS production and their loss results in the inability to produce EPS (Kojic *et al.*, 1992, Neve *et al.*, 1988, Vedamuthu and Neville, 1986, Vescovo *et al.*, 1989, von Wright and Tynkkynen, 1987). Genetic transfer of two of these lactococcal plasmids was achieved and resulted in an EPS-producing phenotype for the recipient strain (Vedamuthu and Neville, 1986, von Wright and Tynkkynen, 1987). For thermophilic LAB, EPS production has not been found to be linked to plasmids (Cerning, 1990, Vescovo *et al.*, 1989). For these strains genetic instability may be the result of disruption of essential gene activity by mobile genetic elements or generalised genomic instability caused by deletions and rearrangements (Gancel and Novel, 1994).

Genes directing EPS synthesis in LAB strains were first described for *S. thermophilus* Sfi6 (Stingele *et al.*, 1996). The 14.5-kb *eps* gene cluster comprises 13 genes that seem to be organised in a single operon (*epsABCDEFGHIJKLM*). The organisation of the *eps* gene cluster is comparable to that of *S. pneumoniae* and *Streptococcus agalactiae cps* gene clusters directing CPS synthesis and the *epsABCDE* gene products are highly homologous

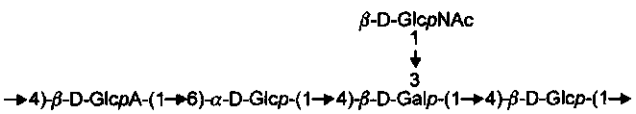
Fig. 6. (Following pages) Molecular structures of EPSs formed by various LAB. (A) *Lb. delbrückii* subsp. *bulgaricus* rr (Gruter *et al.*, 1993) and *Lb. delbrückii* subsp. *bulgaricus* NCFB 2772 (Grobben *et al.*, 1997). (B) *Lb. acidophilus* LMG9433 (Robijn *et al.*, 1996b). (C) *Lb. helveticus* 766 (Robijn *et al.*, 1995b). Staaf *et al.*, 1996. (D) *Lb. helveticus* TY1-2 (Yamamoto *et al.*, 1994). (E) *Lb. helveticus* TN-4 (Yamamoto *et al.*, 1995) and *Lb. helveticus* Lh59 (Stingele *et al.*, 1997). (F) *Lb. helveticus* ssp. (Staaf *et al.*, 1996). (G) *Lb. paracasei* 34-1 (Robijn *et al.*, 1996a). (H) *Lb. sake* 0-1 (Robijn *et al.*, 1995a). (I) *Lb. rhamnosus* C83 (Vanhaverbeke *et al.*, 1998). (J) *S. thermophilus* Sfi6 (Doco *et al.*, 1990, Stingele *et al.*, 1996). (K) *S. thermopilus* Sfi12 (Lemoine *et al.*, 1997). (L) *S. thermophilus* Sfi32 (Lemoine *et al.*, 1997). (M) *S. thermophilus* OR 901 (Bubb *et al.*, 1997), *S. thermophilus* Rs (Faber *et al.*, 1998), and *S. thermophilus* Sts (Faber *et al.*, 1998). (N) *S. thermophilus* MR-1C (Low *et al.*, 1998). (O) *L. lactis* SBT 0495 (Nakajima *et al.*, 1992b), *L. lactis* NIZO B40 (van Casteren *et al.*, 1998). (P) *L. lactis* H414 (Gruter, 1992). Abbreviations: Glc, glucose; Gal, galactose; Rha, rhamnose; GlcNAc, *N*-acetylglucosamine; GalNAc, *N*-acetylgalactosamine; Fuc, fucose; Ac, acetyl. The D- (D) and L- (L) configuration, and pyranose (*p*) and furanose (*f*) structure are indicated.

Lactobacillus:

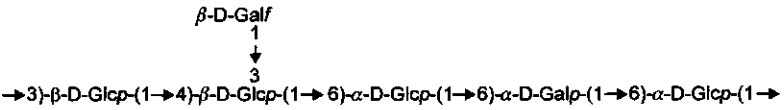
A



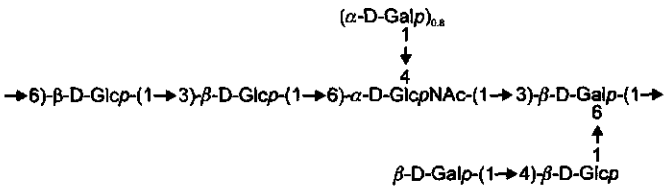
B



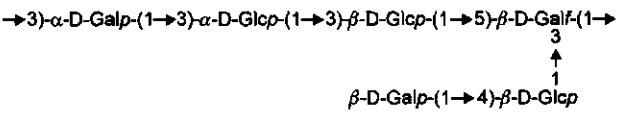
C



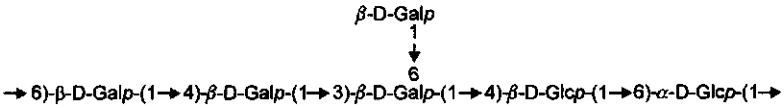
D



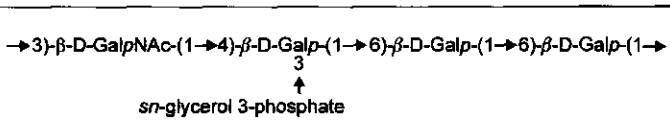
E



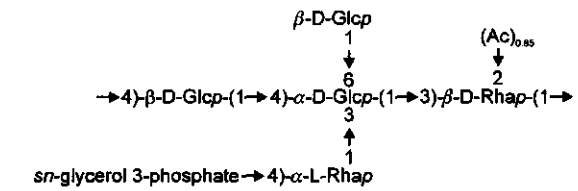
F



G



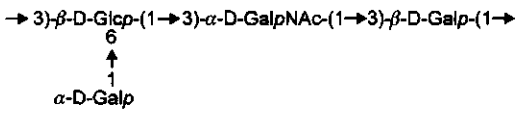
H



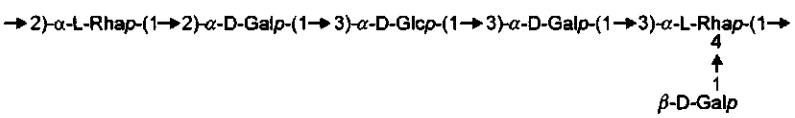
I

Streptococcus thermophilus:

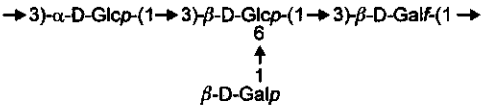
J



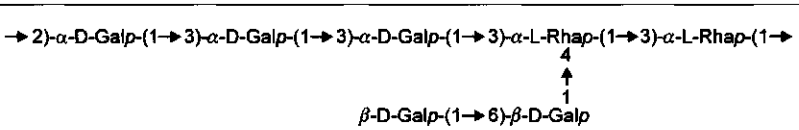
K



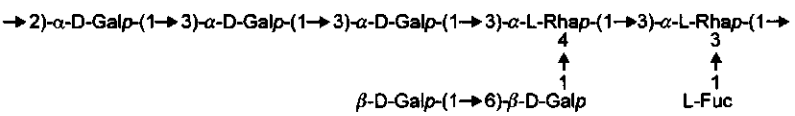
L



M

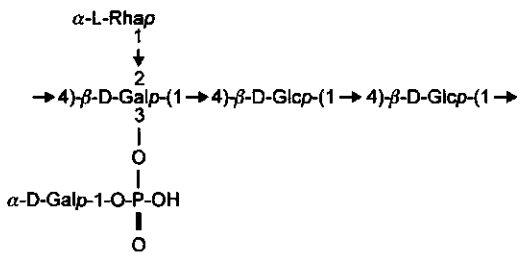


N

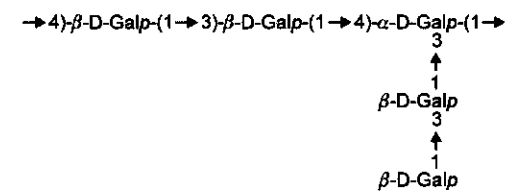


Lactococcus lactis:

O



P



with the pneumococcal *cpsABCDE* gene products (Stingele *et al.*, 1996). Based on homologies, the functions of the different gene products may be regulation (EpsA), chain-length determination and export (EpsC/EpsD), repeating unit synthesis (EpsE/EpsF/EpsG/EpsH/EpsI), and polymerisation and export (EpsJ/EpsK). Griffin *et al.* (1996) cloned part of a putative *eps* gene cluster from *S. thermophilus* NCBF 2393 (*cpsABCDE*) which is almost identical to that of Sfi6. Based on the sequence of the Sfi6 *eps* gene cluster, part of a highly homologous *eps* gene cluster of *S. thermophilus* MR-1C (approximately 98% identity on DNA-level) was cloned using PCR-methods and designated *epsABCDEF* (Low *et al.*, 1998). The EPS produced by this strain has a repeating unit with a different structure compared to that of Sfi6 (see Fig. 6 I and M). Attempts to characterise the MR-1C region from *epsG* to *epsM* by PCR were unsuccessful (Low *et al.*, 1998), indicating that this part of the gene cluster has a lower degree of homology or a different organisation, which may be expected for genes directing the synthesis of a polymer with a different structure.

Outline of the thesis

Both to study the structure-function relation of EPS and for industrial applications, it would be desirable to be able to control the EPS-production level, the chain-length of the EPS molecule, and the primary structure of the EPS repeating unit, enabling polysaccharide engineering. To achieve this, knowledge of *eps* gene expression and its regulation, the mechanism of EPS synthesis, polymerisation and export, and diversity of *eps* genes is essential. At the start of the research described in this thesis, no data on genetics of EPS production were available apart from reports describing its possible plasmid association in *L. lactis* strains. Therefore, a study was initiated to clone and characterise genes involved in EPS production in *L. lactis* strains and use these to engineer the EPS production level and biosynthesis pathway. Chapter 2 describes the identification of the EPS plasmid pNZ4000 from *L. lactis* NIZO B40 and the transfer of the EPS-producing capacity to another *L. lactis* strain by conjugation of this plasmid. The molecular structure of the EPS produced by *L. lactis* NIZO B40 was studied, the *eps* gene cluster involved in EPS biosynthesis was localised on pNZ4000, its DNA-sequence determined, and its transcription analysed. Furthermore, the functional analysis of the priming glycosyltransferase gene by a single-gene disruption and heterologous expression demonstrated its involvement in EPS biosynthesis. Chapter 3 describes the biosynthesis route of the NIZO B40 repeating unit. Homologous and heterologous expression of the glycosyltransferase genes was used for their functional analysis and to determine the substrate specificities of the encoded enzymes. In Chapter 4 the diversity of lactococcal EPSs and *eps* gene clusters is described together with the functional analysis of the parts of two of these *eps* gene clusters encoding glycosyltransferase activity. Furthermore, a system that may enable polysaccharide engineering is described and was demonstrated to be useful for complementing a disruption of the NIZO B40 priming

glycosyltransferase activity by priming glycosyltransferases from different polysaccharide synthesis systems. Chapter 5 describes factors involved in natural transfer of the EPS-producing phenotype and stability of EPS-production. Genes and genetic elements involved in plasmid mobilisation and replication were studied. Replicon stability and transfer to various *L. lactis* strains were analysed. Chapter 6 describes the complete nucleotide sequence analysis of the 42810-bp EPS plasmid pNZ4000. Next to the regions involved in replication, mobilisation, and EPS biosynthesis, two regions putatively involved in transport of divalent cations were identified. Finally, the results of Chapters 2 to 6, combined with additional data are discussed in the general discussion in Chapter 7 focussing on EPS biosynthesis, including glycosyltransferases, polymerisation and export processes, regulation, and practical applications and perspectives of this work.

References

Aldridge, P., F. Bernhard, P. Bugert, D. L. Coplin, and K. Geider. 1998. Characterization of a gene locus from *Erwinia amylovora* with regulatory functions in exopolysaccharide synthesis of *Erwinia* spp. Can. J. Microbiol. **44:**657-666.

Arricau, N., D. Hermant, H. Waxin, C. Ecobichon, P. S. Duffey, and M. Y. Popoff. 1998. The RcsB-RcsC regulatory system of *Salmonella typhi* differentially modulates the expression of invasion proteins, flagellin and Vi antigen in response to osmolarity. Mol. Microbiol. **29:**835-850.

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 synthetase, with a novel mechanism for determining chain length. Mol. Microbiol. **7:**725-734.

Becker, A., H. Küster, K. Niehaus, and A. Pühler. 1995. Extension of the *Rhizobium meliloti* succinoglycan biosynthesis gene cluster: identification of the *exsA* gene encoding an ABC transporter protein, and the *exsB* gene which probably codes for a regulator of succinoglycan biosynthesis. Mol. Gen. Genet. **249:**487-497.

Becker, A., S. Rüberg, H. Küster, A. A. Roxlau, M. Keller, T. Ivashina, H.-P. Cheng, G. C. Walker, and A. Pühler. 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.

Becker, A., F. Katzen, A. Pühler, and L. Ielpi. 1998. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Appl. Microbiol. Biotechnol. **50:**145-152.

Bereswill, S., and K. Geider. 1997. Characterization of the *rcsB* gene from *Erwinia amylovora* and its influence on exopolysaccharide synthesis and virulence of the fire blight pathogen. J. Bacteriol. **179:**1354-1361.

Bronner, D., B. R. Clarke, and C. Whitfield. 1994. Identification of an ATP-binding cassette transport system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumoniae* serotype O1. Mol. Microbiol. **14:**505-519.

Bubb, W. A., T. Urashima, R. Fujiwara, T. Shinnai, and H. Ariga. 1997. Structural characterisation of the exocellular polysaccharide produced by *Streptococcus thermophilus* OR 901. Carbohydr. Res. **301:**41-50.

Chapman, M. R., and C. C. Kao. 1998. EpsR modulates production of extracellular polysaccharide in the bacterial wilt pathogen *Ralstonia (Pseudomonas) solanacearum*. J. Bacteriol. **180:**27-34.

Cerning, J. C. Bouillanne, M. J. Desmazeaud, and M. Landon. 1986. Isolation and characterization of exocellular polysaccharide produced by *Lactobacillus bulgaricus*. Biotechnol. Lett. **8:**625-628.

Cerning, J. C. Bouillanne, M. J. Desmazeaud, and M. Landon. 1988. Exocellular polysaccharide production by *Streptococcus thermophilus*. Biotechnol. Lett. **10:** 255-260.

Cerning, J. 1990. Exocellular polysaccharides produced by lactic acid bacteria. FEMS Microbiol. Rev. **87:**113-130.

Coffey, T. J., M. C. Enright, M. Daniels, J. K. Morona, R. Morona, W. Hryniewicz, J. C. Paton, and B. G. Spratt. 1998. Recombinational exchanges at the capsular polysaccharide biosynthetic locus leads to frequent serotype changes among natural isolates of *Streptococcus pneumoniae.* Mol. Microbiol. **27:**73-83.

Daniels, C., C. Vindurampulle, and R. Morona. 1998. Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy). Mol. Microbiol. **28:**1211-1222.

Degeest, B., S. van de Ven, F. Vanderveken, and L. de Vuyst. 1998. Process optimisation of exopolysaccharide production by *Streptococcus thermophilus. In* Book of abstracts, NIZO 50 years symposium Dairying behind the dikes, p. 80. Arnhem, The Netherlands.

Dierksen, K. P., J. Marks, D. D. Chen, and J. E. Trempy. 1994. Evidence for structural conservation of Lon and RcsA. J. Bacteriol. **176:**5126-5130.

Doco, T., J.-M. Wieruszeski, B. Fournet, D. Carcano, P. Ramos, and A. Loones. 1990. Structure of an exocellular polysaccharide produced by *Streptococcus themophilus.* Carbohydr. Res. **198:**313-321.

Doherty, D., J. A. Leigh, J. Glazebrook, and G. C. Walker. 1988. *Rhizobium meliloti* mutants that overproduce the *R. meliloti* acidic Calcofluor-binding exopolysaccharide. J. Bacteriol. **170:**4249-4256.

Ebel, W., and J. E. Trempy. 1999. *Escherichia coli* RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions to activate its own expression. J. Bacteriol. **181:**577-584.

Faber, E. J., P. Zoon, J. P. Kamerling, and J. F. G. Vliegenthart. 1998. The exopolysaccharides produced by *Streptococcus thermophilus* Rs and Sts have the same repeating unit but differ in viscosity of their milk cultures. Carbohydr. Res. **310:**269-276.

Forsén, R., K. Niskasaari, L. Tasanen, and E.-L. Numiaho-Lassila. 1989. Studies on slimy lactic acid fermentation: detection of lipoteichoic acid containing membrane antigens of *Lactococcus lactis* ssp. *cremoris* strains by crossed immunoelectrophoresis. Neth. Milk Dairy J. **43:**383-393.

Franco, A. V., D. Liu, and P. R. Reeves. 1998. The Wzz (Cld) protein in *Escherichia coli*: amino acid sequence variation determines O-antigen chain length specificity. J. Bacteriol. **180:**2670-2675.

Gancel, F., and G. Novel. 1994. Exopolysaccharide production by *Streptococcus salivarius* ssp. *thermophilus* cultures. 2. Distinct modes of polymer production and degradation among clonal variants. J. Dairy Sci. **77:**689-695.

Gasson, M. J., and W. M. de Vos, eds. 1994. Genetics and biotechnology of lactic acid bacteria. Chapman & Hall, London, UK.

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.

Gottesman, S., and V. Stout. 1991. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. Mol. Microbiol. **5:**1599-1606.

Gray, J. X., M. A. Djordjevic, and B. G. Rolfe. 1990. Two genes that regulate exopolysaccharide production in *Rhizobium* sp. strain NGR234. J. Bacteriol. **172:**193-203.

Gray, J. X., and B. G. Rolfe. 1992. Regulation study of exopolysaccharide synthesis, *exoX* and *exoY* in *Rhizobium* sp. strain NGR234. Arch. Microbiol. **157:**521-528.

Griffin, A. M., V. J. Morris, and M. J. Gasson. 1996. The *cpsABCDE* genes involved in polysaccharide production in *Streptococcus salivarius* ssp. *thermophilus* strain NCBF 2393. Gene **183:**23-27.

Grobben, G. J., W. H. M. van Casteren, H. A. Schols, A. Oosterveld, G. Sala, M. R. Smith, J. Sikkema, and J. A. M. de Bont. 1997. Analysis of the exopolysaccharides produced by *Lactobacillus delbrueckii* subsp. *bulgaricus* NCFB 2772 grown in continuous culture on glucose and fructose. Appl. Microbiol. Biotechnol. **48:**516-521.

Gruter, M., B. R. Leeflang, J. Kuiper, J. P. Kamerling, and J. F. G. Vliegenthart. 1993. Structural characterisation of the exopolysaccharide produced by *Lactobacillus delbrückii* subspecies *bulgaricus* rr grown in skimmed milk. Carbohydr. Res. **239:**209-226.

Huang, J., B. F. Carney, T. P. Denny, A. K. Weissinger, and M. A. Schell. 1995. A complex network regulates expression of *eps* and other virulence genes of *Pseudomonas solanacearum*. J. Bacteriol. **177:**1259-1276.

Ielpi, L., R. O. Couso, and M. A. Dankert. 1993. Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. J. Bacteriol. **175:**2490-2500.

Keenleyside, W. J., and C. Whitfield. 1996. A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar Borreze. J. Biol. Chem. **271:**28581-28592.

Kelly, W. L., and C. Georgopoulos. 1997. Positive control of the two-component RcsC/B signal transduction network by DjlA: a member of the DnaJ family of molecular chaperones in *Escherichia coli*. Mol. Microbiol. **25:** 913-931.

Keller, M., A. Roxlau, W. M. Weng, M. Schmidt, J. Quandt, K. Niehaus, D. Jording, W. Arnold, and A. Pühler. 1995. Molecular analysis of the *Rhizobium meliloti mucR* gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan. Mol. Plant-Microbe Interact. **8:**267-277.

Kelm, O., C. Kiecker, K. Geider, and F. Bernhard. 1997. Interaction of the regulator proteins RcsA and RcsB with the promoter of the operon for amylovoran biosynthesis in *Erwinia amylovora*. Mol. Gen. Genet. **256:**72-83.

Khan, S. A. 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. **61:**442-455.

Kido, N., V. I. Torgov, T. Sugiyama, K. Uchiya, H. Sugihara, T. Komatsu, N. Kato, and K. Jann. 1995. Expression of the O9 polysaccharide of *Escherichia coli:* sequence of the *E. coli* O9 *rfb* gene cluster, characterization of mannosyl transferases, and evidence for an ATP-binding cassette transport system. J. Bacteriol. **177:**2178-2187.

Kitazawa, H., M. Nomura, T. Itoh, and T. Yamaguchi. 1991a. Functional alteration of macrophages by a slime-forming, encapsulated *Lactococcus lactis* ssp. *cremoris*. J. Dairy Sci. **74:**2082-2088.

Kitazawa, H., T. Toba, T. Itoh, N. Kumano, S. Adachi, and T. Yamaguchi. 1991b. Antitumoral activity of slime-forming encapsulated *Lactococcus lactis* subsp. *cremoris* isolated from Scandinavian ropy sour milk "viili". Anim. Sci. Technol. **62:**277-283.

Klee, S. R., B. D. Tzschaschel, K. N. Timmis, and C. Guzmán. 1997. Influence of different rol gene products on the chain length of *Shigella dysenteriae* type 1 lipopolysaccharide O antigen expressed by *Shigella flexneri* carrier strains. J. Bacteriol. **179:**2421-2425.

Kojic, M., M. Vujcic, A. Banina, P. Coconcelli, J. Cerning, and L. Topisirovic. 1992. Analysis of exopolysaccharide production by *Lactobacillus casei* CG11, isolated from cheese. Appl. Environ. Microbiol. **58:**4086-4088.

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.

Kunji, E. R. S., I. Mierau, A. Hagting, B. Poolman, and W. N. Konings. 1996. The proteolytic systems of lactic acid bacteria. Antonie van Leeuwenhoek **70:**187-221.

Leigh, J. A. and G. C. Walker. 1994. Exopolysaccharides of *Rhizobium*: synthesis, regulation and symbiotic function. Trends in Genetics **10:**63-67.

Lemoine, J., F. Chirat, J.-M. Wieuszeski, G. Strecker, N. Favre, and J.-R. Neeser. 1997. Structural characterization of the exocellular polysaccharides produced by *Streptococcus thermophilus* Sfi39 and Sfi12. Appl. Environ. Microbiol. **63:**3512-3518.

Liu, D., R. A. Cole, and P. R. Reeves. 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J. Bacteriol. **178:**2102-2107.

Low, D., J. A. Ahlgren, D. Horne, D. J. McMahon, C. J. Oberg, and J. R. Broadbent. 1998. Role of *Streptococcus thermophilus* MR-1C capsular exopolysaccharide in cheese moisture retention. Appl. Environ. Microbiol. **64:**2147-2151.

Macura, D., and P. M. Townsley. 1984. Scandinavian ropy milk - Identification and characterization of endogenous ropy lactic streptococci and their extracullular excretion. J. Dairy Sci. **67:**735-744.

Marshall, V. M., E. N. Cowie, and R. S. Moreton. 1995. Analysis and production of two exopolysaccharides from *Lactococcus lactis* subsp. *cremoris* LC330. J. Dairy Sci. **62:**621-628.

McGrath, B. C., and M. J. Osborn. 1991. Localization of the terminal steps of O-antigen synthesis in *Salmonella typhimurium.* J. Bacteriol. **173:**649-654.

Morona, R., M. Mavris, A. Falarino, and P. A. Manning. 1994. Characterization of the *rfc* region of *Shigella flexneri.* J. Bacteriol. **176:**733-747.

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.

Nakajima, H., S. Toyoda, T. Toba, T. Itoh, T. Mukai, H. Kitazawa, and S. Adachi. 1990. A novel phosphopolysaccharide from slime-forming *Lactococcus lactis* subspecies *cremoris* SBT 0495. J. Dairy Sci. **73:**1472-1477.

Nakajima, H., Y. Suzuki, H. Kaizu, and T. Hirota. 1992a. Cholesterol-lowering activity of ropy fermented milk. J. Food Sci. **57:**1327-1329.

Nakajima, H., T. Hirota, T. Toba, T. Itoh, and S. Adachi. 1992b. Structure of the extracellular polysaccharide from slime-forming *Lactococcus lactis* subsp. *cremoris* SBT 0495. Carbohydr. Res. **224:**245-253.

Neve, H., A. Geis, and M. Teuber. 1988. Plasmid-encoded functions of ropy lactic acid streptococcal strains from Scandinavian fermented milk. Biochimie **70:**427-442.

Nilsson, R., and G. Nilsson. 1958. Studies concerning Swedish ropy milk. Arch. Mikrobiol. **31:**191-197.

Pickard, D., J. Li, M. Roberts, D. Maskell, D. Hone, M. Levine, *et al.* 1994. Characterization of defined *ompR* mutants of *Salmonella typhi*: *ompR* is involved in the regulation of Vi polysaccharide expression. Infect. Immun. **62:**3984-3993.

Reeves, P. R., M. Hobbs, M. A. Valvano, M. Skurnik, C. Whitfield, D. Coplin, N. Kido, J. Klena, D. Maskell, C. R. H. Raetz, and P. D. Rick. 1996. Bacterial polysaccharide synthesis and gene nomenclature. Trends Microbiol. **4:**495-503.

Reuber, T. L., and G. C. Walker. 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti.* Cell **74:**269-280.

Robijn, G. W., D. J. C. van den Berg, H. Haas, J. P. Kamerling, and J. F. G. Vliegenthart. 1995a. Determination of the structure of the exopolysaccharide produced by *Lactobacillus sake* 0-1. Carbohydr. Res. **276:**117-136.

Robijn, G. W., J. R. Thomas, H. Haas, D. J. C. van den Berg, J. P. Kamerling, and J. F. G. Vliegenthart. 1995b. The structure of the exopolysaccharide produced by *Lactobacillus helveticus* 766. Carbohydr. Res. **276:**137-154.

Robijn, G. W., H. L. J. Wienk, D. J. C. van den Berg, H. Haas, J. P. Kamerling, and J. F. G. Vliegenthart. 1996a. Structural studies of the exopolysaccharide produced by *Lactobacillus paracasei* 34-1. Carbohydr. Res. **285:**129-139.

Robijn, G. W., R. G. Gallego, D. J. C. van den Berg, H. Haas, J. P. Kamerling, and J. F. G. Vliegenthart. 1996b. Structural characterization of the exopolysaccharide produced by *Lactobacillus acidophilus* LMG9433. Carbohydr. Res. **288:**203-218.

Roberts, I. S. 1996. The biochemistry and genetics of capsular polysaccharide production in bacteria. Annu. Rev. Microbiol. **50:**285-315.

Saxena, I. M., R. M. Brown, M. Fevre, R. A. Geremia, and B. Henrissat. 1995. Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. J. Bacteriol **177:**1419-1424.

Seegers, J. F. M., S. Bron, C. M. Franke, G. Venema, and R. Kiewiet. 1994. The majority of lactococcal plasmids carry a highly related replicon. Microbiology **140:**1291-1300.

Sledjeski, D. D., and S. Gottesman. 1996. Osmotic shock induction of capsule synthesis in *Escherichia coli* K-12. J. Bacteriol. **178:**1204-1206.

Staaf, M., G. Widmalm, Z. Yang, and E. Huttunen. 1996. Structural elucidation of an extracellular polysaccharide produced by *Lactobacillus helveticus*. Carbohydr. Res. **291:**155-164.

Steele, J. L., and L. L. McKay. 1989. Conjugal transfer of genetic material in lactococci: a review. J. Dairy Sci. **72:**3388-3397.

Stingele, F., J.-R. Neeser, and B. Mollet. 1996. Identification of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. J. Bacteriol. **178:**1680-1690.

Stingele, F., J. Lemoine, J.-R. Neeser. 1997. *Lactobacillus helveticus* Lh59 secretes an exopolysaccharide that is identical to the one produced by *Lactobacillus helveticus* TN-4, a presumed spontaneous mutant of *Lactobacillus helveticus* TY1-2. Carbohydr. Res. **302:**197-202.

Sugiyama, T., N. Kido, Y. Kato, N. Koide, T. Yoshida, and T. Yokochi. 1998. Generation of *Escherichia coli* O9a serotype, a subtype of *E. coli* O9, by transfer of the *wb** gene cluster of *Klebsiella* O3 into *E. coli* via recombination. J. Bacteriol. **180:**2775-2778.

Sundman, V. 1953. On the protein character of a slime produced by *Streptococcus cremoris* in Finnish ropy sour milk. Acta Chem. Scand. **7:**558-560.

Swartley, J. S., A. A. Marfin, S. Edupuganti, L.-J. Liu, P. Cieslak, B. Perkins, J. D. Wenger, and D. S. Stephens. 1997. Capsule switching of *Neisseria meningitidis*. Proc. Natl. Acad. Sci. USA **94:**271-276.

van Casteren, W. H. M., C. Dijkema, H. A. Schols, G. Beldman, and A. G. J. Voragen. 1998. Characterisation and modification of the exopolysaccharide produced by *Lactococcus lactis* subsp. *cremoris* B40. Carbohydr. Polymers **37:**123-130.

van den Berg, D. J. C., A. Smits, B. Pot, A. M. Ledeboer, K. Kersters, J. M. A. Verbakel, and C. T. Verrips. 1993. Isolation, screening and identification of lactic acid bacteria from traditional food fermentation processes and culture collections. Food Biotechnol. **7:**189-205.

Vanhaverbeke, C., C. Bosso, P. Colin-Morel, C. Gey, L. Gamar-Nourani, K. Blondeau, J. M. Simonet, and A. Heyraud. 1998. Structure of an extracellular polysaccharide produced by *Lactobacillus rhamnosus* strain C83. Carbohydr. Res. **314:**211-220.

van Marle, M. E. 1998. Ph.D.Thesis, University of Twente, Enschede, The Netherlands.

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:**1385-1386.

Vescovo, M., G. L. Scolari, and V. Bottazzi. 1989. Plasmid-encoded ropiness production in *Lactobacillus casei* ssp. *casei*. Biotechnol. Lett. **2:**709-712.

von Wright, A., and Tynkkynen. 1987. Construction of *Streptococcus lactis* subsp. *lactis* strains with a single plasmid associated with mucoid phenotype. Appl. Environ. Microbiol. **53:**1385-1386.

Wang, L., and P. R. 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.

Wang, L, D. Liu, and P. R. Reeves. 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J. Bacteriol. **178:**2598-2604.

Wehland, M., C. Kieker, D. L. Coplin, O. Kelm, W. Saenger, and F. Bernhard. 1999. Identification of an RcsA/RcsB recognition motif in the promoters of exopolysaccharide biosynthetic operons from *Erwinia amylovora* and *Panoea stewartii* subspecies *stewartii*. J. Biol. Chem. **274:**3300-3307.

Whitfield, C., and M. Valvano. 1993. Biosynthesis and expression of cell-surface polysaccharides in Gram-negative bacteria. Adv. Microbiol. Phys. **35:**135-146.

Whitfield, C. 1995. Biosynthesis of lipopolysaccharide O-antigens. Trends Microbiol. **3:**178-185.

Whitfield, C., P. A. Amor, and R. Köplin. 1997. Modulation of the surface architecture of Gram-negative bacteria by the action of surface polymer:lipid A-core ligase and by determinants of polymer chain length. Mol. Microbiol. **23:**629-638.

Whitfield, C., and I. S. Roberts. Structure, assembly and regulation of expression of capsules in *Escherichia coli*. 1999. Mol. Microbiol. **31:**1307-1319.

Xiang, S.-H., M. Hobbs, and P. R. Reeves. 1994. Molecular analysis of the *rfb* gene cluster of a group D2 *Salmonella enterica* strain: evidence for its origin from an insertion sequence-mediated recombination event between group E and D1 strains. J. Bacteriol. **176:**4357-4365.

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.

Yamamoto, Y., T. Nunome, R. Yamauchi, K. Kato, and Y. Sone. 1995. Structure of an exopolysaccharide of *Lactobacillus helveticus* TN-4, a spontaneous mutant strain of *Lactobacillus helveticus* TY1-2. Carbohydr. Res. **275:**319-332.

Molecular Characterisation of the Plasmid-encoded *eps* Gene Cluster Essential for Exopolysaccharide Biosynthesis in *Lactococcus lactis*

Chapter 2

Richard van Kranenburg, Joey D. Marugg, Iris I. van Swam, Norwin, J. Willem, and Willem M. de Vos

Published in Molecular Microbiology (1997) **24:**387-397

Abstract

Lactococcus lactis strain NIZO B40 produces an extracellular phosphopolysaccharide containing galactose, glucose, and rhamnose. A 40-kb plasmid encoding exopolysaccharide production was isolated through conjugal transfer of total plasmid DNA from strain NIZO B40 to the plasmid-free *L. lactis* model strain MG1614 and subsequent plasmid curing. A 12-kb region containing 14 genes with the order *epsRXABCDEFGHIJKL* was identified downstream of an iso-IS*982* element. The predicted gene products of *epsABCDEFGHIJK* show sequence homologies with gene products involved in exopolysaccharide, capsular polysaccharide, lipopolysaccharide, or teichoic acid biosynthesis of other bacteria. Transcriptional analysis of the *eps* gene cluster revealed that the gene cluster is transcribed as a single 12-kb mRNA. The transcription start site of the promoter was mapped upstream of the first gene *epsR*. The involvement of *epsD* in exopolysaccharide (EPS) biosynthesis was demonstrated through a single gene disruption rendering an exopolysaccharide-deficient phenotype. Heterologous expression of *epsD* in *Escherichia coli* showed that its gene product is a glucosyl transferase linking the first sugar of the repeating unit to the lipid carrier.

Introduction

Cell-surface polysaccharides are produced by a wide variety of bacteria. They can be attached to the cell membrane as the O-antigen of lipopolysaccharides (LPSs), form a capsule around the cell as capsular polysaccharides (CPSs), or be completely excreted as exopolysaccharides (EPSs). The biological functions of polysaccharides are diverse as they may play a role in pathogenesis and symbiosis, protect the cell from desiccation or other environmental stresses, or facilitate adherence of bacteria to solid surfaces (Leigh and Coplin, 1992, Whitfield and Valvano, 1993). The biosynthesis of polysaccharides that consist of repeating units shares common features. The repeating units are assembled at the membrane by sequential addition of sugar residues by specific glycosyltransferases from nucleotide sugars to a growing repeating unit that is coupled to an undecaprenylphosphate carrier. Subsequently, the repeating units have to be exported and polymerised to form the cell-surface polysaccharide (Sutherland, 1985, Whitfield and Valvano, 1993).

Recently, several gene clusters have been identified in Gram-positive and Gram-negative bacteria that are involved in the biosynthesis of LPSs (Jiang *et al.*, 1991, Brown *et al.*, 1992, Morona *et al.*, 1994, Allen and Maskel, 1996), CPSs (Rubens *et al.*, 1993, Guidolin *et al.*, 1994, Lin *et al.*, 1994, Arakawa *et al.*, 1995, Kolkman *et al.*, 1996), and EPSs (Glucksmann, *et al.*, 1993, Bugert and Geider, 1995, Huang and Schell, 1995, Stingele *et al.*, 1996). Some of these clusters contain genes involved in the biosynthesis of specific nucleotide sugars (Jiang *et al.*, 1991, Allen and Maskel, 1996). All contain genes for specific glycosyltransferases and genes that are involved in the process of polymerisation and export.

