Genetic Analysis of the *Rhizobium meliloti exoYFQ* Operon: ExoY is Homologous to Sugar Transferases and ExoQ Represents a Transmembrane Protein

P. Müller, M. Keller, W. M. Weng, J. Quandt, W. Arnold, and A. Pühler

Lehrstuhl für Genetik, Fakultät fur Biologie, Universität Bielefeld, Postfach 100131, D-4800 Bielefeld 1, Germany.
Received 3 September 1992. Accepted 3 November 1992.

The nucleotide sequence of a 4.8-kb *Cla*I-*Eco*RI DNA fragment of megaplasmid 2 of *Rhizobium meliloti* Rm2011 involved in succinoglucan (EPS I) synthesis and nodule infection was determined. Four open reading frames (ORFs) were identified on this fragment. A mutational analysis revealed that these ORFs represent genes that were termed *exoX, exoY, exoF, and exoQ*. The locations of transposon insertions in these *exo* genes were determined at the nucleotide level. Plasmid integration mutagenesis revealed that the genes *exoY, exoF,* and *exoQ* are organized in an operon. The *exoX* gene running in opposite direction forms a monocistronic transcriptional unit. The *exoX* gene was shown to negatively influence the amount of EPS I synthesized. The *exoY* gene is coding for a membrane associated protein homologous to the C-terminal part of the *Xanthomonas campestris* glucosyltransferase GumD and the *Salmonella typhimurium* galactose transferase RfbP. ExoF, a probable periplasmatic protein, is nearly identical to the protein encoded by ORF1 of *Rhizobium* sp. strain NGR234. ExoQ is most probably a membrane associated protein as deduced by its hydrophobic structural features. All three genes of the *exoYFQ* operon were shown to be essential for succinoglucan synthesis and nodule infection.

Additional keywords: exopolysaccharide, galactosyltransferase, glucosyltransperase, operon analysis, symbiosis.

The occurrence of nitrogen-fixing root nodules on leguminous plants is the result of a complex symbiotic interaction between the host plant and soil bacteria of the genera *Rhizobium* and *Bradyrhizobium* (Djordjevic *et al.* 1987; Long 1989). Considerable effort has been devoted to the understanding of the mechanisms by which a *Rhizobium* strain and its host might recognize each other, and how the bacteria invade the induced nodules (Halverson and Stacey 1986; Long 1989). For various *Rhizobium* species it was found that exopolysaccharides (EPS) play a role in nodule development. Mutants deficient in EPS synthesis still had retained the ability to induce the formation of root nodules, but they did not penetrate or colonize the central nodule tissue (Chakravorty *et al.* 1982; Chen *et al.* 1985; Finan *et al.* 1985; Müller *et al.* 1988b; Borthakur *et al.* 1988).

The strain SU47 of *Rhizobium meliloti* Dangeard and its derivatives Rm2011 and Rm1021 produce an acidic exopolysaccharide (succinoglucan, EPS I) that can be stained with the dye Calcofluor white (Cfw) (Leigh *et al.* 1985; Müller *et al.* 1988b). Genetic studies in combination with a plate test using Cfw have led to the identification of a number of loci required for EPS I production. Mutations affecting EPS I synthesis were found to be clustered in a 22-kb DNA region of the *R. meliloti* megaplasmid 2 and fell into 13 complementation groups (Long *et al.* 1988; Reuber *et al.* 1991). For example, strains carrying mutations in complementation groups *exoB, exoF,* and *exoQ* failed to produce EPS I and were only able to induce ineffective nodules on alfalfa (Fix^-). Mutants of the adjacent complementation groups *exoG* and *exoJ* produced less EPS I than the wild-type strain and appeared to be symbiotically less efficient than the wild type (Long *et al.* 1988). Recently, Reed *et al.* (1991) published the nucleotide sequence of the *exoG*/*exoJ* region (Fig. 1) and showed the existence of two genes highly homologous to *exoX* and *exoY* of *Rhizobium* sp. strain NGR234. These genes were therefore also termed *exoX* and *exoY*. The *exoJ* gene was shown be an allele of *exoX* and the *exoG* locus is due to insertions in the intergenic region between *exoX* and *exoY* (Reed *et al.* 1991). Sequence analysis revealed that *exoX* and *exoY* are homologous to *psi* and *pss* of *Rhizobium leguminosarum* bv. *phaseoli* Jordan (Borthakur and Johnston 1987; Borthakur *et al.* 1988), respectively (Gray *et al.* 1990; Reed *et al.* 1991). Reed *et al.* (1991) and Zhan *et al.* (1990) showed that the corresponding *exoX* and *exoY* mutants revealed a similar phenotype as *psi* and *pss* mutants. Multiple copies of *psi* or *exoX* reduced the EPS production in the corresponding *Rhizobium* species, whereas mutations in *pss* or *exoY* both resulted in an EPS^- phenotype correlated with defects in symbiosis (Borthakur *et al.* 1985, 1986; Gray *et al.* 1990; Zhan *et al.* 1990).

Corresponding author: M. Keller.
Present address of P. Müller: Fachbereich Biologie/Botanik, Phillips-Universität, Karl-von-Frisch-Str., D-3500 Marburg-Lahn, FRG.
Present address of W. M. Weng: Shanghai Institute of Plant Physiology, Academica Sinica, Shanghai 200032, P.R. China.
Present address of J. Quandt: Faculty of Science, Dept. of Biological Sciences, University of Calgary, Calgary, Alberta, Canada T2N 1N4.
Nucleotide and/or amino acid sequence data have been submitted to GenBank, EMBL, and DDBJ as accession number L05588.

MPMI Vol. 6, No. 1, 1993, pp. 55-65

We have also isolated *R. meliloti* mutants impaired in the infection process (Inf⁻ mutants) and have shown that one class of these mutants was defective in EPS I synthesis (Müller *et al.* 1988b). The mutations were found to be located on a 7.8-kb *Eco*RI DNA fragment of megaplasmid 2 (Hynes *et al.* 1986) that later was shown to be part of the 22-kb *exo* region described by Long *et al.* (1988) (Pühler *et al.* 1988). By transposon mutagenesis of the 7.8-kb *Eco*RI fragment we identified several regions important for EPS I synthesis and nodule infection (Keller *et al.* 1988). By the use of transposon Tn*5*-B20, which carries a promoterless *lacZ* gene we could determine the transcriptional direction of these regions.

