J Ind Microbiol Biotechnol (2004) 31: 70–82
DOI 10.1007/s10295-004-0118-9

ORIGINAL PAPER

Nancy E. Harding · Yamini N. Patel
Russell J. Coleman

Organization of genes required for gellan polysaccharide biosynthesis in *Sphingomonas elodea* ATCC 31461

Received: 30 January 2003 / Accepted: 3 January 2004 / Published online: 6 February 2004

Abstract *Sphingomonas elodea* ATCC 31461 produces gellan, a capsular polysaccharide that is useful as a gelling agent for food and microbiological media. Complementation of nonmucoid *S. elodea* mutants with a gene library resulted in identification of genes essential for gellan biosynthesis. A cluster of 18 genes spanning 21 kb was isolated. These 18 genes are homologous to genes for synthesis of sphingan polysaccharide S-88 from *Sphingomonas* sp. ATCC 31554, with predicted amino acid identities varying from 61% to 98%. Both polysaccharides have the same tetrasaccharide repeat unit, comprised of [→ 4)-α-L-rhamnose-(1 → 3)-β-D-glucose-(1 → 4)-β-D-glucuronic acid-(1 → 4)-β-D-glucose-(1 →]. Polysaccharide S-88, however, has mannose or rhamnose in the fourth position and has a rhamnosyl side chain, while gellan has no sugar side chain but is modified by glyceryl and acetyl substituents. Genes for synthesis of the precursor dTDP-L-rhamnose were highly conserved. The least conserved genes in this cluster encode putative glycosyl transferases III and IV and a gene of unknown function, *gelF*. Three genes (*gelI*, *gelM*, and *gelN*) affected the amount and rheology of gellan produced. Four additional genes present in the S-88 sphingan biosynthetic gene cluster did not have homologs in the gene cluster for gellan biosynthesis. Three of these gene homologs, *gelR*, *gelS*, and *gelG*, were found in an operon unlinked to the main gellan biosynthetic gene cluster. In a third region, a gene possibly involved in positive regulation of gellan biosynthesis was identified.

Keywords *Sphingomonas elodea* · Gellan polysaccharide · S-60 · Sphingan · Gellan biosynthetic genes

N. E. Harding (✉) · Y. N. Patel · R. J. Coleman
Fermentation Sciences,
CPKelco, 8355 Aero Drive,
San Diego, CA 92123, USA
E-mail: nancy.harding@cpkelco.com
Tel.: +1-858-4676673
Fax: +1-858-4676571

Present address: R. J. Coleman
The Dow Chemical Company,
5501 Oberlin Drive, San Diego,
CA 92121, USA

Introduction

The genus *Sphingomonas* [42] includes a growing number of strains that produce capsular polysaccharides with related structures [19, 29, 32]. The backbone of these polymers, collectively referred to as sphingans [48], nearly always consists of a tetrasaccharide repeat comprised of [→ 4)-α-L-rhamnose-(1 → 3)-β-D-glucose-(1 → 4)-β-D-glucuronic acid-(1 → 4)-β-D-glucose-(1 →]. In some polymers (e.g., I-885 [6] and S-7 [43]), 2-deoxyglucuronic acid replaces glucuronic acid. In others, L-rhamnose can be partially (S-88 [15]) or fully (NW-11 [31]) replaced by L-mannose. In addition, many of these polysaccharides contain ester-linked organic acids, and a monomer or dimer side chain comprised of D-glucose, L-mannose or L-rhamnose sugar residues. These structural variations have major effects on the physicochemical characteristics of the polysaccharides, e.g., gelation, thermostability, and cation compatibility [24], and have led to unique commercial food and industrial applications for four of these polysaccharides (gellan, welan, rhamsan and diutan) [19, 29, 39].

Sphingomonas elodea ATCC 31461 (originally designated *Pseudomonas elodea*, also referred to as *Sphingomonas paucimobilis*) was isolated as a Gram-negative bacterium capable of producing a novel gelling polysaccharide, designated S-60 or gellan gum [18, 20]. This polysaccharide is a linear tetrasaccharide repeat lacking any side-chain sugar residues [14, 30], but modified with acetyl and glyceryl substituents, having on average, one glycerate and one-half acetate per repeat unit on the 3-linked glucose residue, as shown in Fig. 1 [23]. While acetyl esters are common on

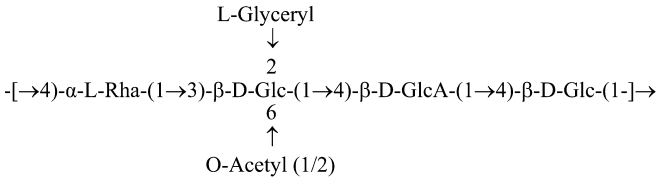


Fig. 1 Structure of the repeat unit of gellan polysaccharide secreted by *Sphingomonas elodea* ATCC 31461. *Glc* Glucose, *GlcA* glucuronic acid, *Rha* rhamnose

exopolysaccharides, the glyceryl group is an unusual substituent that has not been reported as a component of other polysaccharides. The two acyl substituents control the gel-like properties of this polysaccharide. Native or acylated gellan forms elastic gels when heated and cooled. After removal of the acyl groups by alkali treatment, deacylated gellan forms firm brittle gels in the presence of cations [1, 2, 16]. Deacylated gellan is commercially valuable as a microbiology and plant media matrix (Gelrite) and as a texture modifier in food applications (Kelcogel).

The genes involved in the biosynthesis of gellan have not been fully characterized; however, the cluster of genes for synthesis of the related sphingan polysaccharide S-88 produced by *Sphingomonas* sp. ATCC 31554 has been isolated, sequenced and characterized [48]. In addition, a cluster of genes from *Sphingomonas* sp. ATCC 21423, which produces polysaccharide S7, has been identified and shown to increase polymer production when amplified on a multicopy plasmid [43]. Although the sequence of the polysaccharide biosynthetic genes from ATCC 21423 was not reported, the arrangement of genes was shown to be similar, but not identical, to those for S-88 polysaccharide synthesis by hybridization with S-88 gene probes [43]. A gene encoding the β-1,4-glucuronosyltransferase from ATCC 31461 has been cloned, and the enzyme purified and characterized [45]. The genetics and biochemistry of gellan biosynthesis, including identification of several genes, has been recently reviewed [37].

This study reports the isolation of a series of mutants of *S. elodea* ATCC 31461 that are deficient in gellan production, and the subsequent isolation, identification and structural analysis of a cluster of genes for gellan biosynthesis. The genes in the *S. elodea* gellan biosynthetic gene cluster show some differences in organization compared to those for synthesis of the S-88 and S-7 polysaccharides. Three of the genes for polysaccharide synthesis, which are part of the *Sphingomonas* sp. ATCC 31554 S-88 sphingan gene cluster, are not adjacent to the *S. elodea* gellan gene cluster, but are located in a different region of the chromosome. Knowledge of the location and regulation of genes for biosynthesis and modification of the gellan and other sphingan polysaccharides may provide insights into methods of further modification or alteration of the structure of sphingan polysaccharides to provide polymers with unique rheological properties.

Materials and methods

Strains, plasmids, media and culture conditions

Bacterial strains and plasmids used are listed in Table 1. *Escherichia coli* cells were grown in YT or LB medium [28] at 30–37°C. *Sphingomonas* cells were grown in YM (Difco, Detroit, Mich.), YT or LB medium at 30–36°C. *S. elodea* is very mucoid when grown in YM medium, making it difficult to isolate cells free of polysaccharide. Therefore, a modified medium, YEME (0.25% yeast extract, 0.025% malt extract), was developed to reduce gellan production and improve cell suspensions. Antibiotics were used in the following concentrations (μg/ml): rifampicin (100), streptomycin (25), kanamycin (Km; 7–50 for *Sphingomonas*, 50 for *E. coli*), tetracycline (5–10 for *Sphingomonas*, 5–15 for *E. coli*), bacitracin (Bac; 1,000) and ampicillin (100). Plasmid inserts in *E. coli* DH5α strains were identified using X-gal (40 μg/ml) and IPTG (20 μg/ml). For gellan production, cells were grown in S-60 salts medium, modified from that described previously [44], containing per liter: 0.23 g NaCl, 0.165 g $CaCl_2{\cdot}2H_2O$, 2.8 g K_2HPO_4, 1.2 g KH_2PO_4, 1.9 g $NaNO_3$, 1.0 g N-Z-Amine type EKC (Sheffield Products; Quest International, Chicago, Ill.), 36.46 g Star-Dri corn syrup, 2.5 mg $FeSO_4{\cdot}7H_2O$, 24 μg $CoCl_2{\cdot}6H_2O$ and 0.1 g $MgSO_4{\cdot}7H_2O$.

Mutagenesis

Spontaneous rifampicin-resistant colonies were selected on YM agar with 100 μg/ml rifampicin. To isolate nonmucoid mutants, *S. elodea* was mutagenized using ethyl methane sulfonate (EMS, 15 μg/ml) in Tris-HCl buffer (pH 8) for 30 min at 30°C; cells were then washed, grown in YM medium, and plated on YM agar. To select bacitracin-resistant, nonmucoid mutants, an overnight culture of *S. elodea* was spread on YM agar with bacitracin (1 mg/ml).