EPS production is a characteristic of *Lactococcus lactis* strains that are isolated from highly viscous Scandinavian fermented milk products. Various reports describe the involvement of specific plasmids in EPS production in *L. lactis* (Vedamuthu and Neville, 1986, von Wright and Tynkkynen, 1987, Neve *et al.*, 1988). However, none of these EPS plasmids has been further characterised and no evidence for the presence of *eps* genes on these plasmids was presented. In this study we describe the identification, characterisation, and transcriptional analysis of a novel 12-kb *eps* gene cluster located on a 40-kb plasmid, which is essential for EPS biosynthesis.

Materials and methods

Bacterial strains, plasmids, and media. The bacterial strains and plasmids used in this study are listed in Table 1. *L. lactis* subsp. *cremoris* NIZO B40 is a single-colony isolate from *L. lactis* subsp. *cremoris* L416 isolated from Scandinavian ropy milk (Macura and Townsley, 1984). *L. lactis* subsp. *cremoris* SBT 0495 was a gift from Snow Brand European Research Laboratories (The Netherlands). *Escherichia coli* was grown in Luria (L)-broth-based medium at 37°C (Sambrook *et al.*, 1989). *L. lactis* was grown at 30°C in M17 broth (Difco Laboratories) supplemented with 0.5% glucose (GM17) or 0.5% lactose (LM17) or in

a defined medium described by Poolman and Konings (1988). For screening of the lactose-fermenting phenotype *L. lactis* was plated on lactose-indicator agar plates (LIA), containing Elliker medium (Elliker, 1956) supplemented with 0.5% lactose and 0.004% bromocresol purple (Merck). If appropriate, the media contained chloramphenicol (10 μg/ml), erythromycin (10 μg/ml for *L. lactis* and 150 μg/ml for *E. coli*), rifampicin (50 μg/ml), streptomycin (100 μg/ml), or ampicillin (100 μg/ml).

Conjugation and plasmid curing. Conjugation was performed using filter matings. Cells were grown in GM17 broth until an $O.D._{600}$ of 0.4-0.8 was reached. Typically, 2 ml of donor cells and 1 ml of recipient cells were mixed, pelleted and the pellet was plated out on a 0.45 μm filter on a GM17 agar plate. After overnight incubation at 30°C, the cells were recovered from the filter and transconjugants were selected on GM17 agar or LIA plates containing the appropriate antibiotics. EPS production was tested by picking the colonies with a toothpick to reveal ropiness. For temperature-induced plasmid curing, cultures were incubated overnight at 40°C.

EPS purification and characterisation. For large-scale EPS purification *L. lactis* was grown in 1 l reconstituted milk for 24 h at 20°C, trichloroacetic acid was added to a final concentration of 12%, and bacterial cells and precipitated proteins were removed by centrifugation (20 min, 30,000 x *g*, 4°C). The supernatant was adjusted to neutral pH using 10 N NaOH, concentrated by ultrafiltration, dialysed against running tapwater (48 h), and lyophilised. The lyophilised EPS was dissolved in double-distilled water and the contaminating protein was removed by gel filtration on a Sephacryl S-500 (Pharmacia) column (75 x 2.6 cm) by elution with 50 mM NH_4HCO_3 at 0.75 ml/min, monitoring the refractive-index and the absorbance at 280 nm.

For small-scale EPS purification *L. lactis* was grown in 25 ml of defined medium containing 2% of glucose for 16 h at 30°C. Cells were removed by centrifugation (15 min., 6,000 x *g*) and the supernatant was dialysed against running tap water (48 h), and lyophilised. The contaminating protein was removed by gel filtration on a gel permeation HPLC column. The quantity of EPS was established by peak integration using dextran 500 (Serva) as a standard.

Sugar analysis was performed by HPLC analysis of the monosaccharide units after complete hydrolysis with 4 N HCl (van Riel and Olieman, 1991) and phosphor was determined as described by Chen *et al.* (1956).

^{1}H-nuclear magnetic resonance (NMR) and ^{13}C-NMR spectra were taken on a Bruker AM 400 spectrometer operating at 400.13 MHz for ^{1}H and 100.62 for ^{13}C. Experimental conditions were 15 mg/ml EPS, dissolved in D_2O at 70°C.

DNA isolation and manipulation. Isolation of *E. coli* plasmid DNA and standard recombinant DNA techniques were performed as described by Sambrook *et al.* (1989). Large scale isolation of *E. coli* plasmid DNA for nucleotide sequence analysis was performed with Qiagen columns, following the instructions of the manufacturer (Qiagen Inc.). Isolation and transformation of *L. lactis* plasmid DNA were performed as previously described (De Vos *et al.*, 1989).

Table 1. Bacterial strains and plasmids used in this study.

Strain/plasmid	Relevant characteristics[a]	Reference
Strain		
E. coli		
MC1061		Casadaban (1980)
DH5α		Hanahan (1983)
L. lactis		
NIZO B40	Lac^+ Eps^+ multiplasmid strain harboring pNZ4000	This work
SBT 0495	Lac^+ Eps^+	Nakajima *et al.* (1990)
MG1363	Plasmid-free	Gasson (1983)
MG1614	Rif^r Str^r, plasmid-free	Gasson (1983)
NZ4001	Rif^r Str^r Lac^+ Eps^+ derivative of MG1614, obtained by conjugation of NIZO B40 and MG1614	This work
NZ4010	Rrf^r Str^r Lac^- Eps^+ derivative of NZ4001, obtained by plasmid curing	This work
Plasmids		
pNZ4000	40-kb plasmid encoding EPS production	This work
pUC19Ery	Ery^r, 3.8 kb pUC19 containing 1.1 kb *Hin*PI fragment of pIL253 carrying the Ery^R gene	NIZO collection
pUC18Ery	Ery^r, 3.8 kb pUC18 containing 1.1 kb *Hin*PI fragment of pIL253 carrying the Ery^R gene	This work
pNZ4030	Ery^r Eps^+, 27 kb derivative of pNZ4000 carrying the Ery^R gene from pIL253	This work
pNZ4035	Ery^r, 4.2 kb derivative of pUC18Ery containing a 0.4 kb *Sst*I fragment of *epsD* from pNZ4000	This work
pNZ4036	Ery^r, 4.5 kb derivative of pUC18Ery containing a 0.8 kb fragment of pNZ4000	This work
pNZ4040	Cm^r, 5.3 kb derivative of pNZ273 containing a 0.7 kb *Sty*I-*Spe*I fragment carrying the promoter of the *eps* operon	This work
pNZ4050	Amp^r, 4.1 kb derivative of pUC18 containing a 1.4 kb *Sca*I-*Bgl*II fragment carrying *epsD* from pNZ4000	This work

[a] Lac^+, lactose-fermenting phenotype; Eps^+, exopolysacharide-producing phenotype; Rif^r, rifampicin resistant; Str^r, streptomycin resistant; Ery^r, erythromycin resistant; Cm^r, chloramphenicol resistant; Amp^r, Ampicillin resistant.

Nucleotide sequence analysis. Automatic double-stranded DNA sequence analysis was performed on both strands with an ALF DNA sequencer (Pharmacia Biotech). Sequencing reactions were accomplished using the AutoRead sequencing kit, initiated by using fluorescein-labelled universal and reverse primers and continued with synthetic primers in combination with fluorescein-15-dATP following the instructions of the manufacturer (Pharmacia Biotech). Sequence data were assembled and analysed using the PC/GENE program version 6.70 (IntelliGenetics). Hydrophobic stretches within proteins were predicted by the method of Klein *et al.* (1985). The GenBank GB93.0 bacteria library was screened for homologies using TFASTA.

Northern blot and primer extension analysis. Total RNA was isolated from exponentially growing MG1614, NIZO B40, and NZ4010 cultures by the Macaloid method described by Kuipers *et al.* (1993). For Northern blot analysis, RNA was glyoxylated, separated on a 1% agarose gel, blotted and hybridised as described previously (Van Rooijen and De Vos, 1990). The hybridisation probes were radiolabelled with [α-^{32}P]dATP by nick translation.

Primer extension was performed by annealing 20 ng of oligonucleotides to 100 µg of RNA as described (Kuipers *et al.*, 1993). A synthetic 18-mer oligonucleotide, 5'-TTTATTAACTTCTGTAAG-3', complementary to the 5' sequence of *epsR*, was used as a primer.

Plasmid constructions. For plasmid integration plasmid pUC18Ery was constructed from plasmid pUC19Ery. Plasmid pUC19Ery contains a Klenow-treated 1.1-kb *Hin*PI fragment of pIL253 (Simon and Chopin, 1988), carrying the erythromycin resistance (Ery^R) gene, cloned into the *Hin*cII site of pUC19 (Yanisch and Perron, 1985). pUC18Ery was constructed by cloning a 1.1-kb *Acc*I-*Pst*I fragment from pUC19Ery into pUC18 (Yanisch and Perron, 1985), digested with *Acc*I-*Pst*I. For the single gene disruption of *epsD*, plasmid pNZ4035 was constructed by cloning a 415 bp *Sst*I internal gene fragment of *epsD* in the *Sst*I site of pUC18Ery downstream of, and in the same orientation as the Ery^R gene. For demonstration of readthrough from the Ery^R gene, plasmid pNZ4036 was constructed by cloning a polymerase chain reaction (PCR)-generated fragment containing the 3' part of *epsB* and the 5' part of *epsC* (nucleotides 4120 to 4877) in pUC18Ery downstream of, and in the same orientation as the Ery^R gene. The promoter-probe fusion was constructed using pNZ273 (Platteeuw *et al.*, 1994). Plasmid pNZ4040 is pNZ273 digested with *Bgl*II (blunt), containing a 0.7-kb *Sty*I-*Spe*I fragment (blunt). For heterologous expression of *epsD* a 1.4-kb *Sca*I-*Bgl*II fragment was cloned in the *Hin*cII site of pUC18 resulting in plasmid pNZ4050. Apart from plasmid pNZ4050 which was constructed in *E. coli* DH5α, all plasmids were constructed in *E. coli* MC1061.

Glycosyltransferase activity assays and thin-layer chromatography (TLC) analysis. Glycosyltransferase activity assays were performed with *E. coli* cells permeabilised by repeated freeze-thaw cycles. Mid-log cultures (50 ml) were induced with 0.5 mM IPTG and grown for another 2 h at 37°C. Cells were harvested, washed with 10 ml of 20 mM Tris-HCl, pH 8.0, and dissolved in 200 µl of a buffer containing 50 mM Tris-HCl, pH8.0, and 5 mM EDTA. The cell suspension was repeatedly (five times) frozen at -80°C and subsequently thawed to obtain permeabilised cells. The glycosyl transferase activity assays were performed as described by Kolkman *et al.* (1996). The assay mixture contained 25 µl of permeabilised cells (approximately 250 µg of protein), 5 µl of 100 mM $MgCl_2$, 19 µl of a buffer containing 50 mM Tris-HCl, pH8.0, and 1 mM EDTA, and 1 µl (approximately 55,000 cpm) UDP-[^{14}C]glucose or UDP-[^{14}C]galactose. After a 1.5 h incubation at 15°C the reaction was stopped by the addition of 1 ml of chloroform:methanol (2:1). The solution was extracted three times with 200 µl of a solution containing 1.5 ml chloroform, 25 ml methanol, 13.5 ml water, and 275 mM KCl, by removal of the upper phase and interphase. The remaining lipid fraction was vacuum dried. For mild acid hydrolysis, one part was

resuspended in 100 µl of *n*-butanol and 100 µl 0.05 M trifluoroacetic acid (TFA), and incubated for 20 min at 95°C. For complete acid hydrolysis, an equally sized part of the lipid fraction was resuspended in 15 µl H_2O and 85 µl TFA, and incubated for 2.5 h at 95°C. After hydrolysis, the samples were vacuum dried, resuspended in 10 µl 40% isopropanol, and 8 µl was subjected to TLC on HPTLC silica gel (Merck), with 1-butanol-ethanol-water (5:3:2). Glucose and galactose (40 µg in 40% isopropanol) were used as standards. The gel was dried, sprayed with En^{3}Hance spray (DuPont), and exposed for 2-3 d. The sugar standards were visualised by spraying with 5% H_2SO_4in ethanol and heating at 100°C for 15 min.

Results

Isolation of plasmid pNZ4000. The EPS-producing (Eps^+) *L. lactis* strain NIZO B40 is capable of fermenting lactose (Lac^+) and contains at least seven endogenous plasmids (Fig. 1). To test whether EPS production is linked to a transferrable plasmid, conjugal matings were performed between NIZO B40 and the non-mucoid plasmid-free strain MG1614 (Gasson, 1983). As selection for the ropy phenotype is very laborious, lactose fermentation was used for the initial selection. The conjugation experiments showed that this property could be transferred at a low frequency of approximately 10^{-8} to 10^{-9} per donor. Co-transfer of the EPS production was observed to a variable degree, ranging from approximately 10% to unity. Plasmid DNA analysis of Lac^+ Eps^+ transconjugants revealed that various plasmids were co-transferred with the lactose plasmid. One strain with the smallest number of plasmids was used for further studies (Fig. 1). Temperature-induced plasmid curing of this strain, NZ4001, followed by analysis of 143 of the survivors for their capacity to ferment lactose and produce EPS, showed complete segregation of the possible phenotypes, and 99 Lac^+ Eps^+, two Lac^- Eps^+, nine Lac^+ Eps^- and 33 Lac^- Eps^- colonies were obtained. These results indicate that EPS production is encoded by a plasmid which differs from that encoding lactose fermentation. Plasmid analysis showed that both Lac^- Eps^+ strains contained a single 40-kb plasmid species. One of these, strain NZ4010 and its 40-kb plasmid, designated pNZ4000, was subjected to further studies (Fig. 1).

Characterisation of EPS produced by strain NIZO B40. Strain NIZO B40 produced approximately 50 mg/l EPS when cultivated on defined medium. The purified EPS produced by strain NIZO B40 consisted of the monosaccharides galactose, glucose, and rhamnose in the ratio 1.0:1.7:0.8. The molar ratio of carbohydrate and phosphorus was 4.7:1. This resembles the composition of the extracellular phosphopolysaccharide produced by *L. lactis* strain SBT 0495, the structure of which has been elucidated (Nakajima *et al.*, 1990, Nakajima *et al.*, 1992). Comparison of ^{1}H-NMR and ^{13}C-NMR spectra of purified EPS from strains NIZO B40, NZ4010 and SBT 0495 revealed identical spectra (results not shown). These results strongly suggest that the structure of NIZO B40 EPS is identical to that of strain SBT 0495 with a repeating unit consisting of →4)-β-D-Glc*p*-(1→4)[α-L-Rha*p*-(1→2)][α-D-Gal*p*-1-PO_4-3]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→ (Nakajima *et al.*, 1992). As the NMR spectra of

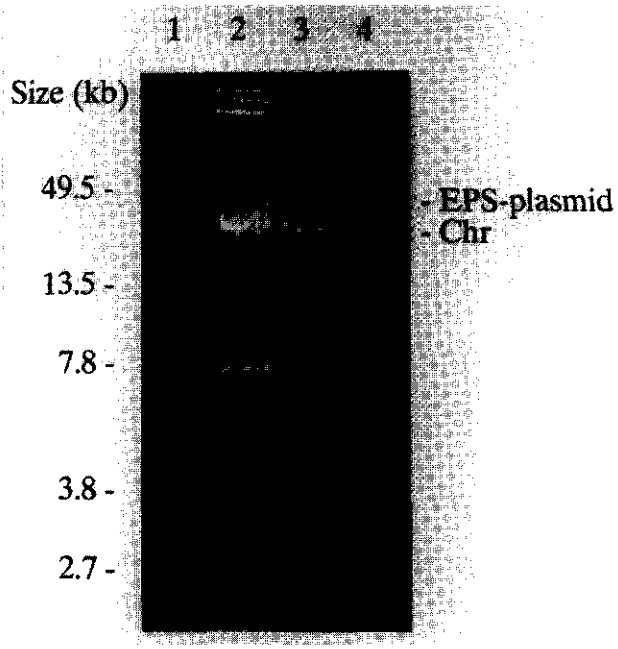


Fig. 1. Identification of EPS plasmid pNZ4000 using a 0.7% agarose gel. Lanes: 1, plasmid markers from strain NCDO 4109 (Gasson, 1983); 2, parent strain NIZO B40 (Lac^+ Eps^+); 3, transconjugant NZ4001 (Lac^+ Eps^+); 4, strain NZ4010, carrying only pNZ4000 (Lac^- Eps^+). The EPS plasmid pNZ4000 and chromosomal DNA (Chr) are indicated on the right. The size of the plasmid markers are indicated on the left.

the EPS from NIZO B40 and NZ4010 were also identical, we conclude that all essential information for the biosynthesis of this specific EPS must be encoded on plasmid pNZ4000.

Localisation and nucleotide sequence analysis of the *eps* genes. Plasmid pNZ4000 contains unique restriction sites for *Xho*I and *Sph*I and double digestion with these enzymes resulted in a 26-kb and a 14-kb fragment. Both fragments were genetically labelled with an Ery^R gene by ligation with a 1.1-kb *Sal*I-*Sph*I fragment derived from pUC18Ery. As both fragments contain at least one functional replicon (van Kranenburg, unpublished results), the resulting plasmids could be stably maintained in *L. lactis*. When the plasmids were transformed to the plasmid-free strain MG1363 (Gasson, 1983), the Ery^R transformants harbouring the largest plasmid (pNZ4030) were Eps^+, and those harbouring the smallest plasmid (pNZ4031) were Eps^-, indicating that the genes involved in EPS biosynthesis are located on the 26-kb *Xho*I-*Sph*I fragment. The complete nucleotide sequence of plasmid pNZ4030 was determined and revealed a 12-kb region downstream of a putative insertion element which contained the *eps* gene cluster (see below) of 14 putative genes. The iso-IS*982* element upstream of the *eps* gene cluster is flanked by a 17 bp perfect inverted repeat and encodes a putative transposase that shares 96.6% identity with the transposase of the lactococcal insertion sequence IS*982* (Yu *et al.*, 1995). Immediately downstream of the *eps* gene cluster a putative gene, *orfY*, was found in the opposite orientation. The N-terminal part

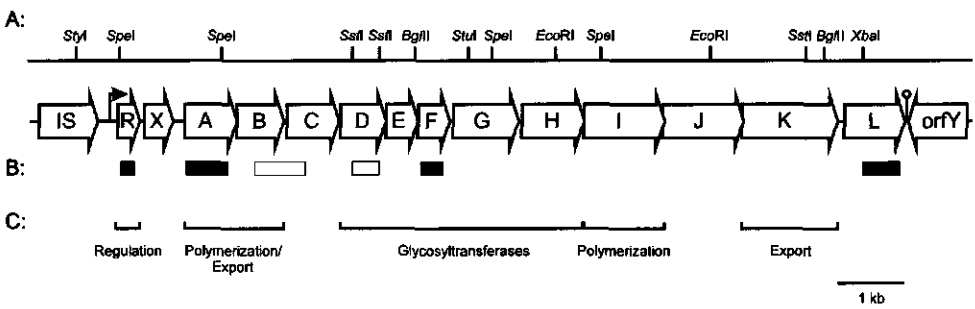


Fig. 2. Genetic organisation of the *eps* gene cluster. (A) Physical map and genetic organisation of the *eps* genes. The complete nucleotide sequence is available in the GenBank under accession number U93364. (B) Fragments of the *eps* genes used for probing of the Northern blot (filled boxes) or integration (open boxes). (C) Putative functions of the *eps* gene products.

of its encoded protein of 300 amino acids shows 95.7% identity with OrfB from *Streptococcus thermophilus* CNRZ368. OrfB is the deduced 208 amino acids product of a partly sequenced gene of unknown function that is located adjacent to an iso-IS*981* element on the chromosome of *S. thermophilus* (Guedon *et al.*, 1995). Furthermore, OrfY shares 24% identity with LytR from *Bacillus subtilis*, an attenuator of the expression of the *lytABC* and *lytR* operons (Lazarevic *et al.*, 1992).

The order of the 14 *eps* genes is shown in Fig. 2 and additional data are listed in Table 2. The G+C content of the gene cluster is 28% which is below the typical G+C content of 38% to 40% reported for *L. lactis* (Holt *et al.*, 1994). This is reflected in the preferential usage of AT-rich codons in the *eps* gene cluster (data not shown). The vast majority of the predicted gene products of the *eps* gene cluster are homologous to proteins involved in polysaccharide biosynthesis (Table 3). However, EpsR does not show significant homology to any protein involved in polysaccharide biosynthesis. In contrast, the N-terminal part of EpsR shows moderate similarity with DNA-binding or regulator proteins: 30% identity in 65 amino acid residues with the N-terminal part of the *B. subtilis* PBSX repressor Xre (Wood *et al.*, 1990), 24% identity in 75 amino acid residues with the N-terminal part of the *Pseudomonas aeruginosa* negative regulator of the pyocin genes PrtR (Matsui *et al.*, 1993), and 22% identity in 75 amino acid residues with the *Erwinia carotovora* DNA-binding protein RdgA (Liu *et al.*, GenBank L32173). Moreover, EpsX and EpsL do not show any significant homology with known proteins and their role in EPS biosynthesis remains to be established.

Transcriptional analysis of the *eps* gene cluster. RNA isolated from strain MG1614, NIZO B40, and NZ4010 was used for Northern blot analysis. Internal gene fragments of *epsR*, *epsA*, *epsF*, and *epsL* were generated by polymerase chain reaction (PCR), labelled and used as DNA probes (Fig. 2B). The size of the transcripts was determined relative to RNA molecular-weight markers. All probes hybridised with a transcript of approximately 12 kb for NIZO B40 and NZ4010 RNA, while there was no reaction with MG1614 RNA (Fig. 3). Smaller transcripts were not detected. These results indicate that the *eps* genes are organised in an operon that is transcribed as a single 12-kb polycistronic mRNA from a promoter

Table 2. Gene positions and predicted properties of the hypothetical proteins encoded by the *eps* gene cluster.

Gene	Putative ribosome binding site[a]	Start and stop at nucleotide	Amino acids of protein	Size of protein (kDa)	pI of protein
epsR	AAAGG	2078-2392	105	12.2	4.96
epsX	GGAGG	2484-2900	139	16.2	9.01
epsA	GGAG	3070-3846	259	28.3	10.09
epsB	AGGAG	3859-4551	231	25.2	6.36
epsC	AGGAG	4609-5370	254	28.4	5.58
epsD	GGAG	5395-6072	226	26.0	9.85
epsE	AAG	6085-6552	156	18.2	9.95
epsF	GGA	6555-7034	160	18.7	9.86
epsG	AGAAAGGA	7092-8039	316	37.5	6.14
epsH	GAAAG	8103-9029	309	36.8	7.57
epsI	AGGA	9032-10219	396	45.2	9.89
EpsJ	AGAAAGGA	10209-11396	396	47.5	6.70
EpsK	GGAG	11386-12801	471	54.0	10.24
EpsL	AAGGA	12902-13798	299	32.9	8.94

[a] Sequence of the 3'-end of the lactococcal 16S rRNA is 3'-UCUUUCCUCC (Chiaruttini and Millet, 1993).

upstream of *epsR*. A putative terminator sequence was detected immediately downstream of *epsL*.

The promoter of the *eps* operon was located on the 0.7-kb *Sty*I-*Spe*I fragment upstream of *epsR* (Fig. 2A) by fusing it to the promoterless *gusA* reporter gene encoding the enzyme β-glucuronidase in pNZ273 (Platteeuw *et al.*, 1994). The resulting plasmid (pNZ4040) was transformed to *L. lactis* MG1363 resulting in a high specific β-glucuronidase activity of 145 nmol/min.mg protein.

The transcription start site of the *eps* operon was determined by primer extension analysis using an antisense oligonucleotide derived from the 5' sequence of the *epsR* gene as a primer, in combination with RNA isolated from strain NIZO B40, or NZ4010. Two transcription start sites were detected at a distance of 2 bp from each other. Upstream of these transcription starts, consensus sequences for the -35 and –10 region could be located (Fig. 4).

The gene *epsD* is essential for EPS biosynthesis. The involvement of the *eps* gene cluster in EPS biosynthesis was studied by the construction of a single gene disruption in selected *eps* genes using the non-replicating integration vector pUC18Ery. When a fragment from the *eps* gene cluster is cloned in this vector and transformed to strain NZ4010 harbouring pNZ4000 (Table 1), Ery^R transformants are only obtained when the pUC18Ery derivative has integrated in pNZ4000 via a single cross-over event between the cloned *eps* gene fragment and the homologous region on pNZ4000. As polar effects could be expected with this approach, the integration plasmid was constructed in such a way that these effects were excluded. This was achieved by driving the expression of the downstream genes by the promoter of the Ery^R gene, the sequence of which does not show a terminator (Martin *et al.*, 1987). Readthrough from the Ery^R gene was illustrated by transformation of strain NZ4010

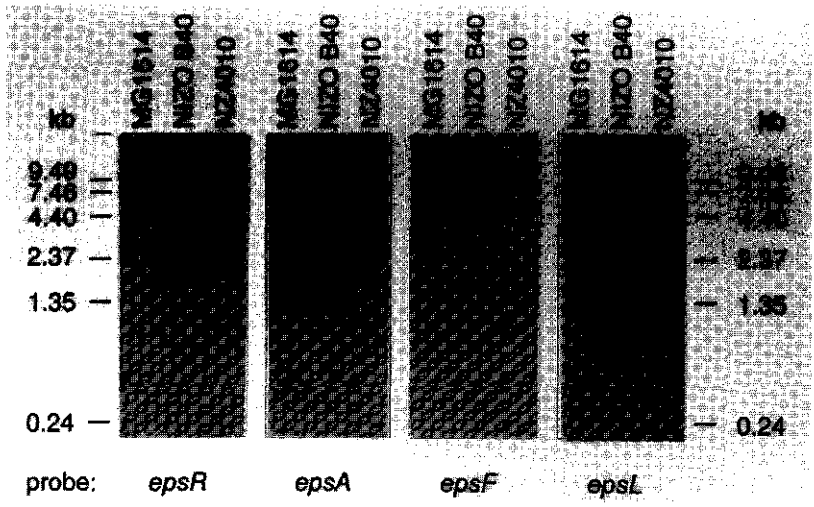


Fig. 3. Northern blot hybridisation of RNA isolated from MG1614, NIZO B40, and NZ4010 probed with internal gene fragments of the *eps* genes *epsR*, *epsA*, *epsF*, or *epsL*.

with plasmid pNZ4036, containing the 3' part of *epsB* and the 5' part of *epsC* (Fig. 2B). The expected single-copy integration in pNZ4000 and the duplication of the cloned gene fragments were verified by PCR and Southern blotting (data not shown), and resulted in ropy Ery^R colonies. This indicated that although transcription of *epsRXAB* remains under control of the promoter upstream of *epsR*, transcription of the downstream genes *epsCDEFGHIJKL* is driven by the promoter of the Ery^R gene.

To show the involvement of *epsD* in EPS biosynthesis, pNZ4035 was constructed, containing an internal fragment of *epsD* (Fig. 2B). Integration of pNZ4035 in pNZ4000 resulted in a non-ropy Ery^R strain. Moreover, no EPS could be detected in cultures of this strain. Thus, the disruption of *epsD* abolished EPS production demonstrating that EpsD is essential for EPS biosynthesis.

The proteins found to be homologous to EpsD (Table 3) are involved in linking the first sugar of the repeating unit to the lipid carrier (Rubens *et al.*, 1993, Kolkman *et al.*, 1996, Reuber and Walker, 1993, Jiang *et al.*, 1991, Wang and Reeves, 1994). Therefore, it may be expected that EpsD will also act as a glycosyltransferase linking either glucose or galactose to the lipid carrier. This was studied by the cloning of *epsD* under control of the *lac* promoter of pUC18 in *E. coli* DH5α resulting in pNZ4050. Permeabilised cells of IPTG-induced *E. coli* DH5α harbouring pNZ4050 were incubated with either UDP-[^{14}C]glucose or UDP-[^{14}C]galactose. *E. coli* DH5α was used as a negative control. The lipid fraction was extracted and subjected to complete acid hydrolysis to determine the sugar specificity of EpsD. The labelled sugars were identified by TLC and autoradiography (Fig. 5). Incubation with UDP-[^{14}C]glucose resulted in a labelled product with the same mobility as glucose. Incubation with UDP-[^{14}C]galactose also resulted in a product with the same mobility as glucose, but

Table 3. Homologies (percentage identity) of the predicted proteins of the *eps* gene cluster from *L. lactis* with those involved in EPS, CPS, LPS, or teichoic acid biosynthesis in other bacteria.

	eps gene product										
	EpsA	EpsB	EpsC	EpsD	EpsE	EpsF	EpsG	EpsH	EpsI	EpsJ	EpsK
S. aureus[a]	CapA (35%)	CapB (42%)	CapC (41%)								CapF (22%)
S. pneumoniae[2]	CpsC (24%)	CpsD (35%)	CpsB (28%)	CpsE[b] (43%)	CpsF (38%)	CpsG (35%)					
S. agalactiae[3]	CpsB (29%)	CpsC (31%)	CpsA (25%)	CpsD[b] (45%)							
S. thermophilus[4]	EpsC (27%)	EpsD (35%)	EpsB (28%)	EpsE (43%)			EpsI (36%)		EpsJ (23%)		EpsM (21%)
B. subtilis[5]							GgaB[a] (32%)			GgaB[b] (31%) RodC (22%)	
R. meliloti[6]	ExoP[a] (21%)	ExoP[b] (30%)		ExoY (34%)							ExoT (22%)
E. amylovora[7]	AmsA[a] (20%)	AmsA[b] (34%)		AmsG[b] (39%)			AmsE (21%)				
K. pneumoniae[8]	Orf6[a] (24%)	Orf6[b] (31%)		Orf14[b] (34%)							Orf11 (22%)
S. enterica[9]				RfbP[b] (41%)					Rfc (21%)		
S. flexneri[10]									Rfc (25%)		RfbX (23%)
E. coli[11]									Rfc (21%)		RfbX (24%)
S. dysenteriae[12]							RfpA (25%)		Rfc (20%)		RfbX (23%)
Y. enterocolitica[13]							Orf10.9 (%) TrsB (34%)	Orf11.8 (24%)			RfbX (23%) TrsA (25%)
B. pertussis[14]				BplG (34%)							
P. solanacearum[15]	EpsB[a] (24%)	EpsB[b] (34%)									
X. campestris[16]				GumD[b] (36%)							
H. influenzae[17]				RfbP[b] (43%)			LgtD (33%)			Orf3[b] (20%)	
Sphingomonas S88[18]		SpsE (25%)		SpsB[b] (34%)	SpsK[a] (28%)	SpsK[b] (24%)					

[a] Homology in N-terminal part of protein.

[b] Homology in C-terminal part of protein.

GenBank accession numbers: 1. U10927. 2. U09239, X85787. 3. L09116. 4. U40830. 5. U13979, X15200. 6. L20758, X16704, Z22636, Z22646. 7. X77921. 8. D21242. 9. U35434, X61917. 10. X71970. 11. U09876. 12. S73325. 13. U46859, Z47767. 14. X90711. 15. U17898. 16. U22511. 17. U32714, U32715, X78559. 18. U51197.

not shown). Incubation of the permeabilised cells with non-labelled UDP-glucose and labelled UDP-galactose reduced the intensity of the product significantly (results not shown). These results indicate that EpsD acts as a glucosyltransferase. Furthermore, by an epimerase activity of *E. coli*, UDP-galactose is turned into UDP-glucose which can be used by EpsD. The lipid fraction was also subjected to mild acid hydrolysis, to test if the linked sugar is a monosaccharide. Mild acid hydrolysis will only disrupt the linkage between the lipid carrier and the saccharides. This treatment resulted in the same product as complete acid hydrolysis, which showed that EpsD is involved in the first step of EPS biosynthesis, linking one glucose moiety to the lipid carrier.

Discussion

We have identified a 40-kb plasmid in *L. lactis* strain NIZO B40 that is involved in the production of EPS. Physical and spectroscopic analysis suggested that EPS produced by strain NIZO B40 is identical to EPS produced by *L. lactis* strain SBT 0495 with a known structure (Nakajima *et al.*, 1992). We have located a gene cluster consisting of 14 genes with the order *epsRXABCDEFGHIJKL*. EPS production could be abolished by a single gene disruption of *epsD*, providing evidence for the involvement of the gene cluster in EPS biosynthesis.

The transcription of the *eps* gene cluster is driven by a promoter upstream of *epsR*. All genes are transcribed as a single polycistronic mRNA of approximately 12 kb. Similarly, large transcripts encompassing all biosynthetic genes for polysaccharide production were also reported for the *ams* gene cluster from *Erwinia amylovora* and the *cps* gene cluster from *Klebsiella pneumoniae* K2 (Arakawa *et al.*, 1995, Bugert and Geider, 1995), and postulated for the *eps* gene cluster of *Pseudomonas solanacearum* (Huang and Schell, 1995). This suggests a co-ordinate control of the genes involved in the biosynthesis of these polysaccharides.

Putative functions for most of the *eps* gene products could be assigned through homologies: regulation (EpsR), polymerisation and export (EpsA, EpsB, EpsI, and EpsK) and biosynthesis of the oligosaccharide repeating unit (EpsD, EpsE, EpsF, EpsG, EpsH) (Fig. 2C).

EpsR is homologous to the regulator proteins Xre, PrtR, and RdgA, that all contain a DNA-binding domain. Similarly, EpsR may be involved in regulation, although we currently have no indication as to the nature of the regulation of EPS production. The first gene of the *S. thermophilus eps* gene cluster has also been implicated in the regulation of EPS production (Stingele *et al.*, 1996). Interestingly, the gene product of the *orfY* gene downstream of the *eps* gene cluster, like EpsA from *S. thermophilus*, is homologous to the regulator protein LytR from *B. subtilis* (Stingele *et al.*, 1996, Lazarevic *et al.*, 1992). However, a function for *L. lactis* OrfY in EPS biosynthesis, if any, remains to be established.

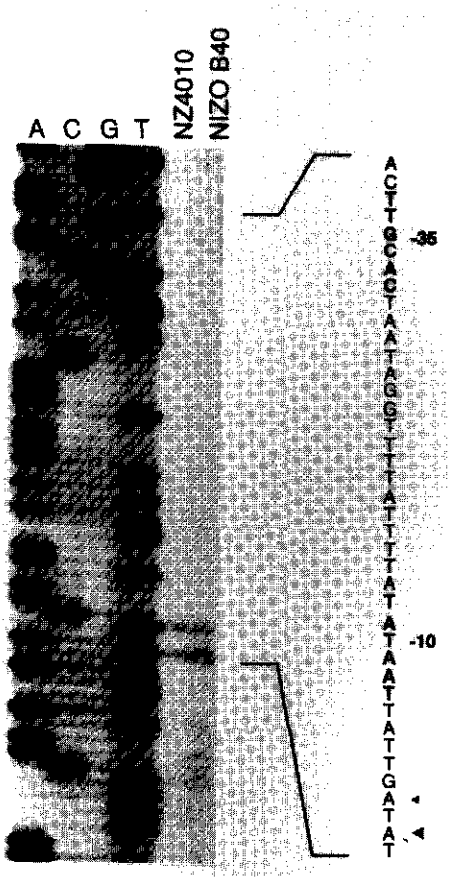


Fig. 4. Determination of the transcription start site of the *eps* gene cluster. RNA from wild-type strain NIZO B40 and conjugant NZ4010.

