Cloning and Characterization of Four Genes of *Rhizobium leguminosarum* bv. *trifolii* Involved in Exopolysaccharide Production and Nodulation

Wilbert A. T. van Workum, Hayo C. J. Canter Cremers, André H. M. Wijffjes, Christa van der Kolk, Carel A. Wijffelman, and Jan W. Kijne

Institute of Molecular Plant Sciences, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands

Received 23 August 1996. Accepted 6 December 1996.

Four different genes of *Rhizobium leguminosarum* bv. *trifolii* strain RBL5599 involved in exopolysaccharide (EPS) production were identified by complementation of Tn5-induced EPS-deficient mutants (Exo mutants) with a cosmid bank. On one cosmid *pssA* was located, which was found to be almost identical to the *pss4* gene from *R. leguminosarum* bv. *viciae* VF39 and highly homologous to a family of glycosyl transferases. Two *pssA* mutants, *exo2* and *exo4*, were characterized and found to produce 19 and 1% of the wild-type amount of EPS, respectively. The three other genes were found to be closely linked on a different complementing cosmid. *pssC* revealed similarity to *exoM* and *exoW* of *R. meliloti*, both encoding glucosyl transferases involved in the synthesis of succinoglycan. A mutation in this gene (mutant *exo50*) did reduce EPS synthesis to 27% of the wild-type amount. We found an operon closely linked to *pssC*, consisting of two overlapping genes, *pssD* and *pssE*, that is essential for EPS production. Homology of *pssD* and *pssE* was found with *cps14F* and *cps14G* of *Streptococcus pneumoniae*, respectively: two genes responsible for the second step in capsule polysaccharide synthesis. Furthermore, *pssD* and *pssE* were homologous to the 5' and 3' parts, respectively, of *spsK* of *Sphingomonas* S88, which encodes a putative glycosyl transferase. Structural analysis of EPS produced by Exo mutants *exo2*, *exo4*, and *exo50* showed it to be identical to that of the parental strain RBL5599, with the exception of acetyl groups esterified to one of the glucose residues being absent.

Under certain culture conditions, *Rhizobium leguminosarum* bacteria produce large amounts of acidic exopolysaccharide (EPS), which are essential for a successful symbiosis with their respective host plants. EPS-deficient mutants (Exo mu-

nants) of *R. leguminosarum* bv. *viciae* and *R. leguminosarum* bv. *trifolii* induce delayed formation of only partially infected nodules in their host plants (Skorupska et al. 1995; van Workum et al. 1995). In contrast, nonpleiotropic Exo mutants of *R. leguminosarum* bv. *phaseoli* induce the timely formation of normal nitrogen-fixing nodules in plants of their cross-inoculation group (Borthakur et al. 1986; Diebold and Noel 1989). This difference in EPS requirement might be due to the fact that *R. leguminosarum* bv. *phaseoli* nodulates host plants forming determinate nodules, whereas *R. leguminosarum* bv. *trifolii* and *R. leguminosarum* bv. *viciae* nodulate host plants forming indeterminate nodules (Diebold and Noel 1989). In general, EPS synthesized by *R. leguminosarum* is a polymer of a conserved octasaccharide repeating unit (Robertson et al. 1981), substituted with *O*-acetyl, pyruvyl, and hydroxybutanoyl groups (Kuo and Mort 1986; Philip-Hollingsworth et al. 1989a, 1989b) (see Figure 1). The distribution of *O*-acetyl and hydroxybutanoyl groups was found to be different for several *R. leguminosarum* strains and no correlation between substitution pattern and the ability to nodulate certain host plants was found (Canter Cremers et al. 1991; Orgambide et al. 1992). Evidence for a role of low molecular weight EPS as a host-specific signaling molecule was obtained by Djordjevic et al. (1987), who found that purified EPS fractions of *R. leguminosarum* bv. *trifolii* could restore nodulation of Exo mutants of this strain, whereas purified EPS fractions of *R. meliloti* could not. This might be due to the fact that the EPS of *R. meliloti*, called succinoglycan, is composed of a different octasaccharide subunit (Fig. 1B).

In *R. meliloti*, the genes involved in biosynthesis of succinoglycan, with the exception of *exoC*, are clustered on a megaplasmid. Recently, most of these genes have been characterized genetically (Becker et al. 1993a, 1993b, 1993c; Buendia et al. 1991; Glucksmann et al. 1993a, 1993b; Müller et al. 1993) and biochemical functions were assigned (Canter Cremers et al. 1990; Leigh et al. 1987; Reuber and Walker 1993a, 1993b). For *R. leguminosarum* strains, biosynthesis of EPS is less well understood. Up to now, one gene has been characterized in both *R. leguminosarum* bv. *viciae* (*pss4*) and *R. leguminosarum* bv. *phaseoli* (*pssA*, previously called *pss2*), which was found to be homologous to *exoY* of *R. meliloti*, and probably encodes the first glycosyl transferase in EPS biosynthesis (Borthakur et al. 1986, 1988; Ivashina et al. 1994).

Corresponding author: Wilbert A. T. van Workum, Institute of Molecular Plant Sciences, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands; Tel. international +31-71-5275082; Fax. international +31-71-5275088; E-mail: workum@rulsfb.LeidenUniv.nl

Nucleotide and/or amino acid sequence data reported in this paper are submitted to the EMBL, GenBank, and DDBJ data bases as accession numbers Y07549 (*pssA*) and X99850 (*pssC*, *pssD*, and *pssE*).

Furthermore, two genes regulating EPS biosynthesis, *psi* and *psr*, were found to be located on the Sym plasmid of *R. leguminosarum* bv. *phaseoli*. *psi* inhibits EPS biosynthesis when present in high copy number, whereas *psr* overcomes the effect of a multicopy *psi* gene (Borthakur et al. 1985; Borthakur and Johnston 1987).

Previously, we described the isolation of several Exo mutants of *R. leguminosarum* obtained by transposon Tn5 mutagenesis (Canter Cremers et al. 1988) as part of our investigation of the relationship between EPS synthesis and symbiotic behavior. We found mutants differing with respect to polysaccharide synthesis and nodulation efficiency (Breedveld et al. 1993; van Workum et al. 1995). In this paper we describe the molecular characterization of four of these mutants.

RESULTS

Isolation and properties of Exo mutants.

When grown on plates, colonies of most Exo mutants of *R. leguminosarum* bv. *trifolii* are nonmucoid, but some of our mutants seemed to have an intermediate phenotype. To quantify EPS production, mutant strains were grown in B⁻ minimal medium and EPS was isolated. It appeared that parental strain RBL5599 produced about 2,400 mg of EPS per liter of culture. Mutants *exo2*, *exo4*, and *exo50* produced 450 mg/liter, 30 mg/liter, and 650 mg/liter, respectively, whereas mutant *exo111* did not produce EPS at all (Table 1). The absence or presence of a *R. leguminosarum* bv. *viciae* or *R. leguminosarum* bv. *trifolii* Sym plasmid did not influence the amount of EPS secreted (data not shown).

NMR spectroscopy was used to analyze the chemical structure of the EPS isolated from mutants *exo2*, *exo4*, and *exo50*. To facilitate spectroscopy, EPS was depolymerized by phage RL38 and purified by column chromatography as described by Djordjevic et al. (1986). Compared with the depolymerized EPS of strain RBL5599, the oligosaccharide re-

peat units isolated from the three Exo mutants yielded the same elution profile with the Sephadex DEAE A25 column used. After further purification, the structure of the depolymerized EPS from the Exo mutant strains was determined by ¹H and ¹³C NMR spectroscopy. Each peak in the ¹³C spectrum of strain RBL5599 (Fig. 2A) was assigned as described previously (Canter Cremers et al. 1991). Apart from signals assigned to *O*-acetyl substitutions, the ¹³C spectra of the depolymerized EPS of mutant strains *exo2*, *exo4*, and *exo50* are indistinguishable from that of parental strain RBL5599 (Fig. 2A,B). Therefore, the basic chemical structure of EPS synthesized by these strains is identical to that of strain RBL5599. Peaks at 70.6, 67.3, 104.7, 72.6, 76.5, 81.1, 78.1, 175.9, and 77.6 ppm in the ¹³C NMR spectrum of RBL5599 were assigned (Canter Cremers et al. 1991) to resonances influenced by acetyl groups attached to the 3-position of residue b (Fig. 1). Peaks of the ¹³C NMR spectrum at 73.3, 73.6, 79.1, and 100.8 ppm

Table 1. Polysaccharide production by Exo mutants of *Rhizobium leguminosarum* bv. *trifolii* strain RBL5599^a

Strain	EPS (mg/liter) ^b	CPS (mg per g of cells) ^c	LPS profile ^d
RBL5599	2,400	93	+
<i>exo2</i> (<i>pssA2::Tn5</i>)	450	105	+
<i>exo4</i> (<i>pssA4::Tn5</i>)	30	98	+
<i>exo50</i> (<i>pssC50::Tn5</i>)	650	101	+
<i>exo111</i> (<i>pssD111::Tn5</i>)	—	112	+
<i>exoB</i> (<i>exoB8::Tn5</i>)	320	11	—

^a Bacteria were grown for 10 days at 28°C. In the table the average values of three experiments are shown. In these three experiments, the values obtained differed by at most 10%.

^b Exopolysaccharide (EPS) concentrations are expressed as milligrams of glucose equivalents per liter of culture.

^c Capsular polysaccharide (CPS) concentrations are expressed as milligrams of extracted material per liter of culture.

^d Lipopolysaccharide (LPS) profile as analyzed by polyacrylamide gel electrophoresis. +, wild-type LPS profile; —, LPS profile lacking the high molecular weight LPS-I.