In this report we analyzed of a 4.8-kb DNA region of the 7.8-kb *Eco*RI DNA fragment carrying genes involved in EPS I production and symbiosis. We present the nucleotide sequence and genetic analysis of this DNA fragment, which completes the DNA sequence between the previously sequenced genes *exoY* (Reed *et al.* 1991) and *exoZ* (Buendia *et al.* 1991).

RESULTS

Sequence analysis of a 4.8-kb DNA region of *R. meliloti* Rm2011 revealed four open reading frames termed *exoX, exoY, exoF,* and *exoQ*.

By Tn*5* and Tn*5*-B20 induced mutations in a 7.8-kb *Eco*RI DNA fragment of megaplasmid 2 we previously identified loci involved in EPS I synthesis and nodule infection (Fig. 1B; Müller *et al.* 1988b; Keller *et al.* 1988). To get further information concerning these regions we sequenced a 4.8-kb *Cla*I-*Eco*RI DNA fragment, located at the right end of the 7.8-kb *Eco*RI fragment (Fig. 1A). Four open reading frames (ORF) could be identified on the sequenced DNA fragment. As will be shown later, all of these ORFs are important for the biosynthesis of the exopolysaccharide EPS I. They therefore represent genes (Fig. 1C). Because of sequence identity to the genes *exoX* and *exoY* of the closely related *R. meliloti* strain Rm1021 (Reed *et al.* 1991), the first two genes were also termed *exoX* and *exoY*. The other two genes presumably comprise the complementation groups *exoF* and *exoQ* described by Long *et al.* (1988) and were termed accordingly. The complete nucleotide sequence of this 4,827-bp *Cla*I-*Eco*RI DNA fragment is presented in Figure 2.

The small, 294-bp coding region located on the left side of the sequenced fragment is identical to the *exoX* gene of *R. meliloti* Rm1021 (Reed *et al.* 1991). In mutant Rm124 (Keller *et al.* 1988), the *exoX* gene is mutated in its 3′ end by a Tn*5*-B20 insertion as was shown by sequencing the leftward junction of the transposon insertion (Fig. 2). The direction of transcription of the transposon Tn*5*-B20 encoded promoterless *lacZ* gene of mutant Rm124 (Keller *et al.* 1988; Fig. 1B) confirmed the direction of transcription of *exoX*. Upstream of the *exoX* gene a putative promoter resembling the σ^{70} promoter of *E. coli* (McClure 1985) was identified (Fig. 2). The −35 region of this promoter shows four out of six matches with the consensus sequence, the −10 region reveals five out of six possible matches.

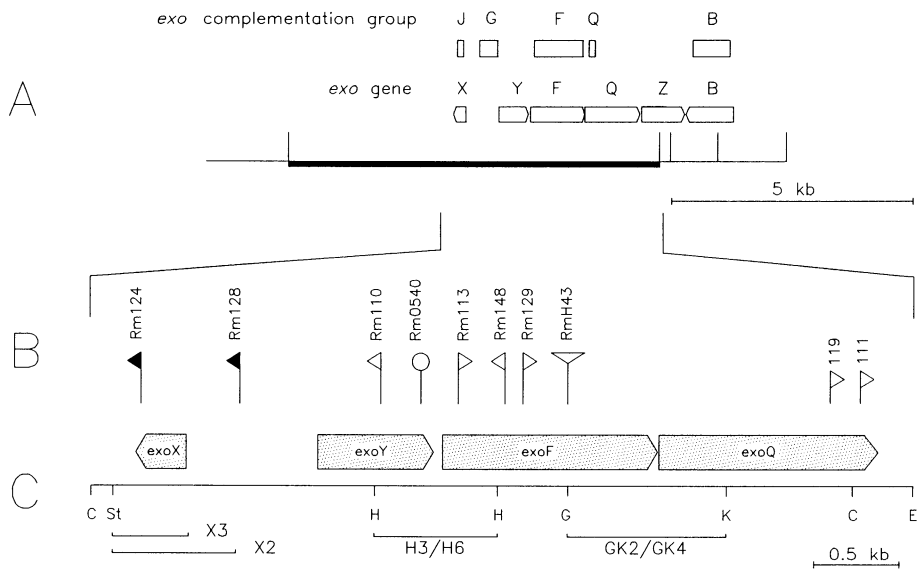

Fig. 1. Genetic and physical map of a DNA region of megaplasmid 2 of *Rhizobium meliloti* 2011 carrying genes for EPS I synthesis and nodule infection. **A,** The *Eco*RI restriction map of a 13-kb DNA region of megaplasmid 2 is shown. On the top of the figure the complementation groups published by Long *et al.* (1988) are shown. The map of *exo* genes published by Buendia *et al.* (1991), Reed *et al.* (1991) and identified in this work are presented below. The previously described 7.8-kb DNA *Eco*RI fragment (Müller *et al.* 1988b) is indicated by a black bar. **B,** Insertion sites of different genetic elements in the 4.8-kb *Cla*I-*Eco*RI DNA fragment are indicated. The insertion of the interposon of mutant RmH43 is marked by a triangle. The locations of the published Tn*5* (pin) and Tn*5*-B20 (flags) insertions and the designation of the corresponding *R. meliloti* mutants are shown. For the Tn*5*-B20 insertions 111 and 119 (small flags) no corresponding *R. meliloti* mutant could be isolated. The flags indicate the transcription direction of the *lacZ* gene of Tn*5*-B20. Black flags indicate an EPS I^+Inf^+ phenotype of the corresponding strain, white markers indicate an EPS I^- Inf^- phenotype. **C,** The restriction map of the sequenced 4.8-kb DNA fragment only containing relevant sites is presented. Above the map the identified coding regions are indicated by bars. Subfragments used for genetic analyses are indicated. They were termed X2, X3, H3/H6, and GK2/GK4. Abbreviations of restriction enzymes: C, *Cla*I; E, *Eco*RI; G, *Bgl*II; H, *Hin*dIII; K, *Kpn*I; St, *Stu*I.

