Biochemical characterization of the β-1,4-glucuronosyltransferase GelK in the gellan gum-producing strain *Sphingomonas paucimobilis* A.T.C.C. 31461

Paula VIDEIRA*, Arsénio FIALHO*, Roberto A. GEREMIA†[1], Christelle BRETON‡ and Isabel SÁ-CORREIA*

*Centro de Engenharia Biológica e Química, Instituto Superior Técnico, Av. Rovisco Pais, 1049–001 Lisboa, Portugal, †Plasticité et Expression des Génomes Microbiens, Université Joseph Fourier and Centre National de la Recherche Scientifique FRE 2383, Bat. CERMO, 460 Rue de la Piscine, 38041 Grenoble cedex 9, France, and ‡Centre de Recherches sur les Macromolécules Végétales, CNRS, BP 53X, 38041 Grenoble, France

Biosynthesis of bacterial polysaccharide-repeat units proceeds by sequential transfer of sugars, from the appropriate sugar donor to an activated lipid carrier, by committed glycosyltransferases (GTs). Few studies on the mechanism of action for this type of GT are available. *Sphingomonas paucimobilis* A.T.C.C. 31461 produces the industrially important polysaccharide gellan gum. We have cloned the *gelK* gene from *S. paucimobilis* A.T.C.C. 31461. GelK belongs to family 1 of the GT classification [Campbell, Davies, Bulone, Henrissat (1997) Biochem. J. **326**, 929–939]. Sequence similarity studies suggest that GelK consists of two protein modules corresponding to the -NH_2 and -CO_2H halves, the latter possibly harbouring the GT activity. The *gelK* gene and the open reading frames coding for the -NH_2 ($GelK_{NH2}$) and -CO_2H ($GelK_{COOH}$) halves were overexpressed in *Escherichia coli*. GelK and $GelK_{NH2}$ were present in both the soluble and membrane fractions of *E. coli*, whereas $GelK_{COOH}$ was only present in the soluble fraction. GelK catalysed the transfer of [^{14}C]glucuronic acid from UDP-[^{14}C]glucuronic acid into a glycolipid extracted from *S. paucimobilis* or *E. coli*, even in the presence of EDTA, and the radioactive sugar was released from the glycolipid by β-1,4-glucuronidase. GelK was not able to use synthetic glucosyl derivatives as acceptors, indicating that the PP_i-lipid moiety is needed for enzymic activity. Recombinant $GelK_{NH2}$ and $GelK_{COOH}$ did not show detectable activity. Based on the biochemical characteristics of GelK and on sequence similarities with N-acetylglucosaminyltransferase, we propose that GT families 1 and 28 form a superfamily.

Key words: bioinformatics, conserved amino acids, exopolysaccharide, N-acetylglucosaminyltransferase, secondary structure prediction.

INTRODUCTION

Extracellular polysaccharides of microbial origin represent an important class of polymeric materials of interest in biotechnology for their unique properties as viscosifiers, stabilizers, emulsifiers and gelling agents. Among them, gellan gum is a new commercial gelling agent produced in high yields by *Sphingomonas paucimobilis* A.T.C.C. 31461. Gellan is composed of a repeating linear tetrasaccharide with D-glucose (Glc), D-glucuronic acid (GlcA) and L-rhamnose residues in a 2:1:1 ratio [1], with glycerate and acetate substitutions [2] (Figure 1). It is known that even subtle changes in the composition/structure of a polysaccharide may drastically affect its properties and subsequently its applications, as clearly evidenced by studies on gellan gum [3,4]. Perhaps the most exciting prospects for microbial polysaccharide modification is genetic engineering, although a prerequisite for progress in this area is the elucidation of the genetics and biochemistry of their synthesis. The gellan biosynthetic pathway has been partially elucidated [5]. The *pgmG* gene, encoding a bifunctional protein with phosphogluco- and phosphomanno-mutase activities, was identified first. PgmG protein may convert Glc 6-phosphate into Glc 1-phosphate, required for the synthesis of the three gellan sugar-nucleotide precursors UDP-D-Glc, dTDP-L-rhamnose and UDP-D-GlcA [5,6]. However, subsequent steps of gellan synthesis and export remain unclear, although the gellan gene cluster was recently identified [6]. The organization of the gellan gene cluster is apparently identical to that previously reported for the sphingan S88 biosynthetic cluster in *Sphingomonas* S88 [7]. These clusters include genes essential for the assembly and secretion of exopolysaccharide and a four-gene operon needed for the synthesis of dTDP-L-rhamnose, but not the *pgmG* gene [6,7]. SpsK is possibly a β-1,4-glucuronosyltransferase, catalysing the addition of GlcA into the glucosyl-α-pyrophosphorylpolyprenol (PPL) intermediate (Figure 1) [8].

One attractive long-term goal in polysaccharide engineering is the controlled modification of sugar composition to produce tailored polysaccharides. This goal requires a deeper understanding of glycosyltransferase (GT) structure and function. To date, our knowledge of GTs is very limited due to the difficulty in obtaining appropriate acceptors. From the sequence point of view, GTs have been classified in 51 families [9,10]. However, the elucidation of the three-dimensional structure of a few GTs, sequence comparison studies and site-directed mutagenesis, suggest that the number of structural folds is much lower. Based on enzyme interaction with nucleotide sugars, the available three-dimensional structures can be classified into two groups. The first group of GTs includes six proteins of prokaryotic and eukaryotic organisms requiring divalent metal cations for activity [11–16], and possess a UDP-binding domain (UDB or SGC domain) involved in interactions with the nucleotide moiety of the nucleotide sugar [14,16,17]. This UDP-binding domain

Abbreviations used: IPTG, isopropyl β-D-thiogalactoside; GT, glycosyltransferase; MurG, UDP-N-acetylglucosamine:N-acetylmuramyl(pentapeptide) pyrophosphoryl-undecaprenol β-1,4 N-acetylglucosaminyltransferase; T4βGT, T4 β-glucosyltransferase; Glc, glucose; GlcA, glucuronic acid; GlcNAc, *N*-acetylglucosamine; PPL, pyrophosphorylpolyprenol; AceA, GDPmannose:cellobiosyl-diphosphopolyprenol α1,3 mannosyltransferase; ExoM, UDPglucose:glucose β-1,4 glucose β-1,4 galactosyl-diphosphopolyprenol β-1,4 glucosyltransferase.

[1] To whom correspondence should be addressed (e-mail roberto.geremia@ujf-grenoble.fr).

[2] The nucleotide sequence data reported appears in DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession number AF305842.

