Proteins Encoded by *Sphingomonas elodea* ATCC 31461 *rmlA* and *ugpG* Genes, Involved in Gellan Gum Biosynthesis, Exhibit both dTDP- and UDP-Glucose Pyrophosphorylase Activities

Elisabete Silva, Ana Rita Marques, Arsénio Mendes Fialho,* Ana Teresa Granja, and Isabel Sá-Correia

Biological Sciences Research Group, Centre for Biological and Chemical Engineering,
Instituto Superior Técnico, Lisbon, Portugal

Received 27 October 2004/Accepted 12 March 2005

The commercial gelling agent gellan is a heteropolysaccharide produced by *Sphingomonas elodea* ATCC 31461. In this work, we carried out the biochemical characterization of the enzyme encoded by the first gene (*rmlA*) of the *rml* 4-gene cluster present in the 18-gene cluster required for gellan biosynthesis (*gel* cluster). Based on sequence homology, the putative *rml* operon is presumably involved in the biosynthesis of dTDP-rhamnose, the sugar necessary for the incorporation of rhamnose in the gellan repeating unit. Heterologous RmlA was purified as a fused His₆-RmlA protein from extracts prepared from *Escherichia coli* IPTG (isopropyl- β -D-thiogalactopyranoside)-induced cells, and the protein was proven to exhibit dTDP-glucose pyrophosphorylase (K_m of 12.0 μ M for dTDP-glucose) and UDP-glucose pyrophosphorylase (K_m of 229.0 μ M for UDP-glucose) activities in vitro. The N-terminal region of RmlA exhibits the motif G-X-G-T-R-X₂-P-X-T, which is highly conserved among bacterial XDP-sugar pyrophosphorylases. The motif E-E-K-P, with the conserved lysine residue (K₁₆₃) predicted to be essential for glucose-1-phosphate binding, was observed. The *S. elodea* ATCC 31461 UgpG protein, encoded by the *ugpG* gene which maps outside the *gel* cluster, was previously identified as the UDP-glucose pyrophosphorylase involved in the formation of UDP-glucose, also required for gellan synthesis. In this study, we demonstrate that UgpG also exhibits dTDP-glucose pyrophosphorylase activity in vitro and compare the kinetic parameters of the two proteins for both substrates. DNA sequencing of *ugpG* gene-adjacent regions and sequence similarity studies suggest that this gene maps with others involved in the formation of sugar nucleotides presumably required for the biosynthesis of another cell polysaccharide(s).

Sphingomonas elodea (formerly *Pseudomonas elodea* and also referred to as *Sphingomonas paucimobilis*) ATCC 31461 synthesizes the exopolysaccharide (EPS) gellan, a gelling agent with applications in the food, pharmaceutical, and other industries (13, 41). This EPS is composed of a repeating linear tetrasaccharide unit consisting of D-glucose (Glc), D-glucuronic acid (GlcA), and L-rhamnose (Rha) in a 2:1:1 ratio, respectively, with glycerate and acetate substituents (22). The significant changes in rheology observed upon the deacylation of gellan are essentially due to the glycerate substituents (17).

The gellan repeat unit is formed by sequential transfer of the sugar nucleotides to a membrane-anchored lipid carrier by committed glycosyltransferases, followed by gellan polymerization and export (34). The pathway leading to the formation of the sugar nucleotides UDP-Glc, UDP-GlcA, and dTDP-L-Rha, donors of the monomers for assembling the gellan tetrasaccharidic repeat unit, was elucidated (27, 34). Eighteen genes, organized in the *gel* cluster (15, 34) and coding for enzymes presumably involved in the synthesis of dTDP-L-Rha, glucosyltransferases, and proteins required for gellan polymerization and export, were identified. The organization and the nucleo-

tide sequence of the *gel* cluster are highly similar to those described for the gene cluster required for the synthesis of the sphingane S88 in *Sphingomonas* strain S88 (*sps* cluster) (15, 45).

The enzymes required for the synthesis of the sugar nucleotides UDP-Glc and UDP-GlcA from glucose-1-phosphate (G1P) are encoded by genes *pgmG*, *ugpG*, and *ugdG*, located outside the *gel* cluster (26, 34, 43).

The other activated sugar essential for the incorporation of L-rhamnose in the gellan repeat unit is dTDP-L-Rha. Rhamnose is a 6-deoxyhexose sugar widely distributed in gram-negative bacteria as part of the lipopolysaccharides (LPS), capsular polysaccharides (CPS), and EPSs. The nucleotide sugar dTDP-L-Rha is formed from G1P in a four-step reaction (21). This reaction involves the enzyme activities of glucose-1-phosphate thymidyltransferase, dTDP-glucose-4,6-dehydratase, dTDP-4-keto-L-rhamnose-3,5-epimerase, and dTDP-L-rhamnose synthase, encoded by genes commonly designated *rmlA*, *rmlB*, *rmlC*, and *rmlD*, respectively (6). These genes are generally grouped together within the LPS, CPS, or EPS gene cluster (24, 34). The *rml* four-gene cluster has been identified for different gram-negative and gram-positive bacterial species; these clusters exhibit a remarkable sequence similarity, although the gene order of the operon-like structures may vary from species to species (11, 21, 24, 30). The two *Sphingomonas* strains *S. elodea* ATCC 31461 and *Sphingomonas* strain S88 exhibit a similar organization (*ACBD*→) and a high range of

* Corresponding author. Mailing address: Biological Sciences Research Group, Centre for Biological and Chemical Engineering, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisbon, Portugal. Phone: 351-21-8417684. Fax: 351-21-8419199. E-mail: afialho@ist.utl.pt.

TABLE 1. Bacterial strains and plasmids used in this work

Strain or plasmid	Genotype and/or phenotype ^a	Source or reference
<i>E. coli</i> strains		
XL1-Blue	<i>recA1 lac</i> [F' <i>proAB lacI^qΔM15</i>] Tn10 (Tc ^r)	Stratagene
BL21(DE3)	F' <i>ompT hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm</i> (DE3)	Novagen
Plasmids		
Hypercos 1	8.8 kb, Ap ^r , double <i>cos</i> , T3, T7	A. Pühler
pWH844	pQE9 derivative, 5.0 kb, <i>lacI^q</i> Ap ^r , T5, His tag	38
p96	Cosmid from <i>S. elodea</i> ATCC 31461 genomic library in Hypercos 1; 35-kb insert	44
pRmlA	<i>rmlA</i> in pWH844, Ap ^r	This work
pUgpG	<i>ugpG</i> in pET29b, Km ^r	26

^a ATCC, American Type Culture Collection; Tc^r, tetracycline resistance; Ap^r, ampicillin resistance; Km^r, kanamycin resistance.

sequence conservation (87 to 96% amino acid identity) (15, 34).