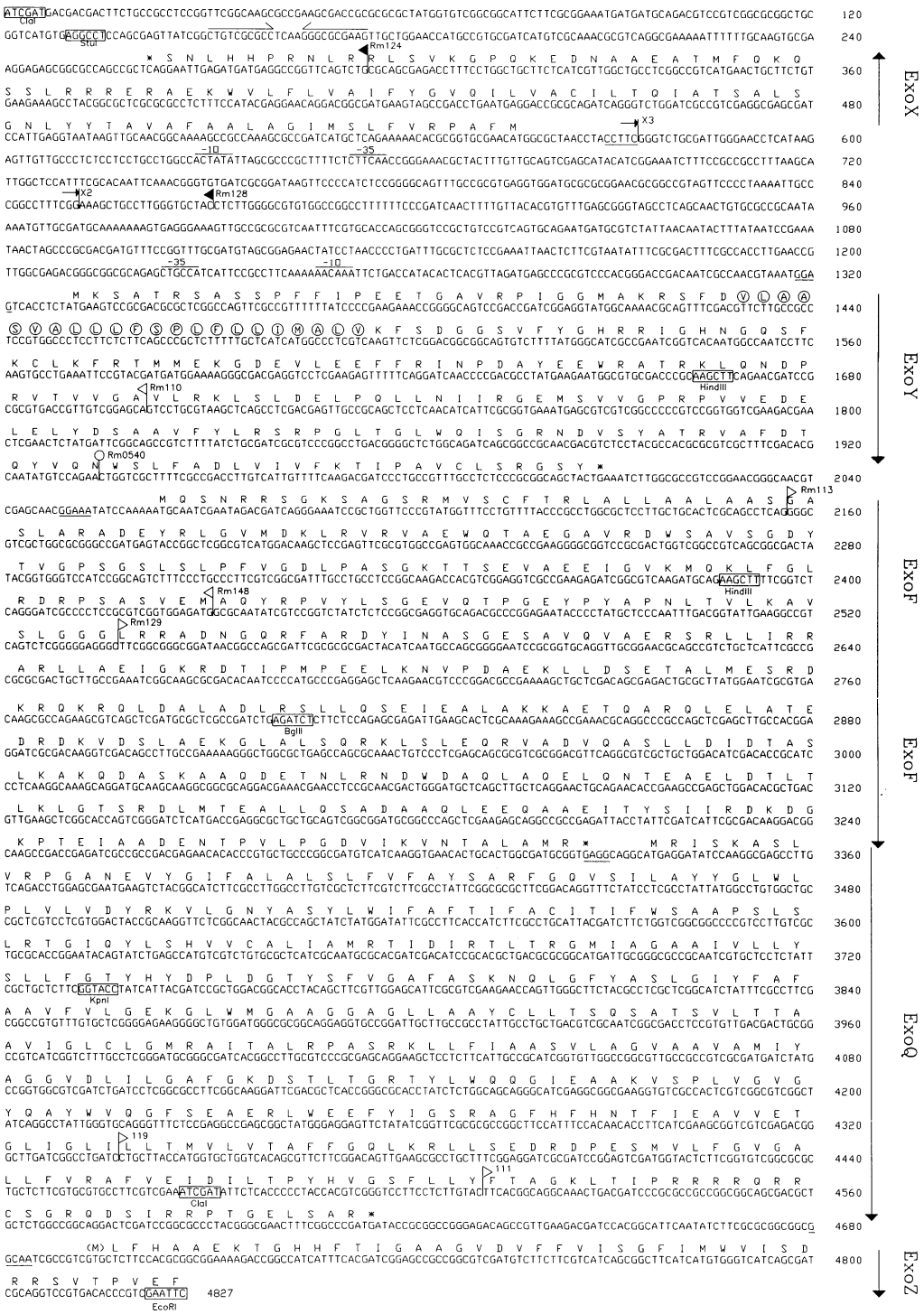

Fig. 2. Nucleotide sequence of a 4.8-kb DNA fragment of the *R. meliloti* megaplasmid 2 containing the genes *exoX, exoY, exoF,* and *exoQ.* The nucleotide sequence of one strand is presented in 5′-3′ direction. The four identified coding regions are indicated by the deduced amino acid sequence written above the nucleotide sequence; asterisks mark stop codons. The arrows at the right indicate the direction of transcription and extension of the genes. Potential ribosome binding sites are underlined, potential σ^{70} promoters are overlined. The numbered pin and flags indicate positions of the leftward junctions of Tn*5* and Tn*5*-B20 insertions, respectively. The flags point in the transcription direction of the promoterless *lacZ* gene. Black flags indicate an EPS I^+Inf^+ phenotype, open symbols indicate an EPS I^-Inf^- phenotype of the corresponding mutant (with the exception of insertions 119 and 111). The endpoints of two subclones (X2 and X3) are marked. The arrows denote a potential stem-loop structure. The hydrophobic amino acid residues of ExoY potentially forming a transmembrane segment are circled. Relevant restriction enzyme sites used for subcloning are boxed.

A putative ribosome binding site for *exoX* was also identified (Fig. 2).

The transposon insertion sites of the *R. meliloti* EPS I⁻ mutants Rm110 and Rm0540 (Keller *et al.* 1988; Müller *et al.* 1988b) are located in the *exoY* coding region, which is 678-bp in size and located more than 700 bp upstream of the *exoX* gene (Fig. 2). The identified *exoY* gene is identical to the partially sequenced *exoY* gene of the strain *R. meliloti* Rm1021 (Reed *et al.* 1991). Upstream of *exoY* we identified sequences with some similarity to the consensus sequences of an *E. coli* σ^{70} promoter (four and three out of six matches with the −35 and the −10 region of the consensus sequence, respectively; Fig. 2). The divergently transcribed *exoX* and *exoY* coding regions are separated by a DNA region of 772-bp with no coding probability (Fig. 2). The Tn*5*-B20 insertion of *R. meliloti* mutant Rm128 located 458-bp upstream the start codon of *exoY* caused a reduction of EPS I production to 30% when compared with the wild-type strain Rm2011 (Keller *et al.* 1988). Therefore, this intergenic region may carry signal sequences for the expression of *exoY*.