EpsA and EpsB are homologous to a family of ExoP-like proteins (Table 3). They consist of a transmembrane domain containing two putative transmembrane helices, and a domain containing an nucleotide-binding motif. Both domains can be combined in one protein (ExoP, AmsA, Orf6, and EpsB), or divided over two proteins (CapA and CapB, CpsC and CpsD, CpsB and CpsC, and EpsC and EpsD). ExoP is supposed to be involved in chain-length determination (Becker *et al.*, 1995). A consensus sequence motif of proteins involved in chain-length determination $SPKX_{11}GX_3G$ (Becker *et al.*, 1995) is present in *L. lactis* EpsA (residues 170 to 188). In *L. lactis* EpsB an A-site, DEGKTT (residues 55 to 60), and a B-site, VVLID (residues 157 to 161), of a nucleotide binding motif could be identified (Fath and Kolter, 1993).

The predicted gene products of *epsD*, *epsE*, *epsF*, *epsG*, and *epsH* from *L. lactis* are homologous to glycosyltransferases. CpsE from *Streptococcus pneumoniae*, ExoY from *Rhizobium meliloti*, and RfbP from *Salmonella enterica* are involved in the linkage of the

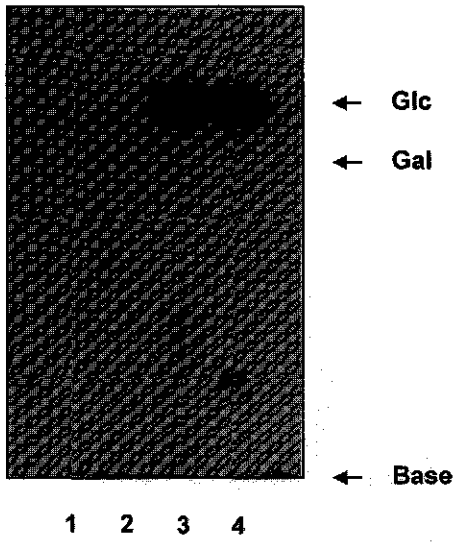


Fig. 5. TLC of ^{14}C-labelled intermediates isolated from the lipid fraction of permeabilised *E. coli* cells. The positions of the standard sugars glucose (Glc) and galactose (Gal) are indicated. Lanes: 1, *E. coli* - complete acid hydrolysis; 2, *E. coli* - mild acid hydrolysis; 3, *E. coli* carrying pNZ4050 - complete acid hydrolysis; 4, *E. coli* carrying pNZ4050 - mild acid hydrolysis.

first sugar of the repeating unit to the undecaprenylphosphate carrier (Rubens *et al.*, 1993, Kolkman *et al.*, 1996, Reuber and Walker, 1993, Jiang *et al.*, 1991, Wang and Reeves, 1994). For RfbP (WbaP), it has been demonstrated that the glycosyltransferase activity is encoded in the C-terminal half of the protein and that the N-terminal half is involved in a later step in biosynthesis (Wang *et al.*, 1996). EpsD from *L. lactis*, like ExoY, EpsE, and BplG, is homologous to the C-terminal part of RfbP and probably contains only glycosyltransferase activity. Disruption of *epsD* should lead to inability to start biosynthesis of the repeating unit, which is in agreement with the Eps$^-$ phenotype of the *epsD* mutant we constructed. The sugar specificity of EpsD was determined by expression of *epsD* in *E. coli*. The backbone of the EPS produced by NIZO B40 consists of one galactose and two glucose moieties. Incubation of the permeabilised *E. coli* cells with UDP-[^{14}C]glucose and UDP-[^{14}C]galactose and subsequent TLC analysis of the hydrolysed lipid fractions showed that UDP-glucose is the substrate for EpsD and that EpsD is involved in the first step of EPS biosynthesis.

For CPS biosynthesis in *S. pneumoniae*, transfer of the second sugar of the repeating unit to the first lipid-linked sugar is performed by the gene products of *cps14F* and *cps14G* (Kolkman *et al.*, 1997). Similar to Cps14F and Cps14G, the *L. lactis* proteins EpsE and EpsF may act as one glycosyltransferase performing the same reaction. A combined function for EpsE and EpsF is supported by their homology to the N-terminal and the C-terminal half of

SpsK, respectively, which is a putative glycosyltransferase from *Sphingomonas* S88 (Yamazaki *et al.*, 1996).

Two glycosyltransferases involved in later steps of biosynthesis of the repeating unit are expected to be encoded by *epsG* and *epsH*, according to their homology with various putative glycosyltransferases (Table 3). The biosynthesis of the repeating unit of NIZO B40 EPS presumably requires five different glycosyltransferase activities. It is postulated that EpsD, EpsE and EpsF are involved in the first two steps, linking the first two sugars to the undecaprenylphosphate carrier, and that EpsG and EpsH act as glycosyltransferases, both linking a sugar to the lipid-linked precursor. A fifth glycosyltransferase still has to be assigned. This could be EpsX, EpsC, EpsJ, or EpsL, since they either share no homology to other proteins (EpsX and EpsL) or their homologous counterparts have unknown functions (EpsC and EpsJ).

EpsI may be the polysaccharide polymerase. The homologous Rfc protein is known to be the polymerase of the O-antigen of *Salmonella* strains (Brown *et al.*, 1992, Morona *et al.*, 1994). Rfc has 11 to 13 predicted transmembrane segments (Morona *et al.*, 1994) and *L. lactis* EpsI is also predicted to be a transmembrane protein with nine predicted membrane-spanning regions (data not shown).

The gene product of *L. lactis epsK* shares moderate homology with RfbX and ExoT; RfbX is involved in export of the O-antigen in *Shigella flexneri* (Macpherson *et al.*, 1995) and *S. enterica* (Liu *et al.*, 1996). In *rfbX* (*wzx*) mutants of *S. enterica* lipid-linked O-units accumulate at the cytoplasmic side of the cytoplasmic membrane, suggesting that RfbX is involved in translocation of the O-unit across the membrane as a flippase (Liu *et al.*, 1996). EpsK from *L. lactis* has a similar hydrophobicity plot as RfbX predicting 12 membrane-spanning domains (data not shown). ExoT from *R. meliloti* is also proposed to be involved in export; mutants are still able to produce complete lipid-linked repeating units, but do not produce succinoglycan (Reuber and Walker, 1993).

The organisation of the *eps* gene cluster from *L. lactis* is similar to that of the gene clusters encoding EPS biosynthesis in *S. thermophilus* and CPS biosynthesis in *Staphylococcus aureus*, *S. pneumoniae*, and *Streptococcus agalactiae* (Stingele *et al.*, 1996, Lin *et al.*, 1994, Guidolin *et al.*, 1994, Kolkman *et al.*, 1996, Rubens *et al.*, 1993). The conservation between the *eps* gene cluster of *L. lactis* and the *cps* gene cluster from *S. pneumoniae* is striking. This conservation may even be extended, because the sequence downstream of the streptococcal *cpsH* remains to be determined. All gene clusters have an operon structure with a high coding density. The genes are located in one orientation and are probably transcribed as a single mRNA. The sequence of the functions of the genes in these Gram-positive cell-surface polysaccharide biosynthesis clusters seems to be: regulation, chain-length determination, biosynthesis of the oligosaccharide of the repeating unit, polymerisation, export (Stingele *et al.*, 1996). In our current studies we are performing functional analysis of the *eps* gene products, initially focussing on the specificity of the glycosyltransferases.

Acknowledgements

We thank Harry Rollema for recording the NMR spectra, Fedde Kingma for the phosphor analysis, Jan van Riel for determining the sugar composition, Ellen Looijesteijn for her advice on growth of ropy *L. lactis* strains. We are grateful to Marc Kolkman for sharing a copy of his submitted manuscript. Finally, we want to thank Michiel Kleerebezem, Roland Siezen, and Mark Smith for critically reading the manuscript. This work was supported by the E.C. research grant 1116/92 1.6.

References

Allen, A. and D. Maskell. 1996. The identification, cloning and mutagenesis of a genetic locus required for lipopolysaccharide biosynthesis in *Bordetella pertussis.* Mol. Microbiol. **19:**37-52.

Arakawa, Y., R. Wacharotayankun, T. Nagatsuka, H. Ito, N. Kato, and M. Ohta. 1995. Genomic organisation of the *Klebsiella pneumoniae cps* region responsible for serotype K2 capsular polysaccharide synthesis in the virulent strain Chedid. J. Bacteriol. **177:**1788-1796.

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.

Bugert, P., and K. Geider. 1995. Molecular analysis of the *ams* operon required for exopolysaccharide synthesis of *Erwinia amylovora.* Mol. Microbiol. **15:**917-933.

Brown, P. K., L. K. Romana, and P. R. Reeves. 1992. Molecular analysis of the *rfb* gene cluster of *Salmonella* serovar muenchen (strain M76): genetic basis of the polymorphism between groups C2 and B. Mol. Microbiol. **6:**1385-1394.

Casadaban, M. J., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli.* J. Mol. Biol. **138:**179-207.

Chen, P.S., T. Y. Toribara, and H. Warner. 1956. Microdetermination of phosphorus. Anal. Chem. **28:**1756-1758.

Chiaruttini, C. and M. Millet. 1993. Gene organization, primary structure and RNA processing analysis of a ribosomal RNA operon in *Lactococcus lactis.* J. Mol. Biol. **230:**57-76.

de Vos, W. M., P. Vos, H. de Haard, H., and I. Boerrigter. 1989. Cloning and expression of the *Lactococcus lactis* subsp. *cremoris* SK11 gene encoding an extracellular serine proteinase. Gene **85:**169-176.

Elliker, P. R., A. Anderson, and G. H. Hannessen. 1956. An agar culture medium for lactic streptococci and lactobacilli. J. Dairy Sci. **39:**1611-1612.

Fath, M. J. and R. Kolter. 1993. ABC transporters: bacterial exporters. Microbiol. Rev. **57:**995-1017.

Gasson, M. J. 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. **154:**1-9.

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.

Guédon, G., F. Bourgoin, M. Pébay, Y. Toussel, 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 intergeneric transfer of insertion sequences occur in lactic acid bacteria co-cultures? Mol. Microbiol. **16:**69-78.

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.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557-580.

Holt, J. G., N. R. Krieg, P. H. A. Sneath, J. T. Staley, and S. T. Williams, S.T. 1994. *Bergey's manual of determinative bacteriology, 9th edition.* Hensyl, W.R. (ed.) Baltimore, USA: Williams and Wilkins, p. 527-558.

Huang, J., and M. Schell. 1995. Molecular characterization of the *eps* gene cluster of *Pseudomonas solanacearum* and its transcriptional regulation at a single promoter. Mol. Microbiol. **16:**977-989.

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. Micriobiol. **5:**695-713.

Klein, P., M. Kanehisa, and C. DeLisi. 1985. The detection and classification of membrane spanning proteins. Biochim. Biophys. Acta. **815:**468-476.

Kolkman, M. A. B., D. A. Morrison, B. A. M. van der Zeijst, and P. J. M. Nuijten. 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: identification of the glycosyl transferase gene *cps14E*. J. Bacteriol. **178:**3736-3741.

Kolkman, M. A. B., B. A. M. van der Zeijst, and P. J. M. Nuijten. 1997. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14. J Biol Chem.**272:**19502-19508.

Kuipers, O. P., M. M. Beerthuyzen, R. J. Siezen, and W. M. de Vos. 1993. Characterization of the nisin gene cluster *nisABTCIPR* of *Lactococcus lactis*. Requirement of expression of the *nisA* and *nisI* genes for development of immunity. Eur. J. Biochem. **216:**281-291.

Lazarevic, V., P. Margot, B. Soldo, D. and 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.

Leigh, J. A., and D. L. Coplin. 1992. Exopolysaccharides in plant-bacterial interactions. Ann. Rev. Microbiol. **46:**307-346.

Lin, W. S., T. Cunnen, and C. Y. Lee. 1994. Sequence analysis and molecular characterization of genes required for biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*. J. Bacteriol. **176:**7005-7016.

Liu, D., R. A. Cole, and P. R. Reeves. 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J. Bacteriol. **178:**2102-2107.

Macpherson, D. F., P. A. Manning, and R. Morona. 1995. Genetic analysis of the *rfbX* gene of *Shigella flexneri*. Gene **155:**9-17.

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.

Martin, B., G. Alloing, V. Méjean, and J.-P. Claverys. 1987. Constitutive expression of erythromycin resistance mediated by the *erm*AM determinant of plasmid pAMβ1 results from deletion of 5' leader peptide sequence. Plasmid **18:**250-253.

Matsui, H., Y. Sano, H. Ishihara, and T. Shinomiya. 1993. Regulation of pyocin genes in *Pseudomonas aeruginosa* by positive (*prtN*) and negative (*prtR*) regulatory genes. J. Bacteriol. **175:**1257-1263.

Morona, R., M. Mavris, A. Fallarino, and P. A. Manning. 1994. Characterization of the *rfc* region of *Shigella flexneri*. J. Bacteriol. **176:**733-747.

Nakajima, H., S. Toyoda, T. Toba, T. Itoh, T. Mukai, H. Kitazawa, and S. Adachi. 1990. A novel phosphopolysaccharide from slime-forming *Lactococcus lactis* subspecies cremoris SBT 0495. J. Dairy Sci. **73:**1472-1477.

Nakajima, H., T. Hirota, T. Toba, T. Itoh, and S. Adachi. 1992. Structure of the extracellular polysaccharide from slime-forming *Lactococcus lactis* subsp *cremoris* SBT 0495. Carbohydr. Res. **224:**245-253.

Neve, H., A. Geis, and M. Teuber, M. 1988. Plasmid-encoded functions of ropy lactic acid streptococcal strains from Scandinavian fermented milk. Biochimie **70:**437-442.

Platteeuw, C., G. Simons, and W. M. de Vos. 1994. Use of the *Escherichia coli* β-glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria. Appl. Environ. Microbiol. **60:**587-593.

Poolman, B. and W. N. Konings. 1988. Relation of growth of *Streptococcus lactis* and *Streptococcus cremoris* to amino acid transport. J. Bacteriol. **170:**700-707.

Reuber, T .L. and G. C. Walker. 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti.* Cell **74:**269-280.

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:**843-855.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Simon, D. and A. Chopin. 1988. Construction of a vector plasmid family and its use for molecular cloning in *Streptococcus lactis.* Biochimie **70:**559-566.

Stingele, 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.

Sutherland. I. W. 1985. Biosynthesis and composition of gram-negative bacterial extracellular and wall polysaccharides. Ann. Rev. Microbiol. **39:**243-270.

Vedamuthu, E. R., and J. M. Neville. 1986. Involvement of a plasmid in production of ropiness (mucoidness) in milk cultures by *Streptocuccus cremoris* MS. Appl. Environm. Microbiol. **51:**677-682.

van Riel, J. and C. Olieman. 1991. Selectivity control in the anion-exchange chromatographic determination of saccharides in dairy products using pulsed amperometric detection. Carbohydr. Res. **215:**39-46.

van Rooijen, R. J. and W. M. de Vos. 1990. Molecular cloning, transcriptional analysis, and nucleotide sequence of *lacR*, a gene encoding the repressor of the lactose phosphotranspherase system of *Lactococcus lactis*. J. Biol. Chem. **265:**18499-18503.

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.

Wang, L. and P. R. 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.

Wang, L., D. Liu, and P. R. Reeves. 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J. Bacteriol. **178:**2598-2604.

Whitfield, C. and M. A. Valvano. 1993. Biosynthesis and expression of cell-surface polysaccharides in gram-negative bacteria. Adv. Microbial. Physiol. **35:**136-246.

Wood, H. E., K. M. Devine, and D. J. McConnell. 1990. Characterisation of a repressor gene (*xre*) and a temperature-sensitive allele from the *Bacillus subtilis* prophage, PBSX. Gene **96:**83-88

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.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33:**103-119.

Yu, W., I. Mierau, A. Mars, E. Johnson, G. Dunny, and L. L. McKay. 1995. Novel insertion sequence-like element IS*982* in lactococci. Plasmid **33:**218-225.

Exopolysaccharide Biosynthesis in *Lactococcus lactis* NIZO B40: Functional Analysis of the Glycosyltransferase Genes Involved in Synthesis of the Polysaccharide Backbone

Chapter 3

Richard van Kranenburg, Iris I. van Swam, Joey D. Marugg, Michiel Kleerebezem, and Willem M. de Vos

Published in Journal of Bacteriology (1999) **181**:338-340.

Abstract

We used homologous and heterologous expression of the glycosyltransferase genes of the *Lactococcus lactis* NIZO B40 *eps* gene cluster to determine the activity and substrate specificities of the encoded enzymes and established the order of assembly of the trisaccharide backbone of the exopolysaccharide repeating unit. EpsD links glucose-1-P from UDP-glucose to a lipid carrier, EpsE and EpsF link glucose from UDP-glucose to lipid-linked glucose, and the EpsG links galactose from UDP-galactose to lipid-linked cellobiose. Furthermore, EpsJ appeared to be involved in EPS biosynthesis as a galactosyl phosphotransferase or an enzyme which releases the backbone oligosaccharide from the lipid carrier.

Introduction

Many bacteria are known to produce polysaccharides, which can either be excreted into the environment as exopolysaccharides (EPSs), form a capsule around the cell as capsular polysaccharides (CPSs), or be attached to the cell membrane as the O antigens of lipopolysaccharides (LPSs). The biosynthesis of polysaccharides that consist of repeating units includes their assembly on a lipid carrier by sequential transfer of monosaccharides from nucleotide sugars by glycosyltransferases and the subsequent polymerisation and export of these repeating units (Sutherland, 1985, Whitfield and Valvano, 1993).

Although numerous bacterial gene clusters involved in cell-surface polysaccharide biosynthesis have been described, only a few of these have been analysed for the function of their glycosyltransferase genes. Homologous expression has been used to study the glycosyltransferase genes involved in O-antigen synthesis from different serogroups of *Salmonella enterica* (Liu *et al.*, 1993). Mutations in the different glycosyltransferase genes involved in *Rhizobium meliloti* EPS biosynthesis have been generated, and the lipid-linked intermediates which accumulated in permeabilised cells of the mutant bacteria were analyzed by thin-layer chromatography (TLC) to infer the biosynthetic step catalysed by each enzyme (Reuber and Walker, 1993). Furthermore, heterologous complementation has been used for the functional analysis of glycosyltransferase genes of *Sphingomonas* ssp. and *Rhizobium leguminosarum* (Pollock *et al.*, 1998). The involvement of streptococcal glycosyltransferase genes in the capsule biosynthesis have been studied for type III of group B *Streptococcus* and *Streptococcus pneumoniae* serotype 14 (Kolkman *et al.*, 1997a, Kolkman *et al.*, 1997b, Rubens *et al.*, 1993). For serotype 14, these genes have been shown to be essential for the synthesis of the repeating unit by their expression in *Escherichia coli* (Kolkman *et al.*, 1997a, Kolkman *et al.*, 1997b). Finally, the *eps* gene cluster of *Streptococcus thermophilus* coding for an unknown number of glycosyltransferases has been demonstrated to be involved in EPS biosynthesis by heterologous expression in a plasmid-free *Lactococcus lactis* strain (Stingele *et al.*, 1996).

Results and discussion

The *L. lactis* NIZO B40 produces an extracellular phosphopolysaccharide with a repeating unit consisting of →4)[α-L-Rha*p*-(1→2)][α-D-Gal*p*-1-PO_4-3]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→ (Nakajima *et al.*, 1992, van Kranenburg *et al.*, 1997). A plasmid-located *eps* gene cluster, including fourteen co-ordinately transcribed genes with the order *epsRXABCDEFGHIJKL* has been implicated in the biosynthesis of this EPS. A single gene disruption and heterologous expression of the *epsD* gene have been used to demonstrate that it is essential for the synthesis of the repeating unit and encodes a glycosyltransferase transferring glucose from UDP-glucose to a lipid-carrier (van Kranenburg *et al.*, 1997). Because of its homology to GumD from *Xanthomonas campestris*, we assume that EpsD, like

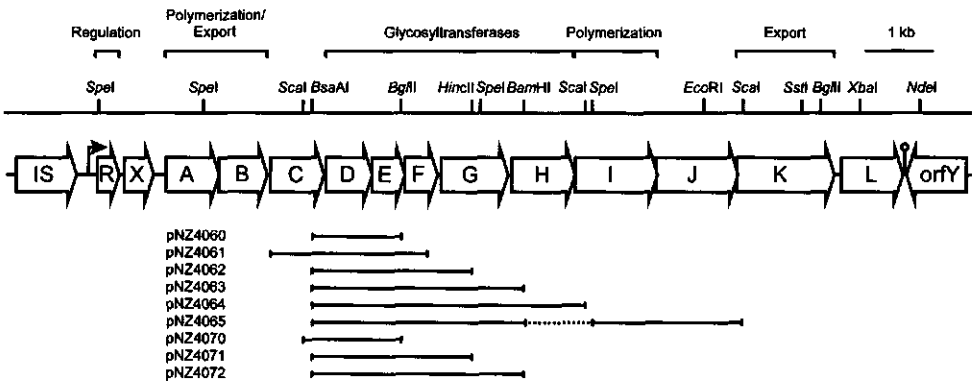


Fig. 1. Physical and genetic map of the *eps* gene cluster of plasmid pNZ4000. For *Bsa*AI, *Eco*RI, *Hin*cII, *Nde*I, *Sca*I, and *Sst*I, only sites relevant for subcloning are indicated. The plasmids used for heterologous and homologous expression of the *eps* genes are listed. Plasmids pNZ4060 through pNZ4065 are pUC19 derivatives carrying the indicated fragments under control of the *lac* promoter. Plasmids pNZ4070, pNZ4071, and pNZ4072 are pNZ8020 derivatives carrying the indicated fragments under control of the *nisA* promoter.

GumD, catalyzes the transfer of glucosyl-1-phosphate from UDP-glucose to undecaprenyl phosphate (Ielpi *et al.*, 1993). To determine the function and substrate specificities of other NIZO B40 glycosyltransferase gene products and to establish the order of assembly of the backbone of the EPS repeating unit, which could not be determined by our previous results, we expressed the relevant glycosyltransferase genes in *E. coli*. Fragments containing the *epsD*, *epsDE*, *epsDEF*, and *epsDEFG* genes (Fig. 1) were cloned under control of the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *lac* promoter in pUC19 (Yanish-Perron *et al.*, 1985) and introduced into *E. coli* DH5α (Hanahan, 1983) which has no background glycosyltransferase activity as was shown previously (van Kranenburg *et al.*, 1997). The fragment containing *epsDE* was generated by PCR with the primers 5'-CCGCGCGGATCCGGGTATAGATGATTATC-3' and 5'-CCGCGCCGAATTCTTGAAA-CGCCCTTGCTATCTC-3' by using the *Bam*HI and *Eco*RI sites of the primers (underlined) for cloning. Since the backbone of the EPS repeating unit is known to contain glucose and galactose, permeabilised cells of IPTG-induced (0.1 ng ml^{-1}) *E. coli* harbouring these plasmids were incubated with UDP-[^{14}C]glucose and UDP-[^{14}C]galactose. The lipid fraction was extracted, subjected to complete and mild acid hydrolysis, and analysed by TLC and autoradiography as described previously (van Kranenburg *et al.*, 1997) to detect the ^{14}C-labelled monosaccharides (complete acid hydrolysis) and oligosaccharides (mild acid hydrolysis), respectively (Fig.2). Expression of *epsDE* showed the same sugar incorporation as expression of *epsD* alone. In contrast, expression of *epsDEF* resulted in the production of a lipid-linked oligosaccharide with the same mobility on TLC as cellobiose, which is the β-D-Glc*p*-(1→4)-β-D-Glc*p* part of the repeating unit. Expression of *epsDEFG* resulted in the production of lipid-linked oligosaccharides containing glucose and galactose. Mild acid hydrolysis of these products yielded two oligosaccharides with a lower mobility on TLC than

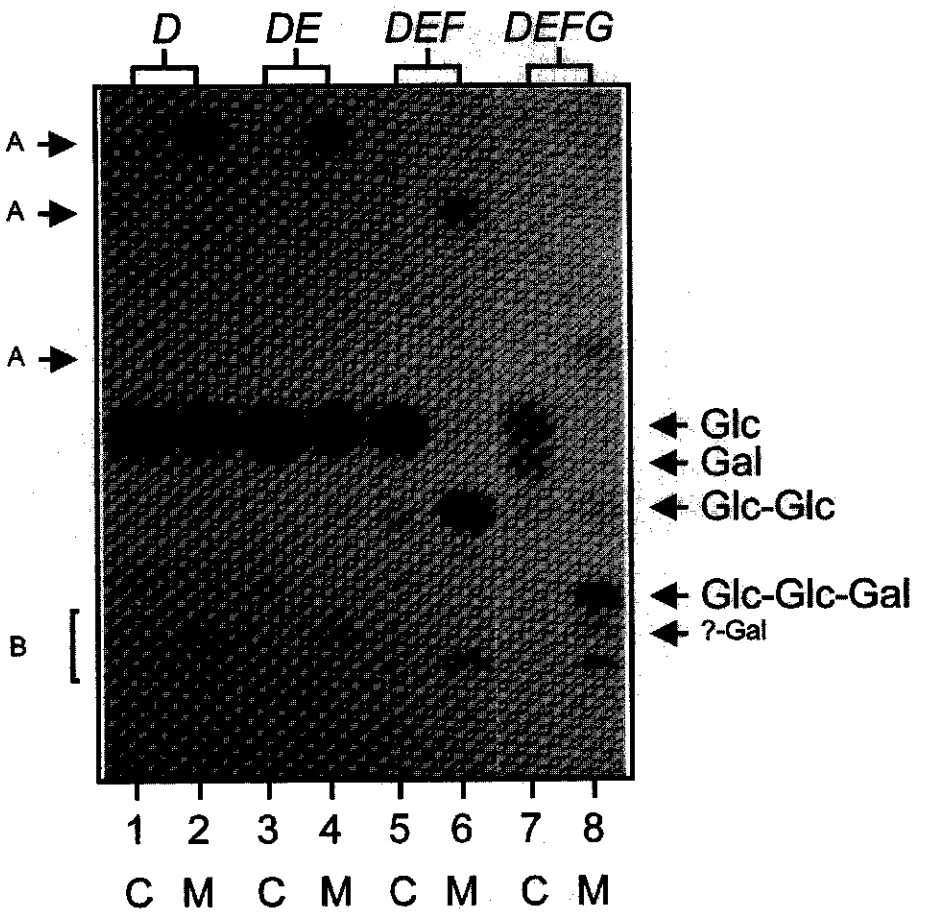


Fig. 2. TLC of ^{14}C-labelled intermediates isolated from the lipid fraction of permeabilised *E. coli* cells. The position of the standard sugars glucose (Glc), galactose (Gal), and cellobiose (Glc-Glc), as well as those of the predicted trisaccharide (Glc-Glc-Gal) and the unknown galactose-containing molecule (?-Gal), are indicated on the right. Additional products of incomplete mild hydrolysis, putative sugar-phosphate (A) and putative lipid-linked sugars (B), are indicated on the left. Lanes: 1 and 2, *E. coli* carrying pNZ4060; 3 and 4, *E. coli* carrying pNZ4061; 5 and 6, *E. coli* carrying pNZ4062; 7 and 8, *E. coli* carrying pNZ4063. Complete (C) and mild (M) acid hydrolysis treatments are indicated.

that of cellobiose that were retrieved from the TLC plate and subjected to complete acid hydrolysis, followed by a second TLC analysis. The oligosaccharide with the higher mobility consisted of glucose and galactose with a ratio of approximately 2:1, as judged from the intensity of the spots on the autoradiograph of the TLC, and is likely to be the β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p* trisaccharide of the backbone of the repeating unit. The other oligosaccharide contained only ^{14}C-labelled galactose, indicating that EpsG may act with a slightly lower affinity on another lipid acceptor than lipid-linked cellobiose in the *E.*

coli membrane. These results indicate that EpsF is the second glycosyltransferase, linking glucose to lipid-linked glucose, and that EpsG is the third glycosyltransferase, linking galactose to lipid-linked cellobiose. As EpsE and EpsF are homologous to the pneumococcal Cps14F and Cps14G, respectively, we assume that, like Cps14F and Cps14G, both lactococcal gene products act together as one glycosyltransferase and that EpsF contains the glycosyltransferase activity while EpsE has an accessory function (Kolkman *et al.*, 1997a).

To test the substrate specificities of the *epsDEFG* gene products in *L. lactis*, fragments of pNZ4000 carrying *epsD*, *epsDEF*, or *epsDEFG* were cloned under control of the *nisA* promoter of pNZ8020, which, when introduced in *L. lactis* NZ3900 (de Ruyter *et al.*, 1996a), allows the use of the NICE (nisin-controlled expression) system (de Ruyter *et al.*, 1996b, Kuipers *et al.*, 1998). Cultures were induced with 0.1 ng nisin A ml^{-1} at an optical density at 600 nm of 0.5, and cells were harvested 2 h after induction. After lysozyme treatment, permeabilised cells were prepared as described previously (van Kranenburg *et al.*, 1997). Incubation of (uninduced) permeabilised, plasmid-free *L. lactis* NZ3900 cells with UDP-[^{14}C]glucose or UDP-[^{14}C]galactose resulted in a high level of incorporation of [^{14}C]glucose in the lipid fraction by an unknown glucosyltransferase activity that may be involved in the biosynthesis of other cell-surface polysaccharides (Fig. 3). Mild acid hydrolysis of the extracted lipid fractions yielded five species of labelled saccharides (Fig. 3): one with the same mobility on TLC as glucose, one migrating slightly slower than cellobiose, and three with a higher mobility than glucose (the latter three are not shown in Fig. 3). This glucosyltransferase activity was not restricted to *L. lactis* subsp. *cremoris* MG1363 derivative NZ3900, but was also observed for *L. lactis* subsp. *lactis* IL1403 and *L. lactis* subsp. *lactis* biovar. *diacetylactis* BU2-60 (data not shown). The background incorporation prevented detection of the activity of EpsD and EpsF, as no additional effect of expression of the *epsD* or *epsDEF* genes was observed (data not shown). However, expression of *epsDEFG* resulted in the formation of a new lipid-linked oligosaccharide. Its complete acid hydrolysis yielded an additional product with the same mobility as galactose. Its hydrolysis by mild acid resulted in a product with the same mobility as the putative trisaccharide detected in *E. coli* expressing *epsDEFG* (Fig. 3). The latter product was retrieved from the TLC-plate and subjected to complete acid hydrolysis and a second TLC analysis. The oligosaccharide consisted of glucose and galactose, identical to the glucose-and-galactose-containing oligosaccharide found in *E. coli* expressing *epsDEFG*. Therefore we conclude that the functions of the *epsDEFG* genes in *E. coli* and *L. lactis* are identical.

It is likely that the subsequent steps of the repeating unit synthesis include the coupling of the side chain sugars rhamnose and galactose-phosphate to the galactose of the backbone. Possible candidates for these activities are EpsH, which is homologous to several glycosyltransferases, and EpsJ, which is homologous to a CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase of *Bacillus subtilis* designated TagH (RodC) (Pooley *et al.*, 1992, van Kranenburg *et al.*, 1997). To test the function of *epsH*, a fragment containing *epsDEFGH* (Fig. 1) was cloned under control of the *lac* promoter in pUC19. Incubation of permeabilised *E. coli* cells expressing *epsDEFGH* with

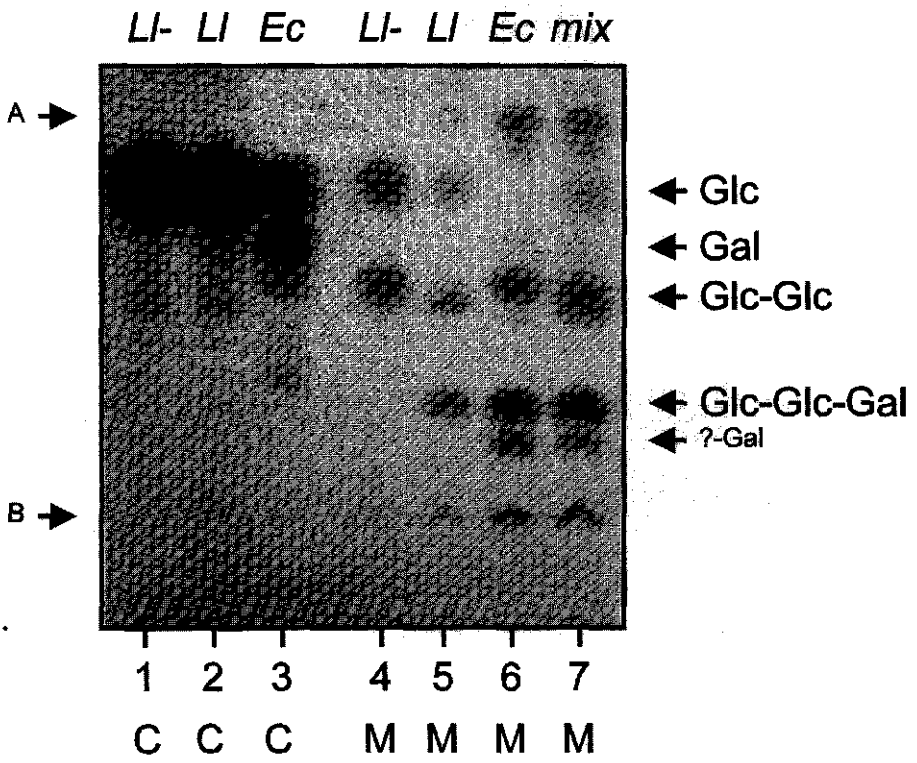


Fig. 3. TLC of ^{14}C-labelled intermediates isolated from the lipid fraction of permeabilised *L. lactis* and *E. coli* cells harbouring plasmids with the *epsDEFG* genes. The position of the standard sugars and the products of incomplete mild hydrolysis are indicated as in Fig. 2. Lanes: 1 and 4, plasmid-free *L. lactis* (*Ll-*); 2 and 5, *L. lactis* carrying pNZ4072 (*Ll*); 3 and 6, *E. coli* carrying pNZ4063 (*Ec*); 7, mixture of mild acid hydrolysates of *L. lactis* carrying pNZ4072 and *E. coli* carrying pNZ4063 (mix). Complete (C) and mild (M) acid hydrolysis treatments are indicated.

UDP-[^{14}C]glucose and UDP-[^{14}C]galactose resulted in the same products as those of cells expressing *epsDEFG*, indicating that EpsH is either not a galactosyltransferase, inactive in this assay, or not expressed. EpsH may be the rhamnosyltransferase, which could not be tested, as its substrate, dTDP-rhamnose, is unstable. The *epsJ* gene was cloned downstream of *epsDEFG* in pNZ4063 (Fig. 1). The resulting plasmid, pNZ4065, was readily obtained. Incubation of permeabilised *E. coli* cells expressing *epsDEFGJ* resulted in a complete loss of incorporation of ^{14}C-labelled sugars from the lipid fraction and no radioactivity on TLC. This strongly indicates that the *epsJ* gene is active and may encode either an enzyme linking galactosyl phosphate to galactose, after which *E. coli* enzymes can release the oligosaccharide from the membrane, or an enzyme releasing the trisaccharide backbone from the lipid carrier. Presently, we can not distinguish between these two alternatives.