Gene library construction

S. elodea chromosomal DNA was purified [44], partially digested with *Pst*I and fractionated by sucrose gradient centrifugation [10–40% gradient, centrifuged in an SW41 rotor at 25,000 rpm (107,000 *g*) for 20 h]. Fractions containing DNA fragments of approximately 20–30 kb were pooled, purified by ethanol precipitation and the isolated DNA ligated into the polylinker of pLAFR3 [41]. The ligation mixture was packaged into λ bacteriophage using packaging extracts (Gigapack Gold Packaging Extract; Stratagene, La Jolla, Calif.) and transduced into *E. coli* strain DH5αMCR (Gibco BRL, Rockville, Md.).

DNA techniques

Plasmid DNA was purified by alkaline lysis [38] or using commercially available kits (Qiagen, Valencia, Calif.). Poor plasmid yield was obtained with DNA purified directly from *S. elodea*, therefore plasmids isolated from *S. elodea* were transformed into *E. coli* (DH5α or SCS1) for further purification and characterization. Restriction digests and ligations were as recommended by the manufacturer (Gibco BRL). Gel electrophoresis was conducted using standard procedures [38]. Plasmids were transferred from *E. coli* into *S. elodea* using the triparental mating system [4]. Cell cultures, grown overnight in YT (minus NaCl) or YEME, were mixed in a 1:1:2 ratio of donor, helper and recipient cells, placed on YT agar, directly or after recovery on filters, and incubated for 6–8 h at 30°C, prior to plating on selective media. Plasmids were transferred into competent or electrocompetent *E. coli* cells (Gibco BRL; Stratagene) using standard procedures [38] or as recommended.

Table 1 Bacterial strains and plasmids. Sm^R Streptomycin resistant, Rif^R rifampicin resistant, Km^R kanamycin resistant, Tc^R tetracycline resistant, Su^R sulfonamide resistant, Amp^R ampicillin resistant, Mob^+ mobilizable, Tra^+ provides transfer functions

Strain or plasmid	Genotype or phenotype	Source or reference
Sphingomonas elodea		
0201 Wild-type (ATCC 31461)	Sm^R	This laboratory
4150 Plasmid cured	Sm^R	This work
4153 Rifampicin resistant #4150	Sm^R, Rif^R	This work
Sphingomonas sp. S-88	Sm^R	This laboratory
E. coli		
JZ279	*recA56 lacY galK galT22 metB1 trpR55 supE44 supF58 hsdR514*	[11]
DH5α	F^- *Φ80dΔlacZM15 Δ(lacZYA-argF)U169 deoR recA1 endA1 hsdR17*(r_K^-, m_K^+) *phoA supE44* λ^- *thi-1 gyrA96 relA1*	Gibco BRL
DH5αMCR	F^- *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dΔlacZM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 phoA supE44* λ^- *thi-1 gyrA96 relA1*	Gibco BRL
DH10B	F^- *mcrA Δ(mrr-hsdRMS-mcrBC)* Φ80d*ΔlacZM15 ΔlacX74 deoR recA1 endA1 araΔ139 Δ(ara, leu)7697 galU galK* l^- *rpsL nupG*	Gibco BRL
SCS1	*recA1 endA1 gyrA96 thi-1 hsdR17* (r_K^- m_K^+) *supE44 relA1*	Stratagene
XLI-blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′*proAB lacI*^{q} *ΔZM15 Tn10* (Tc^R)]	Stratagene
Plasmids		
pJC106	Sm^R Su^R Km^R Mob^+	This work
pLAFR3	Tc^R Mob^+ Tra^-	[41]
pRK2013	ColE1 Mob^+ Tra^+ (RK2) Km^R	[4]
pLO2	Km^R Mob^+ *sacB*	[26]
pBluescript II KS+	ColE1 f1 origin Amp^R MCS *lacZ*αpromoters T3 and T7	Stratagene
pZERO	ColE1 f1 origin Km^R MCS *lacZα*	Invitrogen

PCR amplification for plasmid constructs used to perform chromosomal insertion and deletion mutations was performed using commercial enzymes, reagents and equipment (Roche, Indianapolis, Ind., Qiagen) and primers (Amitof, Boston, Mass.), by methods described by suppliers. Probes for colony blot and Southern hybridizations were labeled with biotin (Schleicher and Schuell, Riviera Beach, Fla.) or by incorporation of digoxigenin-labeled dUTP during PCR (Boehringer Mannheim, Mannheim, Germany). Colony hybridizations were performed according to manufacturer's protocols (Schleicher and Schuell). For evaluation of the gene library, LB-tetracycline plates with 500–1,000 colonies were hybridized with labeled probe. As needed, individual colonies were purified and retested. For Southern hybridization, transfer of DNA, denaturation, hybridization and detection procedures were performed according to procedures described by Schleicher and Schuell or Boehringer Mannheim.

DNA sequencing was performed on restriction fragments subcloned into appropriate vectors. Sequences were assembled using SeqMan software (DNASTAR) and analyzed for open reading frames using Clone Manager (Sci Ed Central, Durham, N.C.) or the GCG program package (http://www.gcg.com). Homologous proteins in the database at the National Center for Biotechnology Information (NCBI) were identified using BlastX, BlastP or Psi Blast programs, or Pfam (http://www.ncbi.nlm.nih.gov).

Insertion mutagenesis and deletion

For insertional inactivation, a region of DNA within the gene was amplified by PCR and cloned into the multiple cloning site of vector pLO2 [26]. Restriction nuclease sites for cloning were added as part of the PCR primers. To construct deletions, two regions of DNA flanking the gene(s) were amplified by PCR and cloned into pLO2. Insertion of the plasmid into the chromosome was selected by kanamycin resistance. Gene deletions were isolated by selecting for loss of sucrose (8%) toxicity encoded by the *sacB* gene of pLO2 [9, 17], followed by screening by PCR for those with the correct excision. Diagnostic PCR using primers outside the deletion or insertion was used to confirm each strain construction.

Characterization of gellan

To evaluate the composition of gellan, *S. elodea* strains were grown in shake flask cultures. An aliquot from a YM overnight culture was transferred to S-60 salts medium for 24 h, then 5 ml of this culture transferred to salts medium (100 ml per 500 ml flask) and incubated with shaking at 350 rpm for 72 h at 36°C. Broth viscosity was measured at 60 rpm with a number 4 spindle in the Brookfield LVT viscometer. Gellan was recovered by heating the broth to 121°C for approximately 15 min, followed by precipitation with two volumes of isopropyl alcohol. After drying at 40°C for 16 h, the amount of total precipitable material was determined gravimetrically. For composition analysis, aliquots of approximately 50 mg were hydrolyzed with trifluoroacetic acid (100°C, about 18 h). The hydrolysate neutral sugars were quantitated by high-performance anion-exchange chromatography with pulsed amperometric detection using Dionex BioLC 300 system, as described by the supplier (Dionex, Sunnyvale, Calif.). The hydrolysate organic acids were quantitated by high-performance ion-exclusion chromatography with chemically suppressed conductivity detection using the same type of instrument (Dionex).

Nucleotide sequence accession numbers

The DNA sequences have been deposited in the NCBI GenBank database and assigned the following accession numbers: region I *gel* gene cluster (AY217008), *gelR*, *gelS*, *gelG* genes (AY220099) and *gelA* (AY220100).

Results

Isolation of mutants of *S. elodea* with increased conjugation frequency

Initial experiments showed that the frequency of conjugal transfer of the IncP type broad-host-range vector pLAFR3 (22 kb) [41] from *E. coli* to *S. elodea* was very low (0.7×10^{-7}) whereas transfer of the IncQ type vector, pJC106 (10 kb), a Km^R derivative of RSF1010 [10] (J. M. Cleary, unpublished), was much higher (2.5×10^{-2}). This result was consistent with observations by others [7], noting a significant reduction in conjugation frequency with increasing plasmid size. The low conjugation frequency of pLAFR3 was not due to capsular polysaccharide, since a low frequency (0.6×10^{-7}) was also obtained with a spontaneous nonmucoid mutant.

To efficiently perform mutant complementation analysis with a *S. elodea* gene library made using the pLAFR3 vector, it was necessary to increase the efficiency of transfer. It was suspected that those few cells that were successful recipients of this plasmid represented mutants in the population with an increased ability to take up this plasmid DNA. Three transconjugants containing pLAFR3 were cured of the plasmid by growing them under nonselective conditions with serial passage for about 30 generations. Three plasmid-cured strains (tetracycline-sensitive derivatives from each of the initial transconjugants) exhibited conjugation frequencies of 4.2×10^{-3}, 0.6×10^{-2} and 1.5×10^{-2}, representing a 10^5-fold increase compared to the wild type strain. The increased conjugation frequency was stable and reproducible. A possible explanation may be the elimination of a restriction nuclease system in *S. elodea* that was causing degradation of incoming DNA; however, this has not been tested. One of these plasmid-cured strains, #4150, and a rifampicin-resistant derivative, #4153 (Table 1) were used for isolation of nonmucoid mutants. However, rifampicin resistance was observed to decrease gellan production by between 10 and 40%, as reported [7].