Fig. 1. A, Chemical structure of repeating unit of exopolysaccharide (EPS) of *Rhizobium leguminosarum* bv. *trifolii* strain RBL5599. Nonstoichiometric substituents, i.e., acetate groups at the 2- and 3-position of residue b and d, and 3-hydroxybutanoate at the 2- and 3-position of residue h, have not been indicated. Fragments obtained by action of phage depolymerase have 4-deoxy- α -L-threo-hex-4-enepyransyluronic acid in place of glucuronic acid residue c. B, chemical structure of succinoglycan (EPS I) of *R. meliloti* Rm1021 (Aman et al. 1981). Abbreviations: Glc, glucose; GlcA, glucuronic acid; and Gal, galactose.

were assigned to resonances influenced by the presence of acetate at the 2-position of residue b (Canter Cremers et al. 1991). These peaks are present in the ^{13}C spectra of the three Exo mutants, indicating that in the EPS of these mutants *O*-acetyl groups are present on this glucuronic acid residue. Small peaks in the ^{13}C spectrum of parental strain RBL5599 (Fig. 2A) at 90.6 and 95.3 ppm showed the presence of acetate at the 2-position of residue d (Fig. 1), while peaks at 93.0 and 96.9 ppm indicated the presence of acetate at the 3-position of the same residue. These peaks were absent in the ^{13}C spectra of the three mutant strains (Fig. 2B). Apparently, EPS of the three mutant strains lacks acetyl esterifications at residue d. In conclusion, the basic structure of the EPS of the three Exo mutants is similar to that of the parental strain RBL5599, but lacks acetyl groups esterified to residue d.

To determine whether the Exo mutants show a pleiotropic phenotype with regard to other cell surface polysaccharides, we tested the Exo mutants for alterations in lipopolysaccharide (LPS) profiles, and for the production of the gelling capsular polysaccharide (CPS) described by Zevenhuizen and van Neerven (1983). Parental strain RBL5599 and mutant *exoB52*

(Canter Cremers et al. 1990) that is defective in EPS, CPS, and LPS synthesis were used as control strains. As shown in Table 1, the four Exo mutants are not affected in both CPS production and LPS profile.

Symbiotic properties of Exo mutants.

R. leguminosarum bv. *trifolii* mutant strains were tested for their nodulation phenotype on Dutch white clover (*Trifolium repens* L.) as described previously (van Workum et al. 1995). Mutant *exo50* nodulated clover with a delay of 3 to 4 days, compared with parental strain RBL5599. The nodules were infected and were fixing nitrogen. Mutant *exo111* induced non-nitrogen-fixing nodules that appeared about a week later than did nodules induced by RBL5599. The nodules were hardly infected by rhizobia, which is characteristic for nodules induced by Exo mutants of *R. leguminosarum* (Skorupska et al. 1995; van Workum et al. 1995). The nodulation characteristics of *exo2* and *exo4* were described previously (van Workum et al. 1995). In short, *exo2* nodulates as efficiently as parental strain RBL5599, whereas *exo4* induces formation of nitrogen-fixing nodules with a delay of about 3 to 4 days.

Fig. 2. ^{13}C NMR spectra from depolymerized exopolysaccharide (EPS) isolated from *Rhizobium leguminosarum* bv. *trifolii* strain RBL5599 (A) and *exo2* (B). Spectra of *exo4* and *exo50* were indistinguishable from that of *exo2*. 1: C4 of 4-deoxy- α -L-threo-hex-4-enepyransyluronic acid (residue c). 2: C1 of glucuronic acid residue b; 2': C1 of 2-*O*-acetyl glucuronic acid residue b. 3: C1 of the carboxyethylidene residue of 4,6-*O*-(1-carboxyethylidene)glucose (residue g). 4: C1 of the carboxyethylidene residue of 4,6-*O*-(1-carboxyethylidene)galactose (residue h). 5: C1 of 4-deoxy- α -L-threo-hex-4-enepyransyluronic acid residue c. 6: C1 of 3-*O*-acetyl- β -glucose residue d; 6': C1 of 2-*O*-acetyl β -glucose residue d; 6'': c1 of β -glucose residue d. 7: C1 of 3-*O*-acetyl- α -glucose residue d; 7': C1 of 2-*O*-acetyl α -glucose residue d; 7'': c1 of α -glucose residue d. 8: C4 of glucuronic acid residue b; 8': c4 of 3-*O*-acetyl glucuronic acid residue b. 9: C3 of 4,6-*O*-(1-carboxyethylidene)glucose residue g. 10: C4 of glucose residue a; 10': c4 of glucose residue a adjacent to 3-*O*-acetyl glucuronic acid residue b. 11: C5 of glucuronic acid residue b; 11': C5 of 3-*O*-acetyl glucuronic acid residue b. 12: C3 of glucuronic acid residue b. 13: Methyl of 4,6-*O*-(1-carboxyethylidene)glucose residue g. 14: Methyl of 4,6-*O*-(1-carboxyethylidene)galactose residue h. 15: Hydroxybutanoate methyl. 16: Acetate methyls. g: C6 of glucose residues a, e, and f.

Complementation of *R. leguminosarum* bv. *trifolii* Exo mutants.

To isolate DNA fragments that complemented the Exo mutants, a pLAFR1 cosmid library of *R. leguminosarum* strain RBL5515 was introduced into the mutants, followed by screening for mucoid colonies on YMB plates. The cosmids of single mucoid colonies were transferred to *Escherichia coli* DH5 α by triparental mating, after which they could be used for complementation of each Exo mutant. This approach resulted in the identification of four groups of cosmids, each complementing different Exo mutants: (i) a group complementing *exoB* (Canter Cremers et al. 1990); (ii) a group complementing *exo5* and *exo370* (van Workum et al. 1995); (iii) a group complementing *exo2* and *exo4*; (iv) and a group complementing *exo50* and *exo111*. Because mutants *exo2*, *exo4*, and *exo50* produced residual amounts of EPS, we decided to further subclone the corresponding genes to search for regulators of EPS synthesis. Below, we describe cloning and characterization of EPS biosynthetic genes from cosmid groups iii and iv.

pssA.

Further subcloning of DNA of a cosmid of group iii in pMP92 yielded pMP2604, harboring a 1.8-kb *Hind*III insert. After introduction of this plasmid in mutant strains *exo2* and *exo4*, the colonies of the resulting strains acquired a mucoid morphology comparable to that of parental strain RBL5599. Furthermore, plasmid pMP2604 was able to complement the nodulation defect of mutant strain *exo4* on clover plants.

To localize the relevant gene(s), we mutagenized plasmid pMP2604 with transposon Tn5. Nine derivatives of plasmid pMP2604 with Tn5 inserted at various sites in the 1.8-kb fragment were obtained, after which the location of Tn5 was determined by Southern blotting (Fig. 3). Four of these derivatives, namely, pMP2604, Ω 1::Tn5, Ω 2::Tn5, Ω 3::Tn5, and Ω 4::Tn5, complemented both mutant strains for the synthesis of EPS and in the case of mutant *exo4* also for the nodulation deficiency. Mutant strain *exo4* harboring plasmids pMP2604, Ω 5::Tn5, or Ω 6::Tn5 was partially complemented for the synthesis of EPS, producing about 10% of the amount of EPS synthesized by parental strain RBL5599. However, the nodulation behavior of mutant *exo4* harboring either plasmid pMP2604, Ω 5::Tn5 or Ω 6::Tn5 was comparable to that of pa-

Fig. 3. Site of insertion of transposon Tn5 in the 1.8-kb *Hind*III DNA insert present in plasmid pMP2604. H, *Hind*III; E, *Eco*RV; Sc, *Sac*I; Sp, *Sph*I; Pv, *Pvu*II.

rental strain RBL5599. Each of the remaining derivatives of plasmid pMP2604, namely, Ω 7::Tn5, Ω 8::Tn5, and Ω 9::Tn5, failed to complement both mutant strains for the synthesis of EPS. Mutant *exo4* harboring either one of these plasmids also showed its normal nodulation deficiency. The gene(s) in which these mutants are affected are thus present in the *Hind*III-*Eco*RV fragment, 1.2 kb in size, present in plasmid pMP2642.

Both strands of this 1.2-kb DNA fragment were sequenced (Fig. 4). A search, using the program FRAME (Bibb et al. 1984) revealed an open reading frame (ORF) that reads toward the *Hind*III site (Fig. 3) and has a high probability of encoding a polypeptide. In the deduced amino acid sequence we could indicate two possible GTG start codons: one at position

```

1
GATATCTTTCATCAGAAAAATTTTCGCTCTGCTGCGTTAAGAAATACGCTTAAGAGCCAT
61
GGGCCAAGTGTGGAACGCGACGTCGCCAGACCCCGCGCAACCCCTCCAAAGGCTAGAA
121
ACTTTGACAAAATTTCAATTATTAACGAAATATCATGAGCTTGCCTCAACTTAACGCAGA
181
TGCGAAAGATTTTGCTGCGATGCGATATTTCTCTGACACCATTTACACAATTTGTCGCAATT
      Ω5      exo2
241
TTCCGTTTTAGTAGGTTACCAAGACCTAAGTGCAGCCCTGCTGACAGGGTTAACCATG
      V T G L T I
      Ω6
301
ATCGCCTATTGCGCGCGCCCTCGAAGACAACAGATACGATCCTTGAGCATCACGG
D R L L P P R P L E D N R V R S L S I T
361
AATTAACAACAGCATTTCGACAGAGTCCTTCGCGCGAGCCGCCAGCAGCCGAGCC
E L N N S I S T E S F R P S R R Q Q P S
421
TGAAGATCCAGACCCCTGTTATCCACAGCGATGCGGCTCTGGCGCGCTTGGTGGATCTGG
L K I Q T P V I H S D A A L A P L V D L
481
TCCTGAAGCGGGCGTTCGACATTGTTTCGCTGTTGAGCGCCCTCTCTCTCGCCCTT
V L K R A F D I V S S L S A L L V L A P
541
TCCTTCTGTTCGCTCGCCCTGCTGATCAAGATCGACAGCCAGGGCGGCTGCTCTTCAAGC
F L L F V A L L I K I D S P G P V L F K
601
AGACCCGCTGGGGCAAGAACTGCAAGGCCATCAAGGTCTACAAGTTCCGTTCCATGCGCA
Q T R W G K N C K A I K V Y K F R S M R
      exo4
661
CCGACCTCTGCGACGCTCTCGGCGGTTCGCCAGACGGTGAACGATCCGCGCGTCACTC
T D L C D V S G V A Q T V M N D P R V T
721
GCGTCGTTGCCATTCTGCGCGCGACGAATGTGATGAGTTGCGCGAGCTGCTGAACGTAC
R V G A I L R R T N V D N F S I L M D M
781
TGCTGGGCGACATGTGCGTCTGCGCCCGCGCTGCCATGCGATCGGCGATGCGCGCGCGG
L L G H M S V V G P R C H A I G M R A G
841
GCGTGCTCTACGAAGAGCTCGTACCGGAATACCATCAGCGCCATGCGATGCGCCCGGCA
G V L Y E E L V P E Y H Q R H A M R P G
901
TGACGGGTCTTGTCTCAGATGCGCGGTCTGCGCGGTCCGACCGATGCTCCGCCAAGGCC
M T G L A Q M R G L R G P T D R P A K A
961
GCGCCCGCATTGCCAGCGATCTCTACTACGTGCGGAATTTTTCGATCTTGATGGACATGC
R A R I A S D L Y Y V G N F S I L M D M
1021
GCATCATTTTCGGAACCGTCTGTCGAGCTGACCCGCGGAAAGGCTTCTAAAGCGCGT
R I I F G T V V S E L T R G K G F *
1081
CGCGATCTTTCAGATTCGCTCCTCGCGCTTTTGTCTTTGTTTTGCGCATGTCGCACAAA
1141
AAGCGATGCATACTTTTTCGCGACATGCCTCCAAAGCTT