In conclusion, our report describes the heterologous and homologous expression of the lactococcal *eps* genes encoding the glycosyltransferases involved in the assembly of the EPS repeating unit. EpsD, EpsE and EpsF, and EpsG link glucose-1-phosphate to a lipid carrier (presumably undecaprenyl phosphate), glucose to lipid-linked glucose, and galactose to lipid-linked cellobiose, respectively. The *epsJ* gene product is active and likely to be a galactosyl phospho-transferase or an enzyme releasing the trisaccharide backbone from the lipid carrier. Furthermore, to the best of our knowledge, this is the first report describing controlled homologous expression of glycosyltransferase genes in Gram-positive bacteria.

Acknowledgements

This work was partly supported by EC research grants 1116/92 1.6 and BIOT-CT96-0498. We thank Roland Siezen and Ingeborg Boels for critically reading the manuscript.

References

De Ruyter, P. G. G. A., O. P. Kuipers, M. M. Beerthuyzen, I. van Alen-Boerrigter, and W.M. de Vos. 1996a. Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. J. Bacteriol. **178:**3434-3439.

De Ruyter, P. G. G. A., O. P. Kuipers, and W. M. de Vos. 1996b. Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. Appl. Environ. Microbiol. **62:**3662-3667.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557-580.

Ielpi, L., R. O. Couso, M. A. and Dankert. 1993. Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. J. Bacteriol. **175:**2490-2500.

Kolkman, M. A. B., B. A. M. van der Zeijst, and P. J. M. Nuijten. 1997a. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J. Biol. Chem. **272:**19502-19508.

Kolkman, M. A. B., W. Wakarchuk, P. J. M. Nuijten, and B. A. M. van der Zeijst. 1997b. 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.

Kuipers, O. P., P. G. G. A. de Ruyter, M. Kleerebezem, and W. M. de Vos. 1998. Quorum sensing-controlled gene expression in lactic acid bacteria. J. Biotechnol. In press.

Liu, D., A. M. Haase, L. Lindqvist, A. A. Lindberg, and P. R. Reeves. 1993. Glycosyltransferases of O-antigen biosynthesis in *Salmonella enterica*: Identification and characterization of transferase genes of groups B, C2, and E1. J. Bacteriol. **175:**3408-3413.

Nakajima, H., T. Hirota, T. Toba, T. Itoh, and S. Adachi. 1992. Structure of the extracellular polysaccharide from slime-forming *Lactococcus lactis* subsp. *cremoris* SBT 0495. Carbohydr. Res. **224:**245-253.

Pollock, T. J., van W. A. T. Workum, L. Thorne, M. J. Mikolajczak, M. Yamazaki, J. W. Kijne, and R. W. Armentrout. 1998. Assignment of biochemical functions to glycosyl transferase genes which are essential for biosynthesis of exopolysaccharides in *Sphingomonas* S88 and *Rhizobium leguminosarum*. J. Bacteriol. **180:**586-593.

Pooley, H. M., F.-X. Abellan, and D. Karamata. 1992. CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase, which is involved in the synthesis of the major wall teichoic acid in *Bacillus subtilis* 168, is encoded by *tagF* (*rodC*). J. Bacteriol. **174:**646-649.

Reuber, T. L., and G. C. Walker. 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. Cell **74:**269-280.

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:** 843-855.

Stingele, 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.

Sutherland, I. W. 1985. Biosynthesis and composition of Gram-negative bacterial extracellular and wall polysaccharides. Annu. Rev. Microbiol. **39:**243-270.

van Kranenburg, R., J. D. Marugg, N. J. Willem, I. I. van Swam, and W. M. de Vos. 1997. Molecular characterization of the plasmid-encoded *eps* gene cluster essential for exopolysaccharide production in *Lactococcus lactis*. Mol. Microbiol. **24:**387-397.

Whitfield, C., and M. A. Valvano. 1993. Biosynthesis and expression of cell-surface polysaccharides in Gram-negative bacteria. Adv. Microbial Physiol. **35:**136-246.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33:**103-119.

Functional Analysis of Glycosyltransferase Genes from *Lactococcus lactis* and Other Gram-Positive Cocci: Complementation, Expression, and Diversity

Chapter 4

Richard van Kranenburg, Harmjan R. Vos, Iris I. van Swam, Michel Kleerebezem, and Willem M. de Vos

Published in Journal of Bacteriology (1999) **181**:6347-6353.

Abstract

Sixteen exopolysaccharide (EPS)-producing *Lactococcus lactis* strains were analysed for the chemical compositions of their EPSs and the locations, sequences and organisation of the *eps* genes involved in EPS biosynthesis. This allowed the grouping of these strains into three major groups, representatives of which were studied in detail. Previously, we have characterised the *eps* gene cluster of strain NIZO B40 (group I) and determined the function of three of its glycosyltransferase (GTF) genes. Fragments of the *eps* gene clusters of strains NIZO B35 (group II) and NIZO B891 (group III) were cloned, and these encoded the NIZO B35 priming galactosyltransferase, the NIZO B891 priming glucosyltransferase, and the NIZO B891 galactosyltransferase involved in the second step of repeating-unit synthesis. The NIZO B40 priming glucosyltransferase gene *epsD* was replaced with an erythromycin resistance gene, and this resulted in loss of EPS production. This *epsD* deletion was complemented with priming GTF genes from gram-positive organisms with known function and substrate specificity. Although no EPS production was found with priming galactosyltransferase genes from *L. lactis* or *Streptococcus thermophilus*, complementation with priming glucosyltransferase genes involved in *L. lactis* EPS and *Streptococcus pneumoniae* capsule biosynthesis could completely restore or even increase EPS production in *L. lactis*.

Introduction

Many gram-positive bacteria produce significant amounts of capsular polysaccharides (CPSs) or exopolysaccharides (EPSs). Most molecular studies have focused on the CPSs from strains of *Streptococcus pneumoniae*, group B streptococci, and *Staphylococcus aureus* (Roberts, 1996). These CPSs have unique structures that determine the serotype and virulence of these pathogens. Their biosynthesis is encoded by large clusters of genes that often show unidirectional organisation, are transcribed into single polycistronic mRNAs, and appear to be co-ordinately expressed (Morona *et al.*, 1997, Muñoz *et al.*, 1997, Ouyang and Lee, 1997, Sau *et al.*, 1997a). In these clusters, the serotype-specific genes encoding the glycosyltransferases (GTFs) are flanked by genes that are common to all serotypes and are likely to be involved in processes like chain length determination, polymerisation and export (Kolkman *et al.*, 1997b, Morona *et al.*, 1997, Munoz *et al.*, 1997, Sau *et al.*, 1997b). Several lactic acid bacteria are known to produce EPSs that are of industrial importance, as they are beneficial for the structure of dairy products (Cerning, 1990). Recently, the genes encoding EPS production in the dairy starters *Streptococcus thermophilus* Sfi6 and *Lactococcus lactis* NIZO B40 were characterised and their organisation was found to be similar to that of the CPS biosynthesis gene clusters of the gram-positive pathogens (Stingele *et al.*, 1996, van Kranenburg *et al.*, 1997). Functional analysis of the NIZO B40 *eps* genes demonstrated that the *epsDEF* genes are functional homologues of the *cps14EFG* genes from *S. pneumoniae* serotype 14 and code for GTFs that are involved in identical steps of the polysaccharide biosynthesis route (van Kranenburg *et al.*, 1999). In general, the GTF involved in linking the first sugar of the repeating unit to the lipid carrier, here referred to as the priming GTF, is highly homologous in gram-positive bacteria, while other GTFs are often unique or have very little homology to others (Kolkman *et al.*, 1997b, Morona *et al.*, 1997, Sau *et al.*, 1997b, Stingele *et al.*, 1996, van Kranenburg *et al.*, 1997).

In spite of the increasing sequence information on the CPS or EPS gene clusters in gram-positive cocci, very little is known about the function of the predicted GTF genes and even less is known about their specificities. By investigation of the GTF genes expressed in *Escherichia coli*, the substrate specificities of GTFs involved in the biosynthesis of *S. pneumoniae* serotype 14, *L. lactis* strain NIZO B40, and *S. thermophilus* Sfi6 were determined (Kolkman *et al.*, 1997b, Stingele and Neeser, 1999, van Kranenburg *et al.*, 1997). However, it was reported that GTF genes expressed in a heterologous host could result in a different composition of the EPS (Stingele *et al.*, 1999). Therefore, we have used a recently developed homologous expression system to demonstrate the substrate specificity of the *epsDEFG* genes of *L. lactis* NIZO B40 (van Kranenburg *et al.*, 1999). Here we describe a screening approach used to identify new GTF genes in *L. lactis* and show the diversity of GTF genes in *L. lactis* and their EPSs, resulting in a classification of three major groups. Two new priming GTF genes were selected, and their function and substrate specificity were determined. Finally, a transcomplementation of a knockout of the NIZO B40 *epsD* gene

encoding the priming GTF was realised by controlled expression of several homologous GTF genes derived from different gram-positive cocci.

Material and methods

Bacterial strains, plasmids, and media. The bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* was grown in L-broth-based medium at 37°C (Sambrook *et al.*, 1989). *L. lactis* was grown at 30°C in M17 broth (Difco Laboratories) supplemented with 0.5% glucose (GM17) or in a chemically defined medium (Poolman and Konings, 1988). If appropriate, the media contained chloramphenicol (10 μg/ml), erythromycin (10 μg/ml), or ampicillin (100 μg/ml).

DNA isolation and manipulation. Isolation of *E. coli* plasmid DNA and standard recombinant DNA techniques were performed as described by Sambrook *et al.* (1989). Large scale isolation of *E. coli* plasmid DNA for nucleotide sequence analysis was performed with JetStar columns by following the instructions of the manufacturer (Genomed). Isolation and transformation of *L. lactis* plasmid DNA were performed as previously described (de Vos *et al.*, 1989). Southern blots were hybridised with *eps* gene probes at 45°C and washed with 0.1 × SSC (1 × SSC is 0.15 M NaCl plus 0,015 M sodium citrate) at 45°C before exposure.

Nucleotide sequence analysis. Automatic double stranded DNA sequence analysis was performed on both strands with an ALF*red* DNA sequencer (Pharmacia Biotech). Sequencing reactions were accomplished by using the AutoRead sequencing kit, initiated by using Cy5-labelled universal and reverse primers, and continued with synthetic primers in combination with Cy5-13-dATP by following the instructions of the manufacturer (Pharmacia Biotech). Sequence data were assembled and analysed using the PC/GENE program version 6.70 (IntelliGenetics).

Construction of plasmids. For expression of the NIZO B35 *eps* genes in *E. coli*, a 1.0 kb *Sca*I-*Hin*cII fragment containing *orfU*, a 2.7-kb *Sca*I-*Kpn*I fragment containing *orfU-epsD*, and a 1.3-kb *Sca*I fragment containing *epsD* were cloned under control of the *lac* promoter in pUC18 or pUC19 (Yanisch-Perron *et al.*, 1985). To express the NIZO B891 *eps* genes in *E. coli*, a 1.0-kb *Sca*I-*Bal*I fragment containing *epsD* and a 1.9-kb *Sca*I-*Eco*RI fragment containing *epsDEF* were cloned under control of the *lac* promoter in pJF119HE (Fürste *et al.*, 1986). For expression of the NIZO B35 and NIZO B891 *epsD* genes in *L. lactis*, a 1.3-kb *Sca*I fragment and a 1.0-kb *Sca*I-*Bal*I fragment, respectively, were cloned under control of the *nisA* promoter in pNZ8020 (de Ruyter *et al.*, 1996b). To express the streptococcal *cps14* GTF genes in *L. lactis*, a 1.3-kb *Xba*I-*Pvu*II fragment containing the GTF part of *cps14E* and a 2.6-kb *Xba*I fragment containing *cps14EFG* were cloned from pMK100 (Kolkman *et al.*, 1996) under control of the *nisA* promoter in pNZ8020. To express the streptococcal *epsE* GTF gene in *L. lactis*, a 1.8 kb *Eco*RV-*Xba*I fragment containing *epsE* was cloned from pFS30 (Stingele *et al.*, 1996) under control of the *nisA* promoter in pNZ8020. To construct a NIZO B40 *epsD* gene disruption, a PCR was used to clone the flanking regions containing

Table 1. Bacterial strains and plasmids used in this work.

Strain or plasmid	Relevant characteristics[a]	Source or reference
Strains		
E. coli DH5α		Hanahan, 1983
L. lactis		
NIZO B35	Lac^+ Eps^+ multiplasmid strain	NIZO collection
NIZO B36	Lac^+ Eps^+ multiplasmid strain	NIZO collection
NIZO B39	Lac^+ Eps^+ multiplasmid strain	NIZO collection
NIZO B40	Lac^+ Eps^+ multiplasmid strain harbouring pNZ4000	van Kranenburg *et al.*, 1997
NIZO B891	Lac^+ Eps^+ multiplasmid strain	NIZO collection
NIZO B1136	Lac^+ Eps^+ multiplasmid strain	NIZO collection
NIZO B1137	Lac^+ Eps^+ multiplasmid strain	NIZO collection
SBT 0495	Lac^+ Eps^+ multiplasmid strain	Nakajima *et al.*, 1990
H414	Lac^+ Eps^+ multiplasmid strain	Gruter *et al.*, 1992
SD8	Lac^+ Eps^+ multiplasmid strain	Neve *et al.*, 1988
SD11	Lac^+ Eps^+ multiplasmid strain	Neve *et al.*, 1988
VI6	Lac^+ Eps^+ multiplasmid strain	Neve *et al.*, 1988
VI8	Lac^+ Eps^+ multiplasmid strain	Neve *et al.*, 1988
MLT1	Lac^+ Eps^+ multiplasmid strain	Quest collection
MLT2	Lac^+ Eps^+ multiplasmid strain	Quest collection
MLT3	Lac^+ Eps^+ multiplasmid strain	Quest collection
NZ3900	*pepN*::*nisRnisK*	de Ruyter *et al.*, 1996a
Plasmids		
pNZ4080	Ap^r, 3.8-kb pUC19 derivative carrying NIZO B35 *orfU*	This study
pNZ4081	Ap^r, 5.5-kb pUC19 derivative carrying NIZO B35 *orfUepsD*	This study
pNZ4082	Ap^r, 4.0-kb pUC18 derivative carrying NIZO B35 *epsD*	This study
pNZ4083	Cm^r, 4.4-kb pNZ8020 derivative carrying NIZO B35 *epsD*	This study
pNZ4085	Ap^r, 6.3-kb p119HE derivative carrying NIZO B891 *epsD*	This study
pNZ4086	Ap^r, 7.2-kb p119HE derivative carrying NIZO B891 *epsDEF*	This study
pNZ4087	Cm^r, 4.2-kb pNZ8020 derivative carrying NIZO B891 *epsD*	This study
pNZ4090	Cm^r, 4.5-kb pNZ8020 derivative carrying *cps14E*	This study
pNZ4091	Cm^r, 5.8-kb pNZ8020 derivative carrying *cps14EFG*	This study
pNZ4055	Ery^r Eps^-, pNZ4000Δ*epsD* carrying *ery* from pIL253	This study
pNZ4030	Ery^r Eps^+, 27-kb pNZ4000 derivative carrying *ery* from pIL253	van Kranenburg *et al.*, 1997
pNZ4070	Cm^r, 4.6-kb pNZ8020 derivative carrying NIZO B40 *epsD*	van Kranenburg *et al.*, 1999

[a] Lac^+, lactose fermenting phenotype; Eps^+, EPS-producing phenotype; Ap^r, ampicillin resistant; Cm^r, chloramphenicol resistant; Ery^r, erythromycin resistant.

epsC (by using the primers 5'-AGCAGCAAGCTTTTCAAGTTATATATTGA-3' and 5'-TTCAGAGGATCCCTCAAAAACTTCCAT-3') and *epsEF* (by using the primers 5'-CTACATGGATCCGATGCTTATTAAAGTAA-3' and 5'-ATTATTGAATTCATCAGAA-TAATTCCCCTA-3') in pUC18, making use of the *Eco*RI, *Bam*HI, and *Hin*dIII sites of the

primers (underlined). The *ery* gene of pIL253 was cloned from pUC18Ery (van Kranenburg *et al.*, 1997) into the *Bam*HI site between the *epsC* and *epsEF* fragments in the same orientation as the *eps* genes. The complete *Eco*RI-*Hin*dIII insert was transferred to pG$^+$host8 (Maguin *et al.*, 1996), resulting in a tetracycline-resistant (Tet^r), erythromycin-resistant (Ery^r) construct containing a temperature sensitive replicon which is not functional at 37°C. The resulting plasmid was transformed to strain NZ4010 harbouring EPS plasmid pNZ4000 (van Kranenburg *et al.*, 1997), and transformants were subsequently cultured at 37°C. A Tet^s Eps$^-$ Ery^r double-cross over mutant of pNZ4000 was obtained in which *epsD* was exchanged for the *ery* gene (pNZ4055). The pUC, pJF119HE, and pG$^+$host derivatives were constructed in *E. coli* DH5α, and the pNZ8020 derivatives were constructed in *L. lactis* NZ3900.

EPS purification and characterisation. *L. lactis* was grown in 50 ml of defined medium containing 2% glucose for 48 h at 30°C, and after pelleting of the cells, EPS was purified by dialysis and lyophilisation and quantified by gel permeation high-performance liquid chromatography (HPLC) analysis using dextran 500 as a standard as described before (van Kranenburg *et al.*, 1997). Sugar analysis of was performed by HPLC analysis of the monosaccharide units after complete hydrolysis with 4 N HCl (van Riel and Olieman, 1991). To analyse the EPS in overproducing strain NZ3900 harbouring pNZ4055 and pNZ8020 derivatives, induction was performed with nisin A at 1 ng ml^{-1} at and an optical density at 600 nm of 0.5 (de Ruyter *et al.*, 1996a).

GTF activity assays and TLC analysis. GTF activity assays and thin-layer chromatography (TLC) analysis were performed with permeabilised *E. coli* cells as described before (van Kranenburg *et al.*, 1997). Permeabilised *L. lactis* cells were prepared like those of *E. coli* after a 30-min incubation with lysozyme (10 mg ml^{-1}) on ice. After incubation with UDP-[^{14}C]glucose and/or UDP-[^{14}C]galactose, the extracted lipid fractions were subjected to complete and mild acid hydrolysis and analysed by TLC and autoradiography to detect ^{14}C-labelled monosaccharides (complete acid hydrolysis) and oligosaccharides (mild acid hydrolysis), respectively.

Nucleotide sequence accession numbers. The nucleotide sequences of the NIZO B35 and NIZO B891 *eps* gene cluster fragments are available under GenBank accession no. AF100297 and AF100298.

Results

Diversity of lactococcal GTF genes and EPSs. In a search for new GTF genes, we screened a collection of 16 different EPS-producing *L. lactis* strains at the genetic and biochemical levels. To localise putative *eps* gene clusters, DNA from the strains was probed with an internal fragment of the *L. lactis* NIZO B40 *epsB* gene (Fig. 1B), which is highly conserved and has homologues in all studied EPS or CPS gene clusters of gram-positive cocci (van Kranenburg *et al.*, 1997). All of the *L. lactis* strains tested contained a single plasmid (>20 kb) that hybridised with the *epsB* probe (results not shown). This confirms

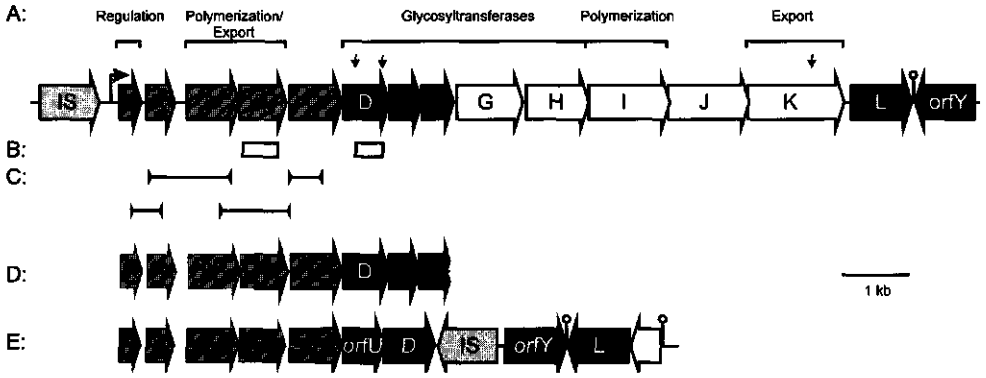


Fig. 1. (A) Genetic map of the *eps* gene cluster of *L. lactis* NIZO B40. The *Sst*I recognition sites are indicated by downward-pointing arrows. The predicted functions of the gene products are depicted above the map (van Kranenburg *et al.*, 1997). (B) DNA fragments of the NIZO B40 *eps* gene cluster used for hybridisation. (C) PCR fragments generated by the primers indicated by the arrowheads used to determine the order of the conserved *eps* genes of various strains (see text). (D and E) Genetic maps of the *eps* gene cluster of *L. lactis* NIZO B891 and NIZO B35 based on DNA sequences of cloned fragments and on PCR analysis.

previous suggestions that EPS production in *L. lactis* is plasmid encoded (Neve *et al.*, 1988, Vedamuthu and Neville, 1986, Von Wright and Tynkkynen, 1987). The diversity of the plasmid-encoded GTF genes was studied by analysing their hybridisation to the NIZO B40 *epsB* and *epsD* genes (Fig. 1B). This *epsD* gene codes for the priming glucosyltransferase and shows homology to other priming GTF genes (van Kranenburg *et al.*, 1997). For this purpose, plasmid DNA of all strains was digested with *Sst*I, which has three sites within the NIZO B40 *eps* gene cluster, two of which are present in the *epsD* gene (Fig. 1A). All strains hybridised with both *epsB* and *epsD* probes, but the sizes of the hybridising *Sst*I bands differed considerably, allowing genetic differentiation (Table 2).

The biochemical diversity of the EPSs isolated from the 16 strains was studied by determining the nature and molar ratio of the sugar monosaccharides (Table 2). No other sugars than glucose, galactose, or rhamnose were present in these polymers. Based on the genetic and biochemical diversity of the putative GTF genes and the EPSs, the *L. lactis* strains could be classified into three main groups (Table 2). Group I contains six strains that produced EPS containing the monosaccharides galactose, glucose, and rhamnose and includes strains SBT 0495 and NIZO B40 which produce EPSs with repeating units consisting of →4)-[α-L-Rha*p*-(1→2)][α-D-Gal*p*-1-PO_4-3]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→ (Nakajima *et al.*, 1992, van Casteren *et al.*, 1999a, van Kranenburg *et al.*, 1997). Group II comprises five strains that produced EPS with only galactose and includes strain H414, the EPS repeating unit of which is known to be →4)-[β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)]-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→ (Gruter *et al.*, 1992). This group shows restriction fragment length polymorphism for *epsD*. Group III contains three strains that produced EPS composed of both galactose and glucose in a molar ratio of

Table 2. Hybridisation patterns of *SstI*-digested lactococcal plasmid DNA and sugar composition of produced EPS.

Strain	Fragment size (kb)[a]		Molar ratio[b]		
	epsB	*epsD*	Gal	Glc	Rha
Group I					
NIZO B40	7.5	0.4	1.00	1.82	0.88
SBT 0495	7.5	0.4	1.00	1.70	0.82
NIZO B1136	7.5	0.4	1.00	1.70	0.82
VI6	7.5	0.4	1.00	1.46	0.73
VI8	7.5	0.4	1.00	1.80	0.89
MLT3	7.5	0.4	1.00	1.46	0.73
Group II					
NIZO B35	7.5	15	1.00		
NIZO B36	7.5	15	1.00		
H414	7.5	15	1.00		
SD8	7.5	13	1.00		
SD11	7.5	13	1.00		
Group III					
NIZO B891	6.5	12	1.00	1.50	
MLT1	6.5	12	1.00	1.46	
MLT2	6.5	12	1.00	1.44	
Unique					
NIZO B39	7.5	12	1.00	0.60	0.55
NIZO B1137	15	18	1.00	1.79	

[a] Approximate sizes of fragments hybridising with NIZO B40 probes are listed.

[b] Abbreviations: Rha, rhamnose; Gal, galactose; Glc, glucose. Molar ratios relative to galactose content are listed.

approximately 2 to 3. In addition to these three major groups, there are two strains (NIZO B39 and NIZO B1137) that show a unique combination of hybridisation pattern and EPS sugar composition.

Genetic variety of *eps* gene clusters. From the three major groups of EPS-producing lactococci, strains NIZO B40, NIZO B35, and NIZO B891 were selected as representatives and further characterised together with the unique strains NIZO B39 and NIZO B1137, as the structure of their EPS is known (NIZO B40) or is being analysed (NIZO B35, NIZO B891) (van Casteren *et al.*, 1999b, van Casteren, personal communication). Plasmid DNA of these strains was analysed by Southern blot analysis with specific probes for each of the genes of the *epsRXABCDEFGHIJKLorfY* operon from NIZO B40 plasmid pNZ4000. The genes *epsR*, *epsX*, *epsA*, *epsB*, *epsC*, and *epsD* hybridised with the EPS plasmids of all five strains, and *epsL* and *orfY* hybridised with those of NIZO B40, NIZO B35, NIZO B39, and NIZO 1137, indicating their conservation in all gene clusters. The other *eps* genes of NIZO B40 only hybridised with NIZO B40 plasmid pNZ4000.

Table 3. Homologies of the *L. lactis* NIZO B35 and NIZO B891 *eps* gene products to those from *L. lactis* NIZO B40.

Strain and gene	No. of amino acids of protein	Proposed function of gene product[a]	Homology (% identity) to		
			NIZO B40	*S. thermophilus* Sfi6	*S.pneumoniae* serotype 14
NIZO B35					
epsC[b]	125	Unknown	EpsC (93.6)	EpsB (33.9)	Cps14B (33.0)
orfU	199	Unknown	EpsD (85.1)	EpsE (42.1)	Cps14E (39.1)
epsD	251	Gal-P-TF	EpsD (39.2)	EpsE (41.9)	Cps14E (36.3)
orf982	296	Transposase	*orf982* (98.0)		
orfY	300	Unknown	OrfY (95.7)		
epsL	300	Unknown	EpsL (88.6)		
NIZO B891					
epsB[b]	155	Chainlength determination	EpsB (93.5)	EpsD (31.8)	Cps14D (33.1)
epsC	230	Unknown	EpsC (96.9)	EpsB (26.7)	Cps14B (29.2)
epsD	228	Glc-P-TF	EpsD (88.1)	EpsE (40.0)	Cps14E (39.1)
epsE	149	Gal-TF[c]	EpsE (40.4)		Cps14F (83.8)
epsF[d]	150	Gal-TF	EpsF (36.5)		Cps14G (53.0)

[a] Gal-P-TF, priming galactosyltransferase; Glc-P-TF, priming glucosyltransferase; Gal-TF, galactosyltransferase.

[b] Incomplete at 5'-end.

[c] Accessory function to EpsF (see text).

[d] Incomplete at 3'-end.

To further determine the organisation of the different *eps* gene clusters, specific primers based on the NIZO B40 *eps* gene cluster were used for PCRs to detect fragments overlapping *epsRX*, *epsXA*, *epsAB*, or *epsBC* (Fig. 1C). For the *epsRX*, *epsAB*, and *epsBC* fragments, all of the strains yielded PCR products identical in size (results not shown). For the *epsXA* fragments, NIZO B39, NIZO B891, and NIZO B1137 yielded PCR products that were 165 bp larger than those of NIZO B35 and NIZO B40 (results not shown). These results confirm the homologies found by the Southern blot analysis and indicate that all of the gene clusters contain a conserved region with the same organisation *i.e.*, *epsRXABC*.

NIZO B35 and NIZO B891 *eps* genes. To study the function of the priming GTF genes, strains NIZO B35 and NIZO B891 were selected, because they represent the two major groups with an EPS structure that differs markedly from that of strain NIZO B40 (Table 2). Overlapping fragments of the *eps* gene clusters of NIZO B35 and NIZO B891 that hybridised with the NIZO B40 *epsD* probe were cloned and sequenced (Fig. 1). The homologies of the deduced gene products are listed in Table 3. Unexpectedly, the NIZO B35 gene cluster contained two different genes that are homologous to NIZO B40 *epsD* (*orfU* and *epsD*, respectively). To test which of these *epsD*-like genes encodes the priming GTF activity, each of these was cloned under control of the isopropyl-β-D-thiogalactopyranoside (IPTG)-

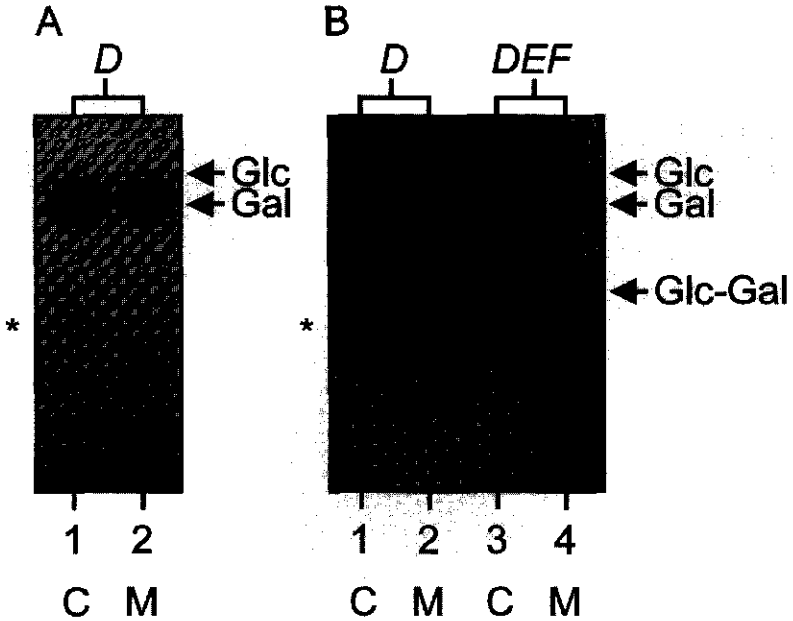


Fig. 2. TLC of ^{14}C-labelled intermediates isolated from the lipid fraction of permeabilised *E. coli* cells. (A) *E. coli* expressing NIZO B35 *epsD* incubated with UDP-[^{14}C]galactose. (B) *E. coli* expressing NIZO B891 *epsD* (1, 2) or NIZO B891 *epsDEF* incubated with a combination of UDP-[^{14}C]glucose and UDP-[^{14}C]galactose (3, 4). The positions of the standard sugars glucose (Glc), galactose (Gal) and lactose (Glc-Gal) are indicated on the right. The products which are non-specific for lactococcal GTF activity are indicated by the asterisk on the left. C, complete acid hydrolysis; M, mild acid hydrolysis.

inducible *lac* promoter in pUC18 and GTF activities were determined in *E. coli*. When NIZO B35 *epsD* was induced in *E. coli*, galactosyltransferase activity could be detected (Fig. 2A). However, when *orfU* was induced, no GTF activity could be detected (data not shown). Simultaneous induction of both *orfU* and *epsD* from NIZO B35 resulted in the same galactosyltransferase activity as that found with NIZO B35 *epsD* alone (data not shown). These results indicate that NIZO B35 *epsD* encodes a priming GTF activity and *orfU* is either not involved in these synthetic steps, poorly expressed, or unstable.

The products of the NIZO B891 *epsD*, *epsE*, and *epsF* genes are expected to be the GTFs involved in the first two steps of EPS biosynthesis in this strain, as they are homologous to NIZO B40 *epsD*, *epsE*, and *epsF*. Fragments containing NIZO B891 *epsD* and *epsDEF* were cloned under control of the *lac* promoter in medium-copy-number expression vector pJF119HE, since attempts to clone them in pUC18 were unsuccessful. When NIZO B891 *epsD* was expressed in *E. coli*, only glucosyltransferase activity could be detected (Fig. 2B, lanes 1 and 2). When *epsDEF* was expressed, both glucosyltransferase and galactosyltransferase activities could be detected (Fig. 2B, lane 3) and the lipid-linked oligosaccharide had the same mobility on TLC as lactose (Fig. 2B, lane 4). The incorporation of ^{14}C-labelled sugars was approximately fivefold lower than that of cells expressing NIZO

B40 or NIZO B35 *eps* genes (data not shown) and this lower GTF activity resulted in an increase in the appearance of a product in the complete acid hydrolysates which is non-specific for lactococcal GTF activity (Fig. 2B, lanes 1 and 3). These results demonstrate that NIZO B891 *epsD* encodes a glucosyltransferase linking glucose to the lipid carrier and *epsE* and/or *epsF* encode a galactosyltransferase linking galactose via a β-1,4 linkage to lipid-linked glucose, resulting in lipid-linked lactose. Methylation analysis of NIZO B891 EPS has confirmed the presence of 1,4-linked glucose and galactose residues (van Casteren *et al.*, 1999b). In analogy to the homologous pneumococcal proteins Cps14F and Cps14G (Kolkman *et al.*, 1997a), EpsF is expected to contain GTF activity while EpsE is expected to have an accessory function.

Homologous and heterologous complementation of a NIZO B40 *epsD* mutant. To analyse the function of the GTFs in a gram-positive host, we constructed pNZ4055, a pNZ4000 derivative in which the *epsD* gene was replaced with an erythromycin resistance (*ery*) gene. This was achieved through a double crossover with a pGhost8 derivative containing the *ery* gene from pIL252 flanked by NIZO B40 *epsC* and *epsEF*. The *ery* gene has no terminator ensuring expression of the downstream genes (van Kranenburg *et al.*, 1997). *L. lactis* harbouring pNZ4055 was erythromycin resistant and produced no EPS. To test whether the *epsD* knockout could be complemented, the pNZ8020 derivative pNZ4070 carrying the NIZO B40 *epsD* gene under control of the lactococcal *nisA* promoter was cotransformed with pNZ4055 into *L. lactis* NZ3900, which allows the use of the NICE (nisin-controlled expression) system (de Ruyter *et al.*, 1996b, Kuipers *et al.*, 1998). Upon induction with nisin A, the EPS production of the resulting heteroplasmid strain was even higher than that of the wild-type strain, demonstrating that controlled overexpression of the *epsD* gene was achieved (Table 4). To test their heterologous complementation ability, various priming GTF genes from *L. lactis*, *S. thermophilus*, and *S. pneumoniae* were cloned in pNZ8020. The EPS produced by cultures of *L. lactis* harbouring pNZ4055 and pNZ8020 derivatives was quantified, and the monosaccharide composition was determined (Table 4). The NIZO B40 and NIZO B891 genes encoding glucosyltransferases were able to complement the EPS-deficient phenotype. While expression of the NIZO B40 *epsD* gene restored EPS production completely, the amount of EPS produced by expression of NIZO B891 *epsD* was dramatically lower. A low GTF activity of the NIZO B891 EpsD compared to that of NIZO B40 EpsD was also found in *E. coli* (see above). In contrast, complete restoration of wild-type EPS production by heterologous complementation was achieved by using the *cps14E* gene of *S. pneumoniae* type 14 (Table 4). This gene is involved in pneumococcal capsule synthesis, encoding the priming glucosyltransferase (Kolkman *et al.*, 1997a), and is homologous to NIZO B40 *epsD* (van Kranenburg *et al.*, 1997). Expression of the NIZO B35 *epsD* or the *S. thermophilus* Sfi6 *epsE* gene (Stingele and Neeser, 1999), both encoding a galactosyltransferase, did not complement the EPS-deficient phenotype (Table 4), indicating that a matching sugar specificity is required for transcomplementation. Although expression of *cps14E* restored EPS production completely, complementation with the pneumococcal *cps14EFG* genes resulted in reduced production of wild-type EPS compared

Table 4. Functional expression of streptococcal GTF genes to complement an *epsD* knockout in *L. lactis* NZ3900

Original host	Gene(s)	Specificity[a]	EPS production (mg liter^{-1})[b]
L. lactis NIZO B40	*epsD*	Glc	133
L. lactis NIZO B891	*epsD*	Glc	7.5
L. lactis NIZO B35	*epsD*	Gal	<0.5
S. pneumoniae type 14	*cps14E*	Glc	102
S. pneumoniae type 14	*cps14EFG*	Glc + Gal	3.9
S. thermophilus Sfi6	*epsE*	Gal	<0.5

[a] Glc, glucosyltransferase; Gal, galactosyltransferase.