Isolation and complementation of nonmucoid mutants

Using cured strain #4150 and rifampicin-resistant #4153, 51 nonmucoid (gellan polysaccharide synthesis) mutants (Gps1–51) were isolated from over 100,000 colonies following EMS mutagenesis. An additional 10 nonmucoid mutants of #4150 (Gps52–61) were isolated using selection for bacitracin resistance. The antibiotic bacitracin inhibits the dephosphorylation of C_{55}-isoprenyl pyrophosphate to form C_{55}-isoprenyl phosphate, which is essential for both polysaccharide and peptidoglycan syntheses [34, 40]. Among mutants resistant to bacitracin are those blocked in polysaccharide synthesis, allowing available isoprenyl phosphate to be used for peptidoglycan synthesis for the bacterial cell wall. The nonmucoid mutants were observed to be of two phenotypes. All Bac^R mutants and 24 EMS mutants grew uniformly suspended in liquid cultures, whereas 27 EMS mutants formed clumps.

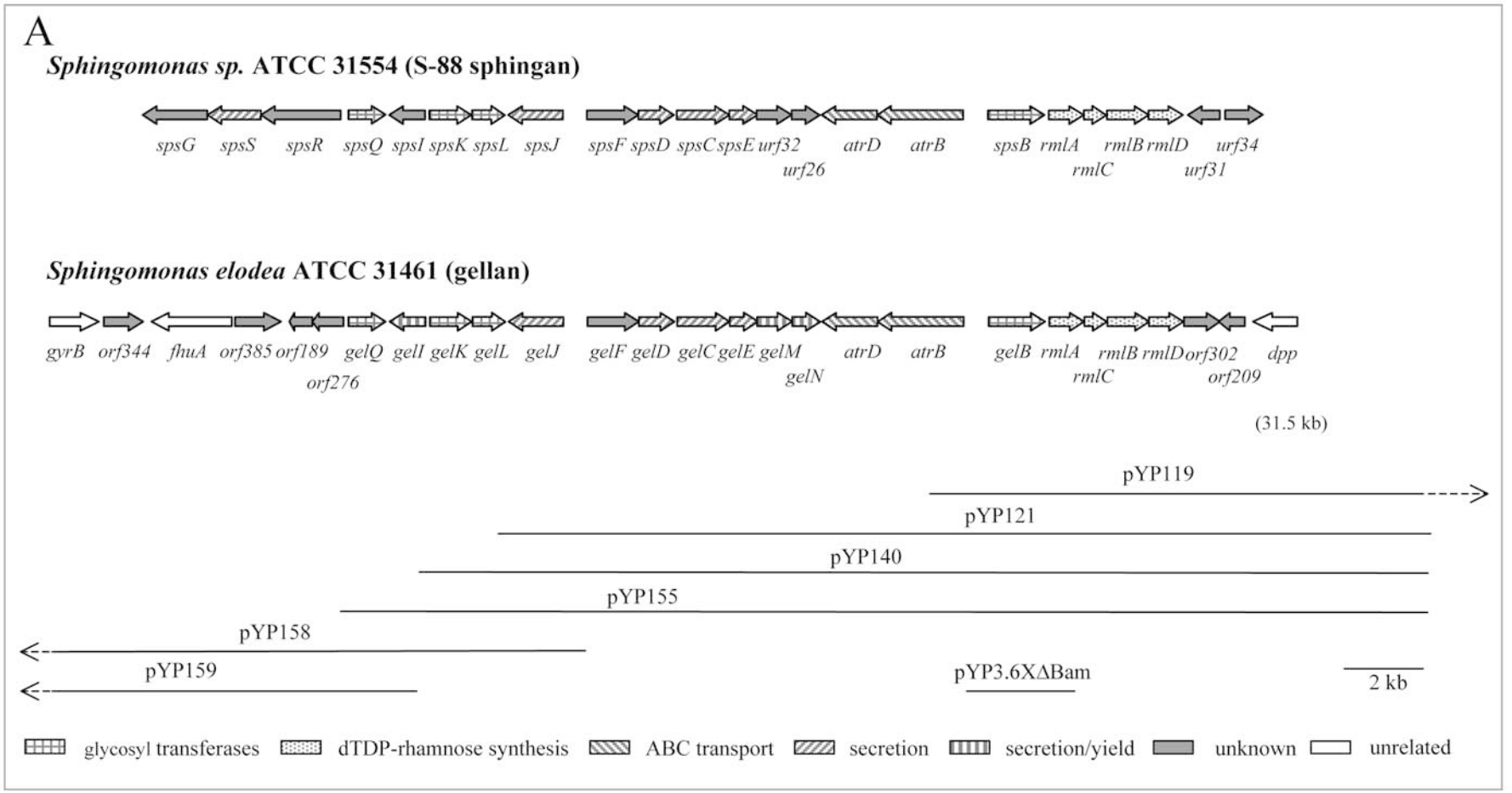
The *S. elodea* gene library was mated into three nonmucoid mutants. Mucoid transconjugants were purified and plasmid DNA isolated and analyzed. Unique plasmids were then mated into all 61 nonmucoid mutants and two major complementation regions were identified. The ten Bac^R mutants and 25 of 51 EMS mutants were complemented only by plasmids with overlapping inserts from a region designated region I. The Bac^R mutants and most (19 of 25) of the other region I mutants were readily suspended in liquid medium, whereas six region I mutants grew in clumps. Most of the mutants not mapping to region I exhibited the clumping phenotype (21 of 26). Two of these mutants (one clumping and one suspended) were used to isolate additional plasmids from the gene library, which were then mated into all remaining (non-region I) mutants. These mutants were restored to full mucoidy by plasmids with overlapping inserts from a region designated region III. These mutants were also partially complemented, (i.e., smaller, less mucoid, colonies than wild-type but more mucoid than the mutants) by plasmids with inserts from region I or another unique region designated region II, or random inserts (data not shown). The region III Gps mutants may produce a low level of polysaccharide and thus different genes can promote increased gellan synthesis. Cell clumping is likely due to polysaccharide, since mutants completely blocked in gellan synthesis do not form clumps but grow in uniform suspension.

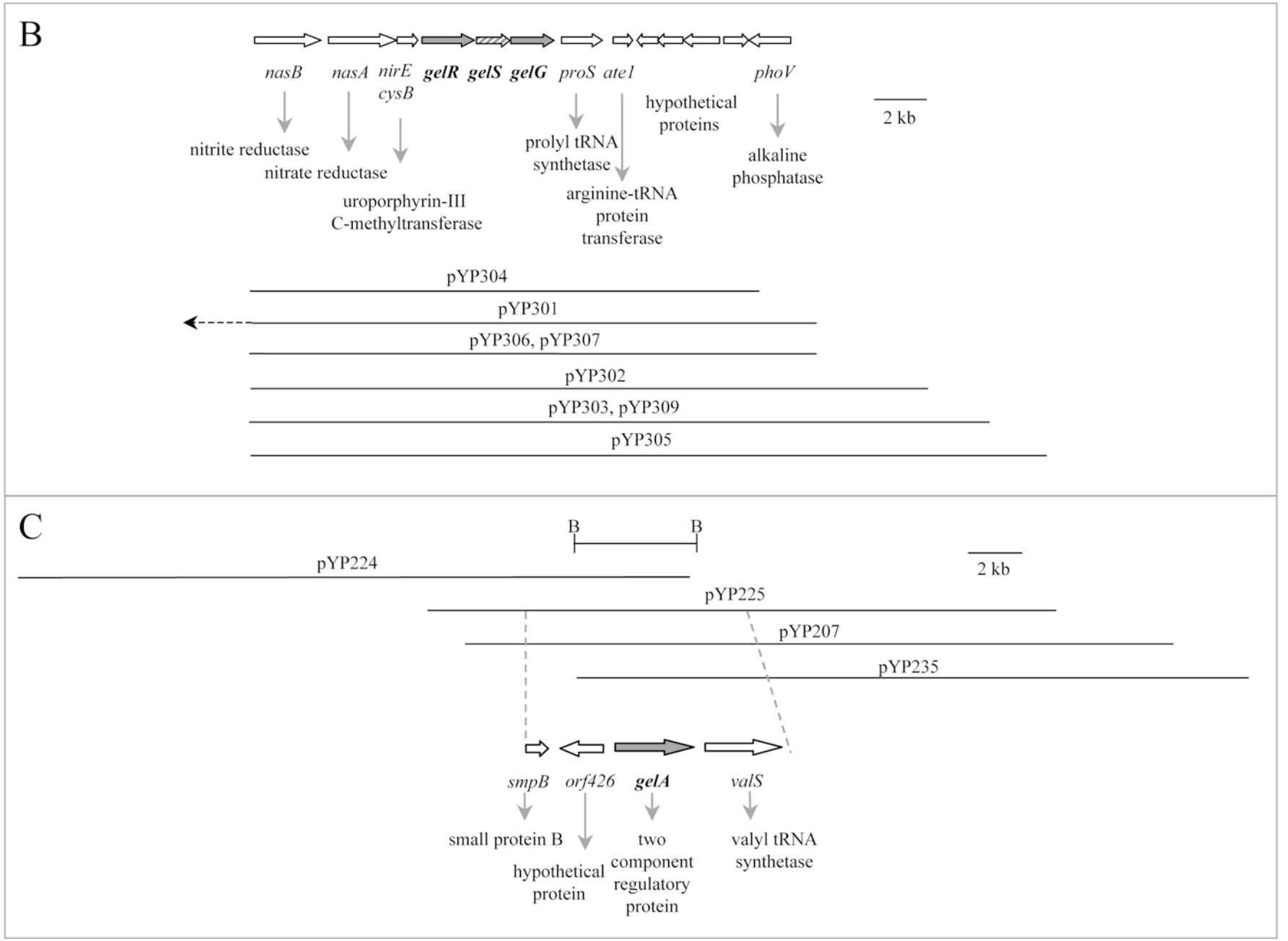
Sequence of genes for gellan synthesis in region I