```

Fig. 4. DNA sequence of 1.2-kb *Eco*RV-*Hind*III fragment present in plasmid pMP2642. *Eco*RV site is found at beginning of sequence (base pair [bp] 1). Predicted amino acid (aa) sequence of *pssA* gene indicated by single letter code. Start of open reading frame is at bp 473, end is at bp 1073. Predicted protein is 22.6 kDa in size. Extension of 63 aa at N terminus of PssA found by Ivashina et al. (1994) is depicted in italics. Respective GTG start codons are in boldface. ▼, sites of insertion of Tn5s present in chromosomal DNA fragments of strains *exo2* and *exo4*, and in the different mutagenized pMP2604 plasmids.

282 and one at position 471 of pMP2642 (Fig. 4). The longest predicted amino acid sequence of this ORF was compared with the translated DNA sequences (each of six possible frames) in the GenBank and EMBL data bases, using the program BLAST. The ORF showed 99.2% identity to *pss4* of *R. leguminosarum* bv. *viciae* and 96.5% identity to *pssA* (= *pss2*) of *R. leguminosarum* bv. *phaseoli*, and showed moderate homology (31.5 to 42.6% identity) to a family of bacterial glycosyl transferases performing the first step of polysaccharide biosynthesis in their respective hosts (Table 2). In order to test whether the gene affected in mutant strains *exo2* and *exo4* was also functionally homologous to the *pssA* gene, described by Borthakur and co-workers (1986, 1988), we introduced plasmid pMP2642 in *R. leguminosarum* bv. *phaseoli* mutant strain 8002,*pss2*::Tn5. The mutant was fully complemented by this plasmid for the synthesis of EPS, and therefore we concluded that the gene affected in mutants *exo2* and *exo4* is functionally homologous to the *pssA* gene of *R. leguminosarum* bv. *phaseoli*.

As can be seen in Table 2, the lengths of the various homologous genes are quite different and it should be noted that *pss4* of *R. leguminosarum* bv. *viciae* has an extension of 63 aa at the N terminus, compared with *pssA* of *R. leguminosarum* bv. *phaseoli* (Ivashina et al. 1994). This corresponds to a GTG start codon at either position 282 or 471 of the *R. leguminosarum* bv. *trifolii* *pssA* gene (Fig. 4). In order to determine the most probable GTG start codon of the *pssA* gene in *R. leguminosarum* bv. *trifolii* strain RBL5599, we localized the Tn5 insertions in the normally nodulating mutant *exo2* and the poorly nodulating mutant *exo4*, as well as in plasmids pMP2604, Ω 5::Tn5 and Ω 6::Tn5, which partly complement EPS synthesis of *exo4*. The position of the Tn5 insertion of mutant *exo2* was found to be downstream of the first GTG initiation codon at bp 294, while the Tn5 of *exo4* is downstream of both putative GTG start codons at bp 696. The Tn5s present in plasmid pMP2604, Ω 5::Tn5 and Ω 6::Tn5, are located in the region immediately downstream of the first GTG

codon, next to the site of insertion of the Tn5 present in the nodulating mutant *exo2*, at positions 285 and 340, respectively.

Taken together, the differences in EPS production between *exo2* and *exo4*, and the location of the Tn5s present in *exo2*, *exo4*, pMP2604, Ω 5::Tn5, and Ω 6::Tn5, strongly suggest that the region between the putative GTG initiation codons is part of the promoter. Therefore, we conclude that the *pssA* gene of *R. leguminosarum* bv. *trifolii* strain RBL5599 starts at position 471, encoding a polypeptide of 200 aa, 22.6 kDa in size.

pssC, pssD, and pssE.

DNA fragments of one cosmid of group iv were subcloned in pMP92, and subsequently tested for complementation in mutants *exo50* and *exo111*. By this approach we obtained pMP3030, a plasmid harboring a 2.8-kb *Pst*I insert, introduction of which in mutant strains *exo50* and *exo111* gave mucoid colonies like those of parental strain RBL5599. Furthermore, plasmid pMP3030 completely restored the nodulation deficiencies of both strains. To determine whether mutants *exo50* and *exo111* are affected in the same gene, we performed a more detailed complementation study (Fig. 5). Subclones of pMP3030 formed two complementation groups: one group complementing *exo50* and the other complementing *exo111*, both for EPS synthesis and nodulation deficiency.

The smallest subclone that complements *exo50* is pMP3032 ranging from the *Pst*I site at position 1 to the *Bst*BI site at position 1317. Deletion of 187 bp at the 3' end yielded a clone (pMP3033) that was unable to complement *exo50* (Fig. 5). Sequencing of the 1.3-kb DNA fragment in pMP3032 revealed an ORF that reads toward the *Bst*BI site and that was named *pssC*. The *pssC* gene starts at an ATG codon (position 183) that is preceded by a putative ribosome-binding site, ends at position 1163, and encodes a 37.1-kDa protein (Fig. 6). We searched the available databases with BLASTP (Altschul et al. 1990) and found homology with *exoM* (27.0% identity) and *exoW* (22.8% identity) of *R. meliloti*, genes encoding glucosyl transferases

Table 2. Percentages of identity and similarity between *pssA* from *Rhizobium leguminosarum* bv. *trifolii* and the amino acid (aa) sequences of various bacterial glycosyl transferases

Gene	Organism	Length (aa) ^a	Identity (%)	Similarity (%)	<i>pssA</i> region (aa) ^b	Reference or accession number
<i>pss4</i>	<i>R. leguminosarum</i> bv. <i>viciae</i>	263	99.2	99.6	1 to 263	Ivashina et al. 1994
<i>pssA</i>	<i>R. leguminosarum</i> bv. <i>phaseoli</i>	200	96.5	99.0	64 to 263	Borthakur et al. 1988
<i>aceA</i>	<i>Acetobacter xylinum</i>	532	42.6	60.1	1 to 263	Griffin et al. 1996
<i>rfbP</i>	<i>Campylobacter hyoilei</i>	200	41.4	66.9	63 to 262	X91081
<i>spsB</i>	<i>Sphingomonas</i> S88	470	40.9	66.0	1 to 263	Yamazaki et al. 1996
<i>exoY</i>	<i>R. meliloti</i>	226	40.0	61.0	25 to 261	Reed et al. 1991
<i>rfbP</i>	<i>Klebsiella pneumoniae</i>	447	39.6	63.1	5 to 263	Arakawa et al. 1995
<i>bplG</i>	<i>Bordetella pertussis</i>	197	38.8	63.9	67 to 255	Allen and Maskell 1996
<i>ORF10</i>	<i>Shigella sonnei</i>	129	37.5	60.2	93 to 228	U34305
<i>rfbP</i>	<i>Haemophilus influenzae</i>	471	37.5	61.3	3 to 263	Fleischmann et al. 1995
<i>ORF7</i>	<i>Vibrio cholerae</i>	186	37.4	62.6	63 to 262	Comstock et al. 1996
<i>exoY</i>	<i>Agrobacterium radiobacter</i>	226	37.1	57.1	41 to 261	X95394
<i>gumD</i>	<i>Xanthomonas campestris</i>	484	36.2	56.5	1 to 263	U22511
<i>epsE</i>	<i>Streptococcus thermophilus</i>	227	36.2	63.8	42 to 262	Stingle et al. 1996
<i>amsG</i>	<i>Erwinia amylovora</i>	477	35.8	58.5	1 to 263	Bugert and Geider 1995
<i>rfbP</i>	<i>Salmonella enterica</i>	125	35.2	60.0	105 to 263	Wang et al. 1992
<i>exoY</i>	<i>Rhizobium</i> strain NGR234	226	34.9	58.4	41 to 261	Gray et al. 1990
<i>cpsD</i>	<i>Streptococcus agalactiae</i>	274	33.3	58.8	1 to 180	Rubens et al. 1993
<i>wbaP</i>	<i>Salmonella enterica</i>	476	32.4	55.6	1 to 263	Wang et al. 1996
<i>cps14E</i>	<i>Streptococcus pneumoniae</i>	455	32.3	57.4	1 to 262	Kolkman et al. 1996
<i>cpsD</i>	<i>Streptococcus pneumoniae</i>	455	31.5	56.2	1 to 262	Guidolin et al. 1994

^a Of protein.