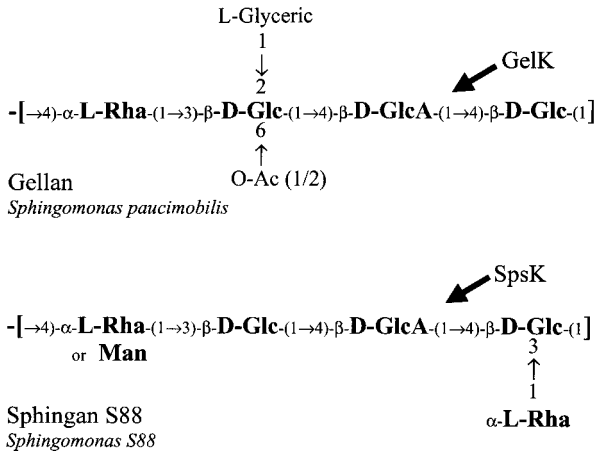


Figure 1 Structure of the repeat unit of gellan gum produced by *S. paucimobilis* A.T.C.C. 31461 and of sphingan S88 produced by *Sphingomonas* S88

The glucuronosyltransferases SpsK and GelK catalyse the formation of the linkage indicated by the arrows, probably using Glc-PPL as acceptor.

presents a motif (DXD or DD motif), that is involved both in recognition of the -OH moiety of the ribose and bridging the PP_i through a divalent metal cation (reviewed in Ünligil and Rini [17]). The second structural group of GTs includes the *Escherichia coli* UDP-N-acetylglucosamine:N-acetylmuramyl(pentapeptide) pyrophosphoryl-undecaprenol β-1,4 N-acetylglucosaminyltransferase (MurG), [18] and the phage T4 β-glucosyltransferase (T4βGT) [19,20]. In this structural group a positively charged residue (arginine) interacts with the PP_i group, whereas a hydroxy group of the ribose is recognized by a glutamate residue.

A systematic survey using different acceptors to determine the acceptor specificity of MurG showed that the PP_i moiety of the undecaprenyl-PP_i-muramic acid acceptor (and not the lipid moiety) is essential for activity [21]. Moreover, it was proposed that a structural 'G-loop' motif of MurG may be involved in recognition of the acceptor's PP_i [18]. The mannosyltransferase GDPmannose:cellobiosyl-diphosphopolyprenol α1,3 mannosyltransferase (AceA) from *Acetobacter xylinum* is not active towards free oligosaccharidic acceptor [22]. For the glucosyltransferase UDPglucose:glucose β-1,4 glucose β-1,4 galactosyldiphosphopolyprenol β-1,4 glucosyltransferase (ExoM) from *Sinorhizobium meliloti*, free oligosaccharides do not inhibit transfer towards the lipid-bound oligosaccharides [23]. The emerging common theme is that most of the GTs involved in repeat-unit biosynthesis require features of the PP_i-polyprenol anchor for activity.

In the present study we report the cloning of *gelK* gene and the biochemical characterization of the encoded protein. We show that GelK is a UDP-glucuronate:glucosyl-PPL β-1,4-glucuronosyltransferase that is not inhibited by EDTA and requires features of the PPL moiety for activity. Sequence similarity studies and biochemical features suggest that GelK and, by extension, all family 1 GTs are structurally related to MurG and, by extension, to all family 28 GTs.

MATERIALS AND METHODS

Bacterial strains, plasmids and media

The strains and plasmids used are listed in Table 1. *E. coli* strains were grown in Lennox broth (Sigma) at 37 °C with orbital agitation. *S. paucimobilis* A.T.C.C. 31461 was grown in S medium [6] at 30 °C. When required, antibiotics were used at the following concentrations in the selective media: ampicillin, (150 mg/l), chloramphenicol (25 mg/l) and kanamycin (100 mg/l). The oligonucleotides were purchased from Amersham Pharmacia Biotech. Restriction enzymes were from Gibco BRL, DNA polymerases from Boerhinger GmbH (Mannheim, Germany), and UDP-[^{14}C]GlcA and UDP-[^{14}C]Glc were from NEN Life Science Products (Boston, MA, U.S.A.). Glcα/β-O-$(CH_2)_3$-NH-CO-$(CH_2)_4$-NH-biotin was kindly provided by N. V. Bovin (Syntesome GmbH, Munich, Germany). All other chemicals were from Sigma.

DNA manipulations

E. coli XL1-Blue cells were used as the host for recombinant plasmids. Plasmid DNA was purified from *E. coli* cultures by the alkaline lysis procedure [24] or using the QIAprep spin plasmid kit (Qiagen). DNA restriction, agarose gel electrophoresis, cloning procedures, colony hybridization and other related techniques were carried out following established protocols [25].

Cloning of total and partial *gelK* sequences

To obtain by PCR amplification an internal region of the *S. paucimobilis gelK* gene, synthetic oligonucleotides A and B were designed based on the nucleotide sequence of two random regions of *spsK* gene of *Sphingomonas* S88, at the positions indicated in Figure 2. The PCR reaction mixture contained, in a final volume of 50 μl, 200 ng of *S. paucimobilis* A.T.C.C. 31461 chromosomal DNA, 10 μM of each deoxynucleoside triphosphate (Promega), 300 pmol of each primer and 2.5 units of *Taq* DNA polymerase, in the presence of *Taq* DNA polymerase buffer and 2.5 mM of $MgCl_2$. PCR amplification was performed using a PTC-100 thermocycler (MJ Research, Watertown, MA, U.S.A.) under the following conditions: 30 amplification cycles consisting of 60 s at 95 °C, 40 s at 65 °C and 60 s at 72 °C. The predicted 567 bp PCR product was recovered from a 1% (w/v) low-melting agarose gel (FMC Bioproducts, Rockland, ME, U.S.A.), purified by using WizardPCR Preps DNA purification kit (Promega) and sequenced to confirm the identity. This PCR product was used as a probe to screen, by colony hybridization, the 1200 clones of the *S. paucimobilis* A.T.C.C. 31461 genomic library [6]. One of the colonies that hybridized with the probe, clone p96, was used to clone the complete *gelK* gene and two partial sequences corresponding to the carboxyl ($GelK_{COOH}$) and amino ($GelK_{NH2}$) regions of GelK protein. These DNA regions were obtained by PCR amplification (carried out using *Pwo* DNA polymerase and 30 amplification cycles consisting of 60 s at 95 °C, 40 s at 64 °C and 60 s at 72 °C) using cosmid p96 as template, and were cloned into pET29 vector leading to pGelK, $pGelK_{COOH}$ and $pGelK_{NH2}$ constructs. The oligonucleotide primers used to amplify the complete and partial gene sequences were designed based on the nucleotide sequence of *spsK* gene; their sequences and positions in the *spsK* nucleotide sequence are indicated in Figure 2. The different DNA fragments thus obtained were sequenced to verify the fidelity of PCR amplification.