[b] Amounts of EPS are the mean values of data from two independent cultures. The EPS production of *L. lactis* NZ3900 harbouring pNZ4030 (wildtype) is 113 mg liter^{-1} and that of NZ3900 harbouring pNZ4055 (Δ*epsD*) is <0.5 mg liter^{-1}. All EPSs had a monosaccharide composition identical to that produced by the wildtype.

to complementation with *cps14E* alone. The gene products of *cps14F* and *cps14G* are involved in the second step of serotype 14 CPS biosynthesis linking galactose to lipid-linked glucose (Kolkman *et al.*, 1997a). Therefore, it is likely that they will compete for the lipid-linked glucose as the acceptor molecule with the products of the NIZO B40 *epsE* and *epsF* genes that link glucose to it, resulting in lipid-linked cellobiose (van Kranenburg *et al.*, 1999). If so, it may be assumed that the lipid-linked lactose resulting from Cps14F and Cps14G activity cannot be used for NIZO B40 EPS biosynthesis, hence lowering NIZO B40 EPS production. These results demonstrate that functional expression of gram-positive GTFs in *L. lactis* is possible and may result in heterologous complementation when the enzymes are alike in sugar-specificity.

Discussion

We have analysed the diversity of GTF genes of 16 different ropy *L. lactis* strains and the EPSs they produced, allowing division into three major groups and two individual strains. The grouping observed is in agreement with the known structural EPS information, as the EPSs produced by group I strains NIZO B40 and SBT 0495 are identical and differ from those of strains H414 (group II) and NIZO B891 (group III) (Gruter *et al.*, 1992, Nakajima *et al.*, 1992, van Casteren *et al.*, 1999b, van Kranenburg *et al.*, 1997). Furthermore, methylation analysis of the EPS produced by strain NIZO B35 (group II) demonstrated that it contains the same galactose linkages as the H414 EPS and it is expected to have an identical EPS repeating unit (van Casteren, personal communication). The sugar specificity of the GTFs needed for EPS biosynthesis in the different groups can be predicted according to the sugars present in the EPSs. The results suggest that EPS biosynthesis in all groups requires active galactosyltransferases, while groups I and III also need glucosyltransferases and only group I needs rhamnosyltransferases.

The genetic organisation of the lactococcal *eps* gene clusters is conserved with respect to the first genes *epsRXABC*, which seem to be highly homologous for all strains. Furthermore, these genes share the most homology with those of other gram-positive polysaccharide biosynthesis gene clusters, including those of *S. aureus*, *S. pneumoniae*, *S. agalactiae*, and *S. thermophilus* (van Kranenburg *et al.*, 1997). These homologies are confirmed for the NIZO B891 *epsB* and *epsC* and NIZO B35 *epsC* gene products by analysis of their nucleotide sequences, demonstrating that these genes are common to gene clusters involved in the biosynthesis of many gram-positive polysaccharide types (Table 3). It is likely that they will be involved in general functions and not directly related to the composition of the polymer produced (Munoz *et al.*, 1997, Sau and Lee, 1996, Stingele *et al.*, 1996).

The *epsL* and *orfY* genes have homologues in all of the lactococcal gene clusters tested. The function of these genes is unknown. OrfY is homologous to the regulator protein LytR from *Bacillus subtilis* (van Kranenburg *et al.*, 1997). NIZO B40 *epsL* can be disrupted by single crossover using an internal gene fragment or overproduced without any effect on EPS production (van Kranenburg, unpublished results). Nonetheless, *epsL*- and *orfY*-like genes are also found at the end of the *eps* gene cluster from *S. thermophilus* CNRZ368 adjacent to an IS element (Bourgoin, 1997).

The genetic organisation of the NIZO B35 *eps* gene cluster differs from that of NIZO B40 and NIZO B891 by an interruption of the gene cluster by an IS*982* element after the first GTF gene. An almost identical IS element is located upstream of the NIZO B40 *eps* gene cluster (Fig. 1A). Furthermore, the NIZO B35 gene cluster differs by containing two *epsD*-like genes, of which only one is actively involved in the first step of EPS biosynthesis, as was shown by the analysis of the products formed in the GTF activity assays of *E. coli* cells expressing NIZO B35 *epsD*, *orfU*, or *epsD* and *orfU*. A possible explanation for the differences in organisation of the NIZO B35 *eps* gene cluster is that it has undergone rearrangement mediated by the IS element and received an additional *epsD* gene from another *eps* gene cluster. Horizontal gene transfer of parts of polysaccharide gene clusters has been observed in various bacteria, including *S. pneumoniae* (Coffey *et al.*, 1998).

All 16 of the *L. lactis* strains studied carry an *epsD* homologue which was cloned and subjected to functional analysis for strains NIZO B35 and NIZO B891. The product of the NIZO B891 *epsD* gene is a glucosyltransferase that is more homologous to NIZO B40 EpsD than to the product of the NIZO B35 *epsD* gene, which is a galactosyltransferase (Table 3). Sequence alignment of several EpsD-like proteins from different polysaccharide biosynthesis systems with known glucosyl- or galactosyltransferase activity showed three blocks that are conserved in all of the proteins (Wang *et al.*, 1996). An alignment of the EpsD-like gram-positive GTFs with known sugar specificity shows that the three blocks are also conserved in these proteins (Fig. 3). Blocks A and B are predicted to interact with the lipid carrier and block C is supposed to contain specific conserved residues for each type of transferase (Wang *et al.*, 1996). From these, only a galactosyltransferase-specific tyrosine was observed (Fig. 3) and different residues appeared to be conserved for the gram-positive GTFs, demonstrating that the previously reported residues are not critical in determining sugar

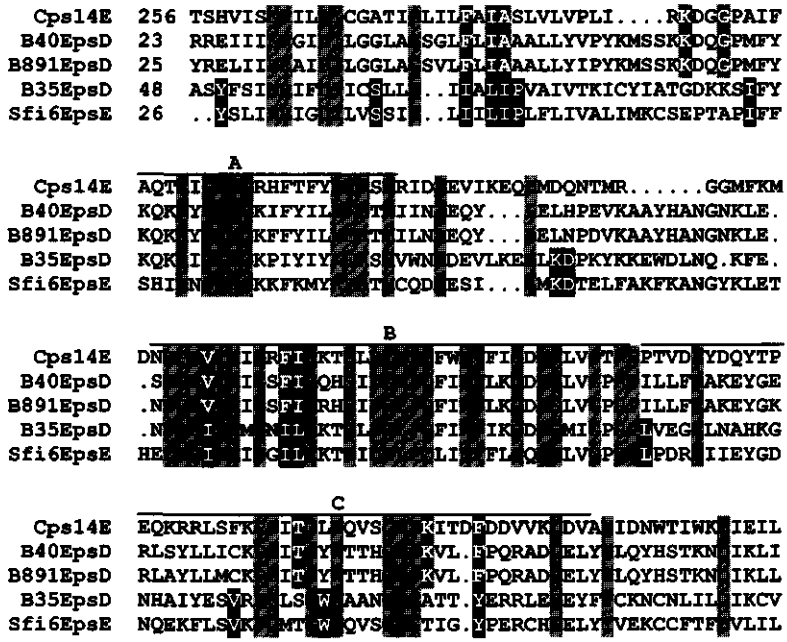


Fig. 3. Multiple-sequence alignment of priming GTFs with known sugar specificity from gram-positive bacteria. Cps14E, B40EpsD, and B891EpsD are glucosyltransferases from *S. pneumoniae* serotype 14 (Kolkman *et al.*, 1997a) and *L. lactis* NIZO B40 and NIZO B891, respectively. B35EpsD and Sfi6EpsE are galactosyltransferases from *L. lactis* NIZO B35 and *S. thermophilus* Sfi6 (Stingele *et al.*, 1999), respectively. Residues conserved in all five sequences, residues conserved only in glucosyltransferases, and residues conserved only in galactosyltransferases are shaded light grey, dark grey, and black, respectively. The three conserved blocks (A, B, and C) described by Wang *et al.* (1996) are indicated.

specificity. GTF activity involves amino acids that can catalyse an acid-base reaction. Hydrophobic cluster analysis of various β-GTFs has shown two aspartic acid residues with a spacing of approximately 50 amino acids to be conserved, and these are predicted to be the catalytic residues (Saxena *et al.*, 1995). Four conserved aspartate residues (D) and two conserved glutamate residues (E) were found for the gram-positive GTFs (Fig. 3), two of which are likely to be the catalytic residues. Two possible candidates are the conserved E residue in block C in combination with the conserved D residue in the C terminus just outside block C, which are separated by 50 amino acids (51 in Cps14E). The amino acid sequence of NIZO B35 OrfU lacks 30 amino acids at its C terminus compared to the other priming GTFs, including this conserved aspartate.

Disruption of the NIZO B40 *epsD* gene could be complemented by homologous expression of NIZO B40 *epsD* and heterologous expression of NIZO B891 *epsD* or the streptococcal capsule biosynthesis gene *cps14E*, which is known to be involved in a similar reaction Kolkman *et al.*, 1996). The use of a controlled expression system enabled the expression of GTFs that did not complement the mutation and could be toxic to the cell as a

result of the accumulation of lipid-linked intermediates (NIZO B35 *epsD*, *S. thermophilus epsE*, and *S. pneumoniae cps14EFG*) as has been reported for the heterologous expression of several gram-negative GTFs (Pollock *et al.*, 1998). Moreover, to the best of our knowledge, this is the first demonstration of functional heterologous expression of a GTF gene in a gram-positive host allowing the expression of GTF genes from different origins by the shotgun or directed-cloning approach in *L. lactis*. Furthermore, these results demonstrate that the enzymes involved in the biosynthesis of different polysaccharides can be functionally coupled, although the *eps* genes are located on different transcriptional units. The possibility of constructing clean deletion mutations in the lactococcal *eps* gene cluster combined with the use of the NICE expression system, enabling induced expression of GTF genes, opens the way to polysaccharide engineering in *L. lactis* and provides a new approach to the study of polysaccharide biosynthesis genes of gram-positive cocci.

Acknowledgements

We thank Ingeborg Boels for assistance with the cloning of Sfi6 *epsE* and EPS quantifications. We are grateful to Peter Vandenbergh for providing us with strains MLT1, MLT2, and MLT3; Marc Kolkman for providing us with plasmid pMK100; Francesca Stingele for providing us with plasmid pFS30; and Willemiek van Casteren for sharing her results prior to publication. We thank Dick van den Berg and Roland Siezen for critically reading the manuscript.

Part of this work was supported by EC research grants 1116/92 1.6 and BIOT-CT96-0498.

References

Bourgoin, F. 1997. Ph.D. thesis. Université Henri Poincaré Nancy I, Vandœvre-lès-Nancy, France.

Cerning, J. 1990. Exocellular polysaccharides produced by lactic acid bacteria. FEMS Microbiol. Rev. **87:**113-130.

Coffey, T. J., M. C. Enright, M. Daniels, J. K. Morona, R. Morona, W. Hryniewicz, J. C. Paton, and B. G. Spratt. 1998 Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae.* Mol. Microbiol. **27:**73-83.

de Ruyter, P. G. G. A., O. P. Kuipers, M. M. Beerthuyzen, I. J. van Alen-Boerrigter, and W. M. de Vos. 1996. Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis.* J. Bacteriol. **178:**3434-3439.

de Ruyter, P.G.G.A., O. P. Kuipers, and W. M. de Vos. 1996. Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. Appl. Environ. Microbiol. **62:**3662-3667.

de Vos, W. M., P. Vos, H. de Haard, and I. Boerrigter. 1989. Cloning and expression of the *Lactococcus lactis* ssp. *cremoris* SK11 gene encoding an extracellular serine protease. Gene **85:**169-176.

Fürste, J.P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarian, and E. Lanka. 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. Gene **48:**119-131.

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.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol **166:**557-580.

Kolkman, M. A. B., D. A. Morrison, B. A. M. van der Zeijst, and P. J. M. Nuijten. 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: identification of the glycosyltransferase gene *cps14E.* J. Bacteriol. **178:**3736-3741.

Kolkman, M. A. B., B. A. M. van der Zeijst, and P. J. M. Nuijten. 1997. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J. Biol. Chem. **272:**19502-19508.

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.

Kuipers, O. P., P. G. G. A. de Ruyter, M. Kleerebezem, and W. M. de Vos. 1998. Quorum sensing-controlled gene expression in lactic acid bacteria. J. Biotechnol. **64:**15-21.

Maguin, E., H. Prévost, S. D. Ehrlich, and A. Gruss. 1996. Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. J. Bacteriol. **178:**931-935.

Morona, J. K., R. Morona, and J. C. Paton. 1997. Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthetic pathway. Mol. Microbiol. **23:**751-763.

Muñoz, R., M. Mollerach, R. López, and E. García. 1997. Molecular organization of the genes required for the synthesis of type 1 capsular polysaccharide of *Streptococcus pneumoniae*: formation of binary encapsulated pneumococci and identification of cryptic dTDP-rhamnose biosynthesis genes. Mol. Microbiol. **25:**79-92.

Nakajima, H., S. Toyoda, T. Toba, T. Itoh, T. Mukai, H. Kitazawa, and S. Adachi. 1990. A novel phosphopolysaccharide from slime-forming *Lactococcus lactis* subspecies *cremoris* SBT 0495. J. Dairy Sci. **73:**1472-1477.

Nakajima, H., T. Hirota, T. Toba, T. Itoh, and S. Adachi. 1992. Structure of the extracellular polysaccharide from the slime-forming *Lactococcus lactis* subsp. *cremoris* SBT0495. Carbohydr. Res. **224:**245-253.

Neve, H., A. Geis, and M. Teuber. 1988. Plasmid-encoded functions of ropy lactic acid streptococcal strains from Scandinavian fermented milk. Biochimie **70:**437-442.

Ouyang, S., and C. Y. Lee. 1997. Transcriptional analysis of type 1 capsule genes in *Staphylococcus aureus.* Mol. Microbiol. **23:**473-482.

Pollock, T. J., W. A. T. van Workum, L. Thorne, M. J. Mikolajczak, M. Yamazaki, J. W. Kijne, and R. W. Armentrout. 1998. Assignment of biochemical functions to glycosyltransferase genes which are essential for biosynthesis of exopolysaccharides in *Sphingomonas* strain S88 and *Rhizobium leguminosarum.* J. Bacteriol. **180:**586-593.

Poolman, B., and W. N. Konings. 1988. Relation of growth of *Streptococcus lactis* and *Streptococcus cremoris* to amino acid transport. J. Bacteriol. **170:**700-707.

Roberts, I.S. 1996. The biochemistry and genetics of capsular polysaccharide production in bacteria. Annu. Rev. Microbiol. **50:**285-315.

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.

Sau, S., and C. Y. Lee. 1996. Cloning of type 8 capsule genes and analysis of gene clusters for the production of different capsular polysaccharides in *Staphylococcus aureus.* J. Bacteriol. **178:**2118-2126.

Sau, S., J. Sun, and C. Y. Lee. 1997. Molecular characterization and transcriptional analysis of type 8 capsule genes in *Staphylococcus aureus.* J. Bacteriol. **179:**1614-1621.

Sau, S., N. Bhasin, E. R. Wann, J. C. Lee, T. J. Foster, and C. Y. Lee. 1997. The *Staphylococcus aureus* allelic genetic loci for serotype 5 and 8 capsule expression contain the type-specific genes flanked by common genes. Microbiology **143:**2395-2405.

Saxena, I. M., R. M. Brown, Jr., M. Fevre, R. A. Geremia, and B. Henrissat. 1995. Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. J. Bacteriol. **177:**1419-1424.

Stingele, 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.

Stingele, F. , S. J. F. Vincent, E. J. Faber, J. W. Newell, J. P. Kamerling, and J.-R. Neeser. 1999. Introduction of the exopolysaccharide gene cluster from Streptococcus thermophilus Sfi6 into Lactococcus lactis MG1363: production and characterization of an altered polysaccharide. Mol. Microbiol. **32:**1287-1295.

Stingele, F., and J.-R. Neeser. 1999. Glycosyltransferase genes epsE and epsG code for a galactosyl-1-phosphate transferase and an α-N-acetylgalactosaminyltransferase in *Streptococcus thermophilus* Sfi6. J. Bacteriol. **181:** in press.

van Casteren, W.H.M., M. A. Kabel, C. Dijkema, H. A. Schols, G. Beldman, and A. G. J. Voragen. 1999. EndoglucanaseV and a phosphatase from *Trichoderma viride* are able to act on modified exopolysaccharide from *Lactococcus lactis* subsp. *cremoris* B40. Carbohydr. Res. **137:**131-144.

van Casteren, W. H. M., C. Dijkema, H. A. Schols, G. Beldman, and A. G. J. Voragen. 1999. Structural characterization and modification of the exopolysaccharide from *Lactococcus lactis* subsp. *cremoris* B891. Manuscript in preparation.

van Kranenburg, R., J. D. Marugg, N. J. Willem, I. I. van Swam, and W. M. de Vos. 1997. Molecular characterization of the plasmid-encoded *eps* gene cluster essential for exopolysaccharide production in *Lactococcus lactis*. Mol. Microbiol. **24:**387-397.

van Kranenburg, R., I. I. van Swam, J. D. Marugg, M. Kleerebezem, and W. M. de Vos. 1999. Exopolysaccharide biosynthesis in *Lactococcus lactis* NIZO B40: functional analysis of the glycosyltransferase genes involved in the synthesis of the polysaccharide backbone. J. Bacteriol. **181:**338-340.

van Riel, J., and C. Olieman. 1991. Selectivity control in the anion-exchange chromatographic determination of saccharides in dairy products using pulsed amperometric detection. Carbohydr. Res. **215:**39-46.

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.

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.

Wang, L., D. Liu, and P. R. Reeves. 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J. Bacteriol. **178:**2598-2604.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13*mp*18 and pUC19 vectors. Gene **33:**103-119.

Characterisation of Multiple Regions Involved in Replication and Mobilisation of Plasmid pNZ4000 Coding for Exopolysaccharide Production in *Lactococcus lactis*

Chapter 5

Richard van Kranenburg, and Willem M. de Vos

Published in Journal of Bacteriology (1998) **180:**5285-5290.

Abstract

We characterised the regions involved in replication and mobilisation of the 40-kb plasmid pNZ4000, encoding exopolysaccharide (EPS) production in *Lactococcus lactis* NIZO B40. The plasmid contains four highly conserved replication regions with homologous *rep* genes (*repB1*, *repB2*, *repB3*, and *repB4*) that belong to the lactococcal theta replicon family. Subcloning of each replicon individually showed that all are functional and compatible in *L. lactis*. Plasmid pNZ4000 and genetically labelled derivatives could be transferred to different *L. lactis* strains by conjugation and pNZ4000 was shown to be a mobilisation plasmid. Two regions involved in mobilisation were identified near two of the replicons; both included an *oriT* sequence rich in inverted repeats. Conjugative mobilisation of the nonmobilisable plasmid pNZ124 was promoted by either one of these *oriT* sequences, demonstrating their functionality. One *oriT* sequence was followed by a *mobA* gene, coding for a *trans*-acting protein, which increased the frequency of conjugative transfer 100-fold. The predicted MobA protein and the *oriT* sequences show protein and nucleotide similarity, respectively, with the relaxase and with the inverted repeat and *nic* site of *oriT* from the *Escherichia coli* plasmid R64. The presence on pNZ4000 of four functional replicons, two *oriT* sequences, and several insertion sequence-like elements, strongly suggests that this EPS plasmid is a naturally occurring cointegrate.

Introduction

Lactococci are known to harbour conjugative plasmids that are used for industrial strain improvement since they encode important metabolic traits such as lactose fermentation, protease activity, bacteriophage resistance, or production of exopolysaccharide (EPS). Therefore, these plasmids are studied for their functional properties as well as for their mode of replication and transfer capacities. Two different mechanisms of replication are known to operate in *Lactococcus lactis*: rolling circle and theta replication. Rolling circle replication seems to be restricted to relatively small lactococcal plasmids with cryptic functions (Khan, 1997). Two of these, the related promiscuous plasmids pWV01 and pSH71, have been developed into widely used cloning and expression vectors (de Vos and Simons, 1994). The replication regions of several theta replicating lactococcal plasmids that encode metabolic functions have been analysed and all are members of a family of highly related, compatible theta replicons as first identified for plasmid pCI305 (Hayes *et al.*, 1991, Seegers *et al.*, 1994). They all contain a homologous *repB* gene encoding the replication protein. The conserved region upstream of *repB* is likely to include the origin of replication and also contains 22-bp repeats which have a replicon-specific regulatory role in plasmid replication and an inverted repeat overlapping the *repB* promoter which is a RepB binding site (Foley *et al.*, 1996).

The capacity for conjugal transfer is an important characteristic of some lactococcal plasmids. Self-transmissible conjugative plasmids have the ability to form effective cell-to cell contact, while mobilisation plasmids are only able to prepare their DNA for transfer (Steele and McKay, 1994). The conjugation process in gram-negative bacteria is initiated at the origin of transfer (*oriT*) by the formation of a relaxosome, usually containing a relaxase and accessory DNA-binding proteins. The relaxase catalyses the cleavage of a specific phosphodiester bond at the *nic* site in the *oriT*, after which it is covalently linked to the 5' end of the cleaved strand through a tyrosyl residue. Single-stranded DNA is transferred to the recipient cell and subsequently ligated through the cleaving-joining activity of the relaxase, resembling the process of leading strand replication by rolling circle replication (Lanka and Wilkins, 1995). To date, very little is known about genes required for conjugation in lactococci and other gram-positive bacteria (Gasson *et al.*, 1995). The chromosomally encoded sex factor and the homologous conjugative element pRS01 of *L. lactis* 712 and ML3, respectively, can mediate a high-frequency transfer of nonconjugative lactose plasmids and confer a cell aggregation (Clu) phenotype (Gasson and Davies, 1980, Anderson and McKay, 1984). The sex factor *cluA* gene encodes a protein that is involved in cell aggregation during conjugation (Godon *et al.*, 1994). On the bacteriophage resistance plasmid pCI528, a 2-kb region involved in conjugative mobilisation has been identified. It contains a putative *oriT* and a *mobA* gene which is predicted to encode a protein involved in mobilisation (Lucey *et al.*, 1993b).

While EPS production by lactococci has long been known to be a plasmid-encoded trait, it was only recently established that structural genes involved in EPS biosynthesis are located

Table 1. Strains and plasmids used in this study

Strain or plasmid	Relevant characteristics[a]	Reference
Strains		
E. coli DH5α		Hanahan, 1983
L. lactis NIZO B40	Lac$^+$ Eps$^+$ multiplasmid strain harbouring pNZ4000	van Kranenburg *et al.*, 1997
L. lactis MG1363	plasmid-free	(Gasson, 1983
L. lactis MG1614	Rif^r Str^r, plasmid-free	(Gasson, 1983)
L. lactis IL1403	plasmid-free	Chopin *et al.*, 1984
L. lactis NZ4010	Rif^r Str^r Eps$^+$, MG1614 harbouring pNZ4000	van Kranenburg *et al.*, 1997
Plasmids		
pCI182	Tet^r, 8.0-kb pBR322 derivative carrying the Tn*919 tetM* gene	Hill et al., 1988
pUC19Ery	Ery^r, 3.8-kb pUC19 carrying the Ery^R gene of pIL253	van Kranenburg *et al.*, 1997
pUC18Ery	Ery^r, 3.8-kb pUC18 carrying the Ery^R gene of pIL253	van Kranenburg *et al.*, 1997
pNZ4000	40-kb plasmid encoding EPS production	van Kranenburg *et al.*, 1997
pNZ4001	Ery^r, 7.1-kb derivative of pUC19Ery carrying *repB1*	This study
pNZ4002	Ery^r, 7.2-kb derivative of pUC19Ery carrying *repB2*	This study
pNZ4003	Ery^r, 7.5-kb derivative of pUC18Ery carrying *repB3*	This study
pNZ4004	Ery^r, 6.7-kb derivative of pUC19Ery carrying *repB4*	This study
pNZ4006	Ery^r, 5.0-kb derivative of pUC19Ery carrying a 1.2-kb *Eco*RI-*Xba*I fragment of pNZ4000 with *orfD1*	This study
pNZ4007	Ery^r, 7.9-kb derivative of pUC19Ery carrying a 4.1-kb *Eco*RI-*Xba*I fragment of pNZ4000 with *orfD2*	This study
pNZ4010	Ery^r, 45-kb, pNZ4000 containing an integrated copy of pNZ4006	This study
pNZ4017	Ery^r, 48-kb, pNZ4000 containing an integrated copy of pNZ4007	This study
pNZ124	Cm^r, 2.8-kb pSH71 replicon	Platteeuw *et al.*, 1993
pNZ4021	Cm^r, 3.6-kb derivative of pNZ124 carrying *oriT1*	This study
pNZ4022	Cm^r, 3.4-kb derivative of pNZ124 carrying *oriT2*	This study
pNZ4023	Cm^r, 5.8-kb derivative of pNZ124 carrying *oriT1* and *mobA*	This study
pNZ4025	Tet^r, 10.2-kb derivative of pUC18 carrying *repB1* and *tetM*	This study
pNZ4026	Tet^r, 11.4-kb derivative of pCI182 carrying *repB2*	This study
pNZ4027	Tet^r, 10.9-kb derivative of pCI182 carrying *repB4*	This study

[a] Lac$^+$, lactose fermenting; Eps$^+$, EPS producing; Rif^r, rifampin resistant; Str^r, streptomycin resistant; Tet^r, tetracycline resistant, Ery^r, erythromycin resistant; Cm^r, chloramphenicol resistant.

on these plasmids (van Kranenburg *et al.*, 1997, 1999). The best-characterised EPS plasmid to date is the 40-kb pNZ4000 from *L. lactis* NIZO B40, which contains a 12-kb gene cluster encoding EPS biosynthesis (van Kranenburg *et al.*, 1997). Furthermore, it contains multiple

replicons, since we were able to separate pNZ4000 in two *Xho*I-*Sph*I fragments that upon labelling with an erythromycin resistance (Ery^r) marker could each replicate in *L. lactis* (van Kranenburg *et al.*, 1997). In this study, we report the identification and characterisation of the regions involved in plasmid replication and mobilisation of this EPS plasmid. Plasmid pNZ4000 contains four functional replicons and two regions involved in mobilisation; one codes for an active *trans*-acting mobilisation protein, and both contain a *cis*-acting *oriT* region.

Material and methods

Bacterial strains, plasmids, and media. The bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* was grown in L-broth-based medium at 37°C (Sambrook *et al.*, 1989). *L. lactis* was grown at 30°C in M17 broth (Difco Laboratories) supplemented with 0.5% glucose (GM17). If appropriate, the media contained chloramphenicol (10 μg/ml), erythromycin (10 μg/ml for *L. lactis* and 150 μg/ml for *E. coli*), rifampin (50 μg/ml), streptomycin (100 μg/ml), tetracycline (25 μg/ml), or ampicillin (100 μg/ml).

DNA isolation, manipulation, and transfer. Isolation of *E. coli* plasmid DNA and standard recombinant DNA techniques were performed as described by Sambrook *et al.* (1989). Large scale isolation of *E. coli* plasmid DNA for nucleotide sequence analysis was performed with Qiagen columns, as instructed by the manufacturer. Isolation and transformation of *L. lactis* plasmid DNA were performed as previously described (de Vos *et al.*, 1989). For whole cell lysates of *L. lactis*, 1.5 ml of a late log-phase culture was harvested and suspended in 100 μl of a buffer containing 30 mM Tris-HCl (pH 8.0), 3 mM $MgCl_2$, 25% sucrose, 10 μg of lysozyme ml^{-1}, and 0.1 mg of RNase ml^{-1}. This suspension was incubated at 37°C for 30 min. Lysis was achieved by addition of 100 μl of 2% sodium dodecylsulfate and vortexing at top speed for 1 min, after which the lysate was treated with 20 μg of proteinase K ml^{-1} at 37°C for 30 min. Conjugation was performed by filter matings as described before (van Kranenburg *et al.*, 1997). The ratio of donor and recipient was 2:1.

Nucleotide sequence analysis. Automatic double stranded DNA sequence analysis was performed on both strands with an ALF DNA sequencer (Pharmacia Biotech). Sequencing reactions, performed with an AutoRead sequencing kit, were initiated by using fluorescein-labelled universal and reverse primers and continued with synthetic primers in combination with fluorescein-15-dATP, following the instructions of the manufacturer (Pharmacia Biotech). Sequence data were assembled and analysed using the PC/GENE program version 6.70 (IntelliGenetics). The GenBank Bacteria library (February 1998 release) was screened for homologies using TFASTA.

Construction of plasmids. For replicon screening and plasmid integration, the *E. coli* plasmid pUC19Ery or pUC18Ery, carrying the Ery^r gene, or pCI182, carrying the tetracycline

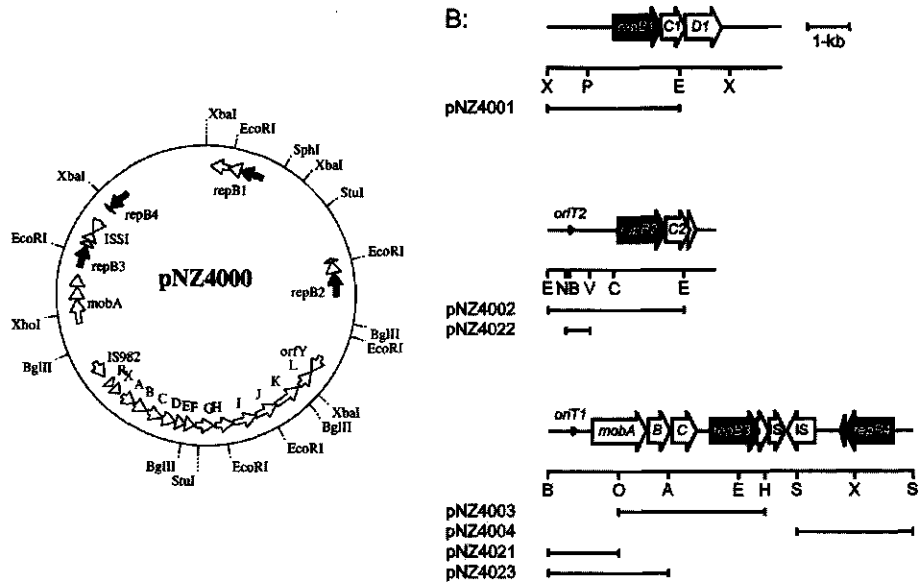


Fig. 1. (A) Physical and genetic map of plasmid pNZ4000. The *eps* gene cluster is located in-between IS*982* and *orfY*. (B) Physical and genetic maps of the replication and mobilisation regions of pNZ4000. The fragments used for functional analysis are depicted below. A, *Acc*I; C, *Cla*I; E, *Eco*RI; H, *Hin*cII; N, *Nco*I; O, *Xho*I; P, *Sph*I; S, *Sau*3AI; V, *Nsp*V; X, *Xba*I. For *Acc*I, *Cla*I, *Hin*cII, *Nco*I, *Nsp*V, and *Sau*3AI, only sites relevant for subcloning are included. Sequences are available under GenBank accession numbers AF036485, AF03686, and AF03687.

resistance (Tet^r) gene were used. For plasmids pNZ4001, pNZ4002, and pNZ4004, a 3.3-kb *Eco*RI-*Xba*I fragment, a 3.4-kb *Eco*RI fragment, and a 2.9-kb *Sau*3AI fragment of pNZ4000 were cloned into pUC19Ery digested with *Eco*RI-*Xba*I, *Eco*RI, and *Bam*HI, respectively. For plasmid pNZ4003, a 3.7-kb *Xho*I-*Hin*cII fragment of pNZ4000 was cloned into *Sal*I-*Sma*I digested pUC18Ery. To construct plasmid pNZ4025, a 3.3-kb *Eco*RI-*Xba*I fragment of pNZ4000 was cloned in pUC18 (Yanisch-Perron *et al.*, 1985) digested with *Eco*RI-*Xba*I, and subsequently the pCI182 *tetM* gene was cloned on a 4.2-kb *Hin*cII fragment in the pUC18 *Hin*cII site. For plasmids pNZ4026 and pNZ4027, a 3.4-kb *Eco*RI and a 2.9-kb *Sau*3AI fragment of pNZ4000 were cloned in pCI182 digested with *Eco*RI or *Bgl*II, respectively.

To obtain Ery^r derivatives of pNZ4000 (pNZ4010 and pNZ4017), plasmids pNZ4006 and pNZ4007 were constructed. Plasmids pNZ4006 and pNZ4007 are pUC19Ery derivatives carrying 1.2- and 4.1-kb *Eco*RI-*Xba*I fragments of pNZ4000, respectively. These plasmids were used for plasmid integration by a single crossover to form pNZ4010 and pNZ4017, respectively.

For functional analysis of the putative *oriT* regions, fragments containing the *oriT1* or *oriT2* sequence were cloned in plasmid pNZ124. Plasmid pNZ4021, carrying *oriT1*, was constructed by cloning a 1.8-kb *Bgl*II-*Xho*I fragment of pNZ4000 in *Bgl*II-*Xho*I digested pNZ124. Plasmid pNZ4022, carrying *oriT2*, was constructed by cloning a Klenow enzyme-treated 0.64-kb *Nsp*V-*Nco*I fragment of pNZ4000 in pNZ124 linearised with *Sca*I. To study

the functionality of *mobA*, plasmid pNZ4023, carrying both *oriT1* and *mobA*, was constructed by cloning a 3.0-kb *Bgl*II-*Acc*I fragment of pNZ4000 with a Klenow enzyme-treated *Acc*I site, in *Bgl*II-*Sca*I-digested pNZ124. All plasmids were constructed in *E. coli*.

Nucleotide sequence accession numbers. The complete nucleotide sequences of the replication and mobilisation regions are available under GenBank accession no. AF03685, AF03686, and AF03687.