The present work focuses on the first gene of the *S. elodea* ATCC 31461 *rml* four-gene cluster, the *rmlA* gene. The heterologous overexpression of *rmlA* in *Escherichia coli*, its purification as a fusion protein, and the biochemical characterization of the encoded enzyme were carried out. The protein RmlA was demonstrated to be a glucose-1-phosphate thymidyltransferase (EC 2.7.7.24) involved in the reversible conversion of G1P and dTTP into dTDP-glucose (dTDP-Glc) and diphosphate. Besides this dTDP-glucose pyrophosphorylase (TGP) activity, RmlA also exhibits a UDP-glucose pyrophosphorylase (UGP) activity (EC 2.7.7.9). In the present study, we also demonstrated that the *S. elodea* UgpG protein, identified before as the enzyme required for the synthesis of UDP-glucose in the same bacterial strain (26), also recognizes both deoxythymidine and deoxyuridine nucleotides as substrates, and the kinetic parameters of these two proteins were compared. Since the *ugpG* gene maps outside the *gel* cluster (26), the adjacent chromosomal regions were sequenced to gain more insight into its physiological function.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. Strains and plasmids used in this study are described in Table 1. *E. coli* strains were grown in Lennox broth (Sigma Chemical Co., St. Louis, Mo.) or in Superbroth medium (36) at 37°C with orbital agitation. When required, the culture medium was supplemented with ampicillin (150 mg/liter) or kanamycin (100 mg/liter). The oligonucleotides were purchased from MWG Biotech (Ebersberg, Germany). Restriction enzymes were from GibcoBRL-Life Technologies (Portugal), and DNA polymerases were from Boehringer GmbH (Mannheim, Germany). UDP-Glc, dTDP-Glc, NADP⁺, glucose-6-phosphate dehydrogenase, phosphoglucosmutase (PGM), glucose-6-P-dehydrogenase (ZWF), and inorganic pyrophosphate (PPi) were obtained from Sigma.

DNA manipulation. All standard procedures, including DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, DNA ligation, and transformation of *E. coli* strains, were performed by conventional methods (36). DNA was recovered from gel slices by using a Concert gel kit (GibcoBRL-Life Technologies). PCR amplifications were performed using *Pwo* polymerase obtained from Roche Diagnostics (Mannheim, Germany) and 200 ng of template DNA.

Sequence analysis. Previous studies have shown the presence of the gellan gene cluster in the p96 cosmid of the genomic library of *S. elodea* ATCC 31461 (44). To confirm the presence of the *rmlA* gene in the p96 insert, an internal region of this gene was amplified by PCR with the synthetic oligonucleotides AU (5'-GCAGCTGCTTCCCGTCTATGA-3') and AL (5'-CGAGCCGGGTGATG TGGAGGTC-3'), designed on the basis of the nucleotide sequence of the homologous gene of *Sphingomonas* strain S88, given that the organization of the sphingon S88 gene cluster is identical to that of the gellan biosynthetic cluster

(45). The PCR product was recovered from a 1% (wt/vol) low-melting-point agarose gel (FMC Bioproducts, Rockland, ME), purified with a Concert gel kit (GibcoBRL), and sequenced with the oligonucleotide primers AU and AL to confirm the identity. The complete sequence of *rmlA* was obtained by directly sequencing the p96 insert with primers designed on the basis of the sequenced internal regions of this gene.

DNA sequencing of the *ugpG* gene-adjacent regions was carried out by using cosmid pC22 as the template (26) and specific primers based on the *ugpG* sequence. Sequence similarity searches were performed using BLAST 2.0 with default settings from the National Center for Biotechnology Information (NCBI), Bethesda, Md. A phylogenetic analysis of the TGP (RmlA) and UGP proteins compared in this study was performed. The manually edited 20-sequence alignment performed with the CLUSTAL X program (16) served to calculate the distance matrix by the TREE-PUZZLE method (39). The phylogenetic tree was constructed by the neighbor-joining method (35) and supported by 1,000 bootstrap steps (12).

Construction of plasmids. To overexpress the *rmlA* gene, the complete sequence of this gene was amplified by PCR with the oligonucleotides TGP1 (5'-AAAGGATCCATGAAGGCATCAT-3') and TGP2 (5'-AAAAAGCTTT CATGCGGCAGCCA-3') and with cosmid p96 as the template DNA. The PCR product (897 bp) was digested with BamHI and HindIII (underlined for the primers used) and cloned into compatible sites of pWH844, generating pRmlA. This recombinant plasmid carries the *rmlA* gene preceded by a sequence coding for six histidine residues present in the pWH844 vector. The insert cloned in pRmlA was then sequenced to confirm the fidelity of DNA amplification.

Protein overexpression and purification. Expression and purification of histidine-tagged proteins from plasmids pRmlA (this study) and pUgpG (26) were carried out as previously described (26).

Transformants of *E. coli* BL21(DE3) harboring plasmid pRmlA or pUgpG or the respective cloning vector, pWH844 or pET29b (controls), were grown in 200 ml of Superbroth and supplemented with the appropriate antibiotic at 37°C until the optical density at 640 nm of the culture reached 0.5 ± 0.1 . The transcription of the *rmlA* and *ugpG* genes was then induced by the addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), followed by an additional 3-h period of cultivation. Cells were harvested by centrifugation, washed with 0.9% (wt/vol) NaCl, and resuspended in 3 ml of cold start buffer (1 mM phosphate buffer [pH 7.4], 50 mM NaCl, and 10 mM imidazole). Cells were sonicated on ice (VibraCell; Sonics Material, Danbury, Conn.) and centrifuged at $18,000 \times g$ for 1 h at 4°C to remove cell debris and insoluble proteins. The overexpressed RmlA and UgpG proteins were purified under native conditions by following the one-step procedure for His-tagged proteins as described by the column manufacturer (Amersham Pharmacia Biotech). After the equilibration of the column with 10 ml of the start buffer, the induced cell crude extract was loaded on the column. The columns were then washed in several steps with the same start buffer containing increasing imidazole concentrations, and the His₆ recombinant protein, RmlA or UgpG, was eluted with a buffer containing 200 mM or 300 mM imidazole, respectively.

The proteins present in the cell crude extracts or in the purified fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (12.5% acrylamide-bisacrylamide gels). Protein bands were revealed by staining with Coomassie brilliant blue, and the molecular masses of RmlA and UgpG were estimated with low-molecular-weight standards (Amersham Pharmacia Biotech). Protein concentrations were determined by the method of Bradford (7), with bovine serum albumin fraction V (Sigma) as the standard.

Enzyme assays. dTDP- and UDP-glucose pyrophosphorylase activities were measured by coupled reaction with the enzymes PGM and ZWF, leading to the formation of NADPH. Enzyme reactions were performed at 30°C in a final volume of 1 ml containing the following activity assay buffer: 100 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 2 mM dTDP-D-glucose or UDP-D-glucose, 1 U PGM, 1 U ZWF, and 1 mM NADP⁺. The reaction was initiated by the addition of the cell crude extracts (1 mg) or the purified proteins (6 μg) and inorganic PPi (2 mM). The rate of formation of NADPH was followed at 340 nm with a double-beam spectrophotometer (model V-200; Hitachi Ltd., Tokyo, Japan). Control assays were performed with enzyme reaction mixtures lacking the extracts or the substrate. Under the assay conditions used, 1 U of enzyme activity was defined as the amount of enzyme needed to reduce 1 μmol of NADP per min. TGP and UGP activities are the means of the values of at least three enzyme assays. The kinetic constants K_m and V_{max} for dTDP-Glc or UDP-Glc, corresponding to the TGP or UGP activities of purified RmlA and UgpG (ranges of 20 to 2,000 μM for RmlA and 3.3 to 25 μM for UgpG), were determined by using the coupled enzyme reaction mixtures mentioned above, in which PGM- and ZWF-catalyzed reaction mixtures were largely saturated. The adequacy of the coupling-enzyme system was verified by the fact that the apparent activity of an enzyme assay