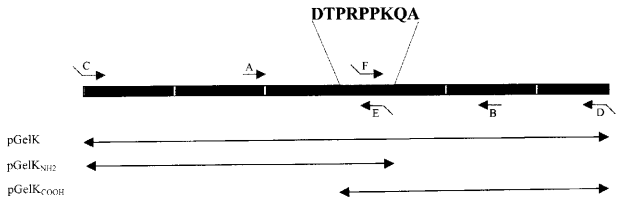
Sequence analysis

A nucleotide sequence of approx. 1600 bp, including the complete *gelK* gene and the upstream (400 nt) and downstream (200 nt) regions, was determined on both strands at the DNA Core Facility, University of Missouri, Columbia, MO, U.S.A.. This

Table 1 Bacterial strains and plasmids used

Abbreviations used: Ap, ampicillin; Kan, kanamycin; r, resistant; s, sensitive; Gel^+, gellan producer. The oligonucleotide sequences of the referred primers are indicated in Figure 2.

Strain or plasmid	Genotype or relevant characteristic	Source or reference
Strains		
E. coli		
XL1-Blue	*rec*A1 lac [F′ proAB *lacI*^{q} *Z* Δ *M15*, Tn*10* (Tet^r)], *thi*	Stratagene
BL21(DE3)	F^- $omp^+ r_B^-$ m_B^- (DE3 is a lysogen bearing *lac*I *lacUV*5 gene 1)	[36]
S. paucimobilis A.T.C.C. 31461	Gel^+	[4]
Plasmids		
Hypercos 1	Col E1 *cos*$^+$ Mob^+ Ap^r	Stratagene
pET29(a)	Kan^r ColE1, P_{T7} and RBS	Novagen
pET29(b)	Kan^r ColE1, P_{T7} and RBS	Novagen
p96	Hypercos 1 clone (with a ~ 35 kb insert) from *S. paucimobilis* A.T.C.C. 31461 genomic bank	This study
pGelK	pET29(a) derivative containing a *Bam*HI-*Eco*RI fragment amplified from p96 using primers C and D	This study
$pGelK_{NH2}$	pET29(a) derivative containing a *Bam*HI-*Eco*RI fragment amplified from p96 using primers C and E	This study
$pGelK_{COOH}$	pET29(b) derivative containing a *Bam*HI-*Eco*RI fragment amplified from p96 using primers F and D	This study



Oligonucleotide	Sequence[a] (5'-3')	Position[b]
A	CGGTGCGGGCGCGGTCTATT	339 to 348
B	GTCTGCGCGATCTCTTCCTGGTGA	872 to 895
C	GGGGATCC ATGGGGGACGGCATGGCA	1 to 18
D	GGGAATTC TAGTGCGGGATGACGACG	1075 to 1092
E	GGGGATCC CCCCGCCCTCCCAAGCAG	540 to 557
F	GGGAATTC CTGCTTGGGAGGGCGGGG	540 to 557

Figure 2 Schematic representation of GelK constructs

Upper panel: the *gelK* open-reading frame showing the position of the primers (arrows) used to generate the various GelK constructs (double-headed arrows) characterized in this paper. Lower panel: nucleotide sequence of primers. Primers A and B were designed based on *Sphingomonas* S88 *spsK* nucleotide sequence (accession number SS:U51197.1). [a]Underlined sequences indicate the target sequences for *Bam*HI or *Eco*RI introduced in the sense and antisense primers respectively. [b]Position of the primers is given using adenine of the initiation codon in *spsK* gene as base one.

sequence was deposited in the GenBank® database under the accession number AF305842.

Protein sequences were retrieved from GenBank® and other databanks and analysed using BLAST programs [26], LALIGN [27], and ClustalW [28]. The sequence alignment displayed in Figure 3 was generated using ClustalW and refined manually from the results obtained using the sensitive hydrophobic cluster analysis method and secondary structure prediction methods.

Expression of *gelK*: protein production and cell fractionation

Competent *E. coli* BL21(DE3) cells were transformed with plasmids pGelK, $pGelK_{NH2}$, $pGelK_{COOH}$ or the cloning vector pET29. Cultures were inoculated with 1/100 of a pre-culture grown overnight. Optimal recombinant protein production was obtained by isopropyl β-D-thiogalactoside (IPTG; 0.1 mM) induction of cultures at a D_{640nm} value ranging from 0.6–0.8, followed by 2 h of cultivation.

The expression of the recombinant proteins by *E. coli* BL21(DE3) was monitored by withdrawing 1.2 D_{640nm} equivalents/ml at the different conditions. Cells were harvested by centrifugation, resuspended in 100 μl of denaturing buffer, and the resulting protein extract was subjected to SDS/PAGE (12% gels) and stained with Coomassie Brilliant Blue R250. The N-terminalS-Tag epitope fusion was detected by protein transfer from SDS/PAGE gels to a PVDF membrane and treatment with the Novagen S-tag Western-blotting kit, using alkaline phosphatase S-protein conjugate according to the manufacturer's instructions.

For monitoring protein localization, *E. coli* BL21(DE3) cells transformed with the recombinant plasmids or the cloning vector were harvested 2 h after IPTG induction, washed once with 70 mM Tris/HCl (pH 8.2), resuspended in 1 ml of 50 mM Tris/HCl, 1 mM EDTA (pH 8.2) buffer at 1.2 D_{640nm} equivalents/ml and broken by two passages through a French press (12.4×10^4 kPa; 18000 lb/in²). Cell debris was removed by centrifugation (12000 ***g*** at 4 °C for 10 min) and the supernatant was submitted to a further centrifugation at 50000 ***g*** for 1 h at 4 °C to recover membrane and soluble fractions. The pellets were resuspended in the same buffer, at a volume equivalent to the starting point.

To assay β-glucuronosyltransferase activity, crude cell extracts or membrane and soluble fractions were obtained following essentially the same procedure, using a more concentrated cell suspension (corresponding to 30 D_{640nm} equivalents/ml) and the following suspending buffer: 50 mM Tris/HCl (pH 8.2), 1 mM EDTA and 0.5 mM PMSF. The different cell fractions were separated in aliquots, which were stored at −20 °C until used.

Preparation of *S. paucimobilis* and *E. coli* glycolipid acceptors

EDTA-permeabilized bacterial cells were obtained using the method described by Tolmasky et al. [29], slightly modified as described below. Cells of *S. paucimobilis* were obtained by cultivation in S medium at 30 °C. Cells were harvested by centrifugation of the culture previously diluted (1:3) to reduce broth viscosity. *E. coli* BL21(DE3) cells carrying either plasmid pGelK or the cloning vector pET29 were grown in 200 ml of Lennox broth at 37 °C, induced with IPTG as described above, and harvested by centrifugation. Bacterial cells were then washed with 70 mM Tris/HCl (pH 8.2) and 10 mM EDTA, and resuspended in 60 ml of the same buffer and submitted to three cycles of freeze–thawing in liquid N_2. Permeabilized cells were