Plasmids complementing region I Gps mutants were analyzed by digestion with *Bam*HI and *Hin*dIII, or *Bam*HI and *Eco*RI, and a restriction enzyme map assembled. Appropriate restriction enzyme fragments were subcloned and sequenced, generating a sequence of about 22 kb. Alignment of this sequence, represented by pYP140 (Fig. 2A), with the published sequence of genes for S-88 sphingan synthesis (GenBank accession number

▶

Fig. 2A–C Genes (*gel*) for gellan biosynthesis. Putative or known gene functions are indicated. Genes of unknown function are designated *orf* (open reading frame) and a number indicative of the encoded amino acids. DNA segments cloned on plasmids are shown. **A** Gene alignment for the cluster of gellan biosynthetic genes (region I, GenBank accession number AY217008) and comparison to S-88 sphingan biosynthetic genes (adapted from Yamazaki et al. [48], GenBank accession number U51197). **B** Location and arrangement of *gelR*, *gelS* and *gelG* genes (AY220099). Plasmids shown were obtained from the gene library by hybridization to a *gelS-gelG* probe. **C** Map of the *gelA* region (AY220100). Plasmids complemented region III Gps (gellan polysaccharide synthesis) mutants. The initial *Bam*HI fragment cloned (B) is indicated


A
Sphingomonas sp. ATCC 31554 (S-88 sphingan)
spsG spsS spsR spsQ spsI spsK spsL spsJ
spsF spsD spsC spsE urf32 urf26 atrD atrB
spsB rmlA rmlC rmlB rmlD urf31 urf34
Sphingomonas elodea ATCC 31461 (gellan)
gyrB orf344 fhuA orf385 orf189 orf276 gelQ gelI gelK gelL gelJ
gelF gelD gelC gelE gelM gelN atrD atrB
gelB rmlA rmlC rmlB rmlD orf302 orf209 dpp
(31.5 kb)
pYP119
pYP121
pYP140
pYP155
pYP158
pYP159
pYP3.6XΔBam
2 kb
glycosyl transferases
dTDP-rhamnose synthesis
ABC transport
secretion
secretion/yield
unknown
unrelated


B
nasB
nasA
nirE
cysB
gelR
gelS
gelG
proS
ate1
phoV
nitrite reductase
nitrate reductase
uroporphyrin-III
C-methyltransferase
prolyl tRNA
synthetase
arginine-tRNA
protein
transferase
hypothetical
proteins
alkaline
phosphatase
2 kb
pYP304
pYP301
pYP306, pYP307
pYP302
pYP303, pYP309
pYP305
C
B
B
pYP224
2 kb
pYP225
pYP207
pYP235
smpB
orf426
gelA
valS
small protein B
hypothetical
protein
two
component
regulatory
protein
valyl tRNA
synthetase

U51197 [48]) suggested that additional DNA was needed. Using mutant Gps31 (a double mutant in region I) as recipient, complementation was performed with the gene library. However, only an additional 2 kb of DNA was isolated, represented by pYP155. Since additional DNA was not obtained by functional complementation, clones containing adjacent DNA were identified by DNA hybridization. Colony hybridization was performed on the *S. elodea* gene library using as a probe a biotin-labeled 280 bp fragment amplified by PCR from the additional 2 kb in pYP155 using primers P17 and P18 (Table 2). Two new plasmids, pYP158 and pYP159 were isolated (Fig. 2A), and additional DNA sequenced. This sequence was added to the original 22 kb, for a total sequence of 31.5 kb (GenBank accession number AY217008).

Comparison of genes for polysaccharide synthesis in region I

The genes for gellan synthesis were compared to the corresponding genes for synthesis of sphingan S-88 [48] (Fig. 2A). Eighteen open reading frames (ORFs)

Table 2 Primers for PCR amplification

Primer	Sequence (5′–3′)	Restriction site	Primer binding site	Purpose
P17	CTGCAGTTCCTGGCCTTTGA	None	pYP155	Probe for additional region I DNA
P18	GATGAAGTCGGTGGAGTGGT	None	pYP155	Probe for additional region I DNA
P21	TT*GAGCTC*GATCTCCACCCCTGG	*Sac*I	3′ end *orf385*	Amplify o*rf189-orf276* flanking region (660 bp)
P22	GC*TCTAGA*TGGACATGCCGACCC	*Xba*I	3′ end *orf189*	Amplify o*rf189-orf276* flanking region (660 bp)
P23	GC*TCTAGA*CATGCTTCTCTCCTC	*Xba*I	5′ end of *orf276*	Amplify o*rf189-orf276* flanking region (699 bp)
P24	AA*CTGCAG*AAGCTCTCGTCGAAG	*Pst*I	5′ end *gelQ*	Amplify o*rf189-orf276* flanking region (699 bp)
P25	AGCACCATCAGGAATAAGGC	None	5′ end *orf385*	Diagnostic for *orf189-orf276* deletion
P26	ATGAAGTCGGTGGAGTGGTA	None	3′ end *gelQ*	Diagnostic for *orf189-orf276* deletion
NH-P17	AC*TCTAGA*CGCTGCAGCGCATCATC	*Xba*I	*gelI*	Amplify *gelI* internal region (416 bp)
NH-P18	AC*GAGCTC*ATCTTGCCGGCGATTTC	*Sac*I	*gelI*	Amplify *gelI* internal region (416 bp)
RC-P17V	ATTCTCACCGGATTCAGTCGTCA	None	pLO2P	Diagnostic for vector integration (pair with NH-P17)
RC-P8	CCAA*GCATGC*ATGTCGACCTCGCCTC AATCCTC	*Sph*I	3′ end *gelE*	Amplify *gelM-gelN* flanking region (494 bp)
RC-P9	CCAA*TCTAGA*GCCCTCCGCGCTCCAGACCT	*Xba*I	5′ end *gelM*	Amplify *gelM-gelN* flanking region (494 bp)
RC-P10	TAGG*TCTAGA*GCCCGCGAGCTGGTCAAG	*Xba*I	3′ end *gelN*	Amplify *gelM-gelN* flanking region (512 bp)
RC-P11	GAA*GAGCTC*GAGGCCGAGGTGATG GTGGAGA	*Sac*I	3′ end *atrD*	Amplify *gelM-gelN* flanking region (512 bp)
RC-P12	CGGCACGCACGGAGCTTCAGTAA	None	3′ end *gelE*	Diagnostics; chromosomal deletion of *gelM-gelN*
RC-P13	AGCGCCAGCAGCCAGTCCGTTCGTA	None	3′ end *atrD*	Diagnostics; chromosomal deletion of *gelM-gelN*
RC-P6	CC*TCTAGA*TGGCGGAGGCACAGACC	*Xba*I	5′ end *urf302*	Amplify *urf302* internal region (455 bp)
RC-P7	TT*GAGCTC*ATCGGCACCACGGCGGAGAC	*Sac*I	3′ end *urf302*	Amplify *urf302* internal region (455 bp)
RC-P6VB	GCTCAGGCGCAATCACGAA	None	pLO2P	Diagnostic for vector integration (pair with RC-P6)
RC-P1	TT*GAGCTC*ACAACCGCGAGATGAT	*Sac*I	5′ end *gelF*	Amplify *gelF* internal region (503 bp)
RC-P2	GC*TCTAGA*TCGGCCTTGGTGTAGTT	*Xba*I	3′ end *gelF*	Amplify *gelF* internal region (503 bp)
P35	TTGGTTGGTCATGGCGACAT	None	*spsG*	Amplification of S-88 *spsG* for probe
P36	CGCCGTGACTACAACCAAAA	None	*spsG*	Amplification of S-88 *spsG* for probe
P37	GGGACTTCTATCACGTCTGA	None	*spsS*	Amplification of S-88 *spsS* for probe
P38	AAATGTCGCCATGACCAACC	None	*spsS*	Amplification of S-88 *spsS* for probe
P39	AACGAGGGTAGAGTACATGC	None	*spsR*	Amplification of S-88 *spsR* for probe
P40	GTCAGACGTGATAGAAGTCC	None	*spsR*	Amplification of S-88 *spsR* for probe
P41	TTCCTAATCCTTGTCGCCAA	None	*gelS*	Amplification of portion of *gelG* and *gelS* for probe
P42	GGGATCCATATGCAAAAAGCT	None	*gelG*	Amplification of portion of *gelG* and *gelS* for probe