TABLE 2. Comparison of amino acid sequences of RmlA, RmlB, RmlC, and RmlD from *S. elodea* ATCC 31461 with other dTDP-L-rhamnose biosynthetic enzymes

Protein ^a	Bacterial species and/or strain	Identity/similarity (%)	Length (no. of amino acids)	Accession no.
RmlA	<i>S. elodea</i> ATCC 31461	100	292	AAO89234
	<i>Sphingomonas</i> strain S88	96/98	292	AAC44072
	<i>P. aeruginosa</i> PAO1	66/80	293	NP_253850
	<i>X. campestris</i> pv. <i>campestris</i> ATCC 33913	60/74	296	NP_636014
	<i>Burkholderia pseudomallei</i>	64/77	297	AAD05455
	<i>N. aromaticivorans</i>	61/76	301	ZP_00095575
	<i>E. coli</i> K-12	63/78	293	NP_416543
	<i>S. enterica</i> subsp. <i>enterica</i> serovar Typhimurium	62/79	292	NP_456644
	<i>L. lactis</i> subsp. <i>cremoris</i>	61/75	289	AAO15542
	<i>S. pneumoniae</i> serotype 18C	62/77	289	AAK20721
RmlB	<i>S. elodea</i> ATCC 31461	100	353	AAP57700
	<i>Sphingomonas</i> strain S88	94/94	353	AAC44074
	<i>P. aeruginosa</i> PAO1	54/66	352	NP_253848
	<i>X. campestris</i> pv. <i>campestris</i> ATCC 33913	54/68	351	NP_636013
	<i>B. pseudomallei</i>	54/67	353	AAD05454
	<i>N. aromaticivorans</i>	49/62	296	ZP_00095572
	<i>E. coli</i> K-12	56/68	361	NP_416545
	<i>S. enterica</i> subsp. <i>enterica</i> serovar Typhimurium	58/70	361	NP_456646
	<i>L. lactis</i> subsp. <i>cremoris</i>	40/56	350	AAO15544
	<i>S. pneumoniae</i> serotype 18C	42/54	349	AAK20723
RmlC	<i>S. elodea</i> ATCC 31461	100	188	AAP57699
	<i>Sphingomonas</i> strain S88	94/96	188	AAC44073
	<i>P. aeruginosa</i> PAO1	49/62	181	NP_253851
	<i>X. campestris</i> pv. <i>campestris</i> ATCC 33913	44/55	185	NP_636015
	<i>B. pseudomallei</i>	48/61	183	AAD05456
	<i>N. aromaticivorans</i>	47/62	181	ZP_00095574
	<i>E. coli</i> K-12	45/58	185	NP_416542
	<i>S. enterica</i> subsp. <i>enterica</i> serovar Typhimurium	48/58	183	NP_456643
	<i>L. lactis</i> subsp. <i>cremoris</i>	29/46	197	AAO15543
	<i>S. pneumoniae</i> serotype 18C	30/45	197	AAK20722
RmlD	<i>S. elodea</i> ATCC 31461	100	288	AAP57701
	<i>Sphingomonas</i> strain S88	87/91	288	AAC44075
	<i>P. aeruginosa</i> PAO1	44/55	302	NP_253849
	<i>X. campestris</i> pv. <i>campestris</i> ATCC 33913	43/56	302	NP_636016
	<i>B. pseudomallei</i>	40/56	298	AAD05457
	<i>N. aromaticivorans</i>	40/52	285	ZP_00095573
	<i>E. coli</i> K-12	41/55	299	NP_416544
	<i>S. enterica</i> subsp. <i>enterica</i> serovar Typhimurium	42/56	299	NP_456645
	<i>L. lactis</i> subsp. <i>cremoris</i>	37/50	299	AAO15545
	<i>S. pneumoniae</i> serotype 18C	35/49	293	AAK20724

^a RmlA, glucose-1-phosphate thymidyltransferase; RmlB, dTDP-glucose-4,6-dehydratase; RmlC, dTDP-4-keto-L-rhamnose-3,5-epimerase; RmlD, dTDP-L-rhamnose synthase.

containing 30 U of pyrophosphorylase/mg protein, at the upper end of the range of data analyzed in Results, was unaffected by a doubling of the coupling-enzyme concentration. The calculation of the kinetic constants was based directly on the Michaelis-Menten equation; the rectangular hyperbolic function was solved by using iterative procedures with the Solver computer program (from Microsoft Excel). The k_{cat} values (min^{-1}) were calculated on the basis of a molecular mass of 31.5 kDa for RmlA and 31.2 kDa for UgpG. The catalytic efficiency was based on the k_{cat}/K_m ratio. All assays were performed at 30°C with 6 μg of purified RmlA or UgpG in the presence of 2 mM PPI.

Nucleotide sequence accession numbers. The nucleotide sequences of the *S. elodea* ATCC 31461 *rmlA* gene and the UgpG genomic region were deposited in the GenBank database under the accession numbers AY247402 and AY873993, respectively.

RESULTS

Sequence analysis of *S. elodea* *rmlACBD* biosynthetic cluster.

A 35-kb chromosomal region of *S. elodea* ATCC 31461, encompassing a cluster of genes involved in gellan synthesis, was

cloned and partially sequenced in a previous study (34, 44). We started the present study by sequencing the first gene (*rmlA*) of the putative *rml* operon, located within the *gel* cluster. More recently, the nucleotide sequence of this four-gene *rml* cluster was deposited in GenBank by Harding et al. (15) and used in this work to deduce the amino acid sequences of the RmlB, RmlC, and RmlD proteins. These sequences were compared with the corresponding sequences from seven gram-negative and two gram-positive organisms (Table 2). The *S. elodea* ATCC 31461 RmlA, -B, -C, and -D proteins showed the highest identities with the four proteins of *Sphingomonas* strain S88 (96, 94, 94, and 87%, respectively) (Table 2). The comparison of the *rml* gene products from the nine bacterial species examined indicated that the *rmlA* gene product exhibited the highest identity values, in particular with the proteins from gram-positive bacteria (Table 2). Based on these sequence

FIG. 1. Diagrammatic representation of putative or demonstrated operonic organization of dTDP-L-rhamnose-synthesizing genes from different bacterial species. Genes are represented by arrows. *rmlA*, *rhsA*, and *rfbA*, glucose-1-phosphate thymidyltransferase; *rmlB*, *rhsB*, and *rfbB*, dTDP-glucose-4,6-dehydratase; *rmlC*, *rhsC*, and *rfbC*, dTDP-4-keto-L-rhamnose-3,5-epimerase; and *rmlD*, *rhsD*, and *rfbD*, dTDP-L-rhamnose synthase.

homology studies, the following putative functions were assigned to the Rml proteins of *S. elodea* ATCC 31461: RmlA is the glucose-1-phosphate thymidyltransferase, RmlB the dTDP-D-glucose-4,6-dehydratase, RmlC the dTDP-4-keto-L-rhamnose-3,5-epimerase, and RmlD the dTDP-L-rhamnose synthase. According to what has been described in the literature for other *rml*-like clusters, it is likely that this *S. elodea* *rml* cluster is involved in dTDP-L-rhamnose synthesis.