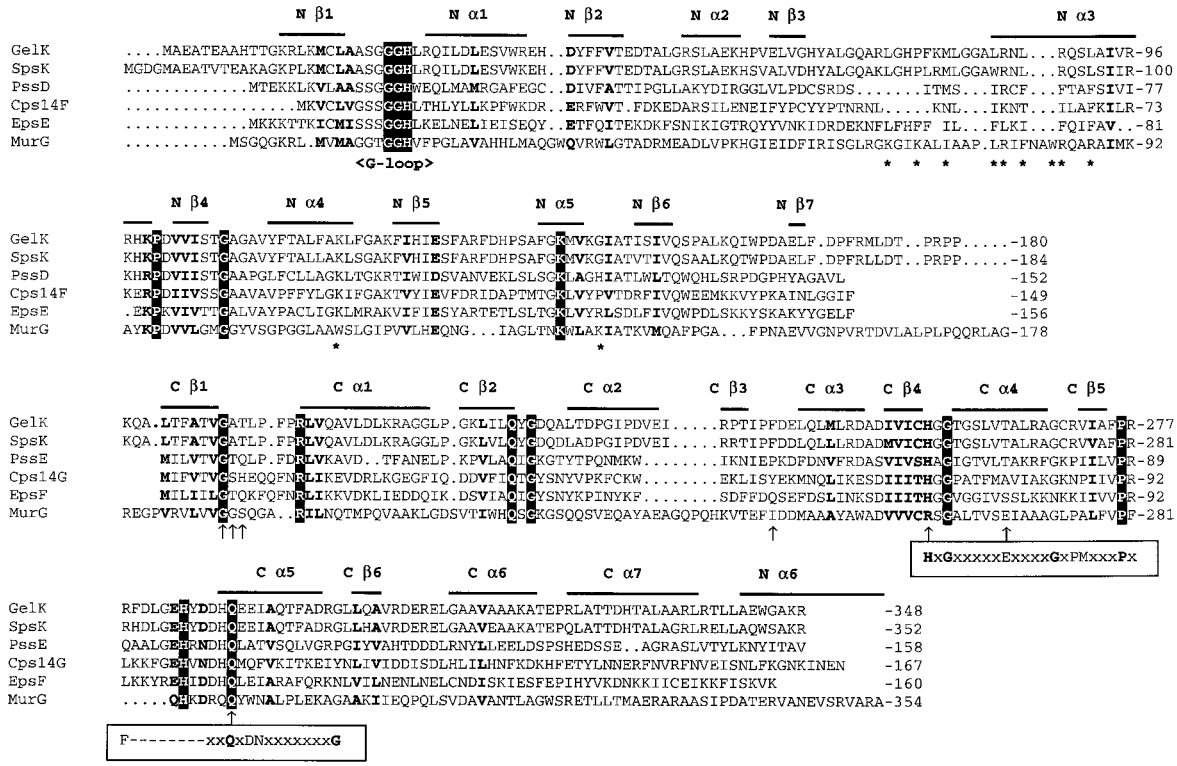


Figure 3 Sequence alignment of GelK with other bacterial homologues and identification of critical residues in GelK

This alignment was generated using ClustalW [26] and refined manually as described in the Materials and methods section. Secondary structure elements were determined from the three-dimensional structure of *E. coli* MurG and are indicated above the sequences. The numbering of α-helices and β-strands is as given in Ha et al. [18]. Invariant amino acids are indicated in white on a black background and the most conserved residues are indicated in **bold** (groups of similar residues are defined as follows: [AILMV], [FWY], [DENQ], [RHK] and [CST]). Boxed consensus sequence delineates the conserved peptide motif found in the large UDP-glucuronosyltransferase family (residues in **bold** represent those that are invariant in GelK). The conserved G-loop in GelK is indicated. Arrows indicate residues in MurG that are predicted to interact with UDP-GlcNAc and asterisks indicate a hydrophobic stretch in MurG proposed to be involved in membrane association [18].

then centrifuged, resuspended in 50 mM Tris/HCl (pH 8.2) and 5 mM EDTA at 200 D_{640nm} equivalents/ml and finally stored in 150 μl aliquots at -70 °C. To prepare glycolipid acceptors, 150 μl of permeabilized cell suspensions were incubated at 16 °C for 30 min in the presence of 2.8 mM UDP-Glc, 72 mM Tris/HCl (pH 8.2) and 8.5 mM $MgCl_2$. The reaction was stopped by the addition of 0.2 ml of 70 mM Tris/HCl (pH 8.2) and 10 mM EDTA, and washed twice with the same buffer. The glycolipid fraction was recovered by extractions with chloroform/methanol/water at a ratio of 1:2:0.3 (by vol.), as previously described [23]. Extracts were dried to a volume of 10 μl and stored at -20 °C until use.

Biochemical characterization of GelK activity

Glucuronosyltransferase assays were performed in a final volume of 100 μl containing: 10 μl of the glycolipid acceptor, 72 mM Tris/HCl (pH 8.2), 8.5 mM $MgCl_2$, 1 mM ATP and 0.12 μCi UDP-[^{14}C]GlcA, the reaction being started by the addition of cell extracts (approx. 300 μg of total protein). After a 1 h incubation at 30 °C, enzyme reactions were stopped by an addition of 200 μl of chloroform/methanol (1:1, v/v). Glycolipids were isolated by solvent extraction and the oligosaccharide moiety synthesized *in vitro* was released by mild acid hydrolysis, as previously described [23]. The radioactivity of aliquots of this oligosaccharide sample was determined by liquid-scintillation counting in a Beckman LS 5000TD scintillation counter, and the rest of the samples were concentrated and spotted (5000–10000 c.p.m. in 10 μl) on a TLC plate (silica gel 60, 0.25 mm; Aldrich). Glc, GlcA and oligomaltose (dimer to tetramer) were used as standards. The TLC plate was developed twice in butan-1-ol/acetic acid/water (2:1:1, by vol.). Photographic films (Amersham) were exposed at -70 °C for 24–96 h. Standards were revealed by charring after treatment with 5% H_2SO_4 in methanol.

Glucuronosyltransferase assays were also carried out using EDTA-permeabilized cells of *E. coli* BL21(DE3)/pGelK or BL21(DE3)/pET29. Enzyme reactions were performed in a final volume of 300 μl containing 72 mM Tris/HCl (pH 8.2), 8.5 mM $MgCl_2$, 0.12 μCi of UDP-[^{14}C]GlcA and a 150 μl cell suspension obtained as described above. Enzyme reactions proceeded for 1 h at 30 °C, followed by the described experimental procedure.

To confirm that the GlcA was incorporated *in vitro*, the oligosaccharides released after mild acid hydrolysis were incubated overnight in 1.85 mM Na_2HPO_4/2.15 mM NaH_2PO_4 buffer pH 6.8, with 1 unit of bacterial β-glucuronidase (G2035; Sigma) at 37 °C, followed by de-lipidation and TLC analysis of the enzyme reaction products.

RESULTS

Cloning of *gelK* gene and sequence computational analysis

The *S. paucimobilis* A.T.C.C. 31461 *gelK* gene was cloned by PCR amplification based on the nucleotide sequence of the equivalent *spsK* gene from *Sphingomonas* S88 [7], as described in the Materials and methods section. Nucleotide sequence data confirmed a high level of identity (84%) of *gelK* and *spsK* genes.