spanning 21 kb, were colinear with the S-88 gene cluster from *spsQ* to *rmlD*. Based on comparison to the S-88 biosynthetic genes [35, 48] and Blast searches of GenBank, four genes (*gelB*, *gelK*, *gelL* and *gelQ*) appear to encode glycosyl transferases for the tetrasaccharide repeat unit of the backbone, four genes (*gelD*, *gelC*, *gelE* and *gelJ*) have DNA sequences homologous to genes for secretion of polysaccharides, two genes (*atrD*, *atrB*) are homologous to genes for ABC transporters of proteins, and four genes (*rmlA*, *rmlC*, *rmlB* and *rmlD*) are involved in the synthesis of the sugar-nucleotide precursor dTDP-L-rhamnose.

A comparison of the protein homologies of the gellan and S-88 sphingan gene products shows a broad range of sequence conservation. The four genes for dTDP-L-rhamnose synthesis are highly conserved (91–98% amino acid identity). Genes encoding the glycosyl transferases [35, 45] are less conserved, e.g., *gelB* and *spsB* encoding transferase I (86% identity); *gelK* and *spsK* encoding transferase II (87%); *gelL* and *spsL*, which likely encode transferase III (77%); and *gelQ* and *spsQ*, possibly encoding transferase IV (67% identity). Homology between *gelB* and *spsB* varies in different parts of the gene. The carboxyl portions (277 amino acids) are very homologous (93% identical), whereas the amino portions (193 amino acids) are more variable (74% identical). The carboxyl portion of GelB has strong homology to glycosyl transferases. As with the S-88 SpsB protein [48], fewer proteins have homology to the amino portion of GelB. Homology extending into the amino portion of the GelB protein was found in Blast search of GenBank for putative glycosyl transferases from *Mesorhizobium loti* and *Sinorhizobium meliloti* and hypothetical proteins from *Novosphingobium aromaticivorans*, *Azotobacter vinelandii* and *Rhodopseudomonas palustris*. Based on translated DNA sequences, proteins for polysaccharide secretion (GelJ, GelD, GelC, and GelE) had homologies relative to S88 proteins of 87%, 79%, 87%, and 88%, respectively. The AtrD and AtrB proteins were 92% and 97% identical to the S-88 proteins. The function of the genes with the least homology, *gelF* and *spsF* (61%), is unknown, and no other genes with significant homology were found in Blast searches of GenBank.

One major difference noted was the absence of genes at one end of the gellan gene cluster corresponding to S-88 genes *spsR*, *spsS* and *spsG*. DNA sequences 7.8 kb beyond *gelQ* and portions of DNA at a distance of up to 14 kb were examined and no homologs to these S-88 genes were found. Among six additional ORFs located adjacent to the gene cluster, at the position where *spsR*, *spsS* and *spsG* are found in the S-88 gene cluster, no homologies with known polysaccharide biosynthetic genes were found by Blast analysis. Two of these ORFs were predicted to encode DNA gyrase subunit B (*gyrB*) and ferrichrome receptor (*fhuA*) homologs (Fig. 2A).

Another difference in gene organization was noted. In the S-88 sphingan gene cluster, the gene *urf31* (an unidentified reading frame encoding a 31 kDa protein) is located between *rmlD* and *urf34* (an unidentified reading frame encoding a 34 kDa protein). No homolog of *urf31* was found in the gellan gene cluster, and the start codon for the adjacent gene designated *orf302* (homolog of *urf34*) corresponding to the start codon of *urf34* is missing, but translation likely starts at a GTG codon corresponding to a position 48 bp downstream from the start in the S-88 sphingan gene cluster, encoding a smaller protein. At the right end of the region sequenced is an unidentified ORF (*orf209*) with weak homology to methyl transferases and part of a gene with homology to *dpp* (encoding a dipeptidylpeptidase), suggesting that the end of the gellan biosynthetic gene cluster has been reached.

The sequenced DNA of region I has a high GC content (66.5%), as expected for *Sphingomonas*. Most of the genes have an ATG start codon, except *gelJ*, *gelC* and *orf302*, which utilize a GTG start codon, as do S-88 *spsJ* and *spsC*. Several genes have overlapping stop/start codons, *gelK-gelL* (ATGA), *gelC-gelE* (TGATG), *atrB-atrD* (ATGA), *rmlA-rmlC* (ATGA), and *rmlB-rmlD* (TGATG), indicating they are cistronically linked. The end of *gelM* overlaps the start of *gelN* by 17 bp. The 3′ ends of *orf302* and *orf206* overlap by 26 bp. On a DNA level, intercistronic regions were less conserved than protein coding regions; however, an 81 bp segment immediately upstream of *gelB* and *spsB* was highly conserved (80 identical residues), although the remainder of the 636 bp intercistronic region was variable. This conservation of sequence suggests that this region may have an important regulatory function.

Insertion inactivation of unidentified genes in region I

Putative functions for most of the genes in region I could be inferred based on homology to other proteins in the databases, or to functions determined by T. Pollock and coworkers [35, 48]. Several genes including three ORFs flanking the gene cluster (*orf189*, *orf276*, *orf302*) and four genes within the gene cluster (*gelI*, *gelF*, *gelM* and *gelN*) were of unknown function. To determine if they affect gellan synthesis or composition, particularly acyl substituents, these genes were inactivated by insertion or deletion mutation. Primers used for construction and confirmation of mutants are shown in Table 2.

orf189–orf276

Of particular interest were two genes adjacent to the gellan gene cluster at the point of divergence from the S-88 genes (Fig. 2A). Gene *orf189* encodes a protein with homology to N-acetyl transferases. Gene *orf276* encodes a protein with homology to numerous conserved hypothetical proteins. A deletion of most of these two genes was constructed. Primers P21 + P22

and P23 + P24 were used to amplify flanking DNA fragments by PCR, which were cloned into pLO2 and used to replace the genes in the chromosome. There were no apparent deleterious effects on growth or gellan production or change in the composition of gellan (Table 3). Thus, these genes are apparently not involved in gellan biosynthesis or modification with *O*-acyl substituents.

gelI

A function for *spsI* of S-88 was not specified by Yamazaki et al. [48]. PCR primers (NH-P17 and NH-P18) were used to amplify an internal fragment from gene *gelI*, which was cloned into plasmid pLO2 and used to inactivate the gene on the chromosome by insertion of the plasmid. The *gelI* mutant had altered colonial morphology (mucoid but runnier), less viscous broth, and 20% lower yield of total precipitable material (Table 3). The composition (glucose, rhamnose, acetyl and glyceryl) was comparable to the control. Thus, the *gelI* gene affects the amount of gellan polysaccharide, but not its composition. Blast analysis showed weak homology (16–22% identity) between the GelI protein and peptidyl-prolyl cis-trans isomerases from *Methylobacillus* sp. 12S, *Caulobacter crescentus*, *Clostridium acetobutylicum*, *Brucella melitensis*, *Brucella suis* and *Pseudomonas aeruginosa* and the *prsA* gene of *Bacillus subtilis*. Peptidyl-prolyl cis-trans isomerase is involved in protein folding and export [22].

gelM–gelN

Functions for *Sphingomonas* S-88 genes *urf32* and *urf26* (unknown reading frames) were not identified, and mutations in these genes did not block sphingan production [48]. A deletion of most of the corresponding *S. elodea* genes *gelM* and *gelN* was constructed by gene replacement, using PCR to amplify a fragment including the end of *gelE* and start of *gelM* (primers RC-P8 and RC-P9), and a fragment with the end of *gelN* and part of *atrD* (RC-P10 and RC-P11). Mutants with deletion of *gelM-gelN* exhibited slightly mucoid, soft colonies compared to the mucoid hard colonies of the wild-type. In shake flask fermentation, the Δ*gelM-gelN* broth was fluid and smooth compared to the more solid, viscous wild-type broth. For the mutant strain, total precipitable yields were 78% and broth viscosity was only 30% (3,000 vs. 10,000 cP) of that of wild-type. The composition of gellan from the mutant strains was comparable to wild-type gellan (Table 3). Thus, one or both of these genes affects the amount of gellan produced. These genes appear to be in an operon with other *gel* genes for gellan secretion, and are therefore designated *gel*. The proteins encoded by genes *gelM* and *gelN* were homologous to morphology proteins 1 and 2 of *Zoogloea ramigera* [25] (28% and 15% identity, respectively), the EpsH and EpsI proteins of *Methylobacillus* sp.12S [49] (33% and 22%, respectively) and to several hypothetical proteins.