At a distance of 52-bp downstream of *exoY* the start codon of the *exoF* coding region is located, which is 1,262-bp in size. In the EPS I⁻ Inf⁻ *R. meliloti* strains Rm113, Rm148, and Rm129 (Keller *et al.* 1988) the *exoF* coding region is mutated by Tn*5*-B20 insertions. By inserting a gentamicin resistance cassette into the single *Bgl*II-site, we obtained the additional EPS I⁻ Inf⁻ mutant RmH43 (Fig. 2). The fourth identified coding region, 1,278-bp long and designated *exoQ,* starts only 7-bp downstream of *exoF*. Two Tn*5*-B20 insertions in *exoQ* (119 and 111; Fig. 1B) were isolated by mutating plasmid pRmPM157.8 in *E. coli* (Keller *et al.* 1988). All attempts to construct the corresponding *R. meliloti exoQ* mutant failed, possibly indicating that some *exoQ* mutants are not viable. No known promoter sequences could be found upstream of *exoF* or *exoQ* supporting the hypothesis that *exoY, exoF,* and *exoQ* form one transcriptional unit.

The putative gene products of the *R. meliloti exoX, exoY,* and *exoF* genes are homologous to other bacterial proteins involved in polysaccharide biosynthesis.

When we carried out a homology search in the EMBL databank with the deduced amino acid sequences of the four coding regions, we found homology to sequences from several bacteria.

The protein encoded by the *R. meliloti exoX* gene was found to be identical for 73% to ExoX of *Rhizobium* sp. strain NGR234 (Gray *et al.* 1990).

We also found a significant homology to the protein encoded by *exoY* to ExoY of *Rhizobium* sp. strain NGR234 (Gray *et* al. 1990; Fig. 3). The homology of the first part of the *R. meliloti* ExoY with the first part of the *Rhizobium* sp. strain NGR234 ExoY was already reported by Reed *et al.* (1991), but their published sequence ended at the *Hin*dIII site in *exoY* (Fig. 2). We have sequenced further downstream and found that the two ExoY proteins are homologous (84% identity) over the entire length of the proteins. The *R. meliloti* ExoY protein consists of 226 amino acids with a molecular mass of 27.4 kDa. Reed *et al.* (1991) and previously Müller *et al.* (1988a) revealed also a significant homology of ExoY to Pss2, which is essential for EPS synthesis in *R. leguminosarum* bv. *phaseoli* (Borthakur *et al.* 1988; Fig. 3). A possible function of ExoY was proposed by Reed *et al.* (1991), since they found homology of ExoY with the C-terminal part of GumD of *Xanthomonas campestris.* GumD is a glucosyltransferase catalyzing the first step of xanthan biosynthesis (Coplin and Cook 1990). The function of ExoY as a sugar transferase is strengthened by a striking homology (51% identity) that exists between the *R. meliloti* ExoY and the C-terminal part of the RfbP protein of *Salmonella thyphimurium.* RfbP is a galactosyltransferase, catalyzing the transfer of UDP-galactose to the C55 lipid carrier (Jiang *et al.* 1991). Several amino acid residues are conserved in all five proteins (marked in Fig. 3), possibly indicating the active sites of the enzymes.

The hydrophobicity plots of ExoY and Pss2 exhibited a nearly identical structure (Keller *et al.* 1990). The hydrophobic regions of these proteins indicate that they are membrane associated. This is further sustained since the first hydrophobic region of ExoY (circled in Fig. 2) is well also conserved in RfbP, where this region was shown

```
RfbP         MDNIDNKYNPQLCKIFLAISDLIFFNLALWFSLGCVYFIFDQV-QRFIPQDQLD    53
GumD    MLLADLSSATYTTSSPRLLSKYSAAADLVLRVFDLTMVVASGLIAYRIVFGTWVPAAPYR    60

RfbP    TRVITHFILSVVCVGWFWIRLRHYTIRKPFWYELKEIFRTI-VIFAIFDL-ALIAFTKWQ    111
GumD    VAIATTLLYSVICFALF--PLYRSWRGRGLLSELVVLGGAFGGVFALFAVHALIVQVGEQ    118

RfbP    FSRYVWVFCWTFALILVPFF--RALTKHLLNKL---GIWKKKTIILGSGQNARGAYSALQ    166
GumD    VSR-GWVGLW-FVGGLVSLVAARTLLRGFLNHLRTQGVDVQRVVVVGLRHPVMKISHYLS    176

RfbP    SEEMMGFDVIAFFDTD---ASDAEINMLPVIKDTEIIWDLNRTGDVHYI-LAYEYTELEK    222
GumD    RNPWVGMNMVGYFRTPYDLAVAEQRQGLPCLGDPDELIEYLKNNQVEQVWISLPLGERDH    236

RfbP    THFWLRELSKHHCRSVTVVPSFRGLPLYNTDMSFIFSHEVMLLRIQNNLAKRSSRFL-KR    281
GumD    IKQLLQRLDRYP-INVKLVPDLFVFGLLNQSAEQIGSVPVINLR-QGGVDRDNYFVVAKA    294
ExoYn                                MKSATRSATTAFFIPQETGAIRPIGG-ISKR     30
ExoY                                 MKSATRSASSPFFIPEETGAVRPIGG-MAKR     30
Pss2                                                        MDL-VLKR      7

RfbP    TFDIVCSIMILIIASPLMIYLWYKVT-RDGGPAIYGHQRVGRHGKLFPCYKFRSMVMNSQ    340
GumD    LQDKILAVIALMGLWPLMLAIAVGVKMSSPGPVFFRQRRHGLGGREFYMFKFRSM-----    349
ExoYn   SFDVLIAILALIALSPLFLLVMGLVKFSDGGSIFYGHRRIGHNGQTFKCLKFRTMMENGD     90
ExoY    SFDVLAASVALLLFSPLFLLIMALVKFSDGGSVFYGHRRIGHNGQSFKCLKFRTMMEKGD     90
Pss2    AFDIFSSLSALLVLAPFLLFVALLIKLDSPGPVLFKQTRWGKNCKAIKVYKFRSM-----     62

RfbP    EVLKELLANDPIARAEWEKDFKLKNDPRITAVGRFIRKTSLDELPQLFNVLKGDMSLVGP    400
GumD    -------RVHDDHGTTIQQAT--KNDTRITRFGSFLRRSSLDELPQIFNVLGGSMSIVGP    400
ExoYn   RVLQEFFKSNPAAYEEWRTTRKLQDDPRVTVVGSVLRKLSLDELPQLLNIIRGEMSIVGP    150
ExoY    RVLEEFFRINPDAYEEWRATRKLQNDPRVTVVGAVLRKLSLDELPQLLNIIRGEMSVVGP    150
Pss2    ---------RTDLCDVSGVAQTVKNDPRITRIGAILRRTNVDELPQLLNVLLGHMSVVGP    113

RfbP    RP---IVSDELERYCDDVDYYLMA---KPGMTGLWQVSG-RNDV----DYDTRVYFDSWY    449
GumD    RP---HAAQHNTHYEKLINHYMQRHYVKPGITGWAQVNGFRGETPELRTMKKRIQYDLDY    457
ExoYn   RP---VVEDELELYDSAAEFYL-RS--RPGLTGLWQISG-RNDV----SYATRVAFDTHY    199
ExoY    RP---VVEDELELYDSAAVFYL-RS--RPGLTGLWQISG-RNDV----SYATRVAFDTQY    199
Pss2    RCHAIGMRAGGMLYEELVPEYHQRHAMRPGMTGLAQMRGLRGPTDRPAKARARIASDLYY    173

RfbP    VKNWTLWNDIAILFKTAKVVLRRDGAY      476
GumD    IRRWSLWLDIRIIVLTAVEVLGQKTAY      483
ExoYn   VQNWSLLADLVIVFKTIPAVCLSRGSY      226
ExoY    VQNWSLFADLVIVFKTIPAVCLSRGSY      226
Pss2    VGNFSIVMDMRIIFGTVVS-ELTRGKGF     200
```