Table 2 Similarity scores of GelK with other bacterial GTs

The accession numbers are from TREMBL, except for MurG and T4BGT, which are from Swiss Prot. NC, Non-classified. N-t, N-terminal; C-t, C-terminal.

Organism	Protein name	Amino acid	Family	Identity (%)	Known or attributed function	Accession number
Sphingomonas S88	SpsK	352	GT1	87	GlcAT	P74819 [8]
Rhizobium leguminosarum	PssD	152	GT1	33 (N-t)	GlcAT	P95624 + P95625
	PssE	158	GT1	34 (C-t)	(PssD + E)	[8]
S. pneumoniae	Cps14F	149	GT1	31 (N-t)	GalT	P72514 + 72515
	Cps14G	167	GT1	27 (C-t)	(Cps14F + G)	[32]
Lactococcus lactis	EpsE	156	GT1	29 (N-t)	GlcT	O06033 + 06034
subsp. *cremoris*	EpsF	160	GT1	26 (C-t)	(EpsE + F)	[33]
E. coli	MurG	354	GT28	24	GlcNAcT	P17443 [31]
Bacteriophage T4	BGT	351	NC	15	Glc	P04547 [37]

The sequences of the DNA regions downstream and upstream the *gelK* gene were highly similar to portions of *spsL* and *spsI* genes respectively. This result indicates that the organization of the *gel* and *sps* gene clusters is identical, at least in this particular region.

The GelK protein deduced from the *gelK* nucleotide sequence is composed of 348 amino acids with a calculated molecular mass of 38077 Da, and is 87% identical to *Sphingomonas* S88 SpsK. Two possible starting codons separated by 12 nt were first considered for GelK but the second codon appears to be the most probable one, as deduced from sequence comparison with SpsK and other related proteins (Figure 3) and by experimental evidences (results not shown).

GelK, like SpsK, belongs to family 1 of the GT classification [9,10]. A sequence motif of UDP-glucuronosyltransferases [30] is conserved in GelK (Figure 3). In Table 2 we show the similarity scores of GelK with other bacterial GTs belonging to GT family 1. The biochemical functions of these proteins have been established based on biochemical assays or genetic complementation studies [8,31–33]. All these GTs have in common the use of the same nucleotide portion (UDP) and of a lipid-linked PP_i as acceptor. Interestingly, several of them are produced in the form of two separate polypeptide chains of similar size (150–170 amino acids). Schematically, one polypeptide (i.e. Cps14F, EpsE, PssD) is homologous to the N-terminal half of GelK, whereas the second one (i.e. Cps14G, EpsF, PssE) is homologous to the C-terminal half. The *Cps14G* gene from *Streptococcus pneumoniae* encodes a polypeptide (167 amino acids) that displays β-1,4-galactosyltransferase activity. It was demonstrated that the upstream gene, *Cps14F*, coding for a polypeptide of 149 amino acids, drastically enhanced the galactosyltransferase activity when coexpressed with Cps14G [32]. Thus it is possible that the C-terminal half of GelK contains the transferase activity. GelK also displays a weak, but significant, overall sequence similarity with bacterial MurG protein. MurG is a β-1,4-N-acetylglucosaminyltransferase involved in the biosynthesis of peptidoglycan [31]. It catalyses the transfer of *N*-acetylglucosamine (GlcNAc) from UDP-GlcNAc to the C-4 hydroxy of a lipid-linked *N*-acetylmuramoyl pentapeptide.

The X-ray crystal structure of *E. coli* MurG has been recently solved [18]. MurG belongs to GT family 28, and contains two similar α/β domains separated by a deep cleft, a topology that is highly similar to T4βGT [19,20]. The structural similarity with T4βGT allowed Ha et al. [18] to dock the UDP-GlcNAc in the C-terminal domain. They also proposed that the acceptor-binding site is located at the N-terminal domain. The sequence alignment of GelK with MurG shows that regions corresponding to structural or sequence key features of the MurG family display significant sequence similarity with GelK (24% identity) (Figure 3). At the N-domain of GelK a G-loop probably involved in acceptor's PP_i binding is conserved, as well as MurG proline-96, glutamate-125 and lysine-135. At the $-CO_2H$ domain, the residues glycine-190 and glutamine-289 of GelK are also invariant, whereas arginine-256 in MurG is replaced by a conservative substitution, histidine. It is interesting to note that all these residues are predicted to interact with the nucleotide sugar, and that the position equivalent to arginine-256 is a histidine as in family 1 GTs. The residues glycine-268, proline-276 and histidine-284 are also invariant, but no precise function was attributed to them.

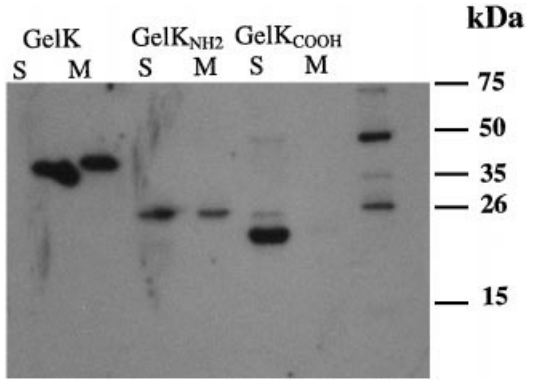


Figure 4 Subcellular localization of GelK, GelK$_{NH2}$ and GelK$_{COOH}$ in *E. coli*

Western blot against S-tag epitope was used to determine the subcellular localization of GelK, GelK$_{NH2}$ and GelK$_{COOH}$ expressed in *E. coli*. Soluble (S) and membrane (M) fractions of crude extracts from *E. coli* BL21(DE3) cells harbouring pGelK, pGelK$_{NH2}$ or pGelK$_{COOH}$ and induced with IPTG were prepared as described in the Materials and methods section.

Expression and subcellular localization of GelK protein and its -NH_2- or -CO_2H-halves in *E. coli*

Expression in *E. coli* BL21(DE3) of the recombinant proteins GelK, GelK$_{NH2}$ and GelK$_{COOH}$ was monitored by SDS/PAGE and subsequent staining with Coomassie Blue. Upon induction with IPTG, bands of approx. 41.9 kDa, 26.8 kDa and 22.6 kDa were detected in cells harbouring pGelK, pGelK$_{NH2}$ and pGelK$_{COOH}$ respectively, but were absent from cells transformed with the cloning vector pET29 (results not shown). These bands have the predicted size of the plasmid-encoded S-tag-fusion proteins. By S-tag Western-blot analysis it was confirmed that

Table 3 Incorporation of [^{14}C]GlcA mediated by GelK

Crude extracts were prepared from *E. coli* BL21(DE3) cells expressing or not the GelK protein and [^{14}C]GlcA. Incorporation of radioactivity was assessed either in the presence or absence of the glycolipid acceptor prepared from *S. paucimobilis* A.T.C.C. 31461, as described in the Materials and methods section.