Table 3 Yield and composition of gellan from mutant strains

Strain	Total precipitable material g/100 ml	Acetyl/ repeat	Glyceryl/ repeat
Wild-type (plasmid cured)	1.42 ± 0.13 ($n = 4$)	0.65	0.62
Δorf189–276	1.38 ± 0.02 ($n = 3$)	0.59	0.55
gelI$^-$	1.13 ± 0.03 ($n = 4$)	0.59	0.71
ΔgelM–gelN	1.11 ± 0.05 ($n = 4$)	0.58	0.76
orf302$^-$	1.42 ± 0.05 ($n = 3$)	0.58	0.80

orf302

Gene *orf302*, which is homologous to, but smaller than, gene *urf34* of S-88 [48], encodes a protein with homology to integral membrane proteins. A mutation in *orf302* was constructed by insertional inactivation, after cloning an internal fragment (PCR primers RC-P6 and RC-P7). The *orf302* mutants were mucoid on plates and had yields of total precipitable material and composition of gellan comparable to wild-type (Table 3). Thus, this gene does not appear to be involved in synthesis of gellan.

gelF

An attempt was made to insert pLO2 into *gelF* by amplification and cloning of an internal region of *gelF* using primers RC-P1 and RC-P2. While insertion could be made in a transferase I-deficient mutant (Gps2), no Km^R colonies were obtained in the wild-type strain. Mutation in *gelF* may be lethal if gellan synthesis has been initiated on the lipid carrier. A broad-host-range plasmid (pYP3.6XΔ*Bam*) containing the transferase I gene was mated into Gps2 and Gps2-*gelF*$^-$. Many mucoid transconjugants were obtained for Gps2 (confluent growth on selection plates), whereas only five colonies were obtained with Gps2-*gelF*$^-$, again suggesting that the mutation in *gelF* is lethal, unless the first step in formation of lipid-linked intermediates is blocked. Such lethality could be due to accumulation of tetrasaccharide repeat units on the lipid carrier that cannot be polymerized, which may impair cell wall synthesis [35]. Alternatively, this mutation in *gelF* may exert a polarity effect on downstream genes, since *gelF* is the first gene in what appears to be an operon of six genes, many of which are involved in polysaccharide secretion. Yamazaki et al. [48] report mutants in *spsF*, *spsD*, *spsC* and *spsE*; however, these mutants may have acquired spontaneous mutations in *spsB*. Many of the mutants described in their reports were double mutants also defective in transferase I [35, 48]. Due to this putative lethality of mutations in genes encoding downstream steps in the biosynthesis of gellan, most of the region I nonmucoid mutants obtained were in gene *gelB* encoding transferase I.

Identification of *gelR*, *gelS*, and *gelG* genes

To determine if genes corresponding to *spsR*, *spsS* and *spsG* are present in *S. elodea*, chromosomal DNA was digested with restriction enzymes (*Bam*HI, *Eco*RI, and *Hin*dIII) and hybridized with probes corresponding to the S-88 genes. DNA fragments were amplified from the S-88 *spsG*, *spsS*, and *spsR* genes using PCR (primer pairs P35/P36, P37/P38, P39/P40, Table 2) and digoxigenin-labeled dUTP. Positive signals on Southern hybridization were obtained for all three probes, confirming that homologous genes are present in *S. elodea*. Hybridizing fragments were 1.4 and 2.4 kb (*Bam*HI), 2.3 kb (*Eco*RI), and 0.9 and 2.2 kb (*Hin*dIII). To make a probe specific for *S. elodea*, genomic *Bam*HI fragments of about 1.4 kb were excised from a gel, ligated into plasmid pZERO and transformed into *E. coli* DH10B. Transformants were tested by hybridization of the *spsS* probe to isolated plasmids. One of 144 transformants had the correct fragment cloned, which was confirmed by DNA sequencing. Specific primers were designed (P41/P42) and the 1.4 kb fragment amplified from the cloned DNA by PCR incorporating digoxigenin-labeled dUTP. This probe was used to identify plasmids from the *S. elodea* cosmid gene library that carried these genes, by colony hybridization. Plasmids that hybridized to the probe were analyzed by restriction enzyme digestion and mapping. Fragments that corresponded to the sizes expected based on the Southern hybridization were sequenced, along with adjacent DNA (Fig. 2B, GenBank accession number AY220099).

The *gelR*, *gelS* and *gelG* genes appear to be in an operon in the same order as in the S-88 *sps* gene cluster, but not adjacent to the genes in the region I gene cluster. The sequence of the *gelR*, *gelS* and *gelG* genes showed that the proteins encoded are more divergent from the S-88 homologous proteins than are those encoded in region I. The GelR protein was somewhat smaller than its S-88 homolog (659 vs 670 amino acids) with 49% identity, and had homology to surface layer proteins and outer membrane proteins. The GelS protein was smaller than SpsS (426 vs 452 amino acids) with 56% identity and 26 fewer amino acids at the amino terminus. GelS had homology to membrane proteins involved in polysaccharide export, and had 22% identity to the protein encoded by an *E. coli wzx* gene (GenBank AAK64374.1), a putative O-antigen flippase, which may be involved in transport of lipopolysaccharide [27]. The GelG protein was larger than SpsG (553 vs 539 amino acids) with 44% identity, and was homologous to several hypothetical proteins. A portion of the protein (amino acids 400–462) was homologous to a consensus sequence for pfam04932 or Wzy, a putative O-antigen polymerase [5].

Genes *gelS* and *gelG* have GTG start codons. The GC content of the *gelR*, *gelS*, and *gelG* genes (58.3%) was distinctly less than that of the overall region cloned on these plasmids (67.4%) and that of region I (66.5%). Flanking the *gelR*, *gelS*, *gelG* genes are a cluster of genes encoding proteins for nitrogen assimilation, and genes encoding prolyl tRNA synthetase, and arginine-tRNA-protein transferase (Fig. 2B).

Regulatory gene affecting gellan synthesis

Of the 51 nonmucoid mutants isolated following EMS mutagenesis, about half (25) mapped to region I, in particular the *gelB* gene encoding transferase I. The other 26 mutants were restored to full mucoidy by region III plasmids. Restriction enzyme maps of these plasmids were compared and 4 kb found to be in common in the inserts. A 4.6 kb *Bam*HI fragment spanning this 4 kb DNA was cloned, sequenced and shown to contain a large ORF encoding a protein of 796 amino acids that was homologous to two component regulatory proteins. This protein had regions homologous to both sensor kinase and response regulator proteins, suggesting that this gene codes for one protein that has both motifs. Thus, this gene may encode a positive regulator for gellan synthesis and is designated *gelA*. Additional DNA adjacent to *gelA* was sequenced from pYP225. Flanking *gelA* are genes encoding an unknown protein, a homolog to valyl tRNA synthetase, and a conserved small protein (Fig. 2C, GenBank accession number AY220100).

Mutants that map to region III appeared nonmucoid on YM agar, but most displayed a clumping phenotype when grown in liquid medium. In contrast, nonmucoid mutants in region I (primarily transferase I-defective) grew uniformly in suspension. The clumping phenotype suggests that a low amount of polysaccharide was produced, which caused the cell association typical of *S. elodea*. The partial complementation observed with plasmids containing region I DNA may be due to increased production of polysaccharide with excess copies of genes in the biosynthetic pathway. Partial complementation of region III mutants was also observed with plasmids containing DNA from another region (II) that has not been fully characterized. Partial sequencing of region II indicates the presence of genes for a sensory transduction histidine kinase, AMP nucleosidase, *S*-adenosyl homocysteine hydrolase, iron receptor protein and three unidentified genes, but whether any of these genes has a role in gellan synthesis has not yet been elucidated. Unique plasmids with DNA from regions I, II and III were mated into mutants of *Xanthomonas campestris* defective in genes encoding UDP-glucose pyrophosphorylase and UDP-glucose dehydrogenase [12], enzymes essential for synthesis of the sugar nucleotides UDP-glucose and UDP-glucuronic acid, required for both gellan and xanthan synthesis. No complementation was observed, suggesting that these genes are not present on the plasmids. Plasmids containing *gelB* encoding transferase I (e.g., pYP3.6XΔ*Bam*, Fig. 2A) did complement a transferase I-deficient mutant of *X. campestris*, (*xps109*) [11], as expected since the first step in polysaccharide biosynthesis, formation of glucose-

isoprenylpyrophosphate, is in common in these organisms [13, 34, 35, 48].