Fig. 3. Alignment of the deduced amino acid sequences of the *Rhizobium meliloti* ExoY protein to amino acid sequences of other bacterial proteins. The proteins used for this alignment are RfbP: *Salmonella thyphimurium* RfbP (Jiang *et al.* 1991), GumD: *Xanthomonas campestris* GumD (Reed *et al.* 1991), ExoYn: *Rhizobium* sp. strain NGR234 ExoY (Gray *et al.* 1990), ExoY: *R. meliloti* ExoY (this work) and Pss2: *R. leguminosarum* bv. *phaseoli* Pss2 (Borthakur *et al.* 1987). Quote marks indicate identical amino acid residues in the N-termini of GumD and RfbP, asterisks indicate amino acid residues identical in all five proteins.

to be a transmembrane segment (Jiang *et al.* 1991).

The *exoF* gene located downstream of *exoY* codes for a putative protein of 45.8-kDa. The *R. meliloti* ExoF protein showed 83% identity with the protein encoded by ORF1 of *Rhizobium* sp. strain NGR234 previously published by Gray *et al.* (1990) (Fig. 4). Only 312 amino acids encoded by ORF1 of *Rhizobium* sp. strain NGR234 have been published (Gray *et al.* 1990), but we assume that the C-terminal amino acids encoded by the *R. meliloti exoF* are also encoded by ORF1 of *Rhizobium* sp. strain NGR234. When we analyzed the amino acid sequence of ExoF for its hydrophobicity (according to Eisenberg *et al.* 1984), we identified a hydrophobic region in the N-terminus preceded by a positively charged region (Fig. 4). These features have been reported for signal peptides (von Heijne

```
                ++ +  +
ExoF   MQSNRRSGKSAGSRMVSCFTRLALLAALAASGASLARADEYRLGVMDKLRVRVAEWQTAE  60
       ||| | |.|||| . |.||||||| ||||.||||.||||||||||||||||||||
ORF1   MQSIRVLGAPAGSRKLLHFARLALCAALVVSGAAVARADDYRLGVMDKLRVRVAEWQTAE  60

ExoF   GAVRDWSAVSGDYTVGPSGSLSLPFVGDLPASGKTTSEVAEEIGVKMQKLFGLRDRPSAS  120
       ||||||||||.||||.|||.|||||||||||.||.|||||||.|||||||||||||||||
ORF1   GAVRDWSAVSGEYTVGASGSVSLPFVGDLPASGRTTTEVAEEIGIKMQKLFGLRDRPSAS  120

ExoF   VEMAQYRPVYLSGEVQTPGEYPYAPNLTVLKAVSLGGGLRRADNGQRFARDYINASGESA  180
       |||||||||||.|||.|||||||||||||||||||.|||||. .|||||||| |.|.|.
ORF1   VEMAQYRPVYLYGEVETPGEYPYAPNLTVLKAVSLSGGLRRGPTGQRFARDYIAANGDSS  180

ExoF   VQVAERSRLLIRRARLLAEIGKRDTIPMPEELKNVPDAEKLLDSETALMESRDKRQKRQL  240
       |||||| |||||||| |||.|.| | .|||||| | .|||.|||||| |||||| |||
ORF1   VQVAERNRLLIRRARLQAEIAKHDKIELPEELKNAPGVDKLLESETALMVSRDKRQDRQL  240

ExoF   DALADLRSLLQSEIEALAKKAETQARQLELATEDRDKVDSLAEKGLALSQRKLSLEQRVA  300
       ||||||.|||||||.|||||||| |||||| ||||||||||||||||||||||||||||
ORF1   DALADLKSLLQSEIESLAKKAETQQRQLELAMEDRDKVDSLAEKGLALSQRKLSLEQRVA  300

ExoF   DVQASLLDIDTASLKAKQDASKAAQDETNLRNDWDAQLAQELQNTEAELDTLTLKLGTSR  360
       |||. ||||||
ORF1   DVQSQLLDIDT  (311)

ExoF   DLMTEALLQSADAAQLEEQAAEITYSIIRDKDGKPTEIAADENTPVLPGDVIKVNTALAMR 421
```

Fig. 4. Alignment of the deduced amino acid sequence of the *Rhizobium meliloti* ExoF protein with that of the protein encoded by ORF1 of *Rhizobium* sp. strain NGR234 (Gray *et al.* 1990). Identical amino acid residues are marked by vertical lines, similar amino acid residues are indicated by dots. The hydophobic domain of the potential signal peptide is overscored in ExoF, positively charged amino acid residues of the signal peptide of ExoF are marked by +.