E. coli strain	Glycolipid acceptor	[^{14}C]GlcA incorporation [pmol/min per mg of protein]
BL21(DE3)/pGelK	+	0.225 ± 0.0150
BL21(DE3)/pET29a	+	0.032 ± 0.0043
BL21(DE3)/pGelK	−	0.083 ± 0.0166
BL21(DE3)/pET29a	−	0.038 ± 0.0047

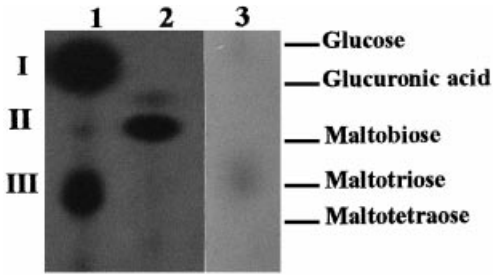


Figure 5 TLC analysis of the lipid-linked oligosaccharides synthesized *in vitro* by the action of GelK

Lane 1, oligosaccharides released when *S. paucimobilis* permeabilized cells were incubated with UDP-[^{14}C]Glc. Lane 2, oligosaccharides released after GelK-mediated transfer of UDP-[^{14}C]GlcA to *S. paucimobilis* glycolipid acceptor. Lane 3, oligosaccharides released after incubation of *S. paucimobilis* permeabilized cells with UDP-[^{14}C]GlcA.

these bands indeed correspond to the expected recombinant proteins (Figure 4). The S-tag epitope at the N-terminal end of complete and truncated GelK and of its parts facilitated the subcellular localization of the recombinant proteins. Cell-free extracts of BL21(DE3)pLysS containing either pGelK, pGelK$_{NH2}$ and pGelK$_{COOH}$ were submitted to differential centrifugation. Interestingly, more than 50 % of GelK and GelK$_{NH2}$ was found associated with membranes, whereas GelK$_{COOH}$ was found exclusively in the 50000 ***g*** supernatant, indicating that the NH_2-half of GelK is involved in protein association to the membrane (Figure 4). Common available sequence analysis softwares failed to predict transmembrane helices in GelK. However, the dense alignment surface method [34] showed the presence of a hydrophobic peptide (A^{108}GAVYKTALFAKL^{120}) whose score is lower than the strict cut-off for the transmembrane helices. Therefore the association of GelK with the membrane is not related to the presence of a clear transmembrane domain. It is worth mentioning that MurG is also associated with membranes, although no transmembrane helix could be identified. Concerning MurG, Ha et al. [18] proposed that a hydrophobic patch, surrounded by basic residues, which is present in the N-terminal domain could be involved in membrane association. Interestingly, some of these residues are conserved in GelK (Figure 3).

Biochemical characterization of GelK and derivatives

The putative acceptor for GelK is Glcα-PPL. Since this product is not commercially available, and the eventual requirement of the PPL moiety is not known, we used as acceptor a glycolipid fraction from *S. paucimobilis* enriched in the glycolipid acceptor. *E. coli* cell extracts, either containing or not GelK were incubated in the presence or absence of *S. paucimobilis* glycolipid with UDP-[^{14}C]GlcA. The glycolipids were extracted and the radioactivity associated with this fraction was determined (Table 3). In the absence of GelK a barely detectable radioactivity value was obtained independently of glycolipid addition, indicating the absence of glucuronosyltransferase in *E. coli*. A low, but significant, increase of radioactivity was detected in the presence of GelK, indicating that *E. coli* contains an endogenous acceptor for GelK. The addition of the *S. paucimobilis* glycolipid fraction to the enzyme assay mixture results in a 5-fold increase of radioactivity in the glycolipid fraction. Replacement of UDP-[^{14}C]GlcA by either UDP-[^{14}C]Glc or UDP-[^{14}C]galactose resulted in loss of sugar incorporation into the glycolipid fraction (results not shown). These results indicate that GelK is a glucuronosyltransferase, which uses as acceptor a glycolipid present in both *S. paucimobilis* and *E. coli*. Since the extract containing the acceptor is a crude extract, and the acceptor is expected to be in low amounts (see below), it was not possible to perform the kinetic characterization of GelK. To overcome this problem, we have investigated the possibility that commercial synthetic glucosylated compounds may be recognized by GelK as acceptors. We performed a series of enzyme assays using Glcβ-O-$(CH_2)_3$-NH-CO-$(CH_2)_4$-NH-biotin and Glcα-O-$(CH_2)_3$-NH-CO-$(CH_2)_4$-NH-biotin as acceptors. After the reaction, unbound UDP-[^{14}C]GlcA was removed and no radioactivity was found associated with the synthetic derivatives. These results suggest that features of the PPL are needed for GelK activity. In the presence of 32.5 mM EDTA, GelK retained 50 % of its activity, thus strongly suggesting that divalent cations are not essential for activity.

In the standard reaction containing GelK glycolipid acceptor and UDP-[^{14}C]GlcA, the incorporation of radioactivity into the glycolipid extract was low (5000 c.p.m.), which may be due to either a low level of GelK activity or low availability of the glycolipid acceptor. Dilutions of GelK (up to 100 times) did not result in a decrease of [^{14}C]GlcA incorporation into the acceptor, indicating that the glycolipid acceptor is the limiting factor. Further studies using synthetic acceptors are needed to better characterize GelK biochemically.

By analogy to the bacterial homologous GTs that are produced as two distinct polypeptide chains (Table 2), we prepared truncated forms of GelK, namely GelK$_{NH2}$ and GelK$_{COOH,}$ and investigated their possible enzyme activity, either separately or mixed in a 1:1 (mol/mol) ratio. None of the assays gave a significant level of [^{14}C]GlcA incorporation into the glycolipid extract as well as the synthetic substrates used as acceptors unlike the findings reported using Cps14G from *Streptococcus pneumoniae* [32].

Since it has previously been reported [8] that the acceptor for the homologous enzyme SpsK is present in *E. coli* BL21(DE3) cells, we have tested this for GelK. Permeabilized BL21 (DE3) cells expressing GelK or not were incubated with UDP-[^{14}C]GlcA. Cells expressing GelK incorporated 10 times more radioactivity in the glycolipid extract than cells containing only the expression vector (results not shown).