Significant differences in the genetic organizations of the *rml* clusters were observed among the gram-negative strains examined (Fig. 1). The two *Sphingomonas* strains, *S. elodea* ATCC

31461 and *Sphingomonas* strain S88, exhibit a similar organization of the *rml* genes (*ACBD*→). However, this organization diverges from the gene order registered for all of the other gram-negative species examined and coincides with that registered for the two gram-positive species (*Lactococcus lactis* and *Streptococcus pneumoniae*) (Fig. 1).

The four *gel* genes of the *S. elodea* ATCC 31461 *rml* cluster are contiguous, and sequence analysis suggests an operon-like structure, like those documented for the equivalent clusters of *L. lactis* (6) and *Salmonella enterica* (23). Putative ribosome-binding site sequences that precede the translational start

TABLE 3. Characteristics of *rml* genes from *S. elodea* ATCC 31461

Gene	Translation start ^a	G+C content (mol%)	Length (no. of amino acids)	Molecular mass (kDa)	pI
<i>rmlA</i>	CGCGCCatgggggATTTTGAatgAAGG	64.1	292	31.5	5.3
<i>rmlC</i>	GCaagctggtGGCTGCCGCatgACCAA	64.5	188	21.0	5.3
<i>rmlB</i>	CTATCAGggctgAGCCatgCAGC	65.4	353	39.5	5.4
<i>rmlD</i>	CGagcggtgGGCCAGACCGCCTGatgCGCA	69.3	288	30.6	4.7

^a Potential ribosome-binding sequences and start codons (bold) are set in lower case. The stop codon of the previous gene (bold) is underlined.

Enzyme/ Microorganism	Accession number	Domain (Activator- binding site)		Domain II (Sugar binding domain)	
RmlA_S.elodea	AAO89234	A G	G S G T R L Y P A T L S I S (22)	V A T S V E E K P	- - A V P K S N W (171)
RhsA_Sphingomonas S88	AAC44072	A G	G S G T R L Y P A T L S I S (22)	V A T S V E E K P	- - A N P K S N W (171)
Tgp_M.lotii	NP_107841	A G	G S G T R L Y P L T L A V S (22)	T A L T I E E K P	- - Q K P K S N W (169)
ExpA7_S.meliloti	NP_437440	A G	G R G T R L Y P V T I S V S (22)	R A E T I E E K P	- - E L A R S S W (169)
Contig30_2874_N.aromaticivorans	ZP_00095575	A G	G S G T R L Y P L T R G V S (31)	R A S S I E E K P	- - E S P K S N Y (177)
Tgp_C.crescentus	NP_419957	A G	G S G T R L H P M T L V T S (22)	T A L S I E E K P	- - K V P K S N W (168)
RmlA_P.aeruginosa	NP_253850	A G	G S G T R L H P A T L A I S (24)	K A I S L E E K P	- - L E P K S N Y (170)
Avin3453_A.vinelandii	ZP_00091744	A G	G S G T R L H P A T L A I S (24)	K A V S L E E K P	- - A Q P K S N Y (192)
RmlA_B.pseudomallei	AAD05455	A G	G S G T R L Y P I T H A V S (24)	R A L S I E E K P	- - A K P R S N Y (170)
RfbA_E.coli	NP_416543	A G	G S G T R L Y P V T M A V S (25)	T A I S L E E K P	- - L E P K S N Y (171)
RfbA_S.enterica	NP_461040	A G	G S G T R L Y P V T M A V S (25)	T A V S L E E K P	- - L Q P K S N Y (171)
RfbA_X.campestris	NP_636014	A G	G S G T R L Y P I T K G V S (36)	K V I D I A E K P	- - E K P R S N Y (172)
RfbA_L.lactis	AAO15542	A G	G S G T R L Y P L T R A A S (22)	N A V S I E E K P	- - E E P K S N Y (168)
RmlA_S.pneumoniae	AAK20721	A G	G S G T R L Y P L T R A A S (22)	N A I S I E E K P	- - E Q P C S N Y (168)
RfbA_P.syringae	NP_790913	A G	G S G T R L H P L T L G V S (27)	R A L S I E E K P	- - A A P K S S Y (173)
Pflu2601_P.fluorescens	ZP_00085347	A G	G S G T R L H P L T L G V S (27)	Q A L S I E E K P	- - L K P K S N Y (173)
RmlA_V.cholerae	AAO88918	A G	G S G T R L Y P I T R G V S (22)	R A I S I E E K P	- - L K P K S N Y (168)
UgpG_S.elodea	AAL79535	V A	G L G T R F L P A T K A M P (28)	E V K G L V E K P	A P G T A P S N L (202)
Contig30_4003_N.aromaticivorans	ZP_00304230	V A	G L G T R F L P A T K A I P (32)	E V K G L V E K P	K R E D A P S N L (207)
Ugp_Z.mobilis	AAD19415	V A	G Q G T R F L P A T K A M P (26)	E V K G L V E K P	- - E K A P S R L (201)
		G X G T R X X P X T		X E K P	

FIG. 2. Multiple-amino-acid alignment of the two characteristic consensus sequences, including the activator binding region (domain I) and the sugar binding region (domain II), of known bacterial TGPs and UGPs. Each consensus sequence delineates the conserved N-terminal peptide motif found in the large family of XDP-sugar pyrophosphorylase enzymes and a lysine residue located within the X-E-K-P motif. Numbers of intervening amino acids are given in parentheses. CLUSTAL X software (16) was used to generate this multiple-sequence alignment. Organisms include *Mesorhizobium lotii*, *Caulobacter crescentus*, *Azotobacter vinelandii*, *Burkholderia pseudomallei*, *Pseudomonas syringae*, *Pseudomonas fluorescens*, and *Zymomonas mobilis*.

codons of the four genes were identified (Table 3). The *rml* genes have overlapping stop/start codons (*rmlA* and *rmlC* [4 bp], *rmlB* and *rmlD* [1 bp]) (Table 3), reinforcing the idea that they may be cistronically linked. A putative *rho*-independent transcription terminator, located downstream of the last gene of the operon, the *rmlD* gene, was also identified. The G+C contents of the four genes, as well as information on the deduced polypeptides, are shown in Table 3. The *rmlACBD* block of genes has a mean G+C content of 65.8 mol%, consistent with the high values described for the *Sphingomonas* genes already characterized (26, 43, 44, 45).