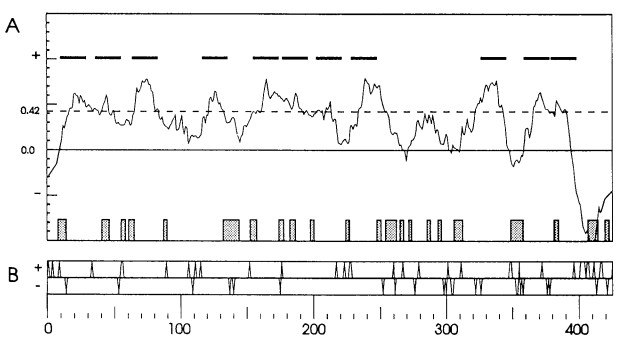

Fig. 5. Analysis of the *Rhizobium meliloti* ExoQ protein as derived from the nucleotide sequence. **A,** Hydropathic profile according to Eisenberg *et al.* (1984) calculated with a window length of 19 amino acids residues. The plot is divided into hydrophobic (+) and hydrophilic (−) regions. Regions with a mean hydrophobicity ≥0.42, indicated by the dotted line, are marked by black bars. Stippled boxes indicate possible β-turns (Chou and Fasman 1978). **B,** Distribution of charged amino acids. The values on the horizontal axis represent amino acid residue positions.

1985). By an analysis following the procedure proposed by von Heijne (1986) we found that the −13/+2 region of ExoY, comprising amino acid 17–31, has a score of 5.5, which is significant for signal peptides. Therefore, it is reasonable to propose that ExoF is a periplasmic protein.

The fourth *R. meliloti* protein ExoQ (46.5 kDa) showed no significant homology to any protein sequence held in the EMBL database. We found only weak homology over the entire length of ExoQ to membrane proteins. Correspondingly, hydrophobicity analysis revealed that ExoQ is hydrophobic with the exception of the positively charged C-terminus (Fig. 5). Several regions of high hydrophobicity, each comprising about 20 amino acids, alternated with regions of lower hydrophobicity. The distribution of hydrophobic regions alternate with β-turns (Fig. 5), which leads us to postulate 11 membrane spanning regions for the ExoQ protein (marked by bars in Fig. 5) following the proposal of Jähnig (1990).

The *R. meliloti exoY, exoF,* and *exoQ* genes are organized in one operon running in opposite direction to the monocistronic *exoX* transcriptional unit.

The *exoX* coding region is reading divergently from the three other identified *exo* coding regions, possibly defining another transcriptional unit. Downstream of *exoX* a possible stem-loop structure with a free energy of −71.9 kJ/M (Tinoco *et al.* 1973) could be identified (Fig. 2), which may serve as a transcription termination signal. Hence,

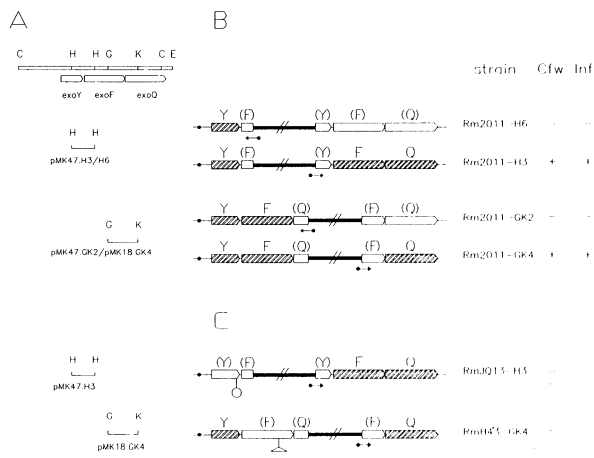

Fig. 6. Plasmid integration mutagenesis to test whether *exoY, exoF,* and *exoQ* form one operon. **A,** The coding regions *exoY, exoF,* and *exoQ* are indicated by arrows. Restriction enzymes are abbreviated as in Figure 1. The DNA fragments shown were cloned into mobilizable vector plasmids in both orientations and used for integration mutagenesis. **B,C,** The genomic structures resulting from different plasmid integrations via single crossing over are presented. The vector is marked by a thick line, coding regions are indicated by boxes and arrows. Abbreviations: Y, *exoY*; F, *exoF*; Q, *exoQ*. Designations of inactivated genes are given in brackets. Hatched arrows indicate intact genes transcribed from promoters located in front of the respective genes. A dot marks the putative *exoY* promoter, a dot with an arrow indicates the plasmid encoded *lacZ* promoter and its transcription direction. The Calcofluor white (Cfw) and the symbiotic phenotype (Inf for nodule infection) of the *Rhizobium meliloti* strains carrying the corresponding plasmid integrated into the genome are listed on the right side. The pin marks the Tn*5* insertion in *exoY* of strain RmJQ13, the triangle the resistance cassette insertion in *exoF* of strain RmH43.

we suppose that *exoX* forms a monocistronic transcriptional unit. This is supported by a Tn*5*-B20 insertion downstream of the *exoX* gene, resulting in mutant Rm154 (Keller *et al.* 1988). In contrast to the EPS I⁻ overproducing *exoX* mutant Rm124 is mutant Rm154 completely EPS I⁻ (Keller *et al.* 1988).

To determine if the coding regions *exoY, exoF,* and *exoQ* form one operon, we carried out a plasmid integration mutagenesis. Therefore, we cloned two subfragments of the 4.8-kb *Cla*I-*Eco*RI DNA region into mobilizable plasmid vectors which cannot replicate in *R. meliloti* (Fig. 6). Selection for the vector-encoded neomycin resistance resulted in integration of the plasmid into the *R. meliloti* genome by a single crossing over and caused a disruption of the operon structure. Transconjugants carrying the integrated plasmid were tested for their Calcofluor phenotype and their symbiotic properties.

Plasmids pMK47.H3 and pMK47.H6, respectively, carry a 724-bp *Hin*dIII DNA fragment with the 3′ end of *exoY*, the intergenic region and the 5′ end of *exoF* in both orientations (Fig. 1C). These plasmids were integrated into the Rm2011 genome via a single crossing over. The resulting strains, Rm2011-H3 and Rm2011-H6, showed different phenotypes, depending on the orientation of the plasmid encoded *lacZ* promoter (Fig. 6). When the *lacZ* promoter was reading in the same orientation as the coding regions a wild-type phenotype of the resulting strain was observed. When the *lacZ* promoter was reading in the opposite direction, the resulting mutant did not fluoresce on Calcofluor-containing plates and was not able to nodulate alfalfa effectively (Fig. 6). These results show clearly that *exoF* has no own promoter and is therefore transcribed by the *exoY* promoter. Additionally, these results demonstrate that the *E. coli lacZ* promoter can mimic an *R. meliloti exo* promoter.