Results and discussion

The EPS plasmid pNZ4000 contains four functional replicons. The 40-kb plasmid pNZ4000 is essential for EPS production in strain NIZO B40 and includes the 12-kb *eps* gene cluster involved in EPS biosynthesis (van Kranenburg *et al.*, 1997). The nucleotide sequence of the EPS plasmid was determined, and analysis of the data revealed the unusual presence of four highly homologous replication regions that belong to a family of lactococcal theta replicons (Seegers *et al.*, 1994) which are located outside the *eps* gene cluster (Fig. 1). DNA fragments carrying these putative replicons were cloned into pUC19Ery or pUC18Ery, which can be used as replicon screening vectors in *L. lactis*. The resulting plasmids (pNZ4001, pNZ4002, pNZ4003, and pNZ4004) were transformed to *L. lactis* MG1363, and in all cases Ery^r transformants that harboured plasmids with the expected configuration were obtained (results not shown). These results indicate that all four replicons are functional in *L. lactis*. Since plasmid replication requires only one of these replicons, pNZ4000 must have derived fragments of several plasmids, which might have formed cointegrates during conjugation processes. This conclusion is corroborated by the presence of complete and truncated copies of IS*S1*-like elements (Fig 1), since it is known that IS*S1* mediates cointegration of the *L. lactis* ML3 lactose plasmid pSK08 with the conjugal plasmid pRS01 (Polzin and Shimizu-Kadota, 1987). In addition, a complete copy of an IS*982*-like element is present on pNZ4000 (Fig. 1).

GenBank analysis of the proteins encoded by the *repB* genes of the four replicons on pNZ4000 showed them to be highly homologous to putative replication proteins of several other lactococcal plasmids which all carry a single replicon and belong to a family of lactococcal theta replicons (Seegers *et al.*, 1994), including pVS40 (86.0% identity with RepB1) (Von Wright and Raty, 1993), pWV04 (98.5% identity with RepB2) (Seegers *et al.*, 1994), pCI528 (99.8% identity with RepB3) (Lucey *et al.*, 1993a), and pFV1201 (99.2% identity with RepB4) (GenBank accession no. X96949). The upstream regions of the *repB* genes of pNZ4000 were highly conserved and corresponded to those found in the other lactococcal replicons as first identified for pCI305 (Hayes *et al.*, 1991). They all contain an A/T rich region that could be the recognition site for host-encoded functions involved in replication (Seegers *et al.*, 1994), a 22-bp sequence repeated 3.5 times which was shown to have a replicon-specific regulatory role in plasmid replication (Foley *et al.*, 1996), and two inverted repeats, one of which overlapped the -35 region of the *repB* promoter (inverted

Table 2. Segregational stability of the four replicons of pNZ4000

Replicon		Fraction of plasmid-containing cells[a]		
Tet^r	Ery^r	Tet^r	Ery^r	$Tet^r + Ery^r$
repB1	*repB2*	0.8	0.9	0.9
repB1	*repB3*	0.9	0.01	0.01
repB4	*repB1*	1.0	1.0	0.9
repB2	*repB3*	0.7	0.05	0.06
repB2	*repB4*	0.8	0.9	0.7
repB4	*repB3*	0.9	0.08	0.07

[a] After 35 generations without selection pressure, cultures containing the Tet^r and Ery^r replicons were plated on medium containing tetracycline, erythromycin, both tetracycline and erythromycin or no antibiotics, and the fraction of antibiotic-resistant colonies was determined.

repeat 1[IR1]) and was found to be a RepB binding site (Foley *et al.*, 1996). The upstream region of *repB1* showed a slightly different architecture and contained only a 2.5-times-repeated 22-bp direct repeat.

All four replicons are compatible but show differences in organisation. The minimal replicons for *repB1*, *repB2*, and *repB4* were labelled with the *tetM* gene to generate pNZ4025, pNZ4026, and pNZ4027, respectively. These plasmids were combined with either pNZ4001, pNZ4002, pNZ4003, or pNZ4004 and transformed to MG1363 to make six strains including all combinations of different replicons carrying a set of Ery^r and Tet^r genes (Table 2). All plasmids had comparable copy numbers, as judged from the intensity of ethidium bromide stained plasmid DNA separated by agarose gel electrophoresis (results not shown). Stable transformants were obtained for all heteroplasmid combinations following selection for Ery^r and Tet^r indicating that these replicons are compatible. The compatibility of the plasmids carrying different replicons was confirmed by determining the segregational stability after growth for 35 generations in medium containing no antibiotics (Table 2). Plasmids carrying the replicons with *repB1*, *repB2*, and *repB4* formed highly stable heteroplasmid combinations. In contrast, the segregational stability of the *repB3* containing replicon was significantly lower than that of the others. This was also observed when this replicon was present as a single replicon in MG1363. After 20 generations without selection pressure, 61% of the population was plasmid containing; after 40 generations 16%, and after 60 generations only 3% of the population contained plasmids. The reason for the difference in stability between the *repB3* containing replicon and the other three highly homologous replicons is unclear. It seems that there is interference with the maintenance functions of the *repB3*-containing replicon which are not directly involved in replication. The *orfC* genes located downstream of and partly overlapping the *repB* genes (Fig. 1) are not likely to be involved in this process. The predicted OrfC proteins are homologous to RepB287 (45 and 43% identity for OrfC1 and OrfC2, respectively) encoded by the *Tetragenococcus halophilus* theta-replicating plasmid pUCL287. RepB287 is not essential for replication, as is OrfC, but its presence reduces the copy number and the segregational stability (Benachour *et al.*, 1997).

The N-terminal parts of the OrfC proteins are highly conserved and contain a helix-turn-helix motif which is probably involved in DNA binding. While *repB1* and *repB2* are followed by almost complete *orfC* genes, *orfC3* and *orfC4* encode only the N-terminal parts of OrfC-like proteins. If the role of the lactococcal OrfC were similar to that of RepB287, we would expect the stability of the replicons containing *repB1* and *repB2* to be lower than that of the replicons containing *repB3* and *repB4*, which is not as we observed.

Downstream of *orfC1* and *orfC2*, we found a partly overlapping third ORF (*orfD1* and *orfD2*, respectively [Fig. 1]). The predicted gene product OrfD1 shows considerable homology (47% identity) to the product of an *hsdS*-like gene from the lactococcal plasmid pIL2614, which encodes the specificity subunit of a type IC restriction-modification system (Schouler *et al.*, 1998). The *hsdS*-like gene is the last of a putative operon of five genes, the first two of which are replication genes homologous to *repB* and *orfC*. These are followed by three genes coding for the endonuclease, methylase, and specificity subunits, respectively, of a type I restriction-modification system (Schouler *et al.*, 1998). This finding indicates that pNZ4000 and pIL2614 contain similarly organised and homologous operons, the one in pNZ4000 lacking the genes encoding the endonuclease and methylase subunits.

The EPS plasmid pNZ4000 is a mobilisation plasmid. We have previously shown that plasmid pNZ4000 can be conjugally transferred together with the lactose plasmid from the *L. lactis* strain NIZO B40 to the recipient strain MG1614 (van Kranenburg *et al.*, 1997). To study the intraspecific conjugative transfer of pNZ4000 in more detail, the Ery^r derivatives pNZ4010 and pNZ4017 were used. These plasmids were transformed to the plasmid-free strain MG1363, and the resulting strains were used as donors in filter matings with strain MG1614. Conjugative transfer of either of these plasmids between these isogenic *L. lactis* subsp. *cremoris* strains occurred at a frequency of 10^{-6} per donor. Plasmid pNZ4017 was also transformed to the plasmid-free *L. lactis* subsp. *lactis* strain IL1403, from which it could be transferred to *L. lactis* subsp. *cremoris* strain MG1614 at a frequency of 10^{-8} per donor. These results demonstrate that pNZ4000 can be mobilised from strains MG1363 and IL1403. It is likely that differences in chromosomal conjugation functions account for the differences in transfer efficiency of the pNZ4000 derivatives from both *L. lactis* subspecies, which are known to share approximately 70 to 80% sequence identity in characterised genes and differ by the presence of a large chromosomal inversion of about half of the genome (Godon *et al.*, 1992, Le Bourgeois *et al.*, 1995). Furthermore, MG1363 harbours the sex factor that encodes conjugative functions (Gasson *et al.*, 1995), which may play a role in mobilisation of pNZ4000.

pNZ4000 contains two functional *oriT* sites. Mobilisation involves a *cis*-acting *oriT* region and a *trans*-acting gene encoding a relaxase (Lanka and Wilkins, 1995). Nucleotide sequence analysis of pNZ4000 revealed the presence of a region upstream of *repB3* (Fig. 1), which is almost identical (98.3% identity) to a 2.0-kb fragment involved in mobilisation of the lactococcal plasmid pCI528 (Lucey *et al.*, 1993b). It contains a *mobA* gene encoding a putative mobilisation protein. The upstream region of the *mobA* gene contains three inverted repeats and a direct repeat and has been postulated to be the *oriT* region (Lucey *et al.*,

Table 3. Transfer frequencies of pNZ124 derivatives from MG1363 to MG1614

Plasmid(s)	Genotype				Transfer frequency[a]
	oriT1	*oriT2*	*mobA* in *cis*	*mobA* in *trans*	
pNZ124	–	–	–	–	$<10^{-10}$
pNZ4021	+	–	–	–	10^{-7}
pNZ4021, pNZ4017	+	–	–	+	10^{-5}
pNZ4022	–	+	–	–	10^{-7}
pNZ4022, pNZ4017	–	+	–	+	10^{-5}
pNZ4023	+	–	+	–	10^{-5}

[a] Number of transconjugants per donor (average of two independent experiments).

1993b). We tested the functionality of the putative *oriT* sequence (*oriT1*) by cloning it in the nonconjugative plasmid pNZ124 and transforming the resulting plasmid pNZ4021 to strain MG1363. This strain was mated with MG1614 and chloramphenicol-resistant transconjugants were selected (Table 3). The 1.8-kb region containing the *oriT1* sequence was sufficient to achieve conjugal transfer of the nonconjugative plasmid pNZ124, showing that the cloned fragment contains a functional *oriT*.

A second *oriT* region sharing 96.6% identity in 417 nucleotides with *oriT1* was found upstream of *repB2*. It was cloned as a 0.64-kb fragment in pNZ124, and the resulting plasmid, pNZ4022, had the same transfer frequency as pNZ4021 (Table 3), indicating the presence of two functional *oriT* sequences on pNZ4000, one upstream of *mobA* (*oriT1*) and one upstream of *repB2* (*oriT2*) (Fig. 1) which are situated in opposite direction.

The *oriT* site of the streptococcal plasmid pMV158 is homologous to sequences of several plasmids from gram-positive hosts (Guzmán *et al.*, 1997). However, no significant homology between the *oriT* regions of pNZ4000 and these sequences could be detected. In contrast, the pNZ4000 *oriT* sequences contain an inverted repeat (IR3) which is highly homologous to that of the *oriT* from IncI1 plasmid R64 (Fig. 2). This includes the R64 mobilisation protein NikA binding site (Furuya and Komano, 1997). Moreover, the homology between the pNZ4000 *oriT* sequences and that of R64 also includes the sequence next to the repeat containing the *nic* site (Fig. 2). In the absence of experimental evidence, we therefore postulate that these sequences may contain the pNZ4000 *nic* sites. The streptococcal plasmid pIP501 and the staphylococcal plasmid pGO1 *oriT* regions are homologous to *oriT* sequences of several gram-negative plasmids. They all contain a conserved sequence with the *nic* site next to a nonconserved inverted repeat centred around the nucleotide sequence 5'-GAA-3' (Climo *et al.*, 1996, Wang and Macrina, 1995). Although no significant homology between these *oriT* regions and those of pNZ4000 could be detected, the IR3 sequence of each of the pNZ4000 *oriT* regions is also situated around a 5'-GAA-3' nucleotide sequence.

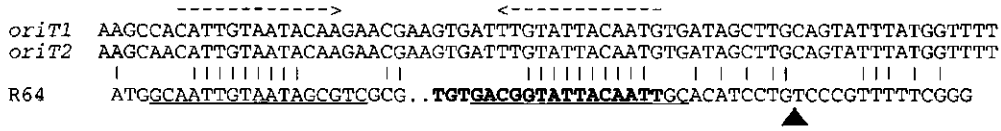


Fig. 2. DNA sequence alignment of the inverted repeat 3 (IR3) part of the *oriT* regions found on pNZ4000 (*oriT1* and *oriT2*) and the sequence of the IncI1 plasmid R64 *oriT*. For R64 *oriT*, the inverted repeat is underlined, the NikA binding site is indicated in bold and the *nic* site is indicated with an arrowhead (Furuya and Komano, 1997).

***mobA* encodes a product *trans*-acting on *oriT*-carrying plasmids.** The involvement of *mobA* in mobilisation was studied by comparing the transfer frequencies of plasmids carrying only *oriT* sequences or carrying *oriT* and *mobA* either in *cis* or in *trans* (Table 3). When *mobA* was provided in *trans* on pNZ4017, the transfer frequencies of pNZ4021 and pNZ4022 increased significantly. The same effect was achieved by when *mobA* was present in *cis* as on plasmid pNZ4023, containing *oriT1* and *mobA*. These results indicate that *mobA* encodes a *trans*-acting element involved in mobilisation.

To verify the relaxation activity of the *mobA* gene product (Novick, 1976), whole-cell lysates of MG1363 harbouring *oriT1*- or *oriT2*-carrying plasmids with or without *mobA* (in *cis* or in *trans*) were separated by agarose gel electrophoresis. The plasmid profiles of pNZ4021 and pNZ4022 showed a significant increase in open circular plasmid DNA only when pNZ4017 was present (approximately half of the *oriT*-carrying plasmids were in the open circular form). Moreover, pNZ4023 carrying *oriT1* and *mobA* showed a similar high degree of open circular DNA (data not shown). These results indicate that the plasmids carrying *oriT* fragments are relaxed by the *trans*-acting *mobA* gene product.

The predicted MobA protein reveals significant homologies (approximately 30% identity) with three mobilisation proteins found on antibiotic resistance plasmids of *Staphylococcus aureus* (Projan and Novick, 1988, Projan *et al.*, 1988, Projan and Archer, 1989) and shares moderate homology (23% identity in 388 amino acids) with the N-terminal part of TraI from the *E. coli* IncPα plasmid RP4. TraI is a relaxase and forms together with TraJ the relaxosome at *oriT* (Pansegrau *et al.*, 1990). TraI contains three conserved regions found in several relaxases (Pansegrau *et al.*, 1994). Motifs I and III are involved in catalysing the cleaving-joining reaction. Motif I contains a conserved tyrosine residue which after nicking is covalently attached to the 5' end of the cleaved DNA. Motif III contains a conserved histidine residue that is likely to activate the tyrosine of motif I by proton extraction. Motif II contains a conserved serine and is thought to be involved in DNA recognition (Pansegrau *et al.*, 1994). Multiple sequence alignment of MobA, the four homologous proteins, and the *E. coli* plasmid R64 relaxase NikB, which is homologous to TraI (Furuya *et al.*, 1991), showed that the three conserved domains and the tyrosine, serine, and histidine residues needed for relaxase activity are present in MobA (Fig. 3). This conservation strongly suggests that the lactococcal MobA is a relaxase which is involved in nicking the *nic* sites of the *oriT*

```
Mob (4-28)   TKLGNTKSASRAINYA...EKRAEEKSG (36-70)  AKSAFKQTRALYGKE..DGIQAHTVIQSFKPGE.VTPE
Rlx (4-28)   TKLGNTKSASRAINYA...EERAEEKSG (36-70)  AKSYFKQTRALYGKE..NGVQAHTVIQSFKPGE.VTAK
Orf1 (4-28)  TKISSTKSTSRAINYA...EKRAEEKSA (36-70)  AKSSFKATREMYGKT..DGNEGHVVIQSFKPNE.VTPE
MobA (5-32)  AKISNGASAASALNYALGQDRPMHEKTE (65-100) AKEQFDVVRQLHNQTK.ESNQVLRITQSFALDE.LNPK
NikB (42-66) AEQPHRSRFSRLVDYA...TRLRNESFV (97-133) AADMEYIARQAHYAKDDTDPVFHYII.SWQSHESPRPE
TraI (8-32)  MRSIKKSDFAELVKYI...TDEQGKTER (48-84)  AVMAEVMATQHGNTRSEADKTYHLLV.SFRAGEKPDAE
                                  Motif I                                  Motif II

Mob (71-121)  QCNQLGLELAEKIAPNHQVAVYTHTDKDH..YHNHIVINSVDLETGKKYQSNK
Rlx (71-121)  ECNEIGLELAKKIAPDYQVAVYTHTDKDH..YHNHIIINSVNLETGNKYQSNK
Orf1 (71-121) QCNQLGLELAEKIAPNHQVAVYTHNDTDH..VHNHIVINSIDLETGKKFNNNK
MobA (107-159) KANDLGVELAENLYPNHQSAVYTHLDGKNHVLHNHIIVNKVNLETGKKLREQK
NikB (140-185) RHTLKSLGLAD.....HQYVSAVHTDTDN..LHVHVAVNRVHPETGYLNRLSW
TraI (91-136)  DRICAGLGFAE.....HQRVSAVHHDTDN..LHIHIAINKIHPTRNTIHEPYR
                                  Motif III
```

Fig. 3. Amino acid sequence comparison of the three conserved regions involved in relaxase activity as determined for TraI (Pansegrau *et al.*, 1994) for the relaxases MobA (Mob), Rlx, and Orf1 from *Staphylococcus aureus* plasmids pC221, pS194, and pC223, respectively, MobA (MobA) from pNZ4000, NikB from *E. coli* plasmid R64 (Furuya *et al.*, 1991), and TraI from the *E. coli* plasmid RP4. The tyrosine (MotifI) and serine (MotifIII) residues involved in cleaving-joining reaction, and the histidine residue (MotifII) involved in DNA binding are indicated in bold.

sequences (Fig. 2), which is corroborated by the formation of open circular DNA of plasmids carrying an *oriT* sequence when *mobA* is present (see above).

On pNZ4000, a second ORF, here designated *mobB*, was found downstream of *mobA*, the putative start codon of which overlaps with the stop codon of *mobA*. This configuration resembles that of the *S. aureus* plasmid pC223, which contains two overlapping mobilisation genes *orf1* and *orf2* (Projan and Novick, 1988, Projan and Archer, 1989). In addition to the homologous Orf1 and MobA proteins, the predicted MobB protein shares moderate homology (24% identity) with Orf2 of pC223. A third ORF, designated *mobC*, was detected 16 bp downstream of the stop codon of *mobB*. Its gene product showed no homology to any protein in the GenBank database and the involvement of *mobC* in the conjugation process remains to be established.

The region on pNZ4000 containing the mobilisation genes and the third replicon has a high degree of homology with the same regions on pCI528. Plasmid pCI528 is a 46-kb plasmid encoding the production of a hydrophilic polymer containing glucose and rhamnose that reduces phage adsorption to its lactococcal host (Lucey *et al.*, 1992). Although pCI528 does not encode EPS production whereas pNZ4000 does, there may be a close relationship between the two plasmids or their ancestors.

In summary, we demonstrated that plasmid pNZ4000 contains four homologous and active replicons, that are compatible with each other. It contains two functional *oriT* sequences. One *oriT* is followed by the *mobA* gene coding for a *trans*-acting protein. The predicted MobA protein and the *oriT* sequences are homologous to the R64 relaxase and the *oriT*. The R64 relaxase is known to nick a site which is also conserved in the *oriT* sequences of pNZ4000.

Acknowledgements

This work was partly supported by European Community research grant 1116/92 1.6.

We thank Norwin Willem and Sónia Mendes for their technical assistance in the sequencing of the replicons. We are grateful to Joey Marugg for advice at the initial stages of this work. We acknowledge Michiel Kleerebezem and Roland Siezen for critically reading the manuscript.

References

Anderson, D. G., and L. L. McKay. 1984. Genetic and physical characterization of recombinant plasmids associated with cell aggregation an high-frequency conjugal transfer in *Streptococcus lactis* ML3. J. Bacteriol. **158:**954-962

Benachour, A., J. Frère, G. Novel, and Y. Auffray. 1997. Molecular analysis of the replication region of the theta-replicating plasmid pUCL287 from *Tetragenococcus* (*Pediococcus*) *halophilus* ATCC33315. Mol. Gen. Genet. **255:**504-513.

Chopin, A., M. C. Chopin, A. Moillo-Batt, and P. Langella. 1984. Two plasmid-determined restriction and modification systems in *Streptococcus lactis.* Plasmid **11:**260-263.

Climo, M. W., V. K. Sharma, and G. L. Archer. 1996. Identification and characterization of the origin of conjugative transfer (*oriT*) and a gene (*nes*) encoding a single-stranded endonuclease on the staphylococcal plasmid pGO1. J. Bacteriol. **178:** 4975-4983.

de Vos, W. M., P. Vos, 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.

de Vos, W. M., and G. F. M. Simons. 1994. Gene cloning and expression systems in *Lactococci*, p. 52-105. *In* M. J. Gasson and W. M. de Vos (ed.), Genetics and biotechnology of lactic acid bacteria. Chapman and Hall, London, England.

Foley, S., S. Bron, G. Venema, C. Daly, and G. F. Fitzgerald. 1996. Molecular analysis of the replication origin of the *Lactococcus lactis* plasmid pCI305. Plasmid **36:**125-141.

Furuya, N., T. Nisioka, and T. Komano. 1991. Nucleotide sequence and functions of the *oriT* operon in IncI1 plasmid R64. J. Bacteriol. **173:**2231-2237.

Furuya, N., and T. Komano. 1997. Mutational analysis of the R64 *oriT* region: requirement for precise location of the NikA-binding sequence. J. Bacteriol. **179:**7291-7297.

Gasson, M. J., and F. L. Davies. 1980. Conjugal transfer of the drug resistance plasmid pAMβ1 in the lactic streptococci. FEMS Microbiol. Lett. **7:**51-53.

Gasson, M. J. 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. **154:**1-9.

Gasson, M. J., J.-J Godon, C. J. Pillidge, T. J. Eaton, K. Jury and C. A. Shearman. 1995. Characterization and exploitation of conjugation in *Lactococcus lactis.* Int. Dairy J. **5:**757-762.

Godon, J.-J., C. Delorme, S. D. Ehrlich, and P. Renault. 1992. Divergence of genomic sequences between *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*. Appl. Environ. Microbiol. **58:**4045-4047.

Godon, J.-J., K. Jury, C. A. Shearman, and M. J. Gasson. 1994. The *Lactococcus lactis* sex-factor aggregation gene *cluA*. Mol. Microbiol. **12:**655-663.

Guzmán, L. M., and M. Espinosa. 1997. The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid *oriT*. J. Mol. Biol. **266:**688-702.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557-580.

Hayes, F., P. Vos, G. F. Fitzgerald, W. M. de Vos, and C. Daly. 1991. Molecular organization of the minimal replicon of novel narrow-host-range, lactococcal plasmid pCI305. Plasmid **25:**16-26.

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. Micriobiol. **54:**1230-1236.

Khan, S. A. 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. **61:**442-455.

Lanka, E., and B. M. Wilkins. 1995. DNA processing reactions in bacterial conjugation. Annu. Rev. Biochem. **64:**141-169.

Le Bourgeois, P., M. Lautier, L. van den Berghe, M. J. Gasson, and P. Ritzenthaler. 1995. Physical and genetic map of the *Lactococcus lactis* subsp. *cremoris* MG1363 chromosome: comparison with that of *Lactococcus lactis* subsp. *lactis* IL1403 reveals a large chromosomal inversion. J. Bacteriol. **177:**2840-2850.

Lucey, M., C. Daly, and G. F. Fitzgerald. 1992. Cell surface characteristics of *Lactococcus lactis* harbouring pCI528, a 46 kb plasmid encoding inhibition of bacteriophage adsorption. J. Gen. Microbiol. **138:**2137-2143.

Lucey, M., C. Daly, and G. F. Fitzgerald. 1993. Identification and sequence analysis of the replication region of the phage resistance plasmid pCI528 from *Lactococcus lactis* subsp. *cremoris* UC503. FEMS Microbiol. Lett. **10:**249-256.

Lucey, M., C. Daly, and G. Fitzgerald. 1993. Analysis of a region from the bacteriophage resistance plasmid pCI528 involved in its conjugative mobilization between *Lactococcus* strains. J. Bacteriol. **175:**6002-6009.

Novick, R. 1976. Plasmid-protein relaxation complexes in *Staphylococcus aureus*. J. Bacteriol. **127:**1177-1187.

Pansegrau, W., G. Ziegelin, and E. Lanka. 1990. Covalent association of the *traI* gene product of plasmid RP4 with the 5'-terminal nucleotide at the relaxation nick site. J. Biol. Chem. **265:**10637-10644.

Pansegrau, W.,W. Schröder, and E. Lanka. 1994. Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. J. Biol. Chem. **269:**2782-2789.

Platteeuw, C., G. Simons, and W. M. de Vos. 1993. Use of the *Escherichia coli* β-glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria. Appl. Environ. Microbiol. **60:**587-593.

Polzin, K. M., and M. Shimizu-Kadota. 1987. Identification of a new insertion element, similar to gram-negative IS*26*, on the lactose plasmid *of Streptococcus lactis* ML3. J. Bacteriol. **169:**5481-5488.

Projan, S. J., and R. Novick. 1988. Comparative analysis of five related staphylococcal plasmids. Plasmid **19:**203-221.

Projan, S. J., Moghazeh, and R. P. Novick. 1988. Nucleotide sequence of pS194, a streptomycin-resistance plasmid from *Staphylococcus aureus*. Nucleic Acids Res. **16:**2179-2187.

Projan, S. J., and G. L. Archer. 1989. Mobilization of the relaxable *Staphylococcus aureus* plasmid pC221 by the conjugative plasmid pGO1 involves three pC221 loci. J. Bacteriol. **171:**1841-1845.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Schouler, C., F. Clier, A. L. Lerayer, S. D. Ehrlich, and M.-C. Chopin. 1998. A type IC restriction-modification system in *Lactococcus lactis*. J. Bacteriol. **180:**407-411.

Seegers, J. F. M., S. Bron, C.M. Franke, G. Venema, and R. Kiewiet. 1994. The majority of lactococcal plasmids carry a highly related replicon. Microbiology **140:**1291-1300.

Steele, J. L., and L. L. McKay. 1989. Conjugal transfer of genetic material in lactococci: a review. J. Dairy Sci. **72:**3388-3397.

van Kranenburg, R., J. D. Marugg, N. J. Willem, I. I. van Swam, and W. M. de Vos. 1997. Molecular characterization of the plasmid-encoded *eps* gene cluster essential for exopolysaccharide production in *Lactococcus lactis*. Mol. Microbiol. **24:**387-397.

van Kranenburg, R., H. R. Vos, I. I. van Swam, M. Kleerebezem, and W. M. de Vos. 1999. Functional analysis of glycosyltransferase genes from Lactococcus lactis and other gram-positive cocci: complementation, expression, and diversity. J. Bacteriol. **181:**6347-6353.

Von Wright, A., and K. Raty. 1993. The nucleotide sequence for the replication region of pVS40, a lactococcal food grade cloning vector. Lett. Appl. Microbiol. **17:**25-28.

Wang, A., and F. L. Macrina. 1995. Streptococcal plasmid pIP501 has a functional *oriT* site. J. Bacteriol. **177:**4199-4206.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. **Gene 33:**103-119.

Nucleotide Sequence Analysis of the Lactococcal EPS Plasmid pNZ4000

Chapter 6

Richard van Kranenburg, Michiel Kleerebezem, and Willem M. de Vos

Submitted for publication

Abstract

The complete 42180-bp nucleotide sequence of the mobilisation plasmid pNZ4000, coding for exopolysaccharide (EPS) production in *Lactococcus lactis*, was determined. This plasmid contains a region involved in EPS biosynthesis, four functional replicons, a region containing mobilisation genes, and three origin of transfer (*oriT*) sequences. Sequences identical to these *oriT* sequences were also found on two other lactococcal plasmids and a plasmid from *Lactobacillus helveticus*. Several complete and partial IS elements were identified on pNZ4000, including iso-ISS*1*, iso-IS*946*, and iso-IS982 sequences. Furthermore, pNZ4000 contains a gene cluster that may encode a cobalt transport system and a gene that encodes a CorA homologue which may function as a magnesium transporter.

Introduction

Lactococcus lactis strains used in food industry are known to harbour several endogenous plasmids. These may encode industrially significant traits like lactose fermentation, proteolytic activity, bacteriophage resistance, bacteriocin production, or exopolysaccharide (EPS) production. Some of these plasmids can be conjugally transferred, and self-transmissible conjugative plasmids, which have the ability to form effective cell-to-cell contact, and mobilisation plasmids, that are only able to prepare their DNA for transfer, can be distinguished (Steele and McKay, 1989).

Complete nucleotide sequences have been determined for a limited number of lactococcal plasmids. Next to several small cryptic plasmids that replicate via the rolling circle mechanism, like pSH71 (2059 bp) and pWV01 (2178 bp), which have been developed into widely used cloning and expression vectors (de Vos and Simons, 1994), only two larger plasmids, which are expected to replicate via the theta mechanism, have been analysed at the sequence level. Plasmid pPF107-3 (29871 bp), isolated from *L. lactis* subsp. *lactis* K214, encodes streptomycin, tetracycline, and chloramphenicol resistance (Perreten *et al.*, 1997) and the self-transmissible plasmid pMRC01 (60232 bp), isolated from *L. lactis* subsp. *lactis* DPC3147, encodes bacteriophage resistance and bacteriocin production (Dougherty *et al.*, 1998).

This report describes the analysis of the complete nucleotide sequence of plasmid pNZ4000 isolated from *L. lactis* subsp. *cremoris* NIZO B40. Previously, we characterised the *eps* gene cluster and the regions involved in mobilisation and replication (van Kranenburg *et al.*, 1997, 1998, 1999). Analysis of the complete sequence revealed the presence of a third origin of transfer (*oriT*) sequence, additional IS-sequences, and several genes that may be involved in transport of divalent cations.

Materials and methods

Bacterial strains and media. *Escherichia coli* MC1061 (Casadaban and Cohen, 1980) was grown in Luria (L)-broth-based medium at 37°C (Sambrook *et al.*, 1989). *L. lactis* NZ4010 harbouring pNZ4000 (van Kranenburg *et al.*, 1997) was grown at 30°C in M17 broth (Difco Laboratories) supplemented with 0.5% glucose (GM17). If appropriate, the media contained rifampicin (50 μg/ml), streptomycin (100 μg/ml), or ampicillin (100 μg/ml).

DNA isolation and manipulation. Isolation of *E. coli* plasmid DNA and standard recombinant DNA techniques were performed as described by Sambrook *et al.* (1989). Large scale isolation of *E. coli* plasmid DNA for nucleotide sequence analysis was performed with Qiagen columns, following the instructions of the manufacturer (Qiagen Inc.). Isolation and transformation of *L. lactis* plasmid DNA were performed as previously described (De Vos *et al.*, 1989).

Nucleotide sequence analysis. Automatic double-stranded DNA sequence analysis was

performed on both strands with an ALF*red* DNA sequencer (Pharmacia Biotech). Sequencing reactions were accomplished using the AutoRead sequencing kit, initiated by using Cy5-labelled universal and reverse primers and continued with synthetic primers in combination with Cy5-13-dATP following the instructions of the manufacturer (Pharmacia Biotech). Sequence data were assembled and analysed using the PC/GENE program version 6.70 (IntelliGenetics). The SWALL and EMBL prokaryote libraries (version 3.2t05 May 1999) were screened for homologies using the Fasta3 WWW service at the European Bioinformatics Institute (EBI) (Pearson, 1990). The multiple sequence alignment was performed using the ClustalW WWW service at the EBI (Thompson *et al.*, 1994).

Plasmid constructions. Overlapping fragments of plasmid pNZ4000 were subcloned in pUC18 or pUC19 for sequence analysis. The region in between *repB4* and *oriT3* was cloned on three overlapping *Xba*I-*Eco*RI fragments generated by PCR using the Advantage genomic polymerase mix (Clontech) with the primer combinations 5'-GGCGCGTCTAGA-ATCTGTCCTCCCTTGTAAACACC-3' and 5'-GGCGCGGAATTCAATTTCCTATAGA-ATGTTGGAGCCA-3', 5'-GGCGCGTCTAGATCCTCTAGTTTGTGTATTCC-3' and 5'-GGCGCGGAATTCGTAAGAACAGGATGGAATGAAGC-3', and 5'-GGCGCGTCTAG-ATCTTCTTCCGAAATCGTCCCTCT-3' and 5'-GGCGCGGAATTCAACTCGATCGTT-TGCCAAAGACC-3' introducing the *Xba*I and *Eco*RI sites (underlined) in the primers.

Nucleotide sequence accession numbers. The complete nucleotide sequence of plasmid pNZ4000 is available under GenBank accession no. AF03685.

Results and discussion

Sequence analysis of pNZ4000. The complete nucleotide sequence of pNZ4000 was determined by the sequencing of overlapping subclones. One of the four *Xba*I sites was arbitrarily designated as bp 1 (Fig. 1). The average G+C content of pNZ4000 is 33%, which is below the typical G+C content of 38-40% reported for *L. lactis* (Holt *et al.*, 1994). This was also observed for the lactococcal plasmids pK214 (32%) (Perreten *et al.*, 1997) and pMRC01 (30%) (Dougherty *et al.*, 1998). All ORFs larger than 60 amino acids were compared to the SWALL database and the intergenic nucleotide sequences to the EMBL prokaryote library. The results of the annotation are depicted in Fig. 1 and listed in Table 1. As several attempts to subclone the region between *repB4* and *oriT3* as a 7051-bp *Xba*I fragment or a 4696-bp *Cla*I fragment were unsuccessful, this part was subcloned on three overlapping PCR-fragments. With this approach *orf212* and *oriT3* were disrupted which, when cloned intact, may have been deleterious or causing instability in *E. coli*.

Regions involved in replication and mobilisation. As reported previously, pNZ4000 contains four functional replicons (van Kranenburg *et al.*, 1998). All four belong to the family of lactococcal theta replicons first identified for plasmid pCI305 (Hayes *et al.*, 1991) and contain (i) a *repB* gene coding for a replication protein preceded by an A/T-rich region that could be the recognition site for host-encoded functions involved in replication (Seegers

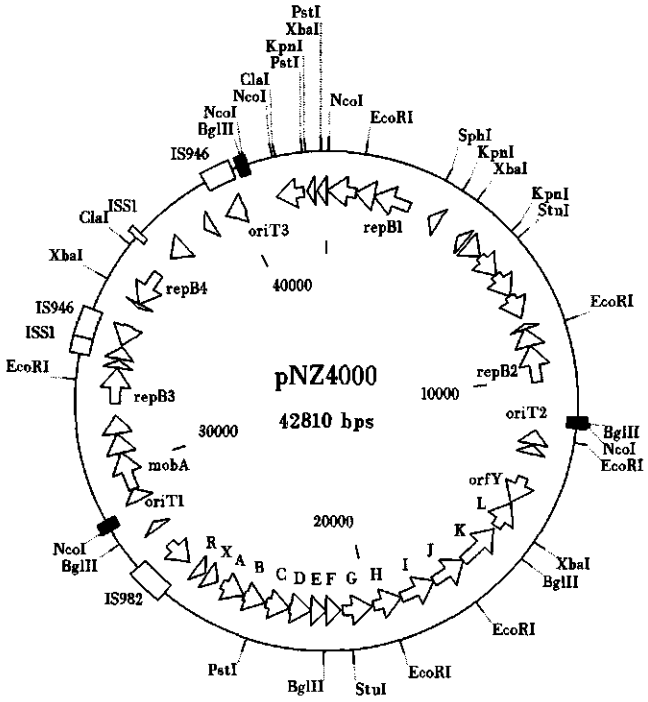


Fig. 1. Plasmid map of pNZ4000. For *ClaI* only relevant sites are shown. Open boxes represent complete or partial IS elements. Filled boxes represent *oriT* sequences.

et al., 1994), (ii) a 22-bp sequence repeated 3.5 times which was shown to have a replicon-specific regulatory role in plasmid replication (Foley *et al.*, 1996), and (iii) two inverted repeats, one of which overlaps the –35 region of the *repB* promoter and is a RepB binding site (Foley *et al.*, 1996).