A complete multiple-amino-acid alignment with the dTDP-glucose pyrophosphorylases (encoded by the *A* genes of the respective clusters) from prokaryotic organisms available in databases was performed using the CLUSTAL X alignment program (16). From all of the sequences examined in this work, only the RmlA proteins from *Pseudomonas aeruginosa* (29) and *Lactococcus lactis* (6) and the RfbA proteins from *E. coli* (46) and *Salmonella enterica* (25) were biochemically characterized. The amino acid sequences of three bacterial UGPs were also included in this study. Two regions (domains I and II) were found to be highly conserved throughout the sequences compared (Fig. 2). The N-terminal region (domain I) of these proteins exhibits the motif G-X-G-T-R-X₂-P-X-T. Domain II, found among the bacterial UGPs and TGPs, is also present in *S. elodea* ATCC 31461 RmlA; this domain has a conserved lysine residue (K₁₆₃), located within the E-E-K-P motif, which has been proven to be essential for glucose-1-phosphate binding in *P. aeruginosa* RmlA (5) (Fig. 2). This motif, or a closely related sequence (A-E-K-P, K-E-K-P, Q-E-K-P, I-E-K-P, or V-E-K-P), is found in many bacterial enzymes which share the ability of catalyzing the synthesis of nucleoside

diphosphate sugars from sugar-1-phosphates and nucleoside triphosphates (5, 42).

The phylogenetic tree prepared with the 17 RmlA proteins (putative or confirmed TGPs) and the three putative or confirmed UGP proteins reveals that RmlA and UGP proteins form two divergent groups (Fig. 3). Within the RmlA group, it is interesting that the RmlA proteins of *S. elodea* and *Sphingomonas* strain S88 are a branch apart from others, including *Novosphingobium aromaticivorans*, that are phylogenetically close to *S. elodea*. This atypical feature, which is not coincident with the phylogenetic distance between the species, may indicate that the corresponding *S. elodea* and *Sphingomonas* strain S88 *rmlA* genes were introduced to their current positions more recently as the result of recombination events.

Overproduction in *E. coli* and purification of heterologous *S. elodea* rmlA. To overproduce the *S. elodea* ATCC 31461 RmlA protein for biochemical characterization, the *rmlA* gene was overexpressed from plasmid pRmlA in *E. coli* BL21(DE3) cells, and the dTDP-glucose pyrophosphorylase activity was assayed in cell crude extracts and compared with that present in the cell crude extract prepared from *E. coli* BL21(DE3) with the cloning vector pWH844. This extract, prepared from the control cells, exhibited a low level of dTDP-glucose pyrophosphorylase activity (21 mU/mg), while those from *E. coli* BL21(DE3) overproducing RmlA protein from plasmid pRmlA exhibited increased levels of dTDP-glucose pyrophosphorylase activity following IPTG induction (73 mU/mg). The fused His₆-RmlA protein was purified to homogeneity from this cell extract by using a His tag purification kit. Under denaturing conditions, the purified fusion protein migrated as a single polypeptide with an estimated molecular mass of approximately 32.8 kDa (31.5 kDa from the native protein, plus

FIG. 3. Phylogenetic tree of the dTDP-glucose pyrophosphorylase (RmlA) and UDP-glucose pyrophosphorylase (Ugp) proteins listed in Fig. 2. The distance matrix was calculated by the TREE-PUZZLE method (39), based on the manual alignment of the complete protein sequences performed with the CLUSTAL X program. The tree was constructed by the neighbor-joining method (35). The numbers for the interior branches are bootstrap percentages. Percentages less than 50 are not shown. A branch length of 0.5 substitution/site is given for phylogenetic distances. Organisms include *Mesorhizobium loti*, *Pseudomonas syringae*, *P. fluorescens*, *Caulobacter crescentus*, *Burkholderia pseudomallei*, *Azotobacter vinelandii*, and *Zymomonas mobilis*.

1.3 kDa corresponding to the His₆ tag) (results not shown). Enzyme assays performed with the purified RmlA protein (2,952 mU/mg) reinforced the idea that RmlA is a dTDP-glucose pyrophosphorylase (Fig. 4).

***S. elodea* RmlA and UgpG exhibit both TGP and UGP activities.** The kinetic analysis of the TGP activity of the RmlA protein for the pyrophosphorolysis reaction showed a typical Michaelis-Menten saturation pattern (Fig. 4). Compared with the other characterized prokaryotic glucose-1-phosphate thymidyltransferase proteins, the estimated apparent K_m (Table 4) showed the highest values for dTDP-Glc (12.0 μ M), compared with 36.0 μ M for *Pseudomonas aeruginosa* RmlA (29),

83.0 μ M for *Salmonella enterica* serovar Typhimurium RfbA (25), and 95.0 μ M for *Escherichia coli* RfbA (46).

With respect to UGP activity, the estimated apparent K_m for *S. elodea* RmlA was 229.0 μ M, 19-fold higher than the K_m value for TGP activity. On the other hand, the V_{max} value for UDP-Glc (37.05 μ mol min⁻¹ mg⁻¹) was approximately 15-fold higher than the corresponding value for dTDP-Glc (Table 4). These results, obtained for the pyrophosphorolysis reactions catalyzed by RmlA in vitro, suggest that for limited concentrations of dTDP-Glc and UDP-Glc, the RmlA enzyme may act preferentially as a glucose-1-phosphate thymidyltransferase, as suggested before for the *P. aeruginosa* RmlA enzyme (29).

FIG. 4. Saturation curves of purified His₆-RmlA and His₆-UgpG with various concentrations of TGP or UGP, conforming to Michaelis-Menten kinetics. The apparent Michaelis parameters (Table 4) were calculated directly based on the Michaelis-Menten equation.

The *S. elodea* UgpG protein was previously characterized in our laboratory as a UGP (or glucose-1-phosphate uridylyl-transferase) and was hypothesized to be involved in the synthesis of UDP-glucose, one of the activated precursors for gellan biosynthesis (26). In the present work, we further demonstrate that, besides being a glucose-1-phosphate uridylyl-transferase, UgpG also exhibits glucose-1-phosphate thymidylyltransferase activity (Fig. 4). The affinity of the two enzyme activities associated with UgpG (K_m values for UDP-glucose and TDP-glucose are, respectively, 7.5 and 9.4 μM) is higher than the affinity estimated for RmlA (K_m values for UDP-glucose and TDP-glucose are, respectively, 229.0 and 12.0 μM), but the $V_{\max}$ values associated with UgpG are lower (Table 4).

Analysis of the nucleotide sequence of the *ugpG* gene-adjacent regions. The apparent redundancy of UgpG and RmlA proteins suggested by the results of the enzyme assays carried out in vitro led us to consider the inspection of the chromosomal region surrounding this putative gellan gene, since the insertional inactivation of the *ugpG* gene in ATCC 31461 was repeatedly tried before, without success (26).

TABLE 4. Kinetic parameters of RmlA and UgpG from *S. elodea* ATCC 31461

Protein	Nucleotide sugar	$V_{\max}$ ($\mu\text{mol min}^{-1}$ mg^{-1})	K_m (μM)	k_{cat} (min^{-1}) ^a	k_{cat}/K_m (min^{-1} μM^{-1})
RmlA	dTDP-Glc	2.57	12.0	81.0	6.8
	UDP-Glc	37.05	229.0	1,167.0	5.1
UgpG	dTP-Glc	1.07	9.4	33.4	3.6
	UDP-Glc	1.28	7.5	39.9	5.3

^a k_{cat} values were calculated on the basis of RmlA and UgpG as 31.5- and 31.2-kDa monomers, respectively.