Discussion

Several *Sphingomonas* strains produce polysaccharides with related structures [e.g., S-7, S-60 (gellan), S-88, S-130 (welan), S-194 (rhamsan), S-198, S-657 (diutan) and NW11] that have unique rheological properties [19, 24, 29, 32]. Recent taxonomic studies (unpublished data) indicate that the *Sphingomonas* strains producing polysaccharides S-88, S-198, and diutan (which all have rhamnose side chains) are closely related to each other and to the strain *Sphingomonas trueperi* based on 16S ribosomal RNA homology. *S. elodea* producing gellan is most closely related to *Sphingomonas pituitosa* (unpublished data). In this report, the genes for gellan biosynthesis were isolated and compared to the sequenced genes for synthesis of S-88 [48]. While many of the genes were similar, differences were noted in gene organization and homology between related proteins. The arrangement of genes is different with respect to the location of homologs of *spsR*, *spsS* and *spsG*, and the lack of a homolog to S-88 gene *urf31*. A gene with homology to S-88 *urf31* is present in the diutan biosynthetic gene cluster (R. J. Coleman, unpublished results). Since diutan, like the S-88 polysaccharide, has a rhamnose side chain [3], *urf31* may be involved in addition of this rhamnose substituent. The biosynthetic gene cluster for S-7, which has glucose rather than rhamnose in the side chain, also lacks a homolog to gene *urf31* [43].

There were variable degrees of homology between genes for synthesis of S-88 and gellan polysaccharides in the main gene cluster (61–98% amino acid identity, based on translated DNA). The most variable genes (61% identity) in the main gene cluster were *gelF* and *spsF*, which encode an unknown function essential for polysaccharide synthesis. Mutations in *gelF* were not obtained when a functional glycosyl transferase I was present, indicating possible lethality of this mutation, presumably due to accumulation of lipid-linked oligosaccharide intermediates [35], suggesting a role for GelF in a downstream step, e.g., polymerization or secretion. Glycosyl transferases III and IV were more variable (77% and 67% identity) than transferases I and II (86% and 87%). For both polysaccharides, transferase III links glucose to glucuronic acid. For transferase IV, the substrate specificity may be different. S-88 can have L-rhamnose or L-mannose in the fourth position, whereas gellan contains only L-rhamnose. For transferase I, the carboxyl portion was more conserved (93% identity) than the amino portion (74%), and had homology to other glycosyl transferases, suggesting that the carboxyl portion likely encodes the enzymatic activity. In *X. campestris*, the carboxyl portion of transferase I is sufficient for enzymatic activity [21]. Gene *gelK* encoding the glucuronosyl transferase is one gene in the gellan and S-88 biosynthetic clusters, but in many other organisms (e.g., *Rhizobium*, *Streptococcus* and *Lactococcus*), homologs to *gelK* are two genes, one encoding a transferase and the other a transferase enhancer protein [45]. GelK protein has been purified and the amino portion shown to be membrane-bound whereas the carboxyl protein was soluble [45]. Thus, the carboxyl portion of these enzymes may encode the enzymatic activity while the amino portion (or a separate protein) encodes a function that could be involved in anchoring the protein in the membrane or in a complex of proteins involved in polysaccharide synthesis. Two genes required for secretion, *gelC* and *gelE*, encode proteins that are encoded by one gene (*wzc*) in many Gram-negative bacteria, where GelC is homologous to the amino portion and GelE to the carboxyl portion. Wzc is a tyrosine autokinase involved in secretion of capsular polysaccharide in *E. coli* and *Klebsiella* [5, 36, 46, 47]. Gram-positive bacteria also encode this activity as two proteins corresponding to the amino and carboxyl domains of the Wzc protein from *E. coli* [47]. Since many of the proteins involved in polysaccharide biosynthesis likely form a complex within the cell membrane, the observed difference in amino acid sequences for homologous proteins may reflect the involvement of these proteins in protein-protein interactions or associations with other components of the cell membrane.

Several genes for which functions were not reported in previous studies [35, 48] have been analyzed by gene inactivation. A deletion of *gelM-gelN* and inactivation of *gelI* resulted in softer, less mucoid, colonies, lower amount of gellan, and lower broth viscosity. It has not been determined whether the decrease in viscosity is due entirely to production of less gellan or to production of gellan with lower molecular weight. GelM and GelN are homologous to morphology proteins 1 and 2 of *Z. ramigera* [25], which affect colony morphology, i.e., rugose or capsular versus smooth. As with the *gelM-gelN* deletion mutants, the variants of *Z. ramigera* produced polysaccharide with the same sugar and organic acid composition. GelM and GelN are also homologous to EpsH (a membrane protein) and EpsI (a putative export protein) involved in exopolysaccharide synthesis in *Methylobacillus* sp.12S [49]. Thus, *gelM* or *gelN* affects the amount of gellan and might affect the length or attachment of the polysaccharide molecules. The enzymatic function of the protein encoded by *gelI* is unknown. Weak homology (16–22%) was detected to peptidyl-prolyl cis-trans isomerase, which is involved in protein folding and export [22]. This homology and the phenotype of decreased gellan production suggest a possible involvement in processing of proteins involved in secretion of the polysaccharide. Inactivation of three genes of unknown function flanking the gellan biosynthetic gene cluster (*orf189*, *orf276* and *orf302*) indicated that these genes did not affect gellan synthesis, composition or O-acyl modification. Genes for addition of acetyl or glyceryl to gellan were not identified in the region I gene cluster.

Genes *gelR*, *gelS* and *gelG* were found in an operon unlinked to the main gene cluster, whereas in S-88 these genes are adjacent to other genes for polysaccharide

synthesis. Differences in the arrangement of the *R*, *S*, *G* genes were also observed for *Sphingomonas* strains ATCC 21423 that produces S-7 polysaccharide [43], and ATCC 53159 that produces diutan (unpublished results). Although a gene homologous to *spsG* is present in the S-7 biosynthetic gene cluster, genes corresponding to *spsS* and *spsR* are not in the cluster [43], but their presence elsewhere on the genome has not been reported. The diutan gene cluster has homologs to *spsG* and *spsR* but lacks a homolog to *spsS*. The *gelR*, *gelS* and *gelG* genes vary considerably in DNA sequence from their S-88 homologs (44–56% identity). Pollock and coworkers reported that inactivation of *spsG* or *spsS* blocked sphingan synthesis, indicating that these genes are essential for polysaccharide production [48]. GelG and GelS are hydrophobic proteins and may function in polymerization and export of gellan polysaccharide. GelS is homologous to the Wzx protein of *E. coli*, which is postulated to be a flippase involved in transport of lipid-linked repeat units across the cytoplasmic membrane [27]. A portion of the GelG protein was homologous to Pfam04932 or Wzy, a putative O-antigen polymerase [5]. The GelR protein had homology to surface proteins and outer membrane proteins. SpsR has been reported to have a secondary structure similar to that of PlyA and PlyB of *Rhizobium leguminosarum*, although there is low sequence similarity [8]. PlyA and PlyB are secreted, but cell-associated, proteins with polysaccharide degradation activity. Mutations of *plyB* result in formation of longer chains of polysaccharide resulting in increased culture viscosity. These proteins are secreted by type I secretion systems, similar to those encoded by *atrB* and *atrD* [8].

S. elodea and other *Sphingomonas* strains [33], exhibit sessile-motile dimorphism. The motile phenotype could be due to hydrolysis of the polysaccharide or release of the polysaccharide from its attachment on the cell surface. GelR and SpsR may be polysaccharide lyases involved in hydrolysis of polysaccharide to liberate the cells from the capsule under conditions of nutrient limitation, promoting cell motility.

The gene *gelA*, encoding a two-component regulatory protein, that was not linked to either the main gellan biosynthetic gene cluster or to the *gelR*, *gelS*, *gelG* operon was also identified. This gene encodes a protein of 796 amino acids that has regions homologous to both sensor kinase and response regulatory proteins, and complemented mutants that were nonmucoid on agar, but produced a low amount of polysaccharide as indicated by their clumping phenotype in liquid culture. This gene may encode a regulatory protein that senses environmental conditions, e.g., nutrient levels, and turns on gellan production under conditions of abundant carbon source. Improved understanding of the genetics, biochemistry and regulation of polysaccharide synthesis in the Sphingomonads may lead to techniques for production of modified polysaccharides with improved functional characteristics.

Acknowledgements We thank Joseph M. Cleary for many useful discussions and his careful review of this manuscript. We appreciate the technical assistance of Alan Kanehl, John McQuown and Brian Mueller, and technical discussions with Tom Ramseier and Stan Bower. Sequencing of DNA was accomplished by the Microchemical Core Facility at San Diego State University; Luis Ielpi's laboratory at the Fundación Instituto Leloir, Buenos Aires, Argentina; Monsanto Company, St. Louis, Mo.; and at Lark Technologies, Houston, Tex.