^b Region over which the similarity occurs.

involved in the biosynthesis of succinoglycan (Becker et al. 1993b; Glucksmann et al. 1993a). The hydrophobicity plot of PssC is nearly identical to that of ExoM (Glucksmann et al. 1993a), indicating that PssC might be attached to the cytoplasmic membrane at the carboxyl terminus (data not shown). The insertion site of the Tn5 transposon in mutant exo50 is located in the coding region of *pssC* between bp 224 and 225 (Fig. 6). However, the mutant still synthesizes about 27% of the amount of EPS produced by the parental strain RBL5599. Possibly, another gene present in *R. leguminosarum* bv. *trifolii* is able to substitute partially for the loss of the *pssC* gene.

Mutant exo111 could be fully complemented, with respect to both EPS synthesis and nodulation deficiency, by pMP3036, a plasmid containing a subclone of pMP3030 ranging from the *Bst*BI site at position 1317 to the *Pst*I site at position 2832. Sequence analysis of this clone revealed not one long, but two short, overlapping ORFs, which were called *pssD* and *pssE*, respectively (Fig. 6). The transposon Tn5 in mutant exo111 was found to be located in *pssD* between bp 1488 and 1489. To test whether both *pssD* and *pssE* are necessary for complementation of mutant exo111, we made additional subclones. As can be seen in Figure 5, a deletion of 173 bp from the *Bst*BI site to the *Sma*I site (pMP3035) resulted in the loss of the complementing ability. The deleted fragment contains the promoter and the 5' part of *pssD*, and therefore we concluded that *pssD* is essential for complementation of

mutant exo111. Likewise, we made some deletions at the other side of the subcloned fragment of pMP3036, which resulted in the construction of plasmids pMP3037 and pMP3038. pMP3037 is the smallest subclone of pMP3030 that is complementing mutant exo111, and contains the complete coding sequences of both *pssD* and *pssE*. In contrast to pMP3037, the 61-bp-shorter subclone pMP3038 lacks a small part of the 3' part of *pssE*, which results in loss of complementing ability. Apparently, next to *pssD* the expression of *pssE* is essential for complementation of mutant exo111. However, a polar effect of the Tn5 insertion on the expression of *pssE* cannot be excluded.

A homology search revealed 24.1% identity of PssD to Cps14F of *Streptococcus pneumoniae* and 32.2% identity to the N-terminal half of SpsK of *Sphingomonas* S88 (aa 14 to 177), whereas PssE is 29.5% identical to Cps14G of *S. pneumoniae* and 27.7% identical to the C-terminal half of SpsK of *Sphingomonas* S88 (aa 188 to 347) (Fig. 7). The *spsK* gene of *Sphingomonas* S88 was shown to be essential for the synthesis of the CPS sphingan S-88, and it was hypothesized that this gene might encode a glycosyl transferase based on local similarities to glucosyl and glucuronosyl transferases (Yamazaki et al. 1996). Cps14F and Cps14G were recently found to be responsible for the second step in the synthesis of the repeating unit of the CPS of *S. pneumoniae*, catalyzing the transfer of galactose to lipid-linked glucose (M. A. B. Kolkman,

Fig. 5. Gene cluster for exopolysaccharide synthesis located at the 2.8-kb *Pst*I fragment from a cosmid of group iv (see text for definition). Subcloned fragments are indicated below genetic map. Complementation results with subcloned fragments in mutants exo50 and exo111 depicted at right side of figure. P, *Pst*I; Sp, *Sph*I; E, *Eco*RV; N, *Nar*I; BB, *Bst*BI; S, *Sma*I; BE, *Bst*EII; M, *Mlu*I.

unpublished results). Both a mutation in *cps14F* and a mutation in *cps14G* abolished the galactosyl transferase activity. This is consistent with our finding that *pssD* and *pssE* form one complementation group. Therefore, we concluded that, like *cps14F* and *cps14G*, *pssD* and *pssE* might encode proteins that act together as a glycosyl transferase. Furthermore, PssE shows an identity of 27.7% to a region of zeaxanthin glucosyl transferase from *Erwinia herbicola*, which is highly conserved in several other glucuronosyl and glucosyl transferases and which was proposed to be a UDP binding site (Hundle et al. 1992). Homology to this motif was also found in Cps14G. Possibly, Cps14G and PssE are responsible for

binding the UDP sugar. However, it should be noted that PssE matches the consensus sequence for only three out of nine residues, and that only 2 aa (glycine at position 70 and proline at position 88) are conserved in all compared glycosyl transferases.

DISCUSSION

In this paper we present the cloning and molecular characterization of four genes of *R. leguminosarum* bv. *trifolii* involved in the biosynthesis of EPS. Each of these genes was found to be homologous to previously identified glycosyl transferases of various bacteria.

```

1
CTGCAGGCACTCGGCAAGAGCCTCGTGGCCGCGGAATTGCCGAAACGACATCGCTTCTG
61
AAGCTCCGGGGCGAAGCTCGCAACACGGTGGCCGCGGTCGGTTCGGCGATGCCGCGCG
121
ACCGAAGTGGCCTTGATGGCGCGCGCAAGCGTTTGGGAATATGCTGCGGAGGACAG
                                exo50
181
TCATGAATCAGCAAGAGACTTTTCCGATTTCATCCGTCGACACAGAAGCAGGCGCGCAC
                                ▼
M N Q Q E T F P H S S V D T E A G A A
241
CGCTGAAGATTGCCGTGGTGTGCTGACCTATCGGCGCCTCGACGGGATCGCCAAGCTGC
P L K I A V G V L T Y R R L D G I A K L
301
TTGACGTCATGACGCGCAGGTGAGGACCCGGCTCGCCCCATCATCTCACCATGGTGA
L D V M T R Q V R H P A R P Y H L T M V
361
TCGTGGATAATGATCGGCGCGCAGCGCAGGACACGGTAGAGAGCTTTGGCCAAACAG
I V D N D A A G S A R T T V E S F G Q T
421
GTGCGCTACGACCTGATCTACGTTCATCGAACAAACAGGGCATAACCTTTGACGCAATC
G A Y D L I Y V I E Q N Q G I P F A R N
481
GCGCACTGGATTACGCGCCTCCCGGCACTGACCTCTTCTGCTTCTCGACGATGACGAAT
R A L D S A P P G T D L F C F L D D D E
541
GGCCGGTTCAGCGGCTGGCTGGATGCGATGCTGGAGACACGGGAGAAAACCGCGCGGATT
W P V D G W L D A M L E T R E K N R A D
601
GCGTCTATGGCCCGCTCAACCGGTCTATCCCGAAAACCGCGGAATATTCATCAAGG
C V Y G P V Q P V Y P E N P F E Y F I K
661
CAAGGGTGTTCGAGCGAAGAAGAAGCAGGACGGTCAACGCATTTCTTATCGGCGCATCGA
A R V F E R K K N E D G Q R I S Y A A S
721
ACAACGTCATGTTGCACTATCCCTGATCCGTTTCATGGAATCTGCGTTTGAAGAGAAGA
N N V M F D Y P L I R S W N L R F E E K
781
TGCGCTTACCGGGGGCAGGATTACCTTTCTTCAACCAAGCCATCCGCGGGGCGATGC
M R F T G G T D Y L F F N Q A I R R G M
841
AGGTCTTCTGGGCTGACAAGCGCTTGTCCATGATATCGTTCCGGCCAGCCGATGACCT
Q V F W A D K A L V H D I V P A S R M T
901
GGAAATGGGTATTGCAAGACAATACCGGCTCGGCAATACCTTCGAGTCAGCGAGGTCT
W K W V L Q R Q Y R L G N T F A V S E V
961
TGCAATGGCAACCTCAAGCGCGGATCTATCGCGCGCCTATGGCGCCACGAGAGTTGTGC
L H G N L K R R I Y R A A Y G A T R V V
1021
TTGGGCTGGCCATGCTGCCGGCGATCTTGATTTACCTTATGGGGCATGCGAGCCCTCA
L G L A M L P A I L I S P Y W G M R A L
1081
CCACGCTTTTACGCGCGCGGGCATGGTAAATGGGATTCTCGGACATGCATACAGGAAT
T H L L R G A G M V N G I L G H A Y Q E
1141
ACAAAGCCCAATGCCGCTCTTTGATATTGTTTACGAGGCGCCCTTGGAAGAGTTTCGGAGC
Y K P N A A L *
1201
GCCATTATGTTGAGCTGAAGCATCACAAAGATTGGCTGTCCATGAACCTCAAGGCTTCA
1261
ATCAACCGGGCCAAAATGGTACAATTCACCTGAACATATCGTGACACAATCGCTGTTTCA
1321
ACCGAAGCGGTTTTCGCCCTTATCCATCGGTCAAATATAGGAGATGCCTGTTTATGACTGA
                                M T E
1381
GAAAAAATGAAAGTTCTCGCTGCCTCGTCCGAGGCGGCCACTGGGAACAGCTGATGGC
K K L K V L A A S S G G G H W E Q L M A