A 4,638-bp region containing adjacent genes upstream and downstream from the *ugpG* gene was thus sequenced to gain insight into the physiological function of the encoded protein that might (also) be involved. This analysis revealed the presence of four open reading frames (ORFs) surrounding the *ugpG* gene (Fig. 5). Results of the comparison of the deduced amino acid sequences of these ORFs with the primary structures of proteins available in databases are shown in Table 5. The deduced amino acid sequence of the partially sequenced region of ORF 1 (362 bp) showed homology with proteins with UDP-*N*-acetyl-D-mannosaminuronic acid dehydrogenase activity, ORF 2 showed strong homology with proteins with UDP-*N*-acetyl-D-glucosamine-2-epimerase activity, ORF 3 showed homology with proteins belonging to the UDP-glucose/GDP-mannose dehydrogenase family, and the sequenced region of ORF 4 (623 bp) showed homology with proteins with UDP-glucuronic acid epimerase activity.

The investigation of the neighboring regions of genes encoding UDP-glucose pyrophosphorylase proteins in other bacterial genomes (*Novosphingobium aromaticivorans*, *Sinorhizobium meliloti*, *Agrobacterium tumefaciens*, *Pseudomonas aeruginosa*, *Xanthomonas campestris*, *Escherichia coli*, *Ralstonia solanacearum*, *Staphylococcus aureus*, and *Vibrio cholerae*) indicated that the gene organization observed in the referred chromosomal region of the gellan-producing strain is distinct. However, orthologues of *S. elodea* ATCC 31461 ORF 1 and ORF 2 appear together in different bacterial species (e.g., *Escherichia coli*, *Ralstonia solanacearum*, *Staphylococcus aureus*, *Vibrio cholerae*, and *Bacteroides fragilis*) (Fig. 6A). These two genes are involved in the synthesis of UDP-*N*-acetyl-D-mannosaminuronic acid from UDP-*N*-acetylglucosamine, required for the incorporation of this activated sugar in the CPS of *S. aureus* (33, 37) and *Bacteroides fragilis* (9) and also for the synthesis of the enterobacterial common antigen (2, 28).

FIG. 5. Genetic organization of the *ugpG* flanking regions in the *S. elodea* ATCC 31461 chromosome (GenBank accession no. AY873993). The nucleotide sequences of ORF 1 and ORF 4 were not completely determined.

The deduced amino acid sequence of ORF 3 showed homology with proteins belonging to the UDP-glucose/GDP-mannose dehydrogenase family, and the sequenced region of ORF 4 (623 bp) showed homology with proteins with UDP-glucuronic acid epimerase activity. The presence of ORF 3 close to ORF 4 suggests that the products of these genes may be necessary for the synthesis of UDP-glucuronic acid or GDP-mannuronic acid. In *Sinorhizobium meliloti* strain 1021, *Agrobacterium tumefaciens* strain C58, and *E. coli* serotype O113, the UDP-glucose-6 dehydrogenase and UDP-glucuronic acid epimerase proteins are also encoded by two adjacent genes (*rpkK* and *lpsL*, *AGR_L_1413* and *AGR_L_1415*, or *ugd* and *wbnF*, respectively) (Fig. 6B). In *S. meliloti*, these two proteins are involved in the synthesis of UDP-galacturonic acid for the incorporation of galacturonic acid in the CPS of this species (20). In *E. coli*, WbnF may be involved in the synthesis of UDP-galacturonic acid or UDP-*N*-acetyl galactosamine, necessary for the incorporation of the respective sugars in the oligosaccharide repeat unit of the *E. coli* O113 O antigen (32).

DISCUSSION

In this work, we carried out a functional analysis of the first gene of the four-gene cluster of *S. elodea* ATCC 31461 with homology to other genes demonstrated to be involved in the synthesis of L-rhamnose for incorporation into bacterial exopolysaccharides (6, 21). As predicted, the encoded RmlA protein was demonstrated to exhibit TDP-glucose pyrophosphorylase (or glucose-1-phosphate thymidyltransferase) activity, thus being involved in the reversible conversion of G1P and dTTP into dTDP-Glc and PPi. *S. elodea* RmlA was found to recognize both deoxythymidine and deoxyuridine nucleotides as sub-

strates, as described before for RmlA homologues (25, 31, 40). These results are in accordance with previously described results for other RmlA enzymes, which were found to catalyze both the formation of dTDP-Glc and UDP-Glc and pyrophosphorolysis (25, 29, 46). For practical reasons, in the present study, the kinetic constants were determined for the pyrophosphorolysis reaction.

The three-dimensional structures of different enzymes with glucose-1-phosphate thymidyltransferase activity and of the complexes they form with their substrate(s), products, and inhibitors were determined for *Pseudomonas aeruginosa* RmlA (5), *Escherichia coli* RmlA (46), and *Salmonella enterica* LT2 RmlA (3). Studies with *P. aeruginosa* RmlA revealed that the arginine residue in conserved domain I, G-X-G-T-R*-X₂-P-X-T (R*₁₃ in the RmlA sequence of *S. elodea* ATCC 31461), is involved in binding the triphosphate group of dTTP (5). It was postulated that the referred conserved domain constitutes at least part of the activator binding site (5, 8, 42), representing a prototypic structure characteristic of the large family of bacterial sugar nucleotidyl transferases (8). The overall conservation of particular residues that were proven to be catalytically important for *P. aeruginosa* RmlA, specifically, K₂₅, D₁₁₀, and D₂₂₅ (5), suggests a common overall three-dimensional structure for dTDP- and UDP-glucose pyrophosphorylase enzymes. According to Blankenfeldt et al. (5), there are three other important conserved residues in the *P. aeruginosa* RmlA protein (G₁₀, Q₈₂, and G₈₇), all of which form hydrogen bonds with the pyrimidine ring (corresponding to G₈, Q₈₀, and G₈₅ in *S. elodea* RmlA) (the last two residues are not shown in Fig. 2).

In the present work, we also found that UgpG, the previously characterized UDP-glucose pyrophosphorylase involved in UDP-glucose formation (reference 26 and our more recent results), also exhibits in vitro dTDP-glucose pyrophosphorylase activity. Similarly, the UGP and RmlA enzymes from *E. coli* (4, 40), *S. enterica* serovar Typhimurium (31), and *S. enterica* (25) also exhibit these two activities. Interestingly, the affinities of UgpG for both substrates are similar (*K_m* of 7.5 μM for UDP-Glc and *K_m* of 9.4 μM for dTDP-Glc), and the *V_{max}* values are also of the same level. The affinities of UgpG for both substrates are also above the affinities estimated for other UGP and RmlA enzymes previously characterized (4, 25, 31, 40). Taken together, the results of the in vitro pyrophosphorolysis

TABLE 5. Comparison of ORFs in UgpG vicinity with primary structures of proteins available in databases