References

1. Baird JK, Talashek TA, Chang H (1992) Gellan gum: effect of composition on gel properties. In: Phillips GO, Williams PA, Wedlock DJ (eds) Gums and stabilizers for the food industry, vol 6. Oxford University Press, New York, pp 479–487
2. Chandrasekaran R, Radha A, Thailambal VG (1992) Roles of potassium ions, acetyl and L-glyceryl groups in native gellan double helix: an X-ray study. Carbohydr Res 224:1–17
3. Chowdhury TA, Lindberg B, Lindquist U, Baird J (1987) Structural studies of an extracellular polysaccharide, S-657, elaborated by *Xanthomonas* ATCC 53159. Carbohydr Res 164:117–122
4. Ditta G, Stanfield S, Corbin D, Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc Natl Acad Sci USA 77:7347–7351
5. Drummelsmith J, Whitfield C (1999) Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). Mol Microbiol 31:1321–1332
6. Falk C, Jansson P-E, Rinaudo M, Heyraud A, Widmalm G, Hebbar P (1996) Structural studies of the exocellular polysaccharide from *Sphingomonas paucimobilis* strain I-886. Carbohydr Res 285:69–79
7. Fialho AM, Monteiro GA, Sa-Correia I (1991) Conjugal transfer of recombinant plasmids into gellan gum-producing and nonproducing variants of *Pseudomonas elodea* ATCC 31461. Lett Appl Microbiol 12:85–87
8. Finnie C, Zorreguieta A, Hartley NM, Downie JA (1998) Characterization of *Rhizobium leguminosarum* exopolysaccharide glycanases that are secreted via a type I exporter and have a novel heptapepide repeat motif. J Bacteriol 180:1691–1699
9. Gay P, Le Coq D, Steinmetz M, Berkelman T, Kado CI (1985) Positive selection procedure for entrapment of insertion sequence elements in Gram-negative bacteria. J Bacteriol 164:918–921
10. Guerry P, Van Embden J, Falkow S (1974) Molecular nature of two nonconjugative plasmids carrying drug resistance genes. J Bacteriol 117:619–630
11. Harding NE, Cleary JM, Cabanas DK, Rosen IG, Kang KS (1987) Genetic and physical analysis of a cluster of genes essential for xanthan gum biosynthesis in *Xanthomonas campestris*. J Bacteriol 169:2854–2861
12. Harding NE, Raffo S, Raimondi A, Cleary JM, Ielpi L (1992) Identification, genetic and biochemical analysis of genes involved in synthesis of sugar nucleotide precursors of xanthan gum. J Gen Microbiol 139:447–457
13. Ielpi L, Couso RO, Dankert MA (1993) Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. J Bacteriol 175:2490–2500
14. Jansson PE, Lindberg B, Sandford PA (1983) Structural studies of gellan gum, an extracellular polysaccharide elaborated by *Pseudomonas elodea*. Carbohydr Res 124:135–139
15. Jansson PE, Kumar NS, Lindberg B (1986) Structural studies of a polysaccharide (S-88) elaborated by *Pseudomonas* ATCC 31554. Carbohydr Res 156:165–172

16. Jay AJ, Colquhoun IJ, Ridout MJ, Brownsey GJ, Morris VJ, Fialho AM, Leitao JH, Sa-Correia I (1998) Analysis of structure and function of gellans with different substitution patterns. Carbohydr Polym 35:179–188
17. Kamoun S, Tola E, Kamdar H, Kado CI (1992) Rapid generation of directed and unmarked deletions in *Xanthomonas*. Mol Microbiol 6:809–816
18. Kang KS, Veeder GT (1982) Polysaccharide S-60 and bacterial fermentation process for its preparation. US Patent 4 326 053
19. Kang KS, Pettitt DJ (1993) Xanthan, gellan, welan, and rhamsan. In: Whistler RL, BeMiller JN (eds) Industrial gums: polysaccharides and their derivatives, 3rd edn. Academic Press, New York, pp 341–398
20. Kang KS, Veeder GT, Mirrasoul PJ, Kaneko T, Cottrell IW (1982) Agar-like polysaccharide produced by a *Pseudomonas* species: production and basic properties. Appl Environ Microbiol 43:1086–1091
21. Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Puhler A, Ielpi L (1998) *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol 180:1607–1617
22. Krontinen VP, Saris P, Sarvas M (1991) A gene (*prsA*) of *Bacillus subtilis* involved in a novel, late stage of protein export. Mol Microbiol 5:1273–1283
23. Kuo MS, Mort AJ, Dell A (1986) Identification and location of L-glycerate, an unusual substituent in gellan gum. Carbohydr Res 56:173–187
24. Lee EJ, Chandrasekaran R (1991) X-ray and computer modeling studies on gellan-related polymers: molecular structures of welan, S-657, and rhamsan. Carbohydr Res 214:11–24
25. Lee S-P, Kim T-R, Sinskey AJ (2000) Cloning and sequencing of genes involved in the morphological change of *Z. ramigera*. J Microbiol Biotechnol 10:161–168
26. Lenz O, Schwartz E, Dernedde J, Eitinger M, Friedrich B (1994) The *Alcaligenes eutrophus* H16 *hoxX* gene participates in hydrogenase regulation. J Bacteriol 176:4385–4393
27. Liu D, Cole RA, Reeves PR (1996) An O antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J Bacteriol 178:2102–2107
28. Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. Moorhouse R (1987) Structure/property relationship of a family of microbial polysaccharides. In: Yalpani M (ed) Industrial polysaccharides: genetic engineering, structure/property relations and applications. Elsevier, Amsterdam, pp 187–206
30. O'Neill MA, Selvendran RR, Morris VJ (1983) Structure of the acidic extracellular gelling polysaccharide produced by *Pseudomonas elodea*. Carboydr Res 124:123–133
31. O'Neill MA, Darvill AG, Albersheim P, Chou KJ (1990) Structural analysis of an acidic polysaccharide secreted by *Xanthobacter* sp. (ATCC 53272). Carbohydr Res 206:289–296
32. Pollock TJ (1993) Gellan-related polysaccharides and the genus *Sphingomonas*. J Gen Microbiol 139:1939–1945
33. Pollock TJ, Armentrout RW (1999) Planktonic/sessile dimorphism of polysaccharide-encapsulated Sphingomonads. J Ind Microbiol Biotechnol 23:436–441
34. Pollock TJ, Thorne L, Yamazaki M, Mikolajczak M, Armentrout RW (1994) Mechanism of bacitracin resistance in Gram-negative bacteria that synthesize exopolysaccharide. J Bacteriol 176:6229–6237
35. Pollock TJ, van Workum WAT, Thorne L, Mikolajczak M, Yamazaki M, Kijne JW, Armentrout RW (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
36. Rahm A, Drummelsmith J, Whitfield C (1999) Conserved organization in the *cps* gene clusters for expression of *Escherichia coli* group 1 K antigens: relationship to the colanic acid biosynthesis locus and the *cps* genes from *Klebsiella pneumoniae*. J Bacteriol 7:2307–2313
37. Sa-Correia I, Fialho AM, Videira P, Moreira LM, Marques AR, Albano H (2002) Gellan gum biosynthesis in *Sphingomonas paucimobilis* ATCC 31461: genes, enzymes and exopolysaccharide production engineering. J Ind Microbiol Biotechnol 29:170–176
38. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
39. Skaggs B, Rakitsky W, Phyfferoen A (2000) Method for improved rheological control in cementitious systems. US Patent 6 110 271
40. Stone KJ, Strominger JL (1971) Mechanism of action of bacitracin: complexation with metal ion and C55-isoprenyl pyrophosphate. Proc Natl Acad Sci USA 68:3223–3227
41. Staskawicz B, Dahlbeck D, Keen N, Napoli C (1987) Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *Glycinea*. J Bacteriol 169:5789–5794
42. Takeuchi M, Hamana K, Hiraishi A (2001) Proposal of the genus *Sphingomonas sensu stricto* and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analysis. Int J Syst Evol Microbiol 51:1405–1417
43. Thorne L, Mikolajczak MJ, Armentrout RW, Pollock TJ (2000) Increasing the yield and viscosity of exopolysaccharides secreted by *Sphingomonas* by augmentation of chromosomal genes with multiple copies of cloned biosynthetic genes. J Ind Microbiol Biotechnol 25:49–57
44. Vartak NB, Lin CC, Cleary JM, Fagan MJ, Saier MH Jr (1995) Glucose metabolism in *Sphingomonas elodea*: pathway engineering via construction of a glucose-6-phosphate dehydrogenase insertion mutant. Microbiology 141:2339–2350
45. Videira P, Fialho A, Geremia RA, Breton C, Sa-Correia I (2001) Biochemical characterization of the β-1,4-glucuronosyltransferase GelK in the gellan gum-producing strain *Sphingomonas paucimobilis* ATCC 31461. Biochem J 258:457–464
46. Vincent C, Doublet P, Grangeasse C, Vaganay E, Cozzone AJ, Duclos B (1999) Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. J Bacteriol 181:3472–3477
47. Wugeditsch T, Paiment A, Hocking J, Drummelsmith J, Forrester C, Whitfield C (2001) Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in *Escherichia coli*. J Biol Chem 276:2361–2371
48. Yamazaki M, Thorne L, Mikolajczak M, Armentrout RW, Pollock TJ (1996) Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. J Bacteriol 178:2676–2687
49. Yoshida T, Ayabe Y, Yasunaga M, Usami Y, Habe H, Nojiri H, Omori T (2003) Genes involved in the synthesis of the exopolysaccharide methanolan by the obligate methylotroph *Methylobacillus* sp strain 12S. Microbiology 149:431–444