```

```

                                exo111
1441
CATGCGCGGCGCATTCGAGGGCTGCGATATCGTCTTTGCAACGACCATCCCGGGACTGCT
M R G A F E G C D I V F A T T I P G L L
1501
TGCGAAATACGATATTCGCGCGGACTAGTCTTCCCGATTGCGAGCCGCGACTCGATTAC
A K Y D I R G G L V L P D C S R D S I T
1561
CATGTCGATCCGGTGCTTTTTCACCGCCTTCTCCATCGTCATCAAGCACAGGCCCGACGT
M S I R C F F T A F S I V I K H R P D V
1621
CATCATCTCGACCGGTGCGGCGACCCGGCCTCTTTGCGCTGCTGGCCGCAAAATTGACGGG
I I S T G A A P G L F C L L A G K L T G
1681
CAAGCGGACGATCTGGATCGATAGCGTCGCCAATGTCGAGAAGCTGCTCCCTATCGGGTAA
K R T I W I D S V A N V E K L S L S G K
1741
ATTGGCCGGCCATATTGCGACGCTCTGGCTCACGCAATGGCAACATCTGTCGAGGCGCGGA
L A G H I A T L W L T Q W Q H L S R P D
1801
TGGCCCCATTACGAGGAGCTGCTCTTTGATCTTGTACCGCTCGGAACGACGAGCTGCCG
G P H Y A G A V L *
                                L I L V T V G T Q L P
1861
TTTGATCGCCTCGTCAAGGCTGTGACACCTTCGCAACGAACTGCCGAAACCGGTCTGTG
F D R L V K A V D T F A N E L P K P V L
1921
GCGCAGATCGGCAAGGGCACCTACACGCCACAGAATATGAAATGGATCAAGAATATCGAG
A Q I G K G T Y T P Q N M K W I K N I E
1981
CCCAAGGACTTTGACAACGCTTTCGCGATGCAAGTGTCTGCTCTCTCATGCGGGGAATT
P K D F D N V F R D A S V I V S H A G I
2041
GGCACTGTGCTTACGCCAAGCGTTTGGAAAGCCGATCATCTCGTTCCCGCAGAGGGCG
G T V L T A K R F G K P I I L V P R Q A
2101
GCTCTCGGCGAGACCCGAAACGATCATCAGCTCGCCACCGTGAGCCAGCTCGTGGGCGGG
A L G E H R N D H Q L A T V S Q L V G R
2161
CCGGGCATCTATGTAGCACACCGATGACGATCTCGGAACATATCTGCTCGAAGAGCTG
P G I Y V A H T D D D L R N Y L L E E L
2221
GATAGCCCATCTCACGAAGATTTCGTCGGAAGCCGCTGGGGCTTCACTGGTCACCTACCTG
D S P S H E D S S E A G R A S L V T Y L
2281
AAAACTATATTACGGCAGTCTGACCCCGAGAGCCGTCATAGCCTCTGACGCGTAAAAGC
K N Y I T A V *
2341
GATTTTTCATTCTCCGCAATGGTTGCCGGCGATACGTTTGACGAGTTCAGGCGCAATGAT
2401
TTTGGAGGCCATGTCGGGCGAGGATGACGCCATCAGGAATTGCCGTGGCCAGGTTCATC
2461
CGGCGTCAGACGCTCCAGTTGGGCGGTGGTGCACAACTCCAGACCAGCTTTTTCGCG
2521
CTCCGCTCATGTGACAGACATAACTGTGACAAAGGAGCGGATCAGCCTACGCCATCC
2581
TGAAAACGGATTTCATCGCCATGACGATGATCCGCGCTTGGCGGACGCCGCTGCTGGAGCTG
2641
GTCGAGGATGTCGCGATATCTTTTCGGCTTTGCGCAAGCGACACGAACCGATGCAACGT
2701
CGCATCATTTGGCGTAGAGCTCAATCAGGATGATATCCGGCTGCTCGGCAACGACGCGATC
2761
AATCTGGGTAGAGCCACAGTGACGTTTTCACCGCTTTTTCGCAACGCTCTCGACCTTCAC
2821
CGGCTTCTGCAG

```

Fig. 6. DNA sequence of 2.8-kb *Pst*I fragment present in plasmid pMP3030. Predicted amino acid (aa) sequence of three open reading frames is indicated by single letter code. ▼, site of insertion of Tn5 present in chromosomal DNA fragments of strains exo50 and exo111. Putative Shine and Dalgarno sequences are in boldface.

pssA is a member of a family of glycosyl transferases involved in the addition of the first sugar to the isoprenoid lipid carrier (Ivashina et al. 1994; Borthakur et al. 1988; Griffin et al. 1996; Yamazaki et al. 1996; Reed et al. 1991; Müller et al. 1993; Arakawa et al. 1995; Allen and Maskell 1996; Fleischmann et al. 1995; Comstock et al. 1996; Stingle et al. 1996; Bugert and Geider 1995; Wang et al. 1992; Gray et al. 1990; Rubens et al. 1993; Wang et al. 1996; Kolkman et al. 1996; Guidolin et al. 1994). Our complementation experiments proved that this gene is functionally homologous to the *R. leguminosarum* bv. *phaseoli* *pssA* (= *pss2*) gene described by Borthakur et al. (1986, 1988). On the basis of the location of the transposon Tn5 in the nodulating mutant *exo2* and in the partially complementing plasmids pMP2604,Ω5::Tn5 and pMP2604,Ω6::Tn5, we propose that the translation of PssA in *R. leguminosarum* bv. *trifolii* strain RBL5599 starts at the second GTG codon in our sequence, which is also the start codon of the *pssA* gene in *R. leguminosarum* bv. *phaseoli* strain 8002. This proposal is inconsistent with the findings of Ivashina and co-workers (Ivashina et al. 1994), who described a nearly identical gene product (Pss4) in *R. leguminosarum* bv. *viciae* strain VF39, which has an extension of 63 aa at the N terminus. This inconsistency might result from their use of an *E. coli* T7-RNA polymerase expression system to determine the size of the ORF, which could yield a translation product different from that normally formed in vivo in *Rhizobium*. On the other hand, it can be hypothesized that in mutant *exo2* a truncated 22.3-kDa PssA protein is formed, which can substitute for the normal 29-kDa protein. Surprisingly, the two *pssA* mutant strains, *exo2* and *exo4*, still produce residual amounts of EPS. NMR analysis of the depolymerized EPS of these strains revealed that the structure of these residual EPS was comparable to that of parental strain RBL5599, except that they lacked acetyl esterifications to residue d. The ability of *exo2* to produce some residual amounts of EPS could be explained by the fact that the transposon in this mutant is inserted in the region immediately upstream of the *pssA* coding sequence. Mutant *exo4*, however, has a Tn5 insertion in the middle of the coding sequence, and nevertheless this mutant is still able to produce about 1% of the amount of EPS synthesized by the parental strain RBL5599. Apparently, some other gene function of *R. leguminosarum* bv. *trifolii* strain RBL5599 is able to partially replace the *pssA* gene. This contrasts with the situation in *R. leguminosarum* bv. *viciae* VF39 (Ivashina et al. 1994) and *R. leguminosarum* bv. *phaseoli* strain 8002 (Borthakur et al. 1986, 1988) for which a mutation of *pss4* and *pssA*, respectively, resulted in total EPS deficiency.

Complementation of mutants *exo50* and *exo111* resulted in cloning and sequencing of three additional genetic loci, *pssC*, *pssD*, and *pssE*, required for the biosynthesis of EPS in *R. leguminosarum* bv. *trifolii*. These genes were found to be closely linked, but not to the *pssA* gene. Furthermore, the existence of a set of mutations unlinked to either *pssA* or *pssC*, *pssD*, and *pssE* (cosmid groups i and ii) suggests that genes essential for EPS synthesis in *R. leguminosarum* bv. *trifolii* strain RBL5599 are not clustered, which is often the case for enzymes of a multi-component complex (Srere 1987). Experiments supporting this observation were recently described by Skorupska and co-workers (1995), who showed that EPS biosynthetic genes of *R. leguminosarum* bv. *trifolii* strain 24.1 are distributed over two large, cryptic megaplasmids and the bacterial chromosome.

Complementation studies showed that *pssC* on the one hand and *pssD* and *pssE* on the other hand formed two complementation groups. PssC was found to be significantly homologous to ExoM and ExoW of *R. meliloti*, both glucosyl transferases catalyzing, respectively, the addition of the fourth and the seventh sugar to the subunit on the lipid carrier (Glucksmann et al. 1993a). Thus, *pssC* probably encodes a glycosyl transferase, but it is impossible to assign sugar-binding specificity based on sequence homology alone, since glycosyl transferases show a very low degree of conservation even for enzymes from a single bacterium that catalyze attachment of the same sugars in identical linkages (Glucksmann et al. 1993a). As mentioned before, *pssC* mutant strain *exo50* still produces residual amounts of EPS. Like EPS from the *pssA* mutant strains *exo2* and *exo4*, this residual EPS was EPS with normal substitutions, except that it lacked acetyl esterifications to residue d. Therefore, it seems likely that another transferase can substitute for the function of PssC, as was postulated for glucosyl transferases ExoO and ExoU of *R. meliloti* (Glucksmann et al. 1993a).

Closely linked to *pssC*, two genes essential for EPS synthesis in *R. leguminosarum* are present, called *pssD* and *pssE*. *pssD* and *pssE* show an overall homology to *cps14F* and

A

PssD	.....MTEKKLVLAASGGGHWQELMAMRGAFEGCDIVPAT	37
Cps14F	.....MKVCLVSGSGHLLTHLYLLKPFKDRERFWVT	32
SpsK	MGDMAEATVTEAKAGKPLKMLAASGGHLLRQILDLESVWKEHYFVVT	50
	TIPGLLAKYDIRGGLVLPDC.....SRDSITMSIRCFFTFAPSIIV	77
	FDKEDARSILENE..IFYPCYYPTN.....RNLKNLKNITLAFKILR	73
	EDTALGRSLAEKHSVALVDHYALGQAKLGHPLRLMLGAWRNLRQSLSIIR	100
	KHRPDIIVISTGAAPGLFCLLAGKLTGKRTIWDISVANVEKLSLSGKLAGE	127
	KERPDIIVSSGAAVAVPPFFYLKIFGAKTVYIEVDFRIDAPTMTGKLVYP	123
	KHKPDVIVISTGAGAVYFTALLAKLSGAKTVHIESFARFDPHPSAFGRMVKG	150
	IATLNLTLQWQHLSR..PDGPHYAGAVL.....	152
	VTDRFIVQWEEMKVVYPKAINLGGIF.....	149
	IATVTIVQSAALKQWPDALFDFPRLLDTPFPKQALTFATVGATLPFP	200