ORF	Organism	Protein	Identity/similarity (%)	Accession no.
1 (362 bp)	<i>N. aromaticivorans</i>	UDP- <i>N</i> -acetyl-D-mannosaminuronate dehydrogenase	50/62	ZP_00094678
	<i>V. cholerae</i>	UDP- <i>N</i> -acetyl-D-mannosaminuronate dehydrogenase	46/59	B82264
	<i>Pasteurella multocida</i>	UDP- <i>N</i> -acetyl-D-mannosaminuronate dehydrogenase	42/60	AAF67264
2	<i>N. aromaticivorans</i>	UDP- <i>N</i> -acetylglucosamine-2-epimerase	63/74	ZP_00094679
	<i>V. cholerae</i>	UDP- <i>N</i> -acetylglucosamine-2-epimerase	50/64	A82264
	<i>R. solanacearum</i>	UDP- <i>N</i> -acetylglucosamine-2-epimerase	49/64	P52641
3	<i>S. meliloti</i>	Nucleotide sugar oxidase protein	58/71	NP_437300
	<i>S. meliloti</i>	UDP-glucose/GDP-mannose dehydrogenase	57/73	NP_437170
	<i>A. tumefaciens</i>	UDP-glucose/GDP-mannose dehydrogenase	56/71	NP_355520
4 (623 bp)	<i>S. meliloti</i>	UDP-glucuronic acid epimerase	57/67	CAA10917
	<i>A. tumefaciens</i>	UDP-glucuronic acid epimerase	56/66	NP_356500
	<i>E. coli</i>	Nucleotide sugar epimerase	52/65	AAD50494

FIG. 6. Genetic organization of homologues of *S. elodea* ATCC 31461 ORF 1 and ORF 2 (A) and ORF 3 and ORF 4 (B) in diverse genomes. Accession numbers are given in parentheses below the strain/plasmid names. nt, nucleotide; str, strain.

assays indicate that these two enzymes, RmlA and UgpG, putatively involved in the gellan biosynthetic pathway, can use both UDP-Glc and dTDP-Glc, at least in vitro. Nevertheless, when plasmid pRmlA was introduced into the UDP-glucose pyrophosphorylase (GalU) mutant strain *E. coli* FF4001, the ability of the mutant to grow on galactose was not restored (results not shown), which was different from earlier conclusions made when UgpG was expressed from plasmid pUgpG in the same genetic background with the *galU* mutation (26). This test monitored the capacity of the recombinant cells to ferment galactose, which requires the expression of UGP activity. These observations suggest that, despite the similarities of the biochemical characteristics examined by the in vitro pyrophosphorolysis assays, RmlA and UgpG proteins are not redundant in vivo.

In order to obtain further insight into the role of UgpG in *S. elodea*, we sequenced the genes in the vicinity of the *ugpG* gene. Interestingly, we found that the genetic organization of the overall region has no homologous regions in other prokaryotes, not even in *N. aromaticivorans*, which is phylogenetically close to *S. elodea*. Nevertheless, the four genes in the chromosomal region surrounding the *ugpG* gene were found in other organisms, organized in genetic clusters involved in the synthesis of capsular polysaccharides (9) or teichuronic acids (1, 10, 14) in gram-positive bacteria and lipopolysaccharides (2, 28, 33, 37) in gram-negative bacteria. This observation raised the question of the putative involvement of this genomic region

in the biosynthesis of a cell wall polysaccharide in *S. elodea*. Little is known about the structure of the cell wall of bacteria belonging to the genus *Sphingomonas*, but it was demonstrated that it does not contain the LPS characteristic of gram-negative bacteria in the outer membrane. Interestingly, this LPS is replaced by two types of glycosphingolipids composed of glucuronic acid (GLS-1) or glucuronic acid, galactosamine, galactose, and mannose in the glycosyl portion (GLS-4) (18, 19). However, no evidence was obtained in this work to support the hypothetical involvement of the chromosomal region where *ugpG* maps in the synthesis of these or other cell glycans.

ACKNOWLEDGMENTS

This work was supported by FEDER and Fundação para a Ciência e a Tecnologia (FCT), Portugal (grants PRAXIS/P/BIO/12020/1998, POCTI/BIO/35733/2000, POCTI/BME/44441/2002, POCTI/BIO/58401/2004, and Ph.D. and postdoctoral grants to E. Silva, A. T. Granja, and A. R. Marques).

REFERENCES

1. Arakawa, H., and E. Ito. 1986. Biosynthetic studies on *N*-acetylmannosaminuronic acid containing teichuronic acid in *Bacillus megaterium*. *Can. J. Microbiol.* 32:822–825.
2. Barr, K., and P. D. Rick. 1987. Biosynthesis of enterobacterial common antigen in *Escherichia coli*. In vitro synthesis of lipid-linked intermediates. *J. Biol. Chem.* 262:7142–7150.
3. Barton, W. A., J. Lesniak, J. B. Biggins, P. D. Jeffrey, J. Jiang, K. R. Rajashankar, J. S. Thorson, and D. B. Nikolov. 2001. Structure, mechanism and engineering of a nucleotidyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* 8:545–551.