Characterization of the GelK product

The low amount of GlcA-glycolipid obtained by GelK-mediated catalysis precluded spectroscopic analysis and only radiolabelled glycolipids were characterized. Mild acid hydrolysis led to the release of the radioactive oligosaccharide from the lipid anchor,

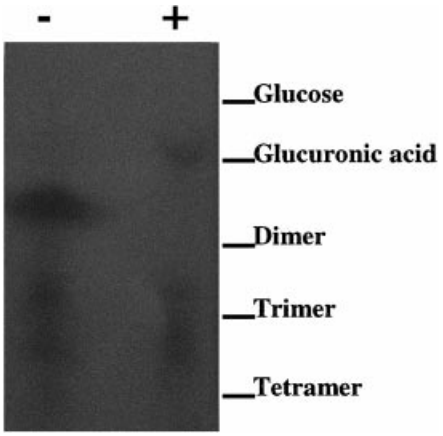


Figure 6 Effect of β-glucuronidase treatment on the oligosaccharide synthesized *in vitro* by GelK

TLC analysis of the oligosaccharide released after GelK-mediated transfer of UDP-[^{14}C]GlcA to glycolipid fraction from *S. paucimobilis* A.T.C.C. 31461 acceptor with (+) or without (−) previous β-glucuronidase treatment.

indicating that sugars are linked through a PP_i linkage. The released oligosaccharide was analysed by TLC (Figure 5). Since GlcAβ-1,4-Glc is not commercially available for use as a TLC standard, we isolated radiolabelled lipid-linked oligosaccharide from *S. paucimobilis* A.T.C.C. 31461 to compare their migration profile with that of the reaction product of GelK. Permeabilized *S. paucimobilis* A.T.C.C. 31461 cells were incubated in the presence of UDP-[^{14}C]Glc or UDP-[^{14}C]GlcA, and the oligosaccharides were released from the glycolipid fraction by mild acid hydrolysis and analysed by TLC (Figure 5). Three radioactive spots (I, II and III) were detected by fluorography after incubation with UDP-[^{14}C]Glc. Spot I co-migrates with Glc and probably represents the first Glc linked to the lipid acceptor in gellan biosynthesis. Spot III probably represents Glc-GlcA-Glc, since no GlcA was added to the incubation media, this may arise either from the cellular pool of UDP-GlcA or by addition of Glc to the endogenous GlcA-Glc-lipid already present in the enzymatic preparation. Since spot II migrates between spot I and III, the major glucosylated products, spot II may represent GlcA-Glc (Figure 5). When permeabilized *S. paucimobilis* cells were incubated with UDP-[^{14}C]GlcA, spot III was mainly obtained (Figure 5). Therefore we speculate that an important UDP-Glc pool is present in the permeabilized cell preparation, and even an extensive washing of the permeabilized cells did not apparently remove this internal pool of UDP-Glc.

The [^{14}C]GlcA-oligosaccharide, obtained by incubation of GelK with either *S. paucimobilis* glycolipid or the *E. coli* endogenous acceptor, co-migrated with spot II, which is probably the expected GlcA-Glc product. Also, the endogenous acceptor of *E. coli* can functionally substitute the one obtained from *S. paucimobilis*. In order to determine if the GlcA was in the terminal position and if the linkage is β(1–4), the [^{14}C]GlcA-oligosaccharide was digested with β-glucuronidase and analysed by TLC. A radioactive spot, co-migrating with GlcA, was released by incubation with β-glucuronidase (Figure 6).

Taken together, these results indicate that the GlcA residue transferred by GelK is β(1–4) linked at the non-reducing end of the glycolipid. Given the size of the oligosaccharide formed upon *in vitro* GelK action and the size of the oligosaccharides formed by permeabilized cells of *S. paucimobilis*, we propose that GelK is a β-1,4-glucuronosyltransferase that probably uses Glc-PPL as acceptor.

DISCUSSION

GTs are enzymes exhibiting both exquisite stereospecificity towards the substrate and precise stereochemistry. The understanding of the basis of such specificity requires the biochemical and structural characterization of GTs with similar or different specificity. Here we report the cloning of the *S. paucimobilis* A.T.C.C. 31461 *gelK* gene and its heterologous expression in *E. coli*. Cell fractionation experiments showed that GelK is present in two forms: one associated with membrane and the other soluble, the membrane target region being located at the N-terminal domain. Based on *in vitro* assays and after characterization of the reaction product, we concluded that GelK is a β-1,4-glucuronosyltransferase capable of using either *S. paucimobilis* or *E. coli* glycolipidic fraction as acceptor. We have also found that features of the PPL moiety are needed for activity.

The substrate requirements for GTs involved in the repeat-unit biosynthesis remained unknown for many years, mainly because reactions were carried out using permeabilized cells relying on endogenous acceptors. Recently developed *in vitro* assays allow the systematic study of structural requirements for acceptors [21–23]. These assays rely on synthetic acceptors or glycolipid fractions containing the acceptors. For several of these GTs, such as AceA from *A. xylinum* and ExoM from *S. meliloti*, it is known that features of the PPL moiety are required for activity [22,23], whereas MurG from *E. coli* requires the presence of the PP_i moiety only [19]. MurG structure shows the presence of two 'G-loops' in the N-terminal domain that are thought to be involved in binding the PP_i moiety of the acceptor [18], bringing together structural and biochemical results about acceptor specificity. Requirements for PPL may be due to the need of assuring a rapid turnover of the lipid for the repeat-unit assembly.

The reduced effect of EDTA on GelK activity strongly suggests that this protein does not belong to the structural group of GTs harbouring the UDP-binding domain, despite the presence of a hypothetical 'DXD' motif ($Asp^{249}Ala^{250}Asp^{251}$). It is worth stressing that a functional DXD motif requires the presence of an α/β/α sandwich (see Gastinel et al. [16] and Ünligil and Rini [17]). Moreover, sequence analysis and secondary structure predictions strongly suggest that GelK belongs to the same structural family as MurG. However, there are two important features that are different in GelK and MurG. First, the equivalent of MurG glutamate-272, probably involved in hydrogen bonding the O2′ and O3′ hydroxy groups of the ribose ring, is not present in GelK. Secondly, the G-loop3 present in MurG, and possibly involved in the binding of the α-PO_4^{3-}, is absent in GelK. The structurally related T4βGT also lacks this G-loop, but arginine-191 located at the beginning of an α-helix takes over its function [19,20]. Based on sequence comparison and site-directed mutagenesis, the same role was proposed for lysine-211 in the mannosyltransferase AceA [35] that belongs to family 4. It is worth mentioning that the conserved residue arginine-197 of GelK is located in the same region where arginine-191 of T4βGT and the G-loop3 of MurG are present. Thus GelK arginine-197 is a candidate for interaction with the α-PO_4^{3-} moiety. In the experiments that we have carried out with GelK isolated N- and C-terminal halves, we did not detect glucuronosyltransferase activity. However, given the sensitivity of our enzyme assay, it may become impossible to detect reduced activities. Assuming that both the overall structure and domain function of GelK are identical to MurG, we do not anticipate that isolated domains may display activity. Structural studies

are, however, necessary to confirm these hypotheses and to gain further insight into the mechanism of action of GelK.

Work at Centro de Engenharia Biológica e Química, Instituto Superior Técnico was supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal (grants: PRAXIS/P/BIO/12020/1998 and POCTI/35733/1999) and a PhD scholarship to P. V. Work at Centre de Recherches sur les Macromolécules Végétales was supported by the program 'Physique et chimie du vivant' (CNRS). This research was carried out within the context of the French-Portuguese Joint Research Program (Action 422,K4/2000; Ambassade de France au Portugal and Instituto de Cooperação Científica e Tecnológica Internacional, Instituto de Cooperação Científica e Tecnológica Internacional, Portugal). We thank Annemarie Lellouch for helpful discussions and Valerie Chazalet for her expert technical assistance.