B

PssE	LILVTVGT..QLPFDRLVKAVDTF..ANLKPFLVLAQIGKGYTFPQNMKI	47
Cps14G	MIFVTVGSHEQQFNRLIKEVDRLKGEFGTQDDVFIQTGYSNYVFKCKWE	50
SpsK	LTFATVGA..TLFPFLVQAVLDLKRAGGLPGKLVLYQGDQDLADPGIPDV	236
UDP-C	.....	
	...KNIEPKDFDNVFRDASVIVSHAGIGTIVLAKRFGKPIILVPRQAALG	94
	...KLISYEKMNQLIKESDIIITHGGPATFMAVIAKGNPIIIVPRLLKFG	97
	EIRRTIIPFDDLQLLRDADMVICHGGTGLSVLTAIRAGCRVVAFFRRHDLG	286
	QXXXXXXXXXXXXXTXGXXXXXXXXLXXXXPXXXXXDXQX	
	EHRNDHQLATVSQLVGRPGIYVAHTDDDLRNVLLELDSPHEDSSEAG.	143
	EEVNDHQMVFVKITKEIYNLVIDDLSHLILHNFKDKHFEYTLNNERF	147
	EEYDDHQEEIAQTADFAGLLHVRDERELGAAV..EAAKATEPQLATTD.	333
	XXXX	
	RASLVTLKNIYITAV.....	158
	NVRFNVEISNLFKGNKINEN	167
	HTALAGRLRELLAQWSAKR.	352

Fig. 7. A, Alignment of PssD with Cps14F (Kolkman et al. 1996) and the N-terminal part of SpsK (Yamazaki et al. 1996). **B**, Alignment of PssE with Cps14G (Kolkman et al. 1996), the C-terminal part of SpsK (Yamazaki et al. 1996), and the consensus of a proposed UDP binding site (UDP-C) of glycosyl transferases (Hundle et al. 1992). Numbers to right are residue numbers for rightmost amino acid on each line. Boldface indicates amino acids conserved in each protein. Identities of one of compared gene products with PssD and PssE indicated with an asterisk.

cps14G, respectively, genes involved in CPS synthesis of the gram-positive bacterium *S. pneumoniae*. These two genes encode enzymes together involved in the addition of galactose to the lipid-linked glucose residue (M. A. B. Kolkman, unpublished results). Like *cps14F* and *cps14G*, *pssD* and *pssE* show an overlap of 4 bp, indicating that expression of these genes might be translationally coupled. The amino acid sequence of *pssE* starts at a TTG codon (position 1828) that overlaps with the stop codon (TGA) of *pssD* and is downstream of a potential ribosome-binding site (AGGAG). The TTG codon is used as a translational initiation codon in about 1% of all known *E. coli* genes (Stormo 1986), and was found as an initiation codon in the *R. leguminosarum* bv. *viciae* nodT gene as well (Surin et al. 1990). Interestingly, *pssD* is also homologous with the 5' part of *spsK* of *Sphingomonas* S88, whereas *pssE* is homologous with the 3' part of the same gene. Therefore, *spsK* may have arisen from a fusion of ancestral genes, bringing together a transferase activity and a yet unknown function. A similar fusion of ancestral genes was proposed for *gumD* of *Xanthomonas campestris* and *rfbP* of *Salmonella enterica* (Wang et al. 1996; Wang and Reeves 1994; Coplin and Cook 1990), two genes of which the 3' part is homologous to *pssA*. Recently, it was shown that the 3' half of *rfbP* performs the galactosyl transferase activity, whereas the 5' half most likely is involved in the release of the undecaprenyl pyrophosphate-linked galactose from RfbP (Wang et al. 1996). Because of the (limited) homology of *cps14G*, *pssE*, and the 3' part of *spsK* to known glycosyl transferases, this entity

most likely performs the transferase activity. Analysis of the hydrophobicity pattern of *PssD* suggests that this protein might be an essential component in the membrane-associated glycosyl transferase complex. Further biochemical analysis of a nonpolar *pssD* mutant is required for elucidation of the actual function of this hydrophobic protein, which might also be informative for other biological systems.

MATERIALS AND METHODS

Strains, plasmids, and media.

R. leguminosarum strain RBL5515 is a Sym plasmid-cured derivative of *R. leguminosarum* bv. *trifolii* strain LPR5 (McNeil et al. 1986). Plasmid pSym5 is the Sym plasmid of *R. leguminosarum* bv. *trifolii* strain RCR5 (Hooykaas et al. 1981), which was marked with *Tn1831* for selection purposes (Pees et al. 1986). A list of the organisms and the characteristics of the plasmids used in this study are shown in Tables 3 and 4. *Rhizobium* spp. were grown in TY (Beringer 1974), YMB (Hooykaas et al. 1977) or B⁻ medium (van Brussel et al. 1977) at 28°C. *E. coli* strains were grown in Luria complete (LC) medium (Maniatis et al. 1982) at 37°C. Tetracycline (Tc) was added to a concentration of 20 and 2 µg/ml for *E. coli* and *R. leguminosarum*, respectively. Carbenicillin (Cb), rifampicin (Rif), and spectinomycin (Spc) were added to final concentrations of 100, 20, and 400 µg/ml, respectively.

Table 3. Strains and their relevant characteristics

Strain	Relevant characteristics ^a	Reference
<i>Rhizobium</i>		
RBL5515	<i>R. leguminosarum</i> bv. <i>trifolii</i> LPR5 cured of its Sym plasmid, <i>str rif</i>	Priem and Wijffelman 1984
RBL5599	<i>R. leguminosarum</i> strain RBL5515 pSym5 <i>Spc51::Tn1831</i>	Pees et al. 1986
RBL5854	<i>R. leguminosarum</i> RBL5599 <i>exoB8::Tn5</i> , <i>str rif</i>	Canter Cremers et al. 1988, 1990
RBL5802	<i>R. leguminosarum</i> RBL5599 <i>pssA2::Tn5</i> (exo2), <i>str rif</i>	Canter Cremers et al. 1988; van Workum et al. 1995
RBL5853	<i>R. leguminosarum</i> RBL5599 <i>pssA4::Tn5</i> (exo4), <i>str rif</i>	Canter Cremers et al. 1988; van Workum et al. 1995
RBL5814	<i>R. leguminosarum</i> RBL5599 <i>pssC50::Tn5</i> (exo50), <i>str rif</i>	Canter Cremers et al. 1988
RBL5862	<i>R. leguminosarum</i> RBL5599 <i>pssD111::Tn5</i> (exo111), <i>str rif</i>	Canter Cremers et al. 1988
<i>Escherichia coli</i>		
DH5α	F ⁺ , <i>supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	

^a *str*, streptomycin resistance; *rif*, rifampicin resistance. *Tn1831* codes for spectinomycin- and streptomycin resistance

Table 4. Plasmids used in this study

Plasmid	Relevant characteristics ^a	Reference
pIC20H	IncColE1, multicopy cloning vector, <i>Cb</i>	Marsh et al. 1984
pLAFR1	IncP, Δ <i>cos</i> Mob ⁺ Tra ⁻ cosmid vector, <i>Tc</i>	Friedman et al. 1982
pRK2013	Mobilizing plasmid used in tripartite matings, <i>Km</i>	Ditta et al. 1980
pMP92	IncP, multiple cloning vector, <i>Tc</i>	Spaink et al. 1987
pMP2604	pMP92 containing <i>pssA</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP2642	pMP92 containing <i>pssA</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3030	pMP92 containing <i>pssC</i> , <i>pssD</i> , and <i>pssE</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3031	pMP92 containing <i>pssC</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3032	pMP92 containing <i>pssC</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3033	pMP92 containing a part of <i>pssC</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3034	pMP92 containing <i>pssD</i> and <i>pssE</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3035	pMP92 containing <i>pssD</i> and <i>pssE</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3036	pMP92 containing <i>pssD</i> and <i>pssE</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3037	pMP92 containing <i>pssD</i> and <i>pssE</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study
pMP3038	pMP92 containing <i>pssD</i> and a part of <i>pssE</i> of <i>R. leguminosarum</i> bv. <i>trifolii</i>	This study

^a *Tc*, tetracycline resistance; *Cb*, carbenicillin resistance; *Km*, kanamycin resistance.

Isolation of EPS.

For standard isolation of EPS, bacteria were grown for 7 days at 28°C on a rotary shaker in Erlenmeyer flasks containing 500 ml of B⁻ medium that had been inoculated from a starter culture to an A₆₆₀ of 0.05. In this period, strain RBL5515 normally reached an A₆₆₀ value of >2.0. The cultures were centrifuged for 20 min at 12,000 × g in a Beckman JA14 rotor. EPS was precipitated from culture supernatants with 3 volumes of ethanol, and quantified by the anthrone-sulfuric acid method (Dubois et al. 1956). For analysis of its chemical structure, EPS was isolated, after the bacterial cells were pelleted, from the culture supernatant by a hollow fiber filtration procedure (Djordjevic et al. 1986). After freeze drying, the crude EPS was dissolved in water and precipitated by adding 2 volumes of 96% ethanol. The precipitated material was collected by centrifugation for 10 min at 15,000 × g, dissolved in water, and freeze dried. In order to depolymerize the EPS, 0.15 g of freeze-dried EPS was incubated for 6 days at 28°C in 25 ml of 25 mM phosphate buffer, pH 7.0, containing 1 mM MgCl₂ and about 10¹⁰ plaque-forming units of phage RL38. After removal of the remaining EPS polymers by ethanol precipitation, concentration by roto-evaporation, and freeze drying, the oligosaccharides were then consecutively purified over Sephadex DEAE A25, Biogel P2, and Dowex 50WX2 columns (Djordjevic et al. 1986). The oligosaccharide samples were dissolved in 99% D₂O and spectra (200 mHz for ¹H, 50.3 mHz for ¹³C) were recorded on a Varian XL-200 spectrometer. The temperature was 20°C, unless otherwise indicated. In the ¹H spectra, the chemical shifts were measured relative to HDO at 4.815 ppm at 20°C. For the ¹³C spectra, methanol at 50.01 ppm was used as an internal standard.

Isolation of CPS.

To test production of CPS, a method for isolation and characterization described by Breedveld et al. (1990) was followed. Routinely, bacteria were grown for 10 days at 28°C in 25 ml of B⁻ medium. The bacterial cells were subsequently collected by centrifugation, and the CPS was extracted from each pellet with 1 M NaOH, precipitated with 1 volume of ethanol, and quantified by the anthrone-sulfuric acid method (Dubois et al. 1956).

Analysis of LPS profile.