4. Bernstein, R. L., and W. P. Robbins. 1965. Control aspects of uridine 5'-diphosphate glucose and thymidine 5'-diphosphate glucose synthesis by microbial enzymes. *J. Biol. Chem.* **240**:391-397.
5. Blankenfeldt, W., M. Asuncion, J. S. Lam, and J. H. Naismith. 2000. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.* **19**:6652-6663.
6. Boels, I. C., M. M. Beerthuyzen, M. H. W. Kusters, M. P. W. Van Kaauwen, M. Kleerebezem, and W. M. de Vos. 2004. Identification and functional characterization of the *Lactococcus lactis* *rfb* operon, required for dTDP-rhamnose biosynthesis. *J. Bacteriol.* **186**:1239-1248.
7. Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
8. Brown, K., F. Pompeo, D. Dixon, D. Mengin-Lecreux, C. Cambillau, and Y. Bourne. 1999. Crystal structure of the bifunctional *N*-acetylglucosamine 1-phosphate uridylyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily. *EMBO J.* **18**:4096-4107.
9. Comstock, L. E., M. J. Coyne, A. O. Tzianabos, and D. L. Kasper. 1999. Interstrain variation of the polysaccharide B biosynthesis locus of *Bacteroides fragilis*: characterization of the region from strain 638R. *J. Bacteriol.* **181**:6192-6196.
10. Deng, L., and J. S. Anderson. 1997. Biosynthesis of teichuronic acid in the bacterial cell wall. Purification and characterization of the glucosyltransferase of *Micrococcus luteus*. *J. Biol. Chem.* **272**:479-485.
11. DeShazer, D., P. J. Brett, and D. E. Woods. 1998. The type II O-antigenic polysaccharide moiety of *Burkholderia pseudomallei* lipopolysaccharide is required for serum resistance and virulence. *Mol. Microbiol.* **30**:1081-1100.
12. Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783-791.
13. Fialho, A. M., L. O. Martins, M.-L. Donval, J. H. Leitão, M. J. Ridout, A. J. Jay, V. J. Morris, and I. Sá-Correia. 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.
14. Gassner, G. T., J. P. Dickie, D. A. Hamerski, J. K. Magnuson, and J. S. Anderson. 1990. Teichuronic acid reducing terminal *N*-acetylglucosamine residue linked by phosphodiester to peptidoglycan of *Micrococcus luteus*. *J. Bacteriol.* **172**:2273-2279.
15. Harding, N. E., Y. N. Patel, and R. J. Coleman. 2004. Organization of genes required for gellan polysaccharide biosynthesis in *Sphingomonas elodea* ATCC 31461. *J. Ind. Microbiol. Biotechnol.* **31**:70-82.
16. Higgins, D. G., and P. M. Sharp. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**:237-244.
17. Jay, A. J., I. J. Colquhoun, M. J. Ridout, G. J. Brownsey, V. J. Morris, A. M. Fialho, J. H. Leitão, and I. Sá-Correia. 1998. Analysis of structure and function of gellans with different substitution patterns. *Carbohydr. Polym.* **35**:179-188.
18. Kawahara, K., U. Seydel, M. Matsuura, H. Danbara, E. T. Rietschel, and U. Zahring. 1991. Chemical structure of glycosphingolipids isolated from *Sphingomonas paucimobilis*. *FEBS Lett.* **292**:107-110.
19. Kawasaki, S., R. Moriguchi, K. Sekiya, T. Nakai, E. Ono, K. Kume, and K. Kawahara. 1994. The cell envelope structure of the lipopolysaccharide-lacking gram-negative bacterium *Sphingomonas paucimobilis*. *J. Bacteriol.* **176**:284-290.
20. Kereszt, A., E. Kiss, B. L. Reuhs, R. W. Carlson, A. Kondorosi, and P. Putnoky. 1998. Novel *rpk* gene clusters of *Sinorhizobium meliloti* involved in capsular polysaccharide production and invasion of the symbiotic nodule: the *rpkK* gene encodes a UDP-glucose dehydrogenase. *J. Bacteriol.* **180**:5426-5431.
21. Köplin, R., G. Wang, B. Hötte, U. B. Priefer, and A. Pühler. 1993. A 3.9-kb DNA region of *Xanthomonas campestris* pv. *campestris* that is necessary for lipopolysaccharide production encodes a set of enzymes involved in the synthesis of dTDP-rhamnose. *J. Bacteriol.* **175**:7786-7792.
22. Kuo, M. S., A. J. Mort, and A. Dell. 1986. Identification and location of L-glycerate, an unusual substituent in gellan gum. *Carbohydr. Res.* **56**:173-187.
23. Li, Q., and P. R. Reeves. 2000. Genetic variation of dTDP-L-rhamnose pathway genes in *Salmonella enterica*. *Gene* **146**:2291-2307.
24. Li, Q., M. Hobbs, and P. R. Reeves. 2003. The variation of dTDP-rhamnose pathway genes in *Vibrio cholerae*. *Microbiology* **149**:2463-2474.
25. Lindquist, L., R. Kaiser, P. R. Reeves, and A. A. Lindberg. 1993. Purification, characterization and HPLC assay of *Salmonella* glucose-1-phosphate thymidyltransferase from the cloned *rfaA* gene. *Eur. J. Biochem.* **211**:763-770.
26. Marques, A. R., P. B. Ferreira, I. Sá-Correia, and A. M. Fialho. 2003. Characterization of the *ugpG* gene encoding a UDP-glucose pyrophosphorylase from the gellan gum producer *Sphingomonas paucimobilis* ATCC 31461. *Mol. Genet. Genomics* **268**:816-824.
27. Martins, L. O., and I. Sá-Correia. 1991. Gellan gum biosynthetic enzymes in producing and nonproducing variants of *Pseudomonas elodea*. *Biotechnol. Appl. Biochem.* **14**:357-364.
28. Meier-Dieter, U., R. Starman, K. Barr, H. Mayer, and P. D. Rick. 1990. Biosynthesis of enterobacterial common antigen in *Escherichia coli*. Biochemical characterization of Tn10 insertion mutants defective in enterobacterial common antigen synthesis. *J. Biol. Chem.* **265**:13490-13497.
29. Melo, A., and L. Glaser. 1965. The nucleotide specificity and feedback control of thymidine diphosphate D-glucose pyrophosphorylase. *J. Biol. Chem.* **240**:398-405.
30. Mitchison, M., D. M. Bulach, T. Vinh, K. Rajakumar, S. Faine, and B. Adler. 1997. Identification and characterization of the dTDP-rhamnose biosynthesis and transfer genes of the lipopolysaccharide-related *rfb* locus in *Leptospira interrogans* serovar Copenhageni. *J. Bacteriol.* **179**:1262-1267.
31. Nakae, T., and H. Nikaide. 1971. Multiple molecular forms of uridine diphosphate glucose pyrophosphorylase from *Salmonella typhimurium*. *J. Biol. Chem.* **246**:4386-4396.
32. Paton, A. W., and J. C. Paton. 1999. Molecular characterization of the locus encoding biosynthesis of the lipopolysaccharide O antigen of *Escherichia coli* serotype O113. *Infect. Immun.* **67**:5930-5937.
33. Portolés, M., K. B. Kiser, N. Bhasin, K. H. N. Chan, and J. C. Lee. 2001. *Staphylococcus aureus* Cap50 has UDP-ManNAc dehydrogenase activity and is essential for capsule expression. *Infect. Immun.* **69**:917-923.
34. Sá-Correia, I., A. M. Fialho, P. Videira, L. M. Moreira, A. R. Marques, and H. Albano. 2002. Gellan gum biosynthesis in *Sphingomonas paucimobilis* ATCC 31461: genes, enzymes and exopolysaccharide production engineering. *J. Ind. Microbiol. Biotechnol.* **29**:170-176.
35. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406-425.
36. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. 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.
38. Schirmer, F., S. Ehrh, and W. Hillen. 1997. Expression, inducer spectrum, domain structure, and function of MopR, the regulator of phenol degradation in *Acinetobacter calcoaceticus* NCIB8250. *J. Bacteriol.* **179**:1329-1336.
39. Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**:502-504.
40. Sivaraman, J., V. Sauvé, A. Matte, and M. Cygler. 2002. Crystal structure of *Escherichia coli* glucose-1-phosphate thymidyltransferase (RffH) complexed with dTTP and Mg²⁺. *J. Biol. Chem.* **277**:44214-44219.
41. Sutherland, I. W. 1998. Novel and established applications of microbial polysaccharides. *Trends Biotechnol.* **16**:41-46.
42. Thorson, J. S., T. M. Kelly, and H.-W. Liu. 1994. Cloning, sequencing, and overexpression in *Escherichia coli* of the α -D-glucose-1-phosphate cytidyltransferase gene isolated from *Yersinia pseudotuberculosis*. *J. Bacteriol.* **176**:1840-1849.
43. Videira, P. A., L. L. Cortes, A. M. Fialho, and I. Sá-Correia. 2000. Identification of the *pgmG* gene, encoding a bifunctional protein with phosphoglucosyltransferase and phosphomannomutase activities, in the gellan gum-producing strain *Sphingomonas paucimobilis* ATCC 31461. *Appl. Environ. Microbiol.* **66**:2252-2258.
44. Videira, P. A., A. M. Fialho, R. A. Geremia, C. Breton, and I. Sá-Correia. 2001. Biochemical characterization of the β -1,4-glucuronosyltransferase GelK in gellan gum producing strain *Sphingomonas paucimobilis* ATCC 31461. *Biochem. J.* **358**:457-464.
45. 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.
46. Zuccotti, S., D. Zanardi, C. Rosano, L. Sturla, M. Tonetti, and M. Bolognesi. 2001. Kinetic and crystallographic analyses support a sequential-ordered bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidyltransferase. *J. Mol. Biol.* **313**:831-843.