REFERENCES

1 O'Neil, M. A., Selvendran, R. R. and Morris, V. J. (1983) Structure of the acidic extracellular gelling polysaccharide produced by *Pseudomonas elodea*. Carbohydr. Res. **124**, 123–133
2 Kuo, M. S., Mort, A. J. and Dell, A. (1986) Identification and location of L-glycerate, an unusual substituent in gellan gum. Carbohydr. Res. **56**, 173–187
3 Jay, A. J., Colquhoun, I. J., Ridout, M. J., Brownsey, G. J., Morris, V. J., Fialho, A. M., Leitão, J. H. and Sá-Correia, I. (1988) Analysis of structure and function of gellans with different substitution patterns. Carbohydr. Polym. **35**, 179–188
4 Fialho, A. M., Martins, L. O., Donval, M. L., Leitao, J. H., Ridout, M. J., Jay, A. J., Morris, V. J. and Sá-Correia, I. (1999) Structures and properties of gellan polymers produced by *Sphingomonas paucimobilis* ATCC 31461 from lactose compared with those produced from glucose and from cheese whey. Appl. Environ. Microbiol. **65**, 2485–2491
5 Martins, L. O. and Sá-Correia, I. (1991) Gellan gum biosynthetic enzymes in producing and nonproducing variants of *Pseudomonas elodea*. Biotechnol. Appl. Biochem. **14**, 357–364
6 Videira, P. A., Cortes, L. L., Fialho, A. M. and Sá-Correia, I. (2000) Identification of the *pgmG* gene, encoding a bifuntional protein with phosphoglucomutase and phosphomannomutase activities, in the gellan gum-producing strain *Sphingomonas paucimobilis* ATCC 31461. Appl. Environ. Microbiol. **66**, 2252–2258
7 Yamazaki, M., Thorne, L., Mikolajczak, M., Armentrout, R. W. and Pollock, T. J. (1996) Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. J. Bacteriol. **178**, 2676–2687
8 Pollock, T. J., van Workum, W. A., Thorne, L., Mikolajczak, M. J., Yamazaki, M., Kijne, J. W. and Armentrout, R. W. (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
9 Campbell, J. A., Davies, G. J., Bulone, V. and Henrissat, B. (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. **326**, 929–939
10 Campbell, J. A., Davies, G. J., Bulone, V. and Henrissat, B. (1998) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities Biochem. J. **329**, 719
11 Charnock, S. J. and Davies, G. J. (1999) Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. Biochemistry **38**, 6380–6385
12 Gastinel, L. N., Cambillau, C. and Bourne, Y. (1999) Crystal structures of the bovine 4-β-galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose. EMBO J. **18**, 3546–3557
13 Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A. and Negishi, M. (2000) Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of human glucuronyltransferase I. J. Biol. Chem. **275**, 34580–34585
14 Ünligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H. and Rini, J. M. (2000) X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. EMBO J. **19**, 5269–5280
15 Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G. and Strynadka, N. C. (2001) Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. Nat. Struct. Biol. **8**, 166–175
16 Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H. and Joziasse, D. H. (2001) Bovine α-1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases. EMBO J. **20**, 638–649
17 Ünligil, U. M. and Rini, J. M. (2000) Glycosyltransferase structure and mechanism. Curr. Opin. Struct. Biol. **10**, 510–517
18 Ha, S., Walker, D., Shi, Y. and Walker, S. (2000) The 1.9 Å crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. Protein Sci. **9**, 1045–1052
19 Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W. and Freemont, P. S. (1999) T4 phage beta-glucosyltransferase: substrate binding and proposed catalytic mechanism. J. Mol. Biol. **292**, 717–730
20 Vrielink, A., Ruger, W., Driessen, H. P. and Freemont, P. S. (1994) Crystal structure of the DNA modifying enzyme β-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. EMBO J. **13**, 3413–3422
21 Ha, S., Chang, E., Lo, M. C., Men, H., Park, P., Ge, M. and Walker, S. (1999) The kinetic characterization of *Escherichia coli* MurG using synthetic substrate analogues. J. Am. Chem. Soc. **121**, 8415–8426
22 Geremia, R. A., Roux, M., Ferreiro, D. U., Dauphin-Dubois, R., Lellouch, A. C. and Ielpi, L. (1999) Expression and biochemical characterisation of recombinant AceA, a bacterial alpha-mannosyltransferase. Mol. Gen. Genet. **261**, 933–940
23 Lellouch, A. C. and Geremia, R. A. (1999) Expression and study of recombinant ExoM, a β1–4 glucosyltransferase involved in succinoglycan biosynthesis in *Sinorhizobium meliloti*. J. Bacteriol. **181**, 1141–1148
24 Birnboim, H. C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**, 1513–1523
25 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning, a laboratory manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402
27 Huang, X. and and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. Adv. Appl. Math. **12**, 337–357
28 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (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
29 Tolmasky, M. E., Staneloni, R. J. and Leloir, L. F. (1982) Lipid-bound saccharides in *Rhizobium melliloti*. J. Biol. Chem. **257**, 6751–6757
30 Kapitonov, D. and Yu, R. K. (1999) Conserved domains of glycosyltransferases Glycobiology **9**, 961–978
31 Mengin-Lecreulx, D., Texier, L., Rousseau, M. and van Heijenoort, J. (1991) The *murG* gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine:N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. J. Bacteriol. **173**, 4625–4636
32 Kolkman, M. A., van der Zeijst, B. A. and Nuijten, P. J. (1997) Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype. J. Biol. Chem. **272**, 19502–19508
33 van Kranenburg, R., van Swam, I. I., Marugg, J. D., Kleerebezem, M. and de Vos, W. M. (1999) Functional analysis of glycosyltransferase genes from *Lactococcus lactis* and other gram-positive cocci: complementation, expression, and diversity. J. Bacteriol. **181**, 338–340
34 Cserzo, M., Wallin, E., Simon, I., von Heijne, G. and Elofsson, A. (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng. **10**, 673–676
35 Abdian, P. L., Lellouch, A. C., Gautier, C., Ielpi, L. and Geremia, R. A. (2000) Identification of essential amino acids in the bacterial alpha-mannosyltransferase AceA. J. Biol. Chem. **275**, 40568–40575
36 Studier, F. W. and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high level expression of clones genes. J. Mol. Biol. **189**, 113–130
37 Tomaschewski, J., Gram, H., Crabb, J. W. and Rueger, W. (1985) T4-induced alpha- and beta-glucosyltransferase: cloning of the genes and a comparison of their products based on sequencing data. Nucleic Acids Res. **13**, 7551–7568

Received 14 March 2001/8 May 2001; accepted 23 May 2001