LPS in crude cell extracts was analyzed by polyacrylamide gel electrophoresis with 18% acrylamide gels with deoxycholic acid as the detergent (Krauss et al. 1988), and visualized by silver staining (Tsai and Frisch 1982).

DNA manipulations and bacterial matings.

Recombinant DNA techniques were carried out as described by Maniatis et al. (1982). Construction of a genomic library of RBL5515 was described by Roest et al. (1995). Plasmids were transferred from *E. coli* to *Rhizobium* strains by triparental conjugal mating as described previously (Ditta et al. 1980) with pRK2013 as a helper plasmid.

DNA sequencing and sequence analyses.

DNA sequencing was performed according to Sanger et al. (1977) with vector pIC20H (Marsh et al. 1984). Sequence reactions were performed with either the T7sequencing kit or

the Deaza G/A T7sequencing kit from Pharmacia LKB (Woerden, The Netherlands) with (α-³⁵S)dATP (Amersham International, Little Chalfont, England) according to the manufacturer's guidelines. The programs FASTA and BLAST (Genetics Computer Group [GCG] package; National Center for Biotechnology Information, Madison, WI), based on the method of Wilburn and Lipman (1983), were used to search data banks for sequences with homology to the different ORFs. The program BESTFIT in the GCG package, which uses the "local homology" algorithm of Smith and Waterman (1981) and GAP, which uses the algorithm of Needleman and Wunsch (1970), were used to perform sequence comparison. The percentages of identity and similarity given below were derived by using BESTFIT with a gap weight of 3.0 and a gap length weight of 0.10. The program PEPLOT (GCG package) (Devereux et al. 1984), on the basis of Kyte and Doolittle (1982), was used to identify potential transmembrane segments.

GenBank accession numbers.

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession numbers Y07549 (*pssA*) and X99850 (*pssC*, *pssD*, and *pssE*).

ACKNOWLEDGMENTS

This research was supported by the Foundation for Life Sciences (SLW), which is subsidized by the Netherlands Organization of Scientific Research (NWO). We thank M. A. B. Kolkman for sharing unpublished results and sequence data with us, and D. Kafetzopoulos and H. P. Spaik for helpful discussions. Furthermore, we greatly acknowledge M. Batley (Macquarie University, Sydney, Australia) for his help in the NMR analyses of the EPS.

LITERATURE CITED

- Allen, A., and Maskell, D. 1996. The identification, cloning and mutagenesis of a genetic locus required for lipopolysaccharide biosynthesis in *Bordetella pertussis*. *Mol. Microbiol.* 19:37-52.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Aman, P., McNeil, M., Franzen, L.-E., Darvill, A. G., and Albersheim, P. 1981. Structural elucidation, using H.P.L.C.-M.S. and G.L.C.-M.S., of the acidic polysaccharide secreted by *Rhizobium meliloti* strain 1021. *Carbohydr. Res.* 95:263-282.
- Arakawa, Y., Wacharotayankun, R., Nagatsuka, T., Ito, H., Kato, N., and Ohta, M. 1995. Genomic organization of the *Klebsiella pneumoniae* *cps* region responsible for serotype K2 capsular polysaccharide synthesis in the virulent strain chedid. *J. Bacteriol.* 177:1788-1796.
- Becker, A., Kleickmann, A., Arnold, W., and Pühler, A. 1993a. Analysis of the *Rhizobium meliloti* *exoH/exoK/exoL* fragment: ExoK shows homology to excreted endo-β-1,3-1,4-glucanases and ExoH resembles membrane proteins. *Mol. Gen. Genet.* 238:145-154.
- Becker, A., Kleickmann, A., Keller, M., Arnold, W., and Pühler, A. 1993b. Identification and analysis of the *Rhizobium meliloti* *exoAMONP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. *Mol. Gen. Genet.* 241:367-379.
- Becker, A., Kleickmann, A., Küster, H., Keller, M., Arnold, W., and Pühler, A. 1993c. Analysis of the *Rhizobium meliloti* genes *exoU*, *exoV*, *exoW*, *exoT*, and *exoI* involved in exopolysaccharide biosynthesis and nodule invasion: *exoU* and *exoW* probably encode glucosyl-transferases. *Mol. Plant-Microbe Interact.* 6:735-744.
- Beringer, J. E. 1974. R factor transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* 84:188-198.
- Bibb, M. J., Findlay, P. R., and Johnson, M. W. 1984. The relationship

- between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* 30:157-166.
- Borthakur, D., Barber, C. E., Lamb, J. W., Daniels, M. J., Downie, J. A., and Johnston, A. W. B. 1986. A mutation that blocks exopolysaccharide synthesis prevents nodulation of peas by *Rhizobium leguminosarum* but not of beans by *R. phaseoli* and is corrected by cloned DNA from *Rhizobium* or the phytopathogen *Xanthomonas*. *Mol. Gen. Genet.* 203:320-323.
- Borthakur, D., Barker, R. F., Latchford, J. W., Rossen, L., and Johnston, A. W. B. 1988. Analysis of *pss* genes in *Rhizobium leguminosarum* required for exopolysaccharide synthesis and nodulation of peas: Their primary structure and their interaction with *psi* and other nodulation genes. *Mol. Gen. Genet.* 213:155-162.
- Borthakur, D., Downie, J. A., Johnston, A. W. B., and Lamb, J. W. 1985. *psi*, a plasmid-linked *Rhizobium phaseoli* gene that inhibits exopolysaccharide production and which is required for symbiotic nitrogen fixation. *Mol. Gen. Genet.* 200:278-282.
- Borthakur, D., and Johnston, A. W. B. 1987. Sequence of *psi*, a gene on the symbiotic plasmid of *Rhizobium phaseoli* which inhibits exopolysaccharide synthesis and nodulation and demonstration that its transcription is inhibited by *psr*, another gene on the symbiotic plasmid. *Mol. Gen. Genet.* 207:149-154.
- Breedveld, M. W., Canter Cremers, H. C. J., Batley, M., Posthumus, M. A., Zevenhuizen, L. P. T. M., Wijffelman, C. A., and Zehnder, A. J. B. 1993. Polysaccharide synthesis in relation to nodulation behavior of *Rhizobium leguminosarum*. *J. Bacteriol.* 175:750-757.
- Breedveld, M. W., Zevenhuizen, L. P. T. M., and Zehnder, A. J. B. 1990. Excessive excretion of cyclic beta-(1-2)-glucan by *Rhizobium trifolii* TA1. *Appl. Environ. Microbiol.* 56:2080-2086.
- Buendia, A. M., Enenkel, B., Köplin, R., Niehaus, K., Arnold, W., and Pühler, A. 1991. The *Rhizobium meliloti* *exoZ/exoB* fragment of megaplasmid2: *ExoB* functions as a UDP-glucose 4-epimerase and *ExoZ* shows homology to NodX of *Rhizobium leguminosarum* biovar *viciae* strain TOM. *Mol. Microbiol.* 5:1519-1530.
- Bugert, P., and Geider, K. 1995. Molecular analysis of the *ams* operon required for exopolysaccharide synthesis of *Erwinia amylovora*. *Mol. Microbiol.* 15:917-933.
- Canter Cremers, H. C. J., Batley, M., Redmond, J. W., Eysdems, L., Breedveld, M. W., Zevenhuizen, L. P. T. M., Pees, E., Wijffelman, C. A., and Lugtenberg, B. J. J. 1990. *Rhizobium leguminosarum* *exoB* mutants are deficient in the synthesis of UDP-glucose 4'-epimerase. *J. Biol. Chem.* 265:21122-21127.
- Canter Cremers, H. C. J., Batley, M., Redmond, J. W., Wijffjes, A. H. M., Lugtenberg, B. J. J., and Wijffelman, C. A. 1991. Distribution of O-acetyl groups in the exopolysaccharide synthesized by *Rhizobium leguminosarum* strains is not determined by the Sym plasmid. *J. Biol. Chem.* 266:9556-9564.
- Canter Cremers, H. C. J., Wijffelman, C. A., Pees, E., Engels, M., Hoogerbruggen, F., Stevens, K., van Dijk, M., Lugtenberg, B. J. J., Bothe, H., de Bruijn, F. J., Newton, W. J., eds. 1988. Chromosomal nodulation genes of *Rhizobium leguminosarum* biovar *viciae*. Page 484 in: *Nitrogen Fixation: Hundred Years After. Gustav Fisher, Stuttgart.*
- Comstock, L. E., Johnson, J. A., Michalski, J. M., Morris, J. G., Jr., and Kaper, J. B. 1996. Cloning and sequence of a region encoding a surface polysaccharide of *Vibrio cholerae* O139 and characterization of the insertion site in the chromosome of *Vibrio cholerae* O1. *Mol. Microbiol.* 19:815-826.
- Coplin, D. L., and Cook, D. 1990. Molecular genetics of extracellular polysaccharide biosynthesis in vascular phytopathogenic bacteria. *Mol. Plant-Microbe Interact.* 3:271-279.
- Devereux, J., Haeblerli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395.
- Diebold, R., and Noel, K. D. 1989. *Rhizobium leguminosarum* exopolysaccharide mutants: Biochemical and genetic analyses and symbiotic behavior on three hosts. *J. Bacteriol.* 171:4821-4830.
- Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 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.
- Djordjevic, S. P., Batley, M., Redmond, J. R., and Rolfe, B. G. 1986. The structure of the exopolysaccharide from *Rhizobium* sp. strain ANU280 (NGR234). *Carbohydr. Res.* 148:87-99.
- Djordjevic, S. P., Chen, H., Batley, M., Redmond, J. W., and Rolfe, B. G. 1987. Nitrogen fixation ability of exopolysaccharide synthesis mutants of *Rhizobium* sp. NGR234 and *Rhizobium trifolii* is restored by the addition of homologous exopolysaccharides. *J. Bacteriol.* 169:53-60.
- Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. 1956. Colorimetric method for determination of sugars and related substances. *Anal. Chem.* 28:350-356.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghegan, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496-512.
- Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J., and Ausubel, F. M. 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene* 18:289-296.
- Glucksmann, M. A., Reuber, T. L., and Walker, G. C. 1993a. Family of glycosyl transferases needed for the synthesis of succinoglycan by *Rhizobium meliloti*. *J. Bacteriol.* 175:7033-7044.
- Glucksmann, M. A., Reuber, T. L., and Walker, G. C. 1993b. Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: A model for succinoglycan biosynthesis. *J. Bacteriol.* 175:7045-7055.
- Gray, J. X., Djordjevic, M. A., and Rolfe, B. G. 1990. Two genes that regulate exopolysaccharide production in *Rhizobium* sp. strain NGR234: DNA sequences and resultant phenotypes. *J. Bacteriol.* 172:193-203.
- Griffin, A. M., Morris, V. J., and Gasson, M. J. 1996. Identification, cloning and sequencing the *aceA* gene involved in acetan biosynthesis in *Acetobacter xylinum*. *FEMS Microbiol. Lett.* 137:115-121.
- Guidolin, A., Morona, J. K., Morona, R., Hansman, D., and Paton, J. C. 1994. Nucleotide sequence analysis of genes essential for capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 19F. *Infect. Immun.* 62:5384-5396.
- Hooykaas, P. J. J., Klapwijk, P. M., Nuti, M. P., Schilperoort, R. A., and Rörsch, A. 1977. Transfer of the *Agrobacterium tumefaciens* Ti plasmid to avirulent agrobacteria and to rhizobia *ex planta*. *J. Gen. Microbiol.* 98:477-484.
- Hooykaas, P. J. J., van Brussel, A. A. N., den Dulk-Ras, H., van Slogteren, G. M. S., and Schilperoort, R. A. 1981. Sym plasmid of *Rhizobium trifolii* expressed in different Rhizobial species and *Agrobacterium tumefaciens*. *Nature (London)* 291:351-353.
- Hundle, B. S., O'Brien, D. A., Alberti, M., Beyer, P., and Hearst, J. E. 1992. Functional expression of zeaxanthin glucosyltransferase from *Erwinia herbicola* and a proposed uridine diphosphate binding site. *Proc. Natl. Acad. Sci. USA* 89:9321-9325.
- Ivashina, T. V., Khmelnsky, M. I., Shlyapnikov, M. G., Kanapin, A. A., and Ksenzenko, V. N. 1994. The *pss4* gene from *Rhizobium leguminosarum* bv. *viciae* VF39: Cloning, sequence and the possible role in polysaccharide production and nodule formation. *Gene* 150:111-116.
- Kolkman, M. A. B., Morrison, D. A., van der Zeijst, B. A. M., and Nuijten, P. J. M. 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: Identification of the glycosyl transferase gene *cps14E*. *J. Bacteriol.* 178:3736-3741.
- Krauss, J. H., Weckesser, J., and Mayer, H. 1988. Electrophoretic analysis of lipopolysaccharides of purple nonsulfur bacteria. *Int. J. Syst. Bacteriol.* 38:157-163.
- Kuo, M.-S. K., and Mort, A. J. 1986. Location and identity of acyl substituents on the extracellular polysaccharides of *R. trifolii* and *R. leguminosarum*. *Carbohydr. Res.* 145:247-265.
- Kyte, J., and Doolittle, R. F. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.
- Leigh, J. A., Reed, J. W., Hanks, J. F., Hirsch, A. M., and Walker, G. C. 1987. *Rhizobium meliloti* mutants that fail to succinylate their calcofluor-binding exopolysaccharide are defective in nodule invasion. *Cell* 51:579-587.