Analogous experiments were carried out with the 931-bp *Bgl*II-*Kpn*I DNA fragment carrying the 3′ end of *exoF*, the intergenic region and the 5′ end of *exo*Q (Fig. 1C). Cloning of this 931-bp DNA fragment in both orientations resulted in plasmids pMK18.GK4 and pMK47.GK2. After integration of these two plasmids into the *R. meliloti exo* region via a single crossing over, the two different strains Rm2011-GK4 and Rm201a-GK2 were obtained (Fig. 6). Mutant Rm2011-GK2 was EPS I⁻ Inf⁻ due to separation of *exo*Q from the *exoY* promoter. In contrast to mutant Rm2011-GK2, strain Rm2011-GK4 was EPS I⁺ Inf⁺. In this strain, the plasmid encoded *lacZ* promoter caused transcription of *exoQ* (Fig. 6B). These results showed that no promoter is located between *exoF* and *exoQ* and that the *lacZ* promoter is able to transcribe *exoQ*. With this plasmid integration mutagenesis we could demonstrate that *exoF* and *exoQ* do not have their own promoters and therefore form a transcriptional unit (operon) with *exoY*. In addition, we could demonstrate by strain Rm2011-GK2 that *exoQ* is necessary for EPS I synthesis. However, Rm2011-GK2 is the only existing mutant of *R. meliloti* Rm2011. We were not able to transfer the previously described Tn*5*-B20 insertions 111 and 119 (Keller *et al. 1988*) located in *exoQ* to the *R. meliloti* genome via double crossing-over.

The *R. meliloti* genes *exoY, exoF,* and *exoQ* are all essential for the synthesis of the exopolysaccharide EPS I.

We could demonstrate that *exoY, exoF,* and *exoQ* are arranged in one operon, which raises the question whether the EPS I⁻ Inf⁻ phenotype of the transposon induced *exoY* and *exoF* mutants is due to a polar effect of the inserted transposons. To solve this problem, we integrated plasmids pMK47.H3 and pMK18.GK4 via single crossing over into the genome of the EPS I− mutants RmJQ13 (*exoY*) and RmH43 (*exoF*), respectively (Fig. 6C). Both resulting strains RmJQ13-H3 and RmH43-GK4 were EPS I⁻, whereas integration of these plasmids into the genome of the wild-type Rm2011 did not affect EPS I production (Fig. 6B). Hence, we conclude that *exoY* and *exoF* are indispensable for EPS I synthesis. Together with the finding that also is needed for EPS I biosynthesis (see previous chapter), it can be summarized that all three genes *exoY exoF,* and *exoQ* organized in one operon are essential for EPS I production.

The *R. meliloti exoX* gene influences the amount of EPS I produced.

Previous experiments (Keller *et al.* 1988) showed that the *exoX* mutant Rm124 produces about three times more EPS I than the wild-type strain Rm2011. Therefore, we subcloned *exoX* and analyzed the influence of *exoX* on EPS I synthesis. Additionally, we wanted to obtain information about the approximate localization of the *exoX* promoter region, and used therefore two appropriate subfragments containing the coding region of *exoX* and upstream sequences. The two fragments were obtained by exonuclease III/S1 nuclease treatment and subsequent digestion with *Stu*I. The extension of the fragments are

Table 1. Effect of different copies of *exoX* on EPS I production in the wild-type strain *Rhizobium meliloti* Rm2011

Strain	Fragment[a] (vector type)[b]	Total no. of *exoX* coding regions	Functional copies of *exoX*	Fluorescence and mucoidity[c]
Rm2011	...	1	1	+
Rm2011-X2	X2 (integrative)	2	2	(+)
Rm2011-X3	X3 (integrative)	2	1	+
Rm2011 (pMK104.X2)	X2 (replicative)	~ 100	~ 100	−
Rm2011 (pMK104.X3)	X3 (replicative)	~ 100	1	+

[a] See Figures 1C and 2; X2 = *exoX* coding region with a potential promoter. X3 = *exoX* coding region without a potential promoter.
[b] Integrative: suicide vector (pSUP202 - basis; pMK202.X2 & pMK202.X3). Replicative: broad host range vector (pSUP104 - basis; pMK104.X2 & pMK104.X3).
[c] Fluorescence and mucoidity of colonies on M98 medium containing Cfw. + wild-type fluorescence, mucoid colonies; (+) slightly reduced fluorescence and mucoidity; − no fluorescence, dry colonies.

marked in Figures 1 and 2. Fragment X2 bears upstream of the start codon of *exoX* a DNA region of 294-bp with the putative σ^{70} promoter, fragment X3 bears only 17-bp upstream of the start codon. These two fragments were cloned into two different vectors, pSUP104 and pSUP202. Plasmid = pSUP104 (Priefer *et al.* 1985) is a broad host range vector with an approximate copy number of 100 in *R. meliloti* (Labes *et al.* 1990), whereas plasmid pSUP202 (Simon *et al.* 1983) is a suicide vector not able to replicate in *R. meliloti*. The resulting plasmids pMK104.X2, pMK104.X3, pMK202.X2, and pMK202.X3 were all transferred to *R. meliloti* Rm2011 (WT). The phenotypes of the resulting strains (Table 1) demonstrated that a high copy number of *exoX* suppressed completely the EPS I production in *R. meliloti*. Similar results have also been published by Zhan and Leigh (1990) and Reed *et al.* (1991) and for *Rhizobium* sp. NGR234 by Gray *et al.* (1990). Additionally, we found that *R. meliloti* strains with a high number of intact copies of *exoX* (plasmid pMK104.X2) are symbiotically ineffective (Inf$^-$) resembling EPS I$^-$ mutants of *R. meliloti*. A high copy number of *exoX* has therefore also negative influence on symbiosis.

Table 1 also shows that fragment X3 (plasmids pMK104.X3 and pMK202.X3) is lacking the *exoX* promoter since, in contrast to fragment X2 no reduction of EPS I production was observed after transfer to *R. meliloti*. These results let us draw the conclusion that the *exoX* promoter is located within the 294-bp fragment upstream the coding sequence of *exoX* resembling possibly the identified σ^{70} promoter marked in Figure 2.