We have demonstrated that pNZ4000 contains two functional *oriT* sequences: one upstream of *repB2* and one upstream of *mobA* (Fig. 1) (van Kranenburg *et al.*, 1998). Careful analysis of the nucleotide sequence revealed the presence of a third *oriT* sequence (*oriT3*; Fig. 1). Highly homologous sequences (over 95% identity) to the pNZ4000 *oriTs* were present on the *L. lactis* phage resistance plasmid pCI528 (Lucey *et al.*, 1993), the *L. lactis* plasmid pSRQ900 encoding the phage abortive infection mechanism Abi900 (Accession no. AF001314), and the *Lactobacillus helveticus* plasmid pLH1 (Accession no. AJ222725). Since all of these contain identical inverted repeats, it is likely that these plasmids also share the same mobilisation mechanism.

The *mobA* gene encodes a *trans* acting protein involved in mobilisation and is likely to be a relaxase involved in nicking the *nic* site of the *oriT* sequences (van Kranenburg *et al.*, 1998). It is preceded by a putative gene (*orf136*) that encodes a protein with homology (33% identity in 113 amino acids) to a mobilisation protein MobC1 from the *Staphylococcus epidermidis* plasmid pIP1630 (Accession no. AF045241). Downstream of *mobA* two genes

are located that may be involved in mobilisation. MobB shares moderate homology to the Orf1 mobilisation protein of the *Staphylococcus aureus* plasmid pC223, but MobC showed no homology to any protein in the SWALL database.

Insertion sequence elements. Plasmid pNZ4000 is likely to be a cointegrate plasmid, as it contains four functional replicons and three *oriT* sequences. IS elements may have been involved in plasmid recombination, mediating cointegration of lactococcal plasmids as has been described for ISS*1* (Polzin and Shimizu-Kadota, 1987). Alternatively, they may form a transposon-like structure, as has been suggested for IS elements on pK214 and pMRC01 (Dougherty *et al.*, 1998, Perreten *et al.*, 1997). Plasmid pNZ4000 contains two copies of an iso-IS*946* element, one of an iso-IS*982* element and two fragments of iso-ISS*1* elements (Fig. 1), which all are known to occur on lactococcal plasmids (Polzin and Shimizu-Kadota, 1987, Romero and Klaenhammer, 1990, Yu *et al.*, 1995). The two 809-bp copies of the iso-IS*946* element differ from each other at one nucleotide position and contain a 14-bp terminal repeat, identical to the terminal parts of the 18-bp repeats of IS*946* and ISS*1* (Romero and Klaenhammer, 1990). Neither of the copies is flanked by 8-bp direct repeats, as described for IS*946* (Romero and Klaenhammer, 1990). However, the IS*946* element downstream of *repB4* is preceded by an 8-bp sequence (5'-GGCGCACT-3') which is identical to the sequence immediately downstream of the other IS*946* element.

pNZ4000 may encode two divalent cation transport systems. Plasmid pNZ4000 contains a putative gene cluster of four overlapping genes, *orf200*, *orf263*, *orf278*, and *orf266*. The encoded proteins are homologous to a family of putative cobalt transport proteins, first identified in *Salmonella typhimurium*. In this organism the *cob* operon involved in vitamin B_{12} biosynthesis contains four genes, *cbiMNQO*, of which the latter three are expected to be involved in cobalt transport (Roth *et al.*, 1993). In *Methanobacterium thermoautotrophicum*, homologues of *cbiMNQO* are located upstream of the *ribC* gene, involved in riboflavin synthesis (Eberhardt *et al.*, 1997). Orf200 and Orf263 are predicted to be transmembrane proteins and share limited sequence identity with CbiM of *Archaeoglobus fulgidus* (24% identity in 171 amino acids) (Klenk *et al.*, 1997) and CbiQ of *M. thermoautotrophicum* (23% in 215 amino acids) (Eberhardt *et al.*, 1997), respectively. Orf278 and Orf266 are both homologous to CbiO, a family of ABC-transporters involved in cobalt transport. Orf278 shares highest identity with a CbiO homologue from *Pyrococcus horikoshii* (35% in 234 amino acids) (Kawarabayasi *et al.*, 1998), and Orf266 with CbiO of *M. thermoautotrophicum* (36% in 274 amino acids) (Eberhardt *et al.*, 1997). Functional analysis is required to demonstrate the involvement of this lactococcal gene cluster in cobalt transport and the reason for the presence of two *cbiO*-like genes.

Another gene that may be involved in metal ion transport is *orf302*. The encoded protein is homologous to the family of CorA Mg^{2+} transport systems and shares highest identity (24% in 295 amino acid residues) with CorA of *Methanococcus jannaschii* (Smith *et al.*, 1998). CorA is found in most bacteria and archaea, and is expected to be their dominant Mg^{2+} uptake system (Smith *et al.*, 1998). CorA of *S. typhymurium* has a large N-terminal periplasmic domain of about 240 amino acids, followed by three C-terminal transmembrane

Table 1. Putative genes and their products, deduced from the nucleotide sequence of pNZ4000

Gene	Putative ribosomebinding site[a]	Start	End	No. of amino acids of protein	Proposed function of gene product
orf100[b]	AAAG	167	42675	100	Unknown
orfD1[b]	AGAAA	1115	183	310	Restriction/modification protein
orfC1[b]	AAAGG	1711	1103	202	Replication protein
repB1[b]	GGAG	2903	1704	399	Replication protein
orf122[b]	GAGG	3989	3621	122	Unknown
orf92[b]	GAAAGGA	5015	4737	92	Unknown
orf200	AAAGGAG	5127	5729	200	Cobalt transport
orf263	AGGA	5729	6520	263	Cobalt transport
orf278	GGAG	6496	7332	278	Cobalt transport
orf266	AAGGAGG	7319	8119	266	Cobalt transport
orfD2[b]	GAAA	8495	8334	53	Restriction/modification protein
orfC2[b]	AAAGG	9070	8483	195	Replication protein
repB2[b]	AAAGGAG	10268	9063	401	Replication protein
orf154[b]	AAAG	12276	11812	154	Unknown
orf98[b]	GAGG	12601	12305	98	Unknown
orfY	GGAG	13400	14302	300	Unknown
epsL[b]	AAGGA	15226	14327	299	Unknown
epsK[b]	GGAG	16742	15324	471	EPS export protein
epsJ[b]	AGAAAGGA	17919	16729	396	Glycosyltranferase
epsI[b]	AGGA	19096	17906	396	EPS Polymerase
epsH[b]	GAAAG	20025	19096	309	Glycosyltransferase
epsG[b]	AGAAAGGA	21036	20086	316	Galactosyltransferase
epsF[b]	GGA	21573	21091	160	Glucosyltransferase
epsE[b]	AAG	22043	21573	156	Accessory function to EpsF
epsD[b]	GGAG	22733	22053	226	Priming glucosyltransferase
epsC[b]	AGGAG	23519	22755	254	Unknown
epsB[b]	AGGAG	24269	23574	231	EPS chain length determination
epsA[b]	GGAG	25058	24279	259	EPS chain length determination
epsX[b]	GGAGG	25644	25225	139	Unknown
epsR[b]	AAAGG	26050	25733	105	Regulator protein
orf982[b]	GAAAG	27231	26341	296	Transposase
orf82	AGGAG	27860	28108	82	Unknown
orf136	AAAGG	28957	29367	136	Mobilization protein
mobA	AAAG	29346	30578	410	Mobilization protein
mobB	GGA	30575	31198	207	Mobilization protein
mobC	GAGG	31215	31817	200	Mobilization protein
repB3	AGGA	32199	33350	383	Replication protein
orfC3	AAAGG	33350	33604	84	Replication protein
orfS1	GAGG	33641	34045	134	Transposase
orf946[b]	GAGG	34760	34080	226	Transposase
orfC4[b]	AGGA	35547	35425	40	Replication protein
repB4[b]	AAAGGAG	36706	35540	388	Replication protein
orf212[b]	AAAGGA	37973	37335	212	Unknown
orf107[b]	AAAGG	39112	38789	107	Unknown
orf946[b]	GAGG	40270	39590	226	Transposase
orf302[b]	GGA	42251	41343	302	Magenesium transport
orf86[b]	AGG	42605	42345	86	Unknown

[a] Sequence of the 3' end of the lactococcal 16SrRNA is 3'-UCUUUCCUCC-5' (Chiaruttini and Millet, 1993).

[b] Encoded by complementary strand.

domains and is expected to form a homo-oligomer for its activity (Smith and Maguire, 1998). The *M. jannaschii* CorA has been identified from the genome sequence by its 22% sequence identity to CorA of *S. typhimurium* and shown to be functional as a Mg^{2+} transporter in a Mg^{2+}-transport deficient *S. typhimurium* strain (Smith *et al.*, 1998). The comparison of CorA from *M. Jannaschii*, CorA from *S. typhimurium*, and Orf302 (Fig. 3) shows that most conservation is within the C-terminal membrane segments, as is found for the other CorA proteins (Smith *et al.*, 1998). Therefore, it is likely that Orf302 has a similar function and may be a Mg^{2+} or other metal ion transporter. Furthermore, Orf302 shares a similar degree of identity (22-24%) with the lactococcal OrfA protein found on the chromosome of *L. lactis* MG1363 and *L. lactis* DB1341 (Arnau *et al.*, 1997). OrfA is of unknown function, but, like Orf302, it seems to belong to the CorA protein family (Fig. 2) and may function as the Mg^{2+} transport system of *L. lactis*.

```
CorA_Mj   MITVIAIAKDGSIVEPKLDEISFEDYRLIWIDCYD-PKDEELYKLSKKIGISVSDLQIGL 59
Orf302    ------------MIKPEKTINGTKWIETIQINAEERATLEDQYGIDEDIIEYVTDNDEST 48
CorA_St   MLSAFQLEKN--RLTRLEVEESQSLIDAVWVDLVE-PDDDERLRVQSELGQSLATRPELE 57
OrfA_Ll   ------MIKNYELSNEKKLISTSEMKNFTYVLN---PTREEIGNISEHYDFPFDYLSGIL 51
                                            .      ..   .

CorA_Mj   DEQEIPRVEEDEDFYLII-YKAPLFEEDIT---TTSLGIYIKNNLLLTIHSDKIKAIGRL 115
Orf302    N--YVYDINED-DQLFIF-LAPYALDKDALRYITQPFGMLLHKGVLFTFNQSGIPEVN-T 103
CorA_St   DIEASARFFEDEDGLHIHSFFFFEDAEDHAG--NSTVAFTIRDGRLFTLRERELPAFR-- 113
OrfA_Ll   DDYENARFETD-DNDNNLILLQYPALSNYGEVATFPYSLVWTKNESVILALNHEIDNG-L 109
          .          : :                .              .          .

CorA_Mj   HKLISTKKPRIVFERGIGFLLYHILNEITRSYSRILMNLEDELEELEDKLLAGYD-REVM 174
Orf302    ALYSALDNP-EV--KSVDAFILETLFTVVVSFIPISRAITKKRNYLDKMLNRKTK-NSDL 159
CorA_St   -LYRMRARSQAMVDGNAYELLLDLFETKIEQLADEIENIYSDLEKLSRVIMEGHQGDEYD 172
OrfA_Ll   IFEREYDYK-----RYKHQLIFQVMYQMTHTFHDYLRDFRTRRRRLEVGIKNSTK-NDQI 163
                              ..   .              .      :    .

CorA_Mj   EKILGLRKTLVYFHKSLIANRDVLVLLKRKYLPITTKED-RENFEDLYYDTLQLIDMSAT 233
Orf302    VSLSYLQQTLTFLSSAVQTNLSELDRLPKTHFGVGADQDKIDLFEDVQIEGEQVQRMFEI 219
CorA_St   EALSTLAELEDIGWKVRLCLMDTQRALNFLVRKARLPGGQLEQAREILRDIESLLPHNES 232
OrfA_Ll   VDLIAIQASLIYFEDALHNNMQVLQNFIDYLREDDED-GFAEKIYDIFVETDQAT-ETKI 221
            .  .                          .              .    ..   .

CorA_Mj   YREVLTSMMDITLSLENIKMNQIMKILTMVTTIFAVPMWITGIYGMNFSYLPLANNPQGF 293
Orf302    ETQVVDRIDHTLNSLANNNLNDTMKFLTIWSLTMAVPTIISGFYGMNVK-LPLAG-MQYA 277
CorA_St   LFQKVNFLMQAAMGFINIEQNRIIKIFSVVSVVFLPPTLVASSYGMNFEFMPELKWSFGY 292
OrfA_Ll   QLKLLENLRDLFSNIVSNNLNIVMKIMTSATFVLGIPAVIVGFYGMNVP-IPGQNFNWMV 280
            .  .  .       .   . :  .:... .  .  :  .   ::::   .:

CorA_Mj   WLVMALMVVIIMIFVYIFRRSGWI-- 317
Orf302    WMLT-LGISVVLIVAMLIMLKVWRKM 302
CorA_St   PGAIIFMILAGLAPYLYFKRKNWL-- 316
OrfA_Ll   WLILVFGILLCVWVTWWLHKKDML-- 304
                .  .   .    .
```

Fig. 2. Multiple sequence alignment of CorA proteins from *M. Jannaschii* (CorA_Mj), CorA from *S. typhimurium* (CorA_St), OrfA from *L. lactis* MG1363, and Orf302. Identical residues (:), and conserved substitutions (.) are indicated. The three C-terminal transmembrane regions determined for *S. typhimurium* CorA (Smith *et al.*, 1993) are underlined.

Although a magnesium transport system is expected to be present in *L. lactis* and may be encoded by the chromosomal *orfA*, the putative divalent cation transport systems of pNZ4000, when functional, may give the strain harbouring pNZ4000 an advantage over cured strains under limiting growth conditions. Very little is known about cobalt requirement in *L. lactis*. Several enzymes, like aminopeptidase P can use Co^{2+} instead of Mn^{2+} (McDonnell *et al.*, 1997). Also for some glycosyltransferases it is known that Co^{2+} can substitute for Mn^{2+} (Gmeiner, 1988, Powell and Brew, 1976) and the cobalt transport may be necessary for enzyme activity of glycosyltransferases involved in EPS biosynthesis. Alternatively, some anionic cell-surface polysaccharides are known to help hold minerals and nutrients near the cell (Sutherland, 1988). Since the EPS encoded by pNZ4000 contains phosphate groups that will be negatively charged, the physiological function of this EPS could be the accumulation of divalent cations in poor environments that are subsequently transported into the cell by the two transport systems.

Acknowledgements

Part of this work was supported by EC research grant 1116/92 1.6.

References

Arnau, J., F. Jorgensen, S. M. Madsen, A. Vrang, H. Israelsen. 1997. Cloning, expression, and characterization of the *Lactococcus lactis pfl* gene, encoding pyruvate formate-lyase. J. Bacteriol. **179:**5884-5891.

Casadaban, M. J., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138:**179-207.

Chiaruttini, C., and M. Millet. 1993. Gene organization, primary structure and RNA processing analysis of a ribosomal RNA operon in *Lactococcus lactis*. J. Mol. Biol. **230:**57-76.

de Vos, W. M., P. Vos, H. de Haard, I. Boerrigter. 1989. Cloning and expression of the *Lactococcus lactis* ssp. *cremoris* SK11 gene encoding an extracellular serine proteinase.Gene **85:**169-176.

de Vos, W. M., and G. F. M. Simons. 1994. Cloning and expression systems in lactococci, p. 52-102. *In* M. J. Gasson and W. M. de Vos (ed.), Genetics and biotechnology of lactic acid bacteria. Chapman and Hall, London, England.

Dougherty, B. A., C. Hill, J. F. Weidman, D. R. Richardson, J. C. Venter, and R. P. Ross. 1998. Sequence and analysis of the 60 kb conjugative, bacteriocin-producing plasmid pMRC01 from *Lactococcus lactis* DPC3147. Mol. Microbiol. **29:**1029-1038.

Eberhardt, S., S. Korn, F. Lottspeich, and A. Bacher. 1997. Biosynthesis of riboflavin: an unusual riboflavin synthase of *Methanobacterium thermoautotrophicum*. J. Bacteriol. **179:**2938-2943.

Foley, S. S. Bron, G. Venema, C. Daly, G. F. Fitzgerald. 1996. Molecular analysis of the replication origin of the *Lactococcus lactis* plasmid pCI305. Plasmid **36:**125-141.

Gmeiner, B. M. 1988. Co^{2+} is able to substitute for Mn^{2+} in some exogenous and endogenous galactosyltransferase reactions. Enzyme **39:**213-219.

Hayes, F., P. Vos, G. F. Fitzgerald, W. M. de Vos, and C. Daly. 1991. Molecular organization of the minimal replicon of novel, narrow-host-range, lactococcal plasmid pCI305. Plasmid **25:**16-26.

Holt, J. G., N. R. Krieg, P. H. A. Sneath, J. T. Staley, and S. T. Williams. 1994. Group 17, Gram-positive

cocci, p. 527-558. *In* W. R. Hensyl (ed.), Bergey's Manual of Determinative Bacteriology. 9th edn. Williams and Wilkins, Baltimore, USA.

Kawarabayasi, Y., M. Sawada, H. Horikawa, Y. Haikawa, Y. Hino, S. Yamamoto *et al.* 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res. **5**:55-76.

Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum *et al.* 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus.* Nature **390**:364-370.

Lucey, M, C. Daly, and G. Fitzgerald. 1993. Analysis of a region from the bacteriophage resistance plasmid pCI528 involved in its conjugative mobilization between *Lactococcus strains.* J. Bacteriol. **175**:6002-6009.

McDonnell, M., R. Fitzgerald, I. Nifhaoláin, P. V. Jennings, and G. O'Cuinn. 1997. Purification and characterization of aminopeptidase P from *Lactococcus lactis* subsp. *cremoris.* J. Dairy Res. **64**:399-407.

Perreten, V., F. Schwarz, L. Cresta, M. Boeglin, G. Dasen, and M. Teuber. 1997. Antibiotic resistance spread in food. Nature **389**:801-802.

Pearson, W. R. 1990. Rapid and Sensitive Sequence Comparison with FASTP and FASTA. Methods Enzymol. **183**:63- 98.

Polzin, K. M., and M. Shimizu-Kadota. 1987. Identification of a new insertion element, similar to gram-negative IS*26*, on the lactose plasmid of *Streptococcus lactis* ML3. J. Bacteriol. **169**:5481-5488.

Powell, J. T., and K. Brew. 1976. Metal ion activation of galactosyltransferase. J. Biol. Chem. **251**:3645-3652.

Romero, D. A., and T. R. Klaenhammer. 1990. Characterization of insertion sequence IS*946*, and iso-ISS*1* element, isolated from the conjugative lactococcal plasmid pTR2030. J. Bacteriol. **172**:4151-4160.

Roth, J. R., J. G. Lawrence, M. Rubenfield, S. Kieffer-Higgins, and G. M. Church. 1993. Characterization of the cobalamin (vitamin B_{12}) biosynthetic genes of *Salmonella typhimurium.* J. Bacteriol. **175**:3303-3316.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Seegers, J. F. M., S. Bron, C. M. Franke, G. Venema, and R. Kiewiet. 1994. The majority of lactococcal plasmids carry a highly related replicon. Microbiology **140**:1291-1300.

Smith, R. L., J. L. Banks, M. D. Snavely, and M. E. Maguire. 1993. Sequence and topology of the CorA magnesium transport systems of *Salmonella typhimurium* and *Escherichia coli.* Identification of a new class of transport protein. J. Biol. Chem. **268**:14071-14080.

Smith, R. L., E. Gottlieb, L. M. Kucharski, and M. E. Maguire. 1998. Functional similarity between archaeal and bacterial CorA magnesium transporters. J. Bacteriol. **180**:2788-2791.

Smith, R. L., and M. E. Maguire. 1998. Microbial magnesium transport: unusual transporter searching for identity. Mol. Microbiol. **28**:217-226.

Steele, J. L., and L. L. McKay. 1989. Conjugal transfer of genetic material in lactococci: a review. J. Dairy Sci. **72**:3388-3397.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673-4680.

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.

van Kranenburg, R., and W. M. de Vos. 1998. Characterization of multiple regions involved in replication and mobilization of plasmid pNZ4000 coding for exopolysaccharide production in *Lactococcus lactis.* J. Bacteriol. **180**:5285-5290.

van Kranenburg, R., I. I. van Swam, J. D. Marugg, M. Kleerebezem, and W. M. de Vos. 1999. Exopolysaccharide biosynthesis in *Lactococcus lactis* NIZO B40: functional analysis of the glycosyltransferase genes involved in synthesis of the polysaccharide backbone. J. Bacteriol. **181**:338-340.

Yu, W., I. Mierau, A. Mars, E. Johnson, G. Dunny, and L. L. McKay. 1995. Novel insertion sequence-like element IS*982* in lactococci. Plasmid **33**:218-225.

General Discussion

Chapter 7

General Discussion

This thesis describes the genetics of exopolysaccharide (EPS) biosynthesis in *Lactococcus lactis* strains. An EPS plasmid designated pNZ4000, encoding EPS biosynthesis in *L. lactis* NIZO B40 was studied in detail. Its complete nucleotide sequence was determined and genes involved in EPS biosynthesis, plasmid replication and mobilisation were analysed. The functions of the glycosyltransferase genes of the *eps* gene cluster were assessed and the order of assembly of the EPS repeating unit backbone was established. Furthermore, the diversity of various *eps* gene clusters (Fig. 1) and the chemical composition of the EPSs encoded by these clusters, was studied and first steps towards polysaccharide engineering were taken.

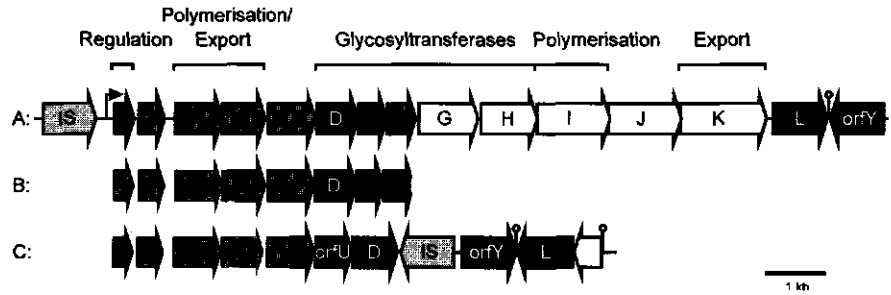


Fig. 1. Genetic organisation of the *eps* gene clusters of *L. lactis* NIZO B40 (A), NIZO B891 (B), and NIZO B35 (C). Predicted functions are listed above.

Biosynthesis of the EPS repeating unit

Lactococcal EPS synthesis occurs via a similar pathway as O-antigen polysaccharide synthesis in *Salmonella enterica* and involves growth of the polymer at the reducing end. *L. lactis* NIZO B40 produces an EPS with an identical repeating unit as that of *L. lactis* SBT 0495 EPS consisting of →4)-[α-L-Rha*p*-(1→2)][α-D-Gal*p*-1-PO_4-3]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→ (Nakajima *et al.*, 1992, Chapters 1 and 2). The genes involved in synthesis of the NIZO B40 EPS repeating unit backbone were identified and the substrate specificity of the encoded glycosyltransferases was determined (Chapters 2 and 3). The priming glycosyltransferase EpsD links glucose, probably as glucose-1-phosphate, to the lipid carrier and subsequently the glycosyltransferases EpsE/EpsF and EpsG link glucose and galactose resulting in the lipid-linked trisaccharide β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-1-PO_4-PO_4-lipid. The repeating unit is completed after addition of rhamnose and galactose-1-phosphate and EpsH and EpsJ are likely to be involved in these processes as they are homologous to glycosyltransferases and phosphocarbohydrate transferases (Chapter 3). Based on these results, a working model for NIZO B40 EPS biosynthesis route can be proposed (Fig. 2).

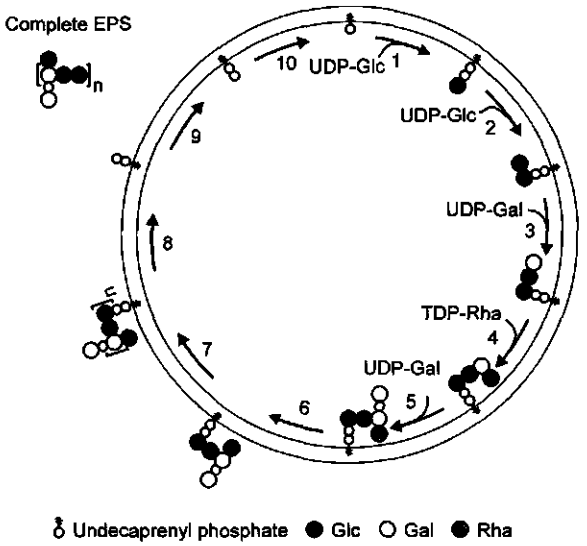


Fig. 2. Working model for NIZO B40 EPS biosynthesis at the lactococcal membrane. (1) EpsD links glucose-phosphate from UDP-glucose (UDP-Glc) to the lipid carrier. (2) EpsE and EpsF add the second glucose moiety (3). EpsG adds galactose from UDP-galactose (UDP-Gal). The repeating unit is completed by the addition of rhamnose (Rha) from TDP-Rha and galactose-phosphate from UDP-Gal, and EpsH (4) and EpsJ (5) are expected to be involved in these steps. Repeating units are predicted to be translocated across the membrane by the activity of EpsK (6) and subsequently polymerised by EpsI, with EpsB and EpsC determining the chain-length (7, 8). The lipid carrier is retranslocated (9) and dephosphorylated (10) to regenerate undecaprenylphosphate.

L. lactis SBT 0495 and NIZO B40 produce an EPS with an identical repeating unit (Nakajima *et al.*, 1992, Chapter 2). Both strains show the same genotyping pattern of their *eps* gene clusters (Chapter 4) and are expected to contain almost identical *eps* gene clusters. Therefore, it is likely that they synthesise their EPSs via the same pathway. Recently, the synthesis of the SBT 0495 EPS was studied biochemically (Oba *et al.*, 1999). Following mild hydrolysis of lipid-linked oligosaccharides isolated from the lactococcal membrane, glucose, lactose (β-D-Gal-(1→4)-β-D-Glc), and the trisaccharide α-l-Rha-(1→2)-β-D-Gal-(1→4)-β-D-Glc were identified amongst other saccharides and designated as biosynthetic intermediates. From these results the proposed order of addition of sugar moieties to the lipid carrier is glucose-1-phosphate, galactose, rhamnose, galactose-1-phosphate, and glucose, respectively, for the SBT 0495 EPS repeating unit (Oba *et al.*, 1999). Nevertheless, with these techniques it can not be established whether the identified saccharides are involved as intermediates in EPS biosynthesis. Therefore, it can be expected that both *L. lactis* NIZO B40 and SBT 0495 synthesise their EPS as is proposed in Fig. 2. However, this study by Oba *et al.* provides experimental evidence for the presence of undecaprenylphosphate as the lipid carrier.

L. lactis NIZO B35 produces an EPS containing only galactose residues which is expected to have an identical repeating unit as that of strain H414 consisting of →4)-[β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)]-α-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→ (van Casteren,

personal communication, Chapter 4). *L. lactis* NIZO B891 produces an EPS containing both glucose and galactose (Chapter 4). Fragments of gene clusters involved in synthesis of NIZO B35 and NIZO B891 EPS were cloned and analysed. The priming glycosyltransferase activities for the synthesis of NIZO B35 and NIZO B891 EPS repeating unit (galactose-1-phosphate and glucose-1-phosphate transfer, respectively) were assigned to the *epsD* gene products of both clusters. NIZO B891 *epsE* and *epsF* are homologous to NIZO B40 *epsE* and *epsF* and encode a galactosyltransferase that links galactose to lipid-linked glucose. Mild hydrolysis treatment of the resulting lipid-linked disaccharide yields a product with the same mobility on thin-layer chromatography gels as lactose, i.e. β-D-Gal*p*-(1→4)-β-D-Glc*p* (Chapter 4).

Glycosyltransferases involved in EPS biosynthesis

Biosynthesis of the lactococcal EPSs studied in this work involves non-processive α- or β-glycosyltransferases. The β-glycosyl transfer reaction resembles that of an inverting-type glycoside hydrolases, while the α-glycosyl transfer reaction can be seen as that of a retaining-type glycosidic hydrolase (see Chapter 1). Hydrophobic cluster analysis (HCA) of non-processive β-glycosyltransferases showed a conserved secondary structure for a region of approximately 100 amino acids of four alternating β-strands separated by three α-helices (Saxena *et al.*, 1995). Two conserved aspartic acid residues with a spacing of approximately 50 amino acids were identified in the C-terminal loops of the β2- and β4-strands, which might be the catalytic amino acids for the nucleophilic substitution reaction involved in β bond formation. The priming glycosyltransferases of *L. lactis* NIZO B40, NIZO B35, NIZO B891, *S. thermophilus* Sfi6, and *S. pneumoniae* serotype 14 have four conserved aspartic acid residues and two conserved glutamic acid residues that could be involved in catalytic activity (Chapter 4). The putative gene product of NIZO B35 *orfU*, which is homologous to these priming glycosyltransferases but shows no activity as priming glucosyl or galactosyltransferase, lacks 30 amino acids at its C-terminus compared to the other priming glycosyltransferases including one of the conserved aspartic acid residues (Chapter 4).

In *L. lactis* NIZO B40 and NIZO B891 the glycosyltransferases that couple the second sugar moiety of the repeating unit to the first, are encoded by two genes (Chapters 3 and 4). They are homologous to *S. pneumoniae* serotype 14 *cps14F* and *cps14G*, which are involved in capsule synthesis (Chapters 2 and 4). For *S. pneumoniae* it has been demonstrated that the second gene, *cps14G*, is essential for the galactosyltransferase activity while the first, *cps14F*, is not and is likely to have an accessory function (Kolkman *et al.*, 1997a). *Rhizobium leguminosarum* has also two genes, *pssD* and *pssE*, that are involved in EPS biosynthesis and encode a single glycosyltransferase coupling glucuronic acid (GlcA) from UDP-GlcA to lipid-linked glucose (Pollock *et al.*, 1998). EpsE and EpsF, CpsF and CpsG, and PssD and PssE are homologous to SpsK, which is involved in EPS synthesis in *Sphingomonas* and catalyses the same reaction as PssD and PssE (Pollock *et al.*, 1998). An alignment of these

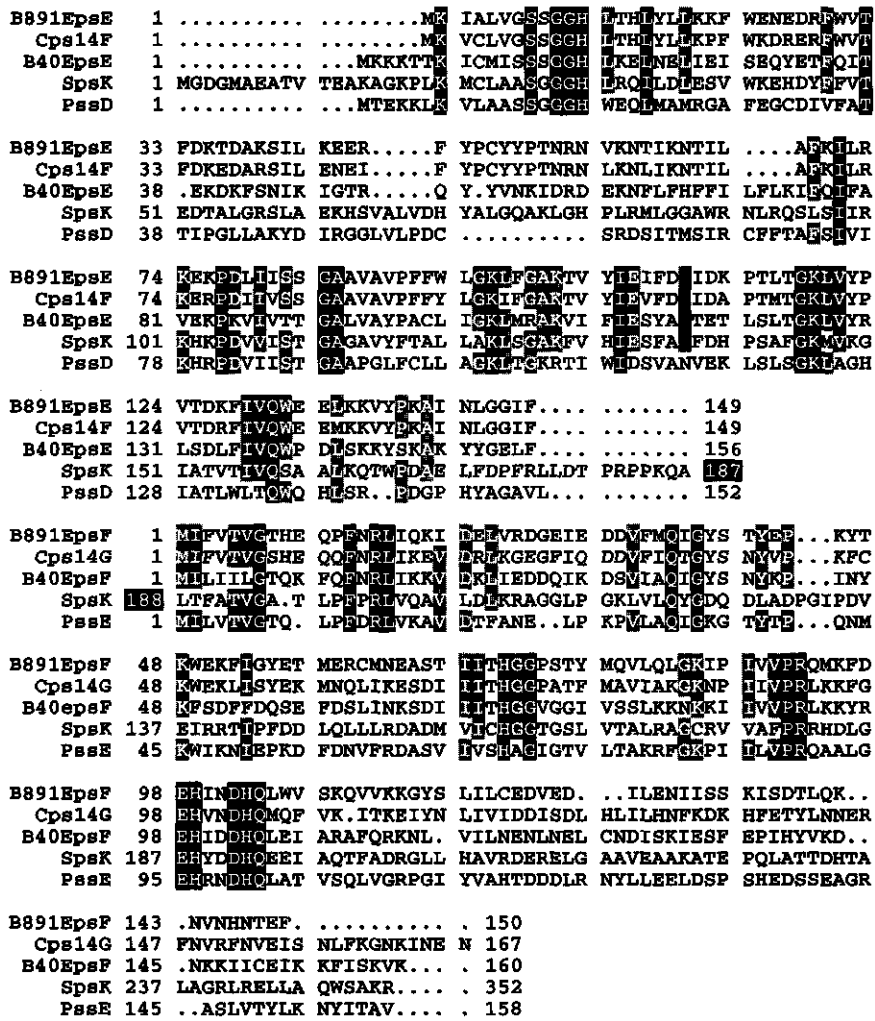


Fig. 3. Multiple sequence alignment of glycosyltransferases involved in addition of the second sugar of the repeating unit to the first from *L. lactis* NIZO B891 (B891EpsE/B891EpsF), *S. pneumoniae* serotype 14 (Cps14F/Cps14G), *L. lactis* NIZO B40 (B40EpsE/B40EpsF), *R. leguminosarum* (PssD/PssE), and *Sphingomonas* (SpsK). Residues conserved in five or four sequences are shaded black and gray, respectively.

proteins shows that EpsE, Cps14F, SpsK, and PssD are highly conserved (Fig. 3). In EpsF, Cps14G, PssE, and SpsK, a conserved glutamic acid (position 98 in B40EpsF) and aspartic acid residue (position 102 in B40EpsF) are found, one of which could be involved in catalytic activity of these β-glycosyltransferases (Fig. 2 of Chapter 1).

Polymerisation and export processes

Polysaccharide repeating units are assembled on a lipid carrier at the cytoplasmic side of the membrane (Fig. 2). For O-antigen synthesis, the transfer across the cytoplasmic membrane is thought to be facilitated by flippase activity of a transmembrane protein Wzx (Liu *et al.*, 1996). The repeating units are polymerised at the reducing end by Wzy and the chain-length is determined by Wzz (Daniels *et al.*, 1998, Morona *et al.*, 1995, Reeves *et al.*, 1996). For succinoglycan EPS biosynthesis in *R. meliloti*, ExoP, ExoQ, and ExoT proteins have roles in polymerisation and export processes. *Rhizobium* strains produce simultaneously high molecular weight succinoglycan and dimers and trimers of the repeating unit, and both forms are polymerised by separate mechanisms that are *exoP/exoQ*-dependent and *exoP/exoT*-dependent, respectively (González *et al.*, 1998). Although ExoT is homologous to Wzx, it is not likely to function as a flippase, as *exoT* mutants are still able to form succinoglycan (González *et al.*, 1998). ExoQ is homologous to Wzy and might serve as a highly processive polymerase that yields high molecular weight succinoglycan (González *et al.*, 1998). ExoP is involved in the formation of both high molecular weight EPS, and dimers and trimers. Its N-terminus is homologous to Wzz, but it has an additional C-terminal domain with an ATP binding domain. It has been suggested that ExoP is involved in chain-length determination and a proline-rich motif ($RX_4PX_2PX_4SPKX_9IXGXMXGXG$) close to the second transmembrane helix in the N-terminal domain is involved in its activity (Becker *et al.*, 1995, Becker and Pühler, 1998). It is also possible that the role of ExoP in succinoglycan synthesis is either to catalyse the formation of dimers of the repeating unit, or to form a complex with ExoT and ExoQ that is essential for the function of these proteins (González *et al.*, 1998).