- Maniatis, T. A., Fritsch, E. F., Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Marsh, J. L., Erfle, M., and Wykes, E. J. 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. *Gene* 32:481-485.
- McNeil, M., Darvill, J., Darvill, A., Albersheim, P., van Veen, R., Hooykaas, P., Schilperoort, R., and Dell, A. 1986. The discernible structural features of the acidic exopolysaccharides secreted by different *Rhizobium* species are the same. *Carbohydr. Res.* 146:307-326.
- Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J. Bacteriol.* 149:114-122.
- Müller, P., Keller, M., Weng, W. M., Quandt, J., Arnold, W., and Pühler, A. 1993. Genetic analysis of the *Rhizobium meliloti* *exoYFQ* operon: ExoY is homologous to sugar transferases and ExoQ represents a transmembrane protein. *Mol. Plant-Microbe Interact.* 6:55-65.
- Needleman, S. B., and Wunsch, C. D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443-453.
- Orgambide, G., Philip Hollingsworth, S., Cargill, L., and Dazzo, F. 1992. Evaluation of acidic heteropolysaccharide structures in *Rhizobium leguminosarum* biovars altered in nodulation genes and host range. *Mol. Plant-Microbe Interact.* 5:484-488.
- Pees, E., Wijffelman, C. A., Mulders, I., van Brussel, A. A. N., and Lugtenberg, B. J. J. 1986. Transposition of Tn1831 to Sym plasmids of *Rhizobium leguminosarum* and *Rhizobium trifolii*. *FEMS Microbiol. Lett.* 33:165-171.
- Philip-Hollingsworth, S., Hollingsworth, R. I., and Dazzo, F. B. 1989a. Host-range related structural features of the acidic extracellular polysaccharides of *Rhizobium trifolii* and *Rhizobium leguminosarum*. *J. Biol. Chem.* 264:1461-1466.
- Philip-Hollingsworth, S., Hollingsworth, R. I., Dazzo, F. B., Djordjevic, M. A., and Rolfe, B. G. 1989b. The effect of interspecies transfer of *Rhizobium* host-specific nodulation genes on acidic polysaccharide structure and *in situ* binding by host lectin. *J. Biol. Chem.* 264: 5710-5714.
- Priem, W. J. E., and Wijffelman, C. A. 1984. Selection of strains cured of the *Rhizobium leguminosarum* Sym plasmid pRL1J1 by using small bacteriocin. *FEMS Microbiol. Lett.* 25:247-251.
- Reed, J. W., Capage, M., and Walker, G. C. 1991. *Rhizobium meliloti* *exoG* and *exoJ* mutations affect the ExoX-ExoY system for modulation and exopolysaccharide production. *J. Bacteriol.* 173:3776-3788.
- Reuber, T. L., and Walker, G. C. 1993a. The acetyl substituent of succinoglycan is not necessary for alfalfa invasion by *Rhizobium meliloti* Rm1021. *J. Bacteriol.* 175:3653-3655.
- Reuber, T. L., and Walker, G. C. 1993b. Biosynthesis of succinoglucan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* 74:269-280.
- Robertson, B. K., Åman, P., Darvill, A. G., McNeil, M., and Albersheim, P. 1981. Host-symbiont interactions. V. The structure of acidic extracellular polysaccharides secreted by *Rhizobium leguminosarum* and *Rhizobium trifolii*. *Plant Physiol.* 67:389-400.
- Rubens, C. E., Heggen, L. M., Haft, R. F., and Wessels, M. R. 1993. Identification of *cpsD*, a gene essential for type III capsule expression in group B streptococci. *Mol. Microbiol.* 8:843-855.
- Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Skorupska, A., Bialek, U., Urbanik-Sypniewska, T., and Van Lammeren, A. 1995. Two types of nodules induced on *Trifolium pratense* by mutants of *Rhizobium leguminosarum* bv. *trifolii* deficient in exopolysaccharide production. *J. Plant Physiol.* 147:93-100.
- Smith, T. F., and Waterman, M. S. 1981. Identification of the common molecular sequences. *J. Mol. Biol.* 147:195-197.
- Spaink, H. P., Okker, R. J. H., Wijffelman, C. A., Pees, E., and Lugtenberg, B. J. J. 1987. Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1J1. *Plant Mol. Biol.* 9:27-39.
- Srere, P. A. 1987. Complexes of sequential metabolic enzymes. *Ann. Rev. Biochem.* 56:89-124.
- Stingle, F., Neeser, J. R., and Mollet, B. 1996. Identification and characterization of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. *J. Bacteriol.* 178:1680-1690.
- Stormo, G. D. 1986. Maximizing gene expression. Pages 195-224 in: Translation Initiations. W. Reznikoff and L. Gold, eds. Butterworth, Boston.
- Surin, B. P., Watson, J. M., Hamilton, W. D. O., Economou, A., and Downie, J. A. 1990. Molecular characterization of the nodulation gene *nodT*, from two biovars of *Rhizobium leguminosarum*. *Mol. Microbiol.* 4:245-252.
- Tsai, C. M., and Frisch, C. E. 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal. Biochem.* 119:115-119.
- van Brussel, A. A. N., Planqué, K., and Quispel, A. 1977. The wall of *Rhizobium leguminosarum* in bacteroid and free-living forms. *J. Gen. Microbiol.* 101:51-56.
- van Workum, W. A. T., van Brussel, A. A. N., Tak, T., Wijffelman, C. A., and Kijne, J. W. 1995. Ethylene prevents nodulation of *Vicia sativa* ssp. *nigra* by exopolysaccharide-deficient mutants of *Rhizobium leguminosarum* bv. *viciae*. *Mol. Plant-Microbe Interact.* 8:278-285.
- Wang, L., Liu, D., and Reeves, P. R. 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J. Bacteriol.* 178:2598-2604.
- Wang, L., and Reeves, P. R. 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica* *rfbP* gene in O-antigen subunit processing. *J. Bacteriol.* 176:4348-4356.
- Wang, L., Romana, L. K., and Reeves, P. R. 1992. Molecular analysis of a *Salmonella enterica* group E1 *rfb* gene cluster: O antigen and the genetic basis of a major polymorphism. *Genetics* 130:429-443.
- Wilburn, W. F., and Lipman, D. J. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA* 80: 726-730.
- Yamazaki, M., Thorne, L., Mikolajczak, M., Armentrout, R. W., and Pollock, T. J. 1996. Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. *J. Bacteriol.* 178: 2676-2687.
- Zevenhuizen, L. P. T. M., and van Neerven, A. R. W. 1983. Surface carbohydrates of *Rhizobium* IV. Gel-forming capsular polysaccharide of *Rhizobium leguminosarum* and *Rhizobium trifolii*. *Carbohydr. Res.* 124:166-171.