For the NIZO B40 *eps* gene cluster (Fig. 1), *epsK* is the *wzx*-homologue which may encode the flippase, and *epsI* the *wzy*-homologue which may encode the polymerase (Chapter 2). EpsA is the Wzz-homologue and may be involved in chain-length determination. It is homologous to the N-terminal domain of ExoP and contains part of the consensus sequence ($SPKX_{11}GX_3G$) of proteins involved in chain-length determination (Chapter 2). EpsB is homologous to the C-terminal domain of ExoP and contains an ATP-binding domain (Chapter 2). Homologues of EpsA and EpsB are encoded by cell-surface polysaccharide gene clusters of *Staphylococcus aureus*, *S. pneumoniae*, *Streptococcus agalactiae*, and *S. thermophilus* (Chapter 2). The *eps* gene clusters of lactococcal strains that produce EPSs with different repeating units have a highly conserved organisation and nucleotide sequence for the first genes comprising *epsA* and *epsB* (Chapter 4). The chain-length of NIZO B891 EPS is approximately 1.5-times that of NIZO B40 EPS, which has a molecular weight of 1.6×10^6 g/mol (Tuinier, 1999). NIZO B891 *epsA* has not been sequenced, but the cloned 3'-half of the *epsB* gene has 96.6% identity with NIZO B40 *epsB*. If the difference in chain-length is caused by activity of EpsA/EpsB, it would probably be the result of a few amino acid changes between both strains. An alternative explanation is that differences in EpsK and EpsI are responsible for a more efficient polymerisation of NIZO B891 EPS, either direct by their

enzymatic activity or indirect by altered interaction with EpsA/EpsB. The *epsK* and *epsI* genes of NIZO B40 do not hybridise with NIZO B891 plasmid DNA and are therefore expected be different (Chapter 4).

Regulation of *eps* gene expression.

To date, no reports have appeared on regulation of lactococcal EPS production under influence of environmental or endogenous factors. Although the EPS-production level may vary when different carbon sources or media are used, no effects on *eps* gene expression have been observed (Looijesteijn, personal communication). In batch fermentations, the *eps* genes were constitutively expressed and expression was independent of the growth phase. This was determined by measuring β-glucuronidase activity of cell-free extracts of *L. lactis* strains harbouring pNZ4040, containing the *epsR* promoter fused to the promoterless *gusA* gene (Chapter 2), in the presence or absence of the EPS plasmid pNZ4000 (van Swam and van Kranenburg, unpublished results).

The first gene of the NIZO B40 *eps* gene cluster, *epsR*, encodes a putative protein which is homologous to regulator proteins containing a DNA-binding domain (Chapter 2). The last gene of the cluster, *orfY*, located downstream and in opposite orientation of the *eps* gene cluster, encodes a putative protein with homology to LytR, an attenuator of the *Bacillus subtilis lytABC* and *lytR* operons (Chapter 2, Lazarevic *et al.*, 1992). For gene clusters involved in capsule synthesis in *S. pneumoniae* and group B streptococci, or EPS biosynthesis in *S. thermophilus*, the first gene is believed to be involved in regulation as it encodes a protein homologous to LytR (Guidolin *et al.*, 1994, Kolkman *et al.*, 1997b, Koskiniemi *et al.*, 1998, Muñoz *et al.*, 1997, Stingele *et al.*, 1996). For group B streptococci, upstream and in opposite orientation of the *lytR*-like gene, *cpsX*, a gene is located, *cpsY*, which encodes a putative protein with homology to the LysR type of transcriptional regulators that, like EpsR, have a helix-turn-helix motif in the N-terminus (Koskiniemi *et al.*, 1998). Both *epsR* and *orfY* seem to be conserved in lactococcal *eps* gene clusters (Chapter 4) but their role in regulation of *eps* gene expression, if any, remains to be established.

Practical applications and perspectives for polysaccharide engineering

The use of lactococcal EPSs for industrial applications may be hampered by the low production yields ranging from 50 to 800 mg/l (Cerning, 1990). Xanthan gum for example, is produced at 10 to 25 g/l and in continuous cultures can convert 60% to 70% of substrate to xanthan (Becker *et al.*, 1998b for a review). EPS production may be increased by overproduction of enzymes needed for EPS biosynthesis, nucleotide sugar biosynthesis, or biosynthesis of the lipid carrier. In Chapter 4 the overproduction of the NIZO B40 priming glycosyltransferase EpsD is described in a strain lacking *epsD* in its *eps* gene cluster. The

overproducing strain has an increased EPS production compared to that of the control strain harbouring a plasmid with the intact *eps* gene cluster. These results demonstrate that EPS production can be increased by overproduction of *eps* genes.

An increase of *eps* gene expression may also be achieved by elevating the copy number of the EPS plasmid. Chapter 5 describes that four active replicons are located on the NIZO B40 EPS plasmid pNZ4000. They are homologous to lactococcal theta replicons which generally have low copy numbers (Frère *et al.*, 1995). Each replicon contains a *repB* gene encoding the replication protein proceeded by a region including the RepB binding site which is overlapping the -35 region of the *repB* gene (Foley *et al.*, 1996). It is likely that RepB autoregulates its own expression and negatively controls the level of RepB protein present in the cell. Attempts have been made to overexpress *repB2* of pNZ4000 with the NICE (nisin-controlled expression) system (Kuipers *et al.*, 1998), in strains harbouring pNZ4030, an erythromycin-resistance encoding derivative of pNZ4000 (Chapter 2), to increase the copy-number of pNZ4030 and elevate the EPS production. For this purpose the *repB2* gene including its own promoter was cloned as a 1.6-kb *Cla*I-*Eco*RI fragment (Fig 1B of Chapter 5) under control of the *nisA* promoter in vector pNZ8020 (de Ruyter *et al.*, 1996b) to form pNZ4024. The final EPS production was determined in duplicate for cultures of *L. lactis* NZ3900 cells (de Ruyter *et al.*, 1996a) harbouring either pNZ4030 alone (115 ± 4 mg/l), pNZ4030 and pNZ4024, with *repB2* under control of its own promoter (140 ± 2 mg/l), or NZ3900 harbouring pNZ4030 and pNZ4024, induced with 1 ng/ml nisin A to activate the *nisA* promoter (174 ± 4 mg/l) (van Kranenburg, unpublished results). These results indicate that the level of *repB2* expression can influence the production level of EPS, most likely via an increased copy number of the *eps* gene cluster.

Plasmid pNZ4000 is a mobilisation plasmid (Chapter 5). It can be transferred from one *L. lactis* strain to another recipient *L. lactis* strain by conjugation, in which process the *mobA* gene and one of the three origins of transfer play a role. The copy number of lactococcal theta plasmids like pNZ4000 is dependent on the host. For pUCL22 derivatives it has been demonstrated that in *L. lactis* MG1614 the copy number is 2 to 3 per chromosome, while in *L. lactis* MMS368 and IL1441 the copy number is 4 to 6 per chromosome (Frère *et al.*, 1995). Therefore, EPS production might already be positively influenced using a 'natural' method, by conjugation of the EPS plasmid to a strain which will give a relatively high copy number. An alternative method is cloning the complete *eps* gene cluster on a plasmid with a higher copy number like pIL253 (Simon and Chopin, 1988). This has been achieved for the *eps* gene cluster from *S. thermophilus* Sfi6 (Stingele *et al.*, 1996).

Polysaccharide engineering is the use of genetic engineering to produce polysaccharides with desired properties. To achieve polysaccharide engineering, one has to know which factors influence these properties and how to direct them. The structure-function relation of polysaccharides for their texturising properties is still poorly understood. Recently, the physical properties of NIZO B40 EPS have been studied and the interactions with different milk components were investigated (Tuinier, 1999). Important factors that influence the intrinsic viscosity of EPS are chain length and chain stiffness, which can be increased by

adjustment of the polymerisation and export processes as described above or by introducing β(1→4) bonds for the less stiffer β(1→3) or α(1→4) bonds (Tuinier, 1999). Chapter 4 describes important progress towards polysaccharide engineering with the construction of a non-polar gene disruption of the priming glycosyltransferase gene and its controlled homologous and heterologous complementation. Although the *eps* genes are co-ordinately expressed, the deletion could be complemented *in trans* by a gene under control of a different promoter. A critical point in polysaccharide engineering is the prevention of accumulating lipid-linked oligosaccharide intermediates. It has been reported that they can have lethal effects for *R. meliloti*, *Sphingomonas* and *X. campestris* (Pollock *et al.*, 1998, Reuber *et al.*, 1991). In the experiments described in Chapter 3, it was also found that accumulation of the lipid-linked trisaccharide by overexpression of *epsDEFG* in *E. coli* and *L. lactis* had severe negative effect on cell growth. In contrast, the lipid carrier can be charged with glucose or cellobiose without any effect for the cells, as overexpression of *epsD* or *epsDEF* had no or only a limited effect on growth (van Kranenburg, unpublished results).

Recently, Stingele *et al.* (1999) showed that production of an EPS with an altered repeating unit but of a similar molecular weight is possible. The *eps* gene cluster of *S. thermophilus* Sfi6 cloned on a plasmid was introduced in *L. lactis* MG1363. The lactococcal strain produced an EPS with an galactose instead of *N*-acetylgalactosamine in the backbone and lacking the side chain sugar (see also Fig. 6 in Chapter 1). These results indicate that successful polysaccharide engineering may be achieved in the near future, as a wide variety of polysaccharide gene clusters is being studied at this moment and will result in a collection of glycosyltransferase genes and a better understanding of polymerisation and export processes.

References

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.

Becker, A., and A. Pühler. 1998. Specific amino acid substitutions in the proline-rich motif of the *Rhizobium meliloti* ExoP protein result in enhanced production of low-molecular-weight succinoglycan at the expense of high-molecular-weight succinoglycan. J. Bacteriol. **180:**395-399.

Becker, A., F. Katzen, A. Pühler, and L. Ielpi. 1998. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Appl. Microbiol. Biotechnol. **50:**145-152.

Cerning, J. 1990. Exocellular polysaccharides produced by lactic acid bacteria. FEMS Microbiol. Rev. **87:**113-130.

Daniels, C., C. Vindurampulle, and R. Morona. 1998. Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy). Mol. Microbiol. **28:**1211-1222.

de Ruyter, P. G. G. A., O. P. Kuipers, M. M. Beerthuyzen, I. van Alen-Boerrigter, and W. M. de Vos. 1996. Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. J. Bacteriol. **178:**3434-3439.

de Ruyter, P. G. G. A., O. P. Kuipers, and W. M. de Vos. 1996. Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. Appl. Environ. Microbiol. **62:**3662-3667.

Foley, S., S. Bron, G. Venema, C. Daly, and G. F. Fitzgerald. 1996. Molecular analysis of the replication origin of *Lactococcus lactis* plasmid pCI305. Plasmid **36:**125-141.

Frère, J., C. Herreman, P. Boutibonnes, M. Novel, G. Novel. 1995. Segregational stability and copynumber of the theta-type lactococcal replicon *rep22* in Lactococcus. Current Microbiol. **30:**33-37.

González, J. E., C. E. Semino, L.-X. Wang, L. E. Castellano-Torres, and G.C. Walker. 1998. Biosynthetic control of molecular weight in the polymerization of the octasaccharide subunits of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti.* Proc. Natl. Acad. Sci. USA **95:**13477-13482.

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.

Hundle, B. S., D. A. O'Brien, M. Alberti, P. Beyer, and J. E. Hearst. 1992. Functional expression of zeaxanthin glucosyltransferase from *Erwinia herbicola* and a proposed uridine diphosphate binding site. Proc. Natl. Acad. Sci. USA. **89:**9321-9325.

Kolkman, M. A. B., B. A. M. van der Zeijst, and P. J. M. Nuijten. 1997. Functional analysis of the glycosyltransferases encoded by the capsular biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J. Biol. Chem. **272:**19502-19508.

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.

Koskiniemi, S., M. Sellin, and M. Norgren. 1998. Identification of two genes, *cpsX* and *cpsY*, with putative regulatory function on capsule expression in group B streptococci. FEMS Immun. Med. Microbiol. **21:**159-168.

Kuipers, O. P., P. G. G. A. de Ruyter, M. Kleerebezem, and W. M. de Vos. 1998. Quorum sensing-controlled gene expression in lactic acid bacteria. J. Biotechnol. **64:**15-21.

Lazarevic, 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.

Liu, D., R. A. Cole, and P. R. Reeves. 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J. Bacteriol. **178:**2102-2107.

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.

Muñoz, R., M. Mollerach, R. López, and E. García. 1997. Molecular organization of the genes required for the synthesis of type 1 capsular polysaccharide of *Streptococcus pneumoniae*: formation of binary encapsulated pneumococci and identification of cryptic dTDP-rhamnose biosynthesis genes. Mol. Microbiol. **25:**79-92.

Nakajima, H., T. Hirota, T. Toba, T. Itoh, and S. Adachi. 1992. Structure of the extracellular polysaccharide from slime-forming *Lactococcus lactis* subsp. *cremoris* SBT 0495. Carbohydr. Res. **224:**245-253.

Oba, T., K. K. Doesburg, T. Iwasaki, and J. Sikkema. 1999. Identification of biosynthetic intermediates of the extracellular polysaccharide viilan in *Lactococcus lactis* subspecies *cremoris* SBT 0495. Arch. Microbiol. **171:**343-349.

Pollock, T. J., W. A. T. van Workum, L. Thorne, M. J. Mikolajczak, J. W. Kijne, and R. W. Armentrout. 1998. Assignment of biochemical functions to glycosyl transferase genes which are essential for biosynthesis of exopolysaccharides in *Sphingomonas* strain S88 and *Rhizobium leguminosarum.* J. Bacteriol. **180:**586-593.

Reeves, P. R., M. Hobbs, M. A. Valvano, M. Skurnik, C. Whitfield, D. Coplin, N. Kido, J. Klena, D. Maskell, C. R. H. Raetz, and P. D. Rick. 1996. Bacterial polysaccharide synthesis and gene nomenclature. Trends Microbiol. **4:**495-503.

Reuber, T. L., S. Long, and G. C. Walker. 1991. Regultation of *Rhizobium meliloti exo* genes in free-living cells and in planta examined using Tn*phoA* fusions. J. Bacteriol. **173:**426-434.

Simon, D., and A. Chopin. 1988. Construction of a vector plasmid family and its use for molecular cloning in *Streptococcus lactis.* Biochimie **70:**559-566.

Saxena, I. M., R. M. Brown, Jr., M. Fevre, R. A. Geremia, and B. Henrissat. 1995. Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. J. Bacteriol **177:**1419-1424.

Stingele, F., J.-R. Neeser, and B. Mollet. 1996. Identification of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. J. Bacteriol. **178:**1680-1690.

Stingele, F., S. J. F. Vincent, E. J. Faber, J. W. Newel, J. P. Kamerling, and J.-R. Neeser. 1999. Introduction of the exopolysaccharide gene cluster from *Streptococcus thermophilus* Sfi6 into *Lactococcus lactis* MG1363: production and characterization of an altered polysaccharide. Mol. Microbiol. **32:**1287-1295.

Tuinier, R. 1999. Ph.D. thesis. Wageningen University and Research Centre, The Netherlands

Summary

Summary

Lactic acid bacteria are Gram-positive bacteria which are used for industrial food fermentation processes. Some have the ability to form exopolysaccharides (EPSs) and these bacteria or the produced EPSs can be used to enhance the structural properties of food products. Furthermore, these EPSs are claimed to be health beneficial. This thesis describes the results of a study on the biosynthesis of these polymers in *Lactococcus lactis* strains.

Chapter 1 provides an overview of the current knowledge of cell-surface polysaccharide biosynthesis, the glycosyltransferases involved, and export and polymerisation processes. Special attention is paid to genetics, regulation, and EPSs produced by LAB.

Chapter 2 describes the characterisation of EPS production by *L. lactis* NIZO B40. The strain produces an extracellular phosphopolysaccharide containing galactose, glucose, and rhamnose. The EPS production is encoded on a 40-kb plasmid, which was isolated after conjugation and subsequent plasmid curing. On this plasmid, a 12-kb region containing 14 genes with the order *epsRXABCDEFGHIJKL* was identified encoding putative gene products which shared sequence homologies with gene products involved in cell-surface polysaccharide biosynthesis of other bacteria. Based on these homologies, predicted functions as regulation (*epsR*), polymerisation and export (*epsA*, *epsB*, *epsI*, *epsK*), or biosynthesis of the repeating unit (*epsD*, *epsE*/*epsF*, *epsG*, *epsH*) could be assigned. The *eps* genes are co-ordinately expressed and transcribed as a single 12-kb mRNA from a promoter upstream of *epsR*. Heterologous expression of *epsD* in *Escherichia coli* showed that its gene product is the so-called priming glucosyltransferase, linking the first sugar of the repeating unit to the lipid carrier.

Chapter 3 describes the functional analysis of the glycosyltransferase genes of the NIZO B40 *eps* gene cluster. The genes were cloned and expressed in *E. coli* and *L. lactis* to determine their function and the sugar-specificity of the encoded enzymes. The EPS consists of repeating units containing a trisaccharide backbone of two glucose and one galactose moieties. The *epsDEFG* gene products are involved in the synthesis of this trisaccharide, linking glucose to a lipid carrier in the membrane (EpsD), glucose to lipid-linked glucose (EpsE/EpsF), and galactose to lipid-linked cellobiose (EpsG), respectively. The *epsJ* gene product was found to be involved in the biosynthesis of EPS and is likely to act either as a galactosyl phosphotransferase or as an enzyme which releases the backbone oligosaccharide from the lipid carrier.

Chapter 4 describes the variety of EPS production by *L. lactis*. Sixteen EPS-producing *L. lactis* strains were analysed and based on the chemical composition of the EPSs they formed and the genotype of their *eps* genes, they were grouped in three major groups and two unique strains. Representatives of the three major groups were studied in detail. Group I comprises strain NIZO B40 which was characterised in the previous chapters. Fragments of the *eps* gene clusters of strains NIZO B35 (group II) and NIZO B891 (group III) were cloned and these encoded the NIZO B35 priming galactosyltransferase, the NIZO B891 priming

glucosyltransferase, and the NIZO B891 galactosyltransferase involved in the second step of repeating unit synthesis.

First successful attempts for genetic engineering of the EPS production were achieved by replacing the NIZO B40 priming glucosyltransferase gene, *epsD*, by an erythromycin resistance gene which resulted in the loss of EPS production and the complementation of the EPS-producing phenotype by controlled expression of priming glycosyltransferase genes from Gram-positive organisms with known function and substrate specificity.

In Chapter 5 the regions involved in replication and mobilisation of the NIZO B40 EPS-plasmid pNZ4000 were characterised. The plasmid contains four highly conserved replication regions that belong to the lactococcal theta replicon family and all are functional and compatible in *L. lactis*. Plasmid pNZ4000 was shown to be a mobilisation plasmid and two regions involved in mobilisation were identified. Both regions contained a functional origin of transfer (*oriT*). One *oriT* sequence was followed by a *mobA* gene, coding for a *trans*-acting protein involved in conjugative transfer and likely to be the relaxase nicking the *nic* sites of the *oriT* sequences.

Chapter 6 describes the complete nucleotide sequence of the EPS-plasmid pNZ4000, which amounts to 42810 bp and represents one of the largest sequenced plasmids in LAB to date. Apart from the regions involved in EPS biosynthesis, replication, and mobilisation, described in Chapters 2 and 5, two regions potentially involved in transport of divalent cations were localised on pNZ4000.

In Chapter 7 the results of the previous chapters are discussed and their implications on practical applications and in particular the perspectives for polysaccharide engineering are described.

Samenvatting

Samenvatting

Melkzuurbacteriën zijn Gram-positieve bacteriën die gebruikt worden in fermentaties voor de voedingsmiddelenindustrie. Sommigen bezitten het vermogen om extracellulaire polysachariden (EPS-en) te vormen. Deze bacteriën, of de door hen gevormde EPS-en, kunnen gebruikt worden om de structuureigenschappen van voedingsmiddelen te verbeteren. Tevens is er gesuggereerd dat EPS-en gezondheidsbevorderend zouden zijn. Dit proefschrift beschrijft de resultaten van een studie naar de biosynthese van deze polymeren in *Lactococcus lactis* stammen.

Hoofdstuk 1 beschrijft de huidige kennis van de biosynthese van polysachariden die door bacteriën uitgescheiden worden en de enzymen betrokken bij de opbouw, polymerisatie en export van deze polymeren. Speciale aandacht wordt gegeven aan de genetica, regulatie van genexpressie en EPS-en die door melkzuurbacteriën gevormd worden.

Hoofdstuk 2 beschrijft de karakterisering van het EPS dat gevormd wordt door de modelbacterie van deze studie, *Lactococcus lactis* stam NIZO B40. Deze stam vormt een EPS dat repeterende eenheden van vijf suikers bevat: twee galactoses, twee glucoses en een ramnose. De genetische informatie voor de EPS productie ligt bij deze stam gecodeerd op een 40-kb plasmide. Een kwart van dit plasmide bevat 14 genen in de volgorde *epsRXABCDEFGHIJKL* die de informatie bevatten voor de aanmaak van enzymen die overeenkomstige eigenschappen hebben als enzymen die in andere bacteriën betrokken zijn bij de biosynthese van polysachariden. Hierdoor is het mogelijk om verschillende functies voor deze genen voor te stellen zoals regulatie (*epsR*), polymerisatie en export (*epsA*, *epsB*, *epsI*, *epsK*) of de aaneenkoppeling van suikers voor de opbouw van de repeterende eenheid (*epsD*, *epsE*/*epsF*, *epsG*, *epsH*). De *eps* genen komen gecoördineerd tot expressie en worden van het DNA afgeschreven als een enkel transcript van 12-kb van een promoter voor *epsR*. De functionaliteit van het *epsD* gen in de EPS biosynthese is aangetoond door dit gen in een andere bacterie, *Escherichia coli*, tot expressie te brengen. Hieruit bleek dat dit gen codeert voor een glycosyltransferase dat de eerste suiker van de repeterende eenheid koppelt aan een celmembraancomponent, de lipidedrager.

Hoofdstuk 3 beschrijft de functionele analyse van de glycosyltransferase genen van het NIZO B40 *eps* gencluster. Deze genen zijn betrokken bij de vervolgstappen in de opbouw van de repeterende eenheid. De suikerspecificiteit van de glycosyltransferases is bepaald door de genen zowel in *E. coli* als in *L. lactis* tot expressie te brengen. De eerste stap in de EPS biosynthese is de koppeling van glucose aan een lipidedrager (EpsD). De tweede stap is de koppeling van glucose aan de eerste glucose (EpsE/EpsF). Vervolgens wordt galactose gekoppeld aan de tweede glucose (EpsG). Tevens is aangetoond dat ook het enzym dat gecodeerd wordt door *epsJ* een rol speelt bij de EPS biosynthese en dat het waarschijnlijk een galactosyltransferase is of een enzym dat de repeterende eenheid losmaakt van de lipidedrager.

Hoofdstuk 4 beschrijft de variatie aan EPS-en die gevormd worden door verschillende *L. lactis* stammen. Zestien EPS producerende *L. lactis* stammen zijn geanalyseerd en op basis

van de suikersamenstelling van het door hen gevormde EPS en de genetische variatie zijn ze onderverdeeld in drie groepen en twee unieke stammen. Van ieder van de drie groepen werd een stam uitgekozen als vertegenwoordiger om in detail bestudeerd te worden. Groep I bevat stam NIZO B40 die in hoofdstukken 3 en 4 beschreven wordt. Fragmenten van de *eps* genclusters van stam NIZO B35 (groep II) en NIZO B891 (groep III) zijn gekloneerd en deze bevatten de genen die coderen voor de galactosyltransferase van NIZO B35 die de eerste suiker aan de lipidedrager koppelt en de glucosyltransferase en galactosyltransferase van NIZO B891 die respectievelijk de eerste en tweede suiker koppelen. Ook staan in dit hoofdstuk de eerste succesvolle stappen beschreven om te komen tot het veranderen van de structuur van het EPS door genetische modificatie. Hiervoor is een glycosyltransferase gen van het NIZO B40 *eps* gencluster (*epsD*) weggenomen en vervolgens is dit gen of een overeenkomstig gen van een andere Gram-positieve bacterie in een gecontroleerd expressie systeem weer in de bacterie teruggebracht om de functie van het ontbrekende gen over te nemen.

Hoofdstuk 5 beschrijft de gebieden van het EPS plasmide pNZ4000 van stam NIZO B40 die betrokken zijn bij de replicatie van het plasmide en de overdracht naar andere bacteriestammen. Het plasmide bevat vier overeenkomstige replicatiegebieden die allen functioneel zijn. Het EPS plasmide is overdraagbaar naar andere *L. lactis* stammen en de betrokkenheid van twee gebieden is bestudeerd. Eén gebied heeft een DNA structuur dat de oorsprong voor overdracht (*origin of transfer*), *oriT*, bevat. Een tweede gebied heeft naast een *oriT* ook nog een *mobA* gen dat codeert voor een eiwit dat betrokken is bij de plasmide overdracht en zeer waarschijnlijk aangrijpt op een specifieke nucleotiden sequentie in de *oriT*.

Hoofdstuk 6 beschrijft de gehele basenvolgorde van het EPS plasmide pNZ4000 dat een grootte heeft van 42810 bp. Dit is momenteel een van de grootste plasmiden van melkzuurbacteriën waarvan de complete basenvolgorde bepaald is. Naast de gebieden die betrokken zijn bij de EPS biosynthese, plasmide replicatie en plasmide overdracht, beschreven in hoofdstukken 2 en 5, zijn er nog twee gebieden gevonden die mogelijk betrokken zijn bij de opname van twee-waardige kationen.

In Hoofdstuk 7 worden de resultaten van de vorige hoofdstukken bediscussieerd en wordt gekeken naar hun waarde voor praktische toepassingen en met name om te komen tot modificatie van polysacharide structuren en eigenschappen.

Dankwoord

Het dankwoord is het misschien wel het meest kritieke gedeelte van het proefschrift om te schrijven. Iedereen leest het en iedereen is op zoek naar zichzelf. De schrijver loopt dus altijd het gevaar om ongewild toch iemand over het hoofd te zien. Daarom wil ik beginnen met iedereen te bedanken die aan het tot standkomen van het proefschrift heeft bijgedragen. Speciale dank gaat uit naar mijn promotor Willem de Vos. Willem, bedankt voor het enthousiasme waarmee je mijn werk begeleid hebt. De afgelopen acht jaar ben je een van de meest bepalende personen geweest voor mijn wetenschappelijke vorming. Joey, ik wil ook jou bedanken voor je begeleiding. Ik begon bij je als student en toen heb je me ingewijd in het sociale NIZO leven. Maar ook bij het begin van mijn promotieperiode heb je een enorme invloed gehad op het werk. Michiel, toen Joey weg was ben jij bij het EPS-werk betrokken geraakt. Ik wil je bedanken voor je input bij het afronden en het opschrijven van dit werk. Ik hoop dat onze prettige samenwerking nog een tijd voort kan duren. Ik wil al mijn collega's van de genetica groep van BFC bedanken voor de aangename tijd op het lab en de sociale gebeurtenissen die buiten werktijd gebeurden: Oscar (zit-ie-erbij?), Ger (wedden voor een doos gebak dat...), Ingrid (geen sloffen meer in de kerstboom stoppen, hoor), Saskia (Wieckse Witte bij de Sub), Hans (ze hebben mijn fiets gestolen...), Roger (vis, vogelsoep en oude koeien), Marke (ik wil het voor mijn promotie nog even over dat archief van je hebben), Jeroen (wiehoe!), Benoît (allez Ajax), Patrick (altijd weer een leuk mailtje), Roland (nu horizontaal geprogrameerd), María ('t is me wat), Hélène (this is ridiculous) en Cindy (altijd het middelpunt van de belangstelling). Speciale aandacht verdienen natuurlijk mijn labgenootjes Pascalle (die van de P_{nis}) en Ingeborg (een mededradentrekker), die altijd een gewillig slachtoffer vormden en mijn zangkunsten wisten te waarderen, mijn studenten Norwin (je blijft voor mij toch Nochi), Sónia en Harmjan (de vlieger), die een belangrijke rol bij het verkrijgen van de resultaten gespeeld hebben, en Iris (per definitie onschuldig), mijn steun en toeverlaat op weg naar de top. Naast Mariska, Martijn, René, Wim, Christel, en Liesbeth wil ik alle studenten en tijdelijke medewerkers in de groep bedanken. Ook de mensen van voorheen BFC-boven bedankt voor de samenwerking en natuurlijk de MI-tjes die nog niet genoemd zijn (Tanja, Maarten, Marjo, Dick, Bert, Carol, Andrea, Ellen, Jeroen). Een speciale vergadergroep binnen NIZO was het EPS-overleg waarin we trachtten een brug tussen verschillende disciplines te slaan. Hiervan heb ik nog niet genoemd: Remco, Nel, Marc, Fedde en Marieke.

Hmmm, dan moet ik hier maar even een aparte alinea wijden aan Evert. Wij vormden namelijk een prima duo. Pascalle en Cindy zijn nooit in de buurt hiervan gekomen (zaterdagamateurs in vergelijking tot eredivisie). Men begreep ons niet, maar dat vonden we

niet erg. Evert, wij kunnen met elkaar tenminste over de essentiële zaken in het leven praten en zijn het dan roerend met elkaar eens: een bovenlader afgewerkt met hout.

Tijdens mijn promotietijd heb ik niet alleen te maken gehad met mensen van het NIZO, maar ook met collega's uit het polysachariden-veld. Daarom, Marc Kolkman, Dick van den Berg (tegenwoordig een gewaardeerd NIZO-collega), en Jan Jore bedankt voor jullie samenwerking en discussie over de polysachariden-genetica.

Aan het einde van dit dankwoord wil ik de mensen bedanken die me het meest dierbaar zijn: Marja en mijn ouders. Marja, bedankt voor je liefde, steun en geduld. Pa en ma, jullie hebben me altijd gesteund en gestimuleerd om het meeste uit mezelf te halen. Ma, we hadden je er graag nog bij willen hebben.

Curriculum Vitae

Richard van Kranenburg werd geboren op 3 februari 1970 in Amsterdam. Nadat hij in 1988 geslaagd was voor zijn diploma van het Sint Oelbert Gymnasium in Oosterhout begon hij aan een studie Biologie, oriëntatie celbiologie, aan de Landbouwuniversiteit Wageningen. Deze studie is cum laude afgesloten met twee afstudeervakken en twee stages. De twee afstudeervakken werden gecombineerd en uitgevoerd bij de vakgroep Microbiologie onder begeleiding van dr. Rik Eggen en bij de vakgroep Biochemie onder begeleiding van Marc Verhagen. De stages werden uitgevoerd bij de Protein Biochemistry Department van Glaxo Group Research in Greenford (Londen) onder begeleiding van dr. Richard Hale en dr. Rob Cooke en bij de Bacteriële Genetica groep van de afdeling Biofysische Chemie van NIZO onder begeleiding van dr. Joey Marugg. Vlak voor zijn afstuderen kreeg hij per 1 november 1993 een aanstelling voor vier jaar als junior onderzoeker bij NIZO. In deze periode werd onder begeleiding van Prof. Willem de Vos het onderzoek verricht dat in dit proefschrift beschreven staat. Sinds 1 februari 1998 is hij werkzaam als postdoctoraal onderzoeker bij NIZO food research binnen de sectie Microbial Ingredients die nu is opgegaan in de afdeling Flavours & Natural Ingredients. Hij voert hier onderzoek uit voor het Wageningen Centre for Food Sciences (WCFS).

List of Publications

Eggen, R. I. L., R. van Kranenburg, A. J. M. Vriesema, A. C. M. Geerling, M. F. J. M. Verhagen, W. R. Hagen, and W. M. de Vos. 1996. Carbon monoxide dehydrogenase from *Methanosarcina frisia* Göl. J. Biol. Chem. **271:**14256-14263.

Marugg, J. D., W. Meijer, R. van Kranenburg, P. Laverman, P. G. Bruinenberg, and W. M. de Vos. 1995. Medium-dependent regulation of proteinase gene expression in *Lactococcus lactis*: control of transcription initiation by specific peptides. J. Bacteriol. **177:**2982-2989.

Marugg, J. D., R. van Kranenburg, P. Laverman, G. A. M. Rutten, and W. M. de Vos. 1996. Identical transcriptional control of the divergently transcribed *prtP* and *prtM* genes that are required for proteinase production in *Lactococcus lactis* SK11. J. Bacteriol. **178:**1525-1531.

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.

de Vos, W. M., P. Hols, R. van Kranenburg, E. Luesink, O.P. Kuipers, J. van der Oost, M. Kleerebezem, and J. Hugenholtz. 1998. Making more of milk sugar by engineering lactic acid bacteria. Int. Dairy J. **8:**227-234.

van Kranenburg, R., and W. M. de Vos. 1998. Characterization of multiple regions involved in replication and mobilization of plasmid pNZ4000 coding for exopolysaccharide production in *Lactococcus lactis*. J. Bacteriol. **180:**5285-5290.

van Kranenburg, R., I. I. van Swam, J. D. Marugg, M. Kleerebezem, and W. M. de Vos. 1999. Exopolysaccharide biosynthesis in *Lactococcus lactis* NIZO B40: functional analysis of the glycosyltransferase genes involved in synthesis of the polysaccharide backbone. J. Bacteriol. **181:**338-340.

van Kranenburg, R., H. R. Vos, I. I. van Swam, M. Kleerebezem, and W. M de Vos. 1999. Functional analysis of glycosyltransferase genes from *Lactococcus lactis* and other Gram-positive cocci: complementation, expression, and diversity. J. Bacteriol. **181:**6347-6353.

Kleerebezem, M., R. van Kranenburg, R. Tuinier, I. C. Boels, P. Zoon, E. Looijesteijn, J. Hugenholtz, and W. M. de Vos. 1999. Exopolysaccharides produced by *Lactococcus lactis*: from genetic engineering to improved rheological properties? Antonie van Leeuwenhoek **76:**357-365.

van Kranenburg, R., I. C. Boels, M. Kleerebezem, and W. M. de Vos. 1999. Genetics and engineering of microbial exopolysaccharides for food: approaches for the production of existing and novel polysaccharides. Curr. Opinion Biotechnol. **10:**498-504.

van Kranenburg, R., M. Kleerebezem, and W. M de Vos. 1999. Nucleotide sequence analysis of the lactococcal EPS plasmid pNZ4000. Submitted for publication.

Fernández, M., W. van Doesburg, G. A. M. Rutten, J. D. Marugg, A. C. Alting, R. van Kranenburg, and O. P. Kuipers. 1999. Molecular and functional analysis of the *metC* gene of *Lactococcus lactis* encoding cystathionine β-lyase. Submitted for publication.