DISCUSSION

The DNA fragment analyzed in this report is part of a large cluster of at least 13 *exo* genes of the *R. meliloti* megaplasmid 2 (Finan *et al.* 1985; Leigh *et al.* 1985; Long *et al.* 1988; Reuber *et al.* 1991). The sequenced 4.8-kb DNA fragment bears four *exo* coding regions which are all involved in EPS I synthesis and nodule infection, as was shown by several transposon and plasmid insertions resulting in EPS I$^-$ overproducing or EPS I$^-$ mutants, respectively. When we compared the restriction map of the *exo* gene region published by Long *et al.* (1988) with the map of the cosmid pRmPM551 (Müller *et al.* 1988a) complementing two EPS I$^-$ mutants of *R. meliloti* Rm2011 (Müller *et al.* 1988b), we found that the sequenced region spans the complementation groups *exoJ, exoG, exoF,* and *exoQ* (Fig. 1A). The complementation group *exoJ* was determined by Reed *et al.* (1991) to be an allele of *exoX*. For the *exoG* complemetation group no gene could be identified, but Reed *et al.* (1991) found that it corresponds to the intergenic region between *exoX* and *exoY*. The complementation group *exoF* (Long *et al.* 1988) matches in its extension and localization with the coding region of the gene *exoF*. In contrast, we found that the *exoQ* gene is much longer than the *exoQ* complementation group.

The sequenced region of *R. meliloti* Rm2011 showed 72% identity on DNA level to a 2,800-bp DNA fragment of *Rhizobium* sp. strain NGR234 recently published by Gray *et al.* (1990). Reed *et al.* (1991) revealed already the high homology (73% identity) of the *R. meliloti* ExoX with ExoX of *Rhizobium* sp. strain NGR234 (Gray *et al.* 1990). The *R. meliloti* ExoX was also found to have homology (24% identity) with Psi of *R. leguminosarum* bv. *phaseoli* (Borthakur and Johnston 1987). A potential functional relationship of the two ExoX proteins and the Psi protein was much more evident when their hydrophobicity plots were compared, since they showed a highly hydrophobic N-terminal part and, after a rapid transition, a hydrophilic C-terminal part (Gray *et al.* 1990). Due to the hydrophobicity of the N-termini, Borthakur *et al.* (1987) and Gray *et al.* (1990) assumed Psi and ExoX to be at least membrane associated. Latchford *et al.* (1991) have confirmed this by using *psi-pho*A fusions. But *psi* and *exoX* mutants differ in their symbiotic properties. *R. leguminosarum* bv. *phaseoli* mutant strains induce Fix$^-$ nodules on *Phaseolus* beans (Borthakur *et al.* 1985), whereas *R. meliloti exoX* mutants are more effective in the symbiosis with the homologous host alfalfa (Keller *et al.* 1988; Zhan and Leigh 1990).

The protein encoded by the *exoY* gene of *R. meliloti* is nearly identical to the protein encoded by the corresponding gene of *Rhizobium* sp. strain NGR234 (84% identity). Furthermore, we found that the first 311 amino acids of the *exoF* gene product are highly homologous to the 311 amino acids encoded by ORF1 of *Rhizobium* sp. strain NGR234 (Gray *et al.* 1990). This points to a close relationship between these two rhizobial species, which was biochemically shown by Zhan *et al.* (1990) and Gray *et al.* (1991). They demonstrated that the *exo* genes of both strains are to some extent interchangable. In addition, the amino acid sequence of the two ExoY proteins revealed a significant homology with the Pss2 protein of *R. l.* bv. *phaseoli*. Because of the homology of ExoY with GumD of *X. campestris,* Reed *et al.* (1991) proposed that ExoY may function as a sugar transferase. GumD is the first transferase in xanthan gum biosynthesis charging the C_{55} lipid carrier with glucose (Coplin and Cook 1990). We found, that ExoY is even more highly homologous to RfbP of *S. typhimurium* (51% identical amino acid residues). RfbP, assumed to be the first transferase in *S. typhimurium* LPS synthesis, is transferring galactose to the lipid carrier. Hence, it can be assumed that ExoY is a galactosyl transferase, especially since the synthesis of the repeating unit of the *R. meliloti* EPS I was found to start with galactose (Tolmasky *et al.* 1982). Borthakur *et al.* (1986) reported that a DNA fragment of *X. campestris* is complementing a *pss*2 mutant. Therefore, it is likely that this fragment encodes *gumD*.

The C-terminal part of RfbP is thought to be the catalytic domain exposed in the cytoplasm (Jiang *et al.* 1991). ExoY and Pss2 lack the hydrophobic N-terminal part of RfbP (or GumD) which anchors the latter in the membrane (Jiang *et al.* 1991). This anchor function may be fulfilled by the first hydrophobic domain of ExoY (circled in Fig. 2) and Pss2. Latchford and co-workers (1991) have demonstrated by *pho*A fusions that Pss2 is associated with the cell surface. Reuber and co-workers (1991) obtained corresponding results for the proteins encoded by the *exoF*/*exoQ* complementation groups. That a galactosyl transferase is located in the cytoplasmic membrane of *R. meliloti* was

already described by Ugalde *et al.* (1986).

By genetic analysis we revealed that the genes *exoY, exoF,* and *exoQ* are arranged in one transcriptional unit and that they are all indispensible for EPS I biosynthesis. Additionally, we could demonstrate that the *E. coli lacZ* promoter is active in *R. meliloti* and that it can mimic an *R. meliloti exo* promoter. This helped us to investigate the operon structure by creating apolar mutations. By insertion of plasmid pMK47.GK2 in the *exo* region disrupting the operon structure, it was possible to create an *exoQ* mutant. This was the only mutant we could obtain since it was not possible to get other mutants via homogenotization of transposon insertions in the second part of the *exoQ* coding region. We have no explanation for this, but we speculate that insertions in the coding region can result in a truncated form of the membrane protein ExoQ, which might be lethal for *R. meliloti.*

Further genetic experiments (Borthakur *et al.* 1988; Gray and Rolfe 1992) indicated that the regulation of EPS I synthesis seems to occur posttranscriptionally in some cases. The data obtained by Gray *et al.* (1990), Zhan *et al.* (1990), Reed *et al.* (1991), and in this work revealed that the ratio of ExoX and ExoY (and possibly of ExoF and ExoQ) is critical for the amount of EPS I produced. Analogous results were reported also for the *R. l.* bv. *phaseoli* Psi (Borthakur *et al.* 1985, 1988; Borthakur and Johnston 1987). Additionally, we found that *exoX* mutants were symbiotically more effective than the wild-type. An analogous result was also reported by Zhan and Leigh (1990), but it remains unknown if the higher amount of EPS I produced by *exoX* mutants is responsible for a more effective nodule infection.

Only little is known about the expression of *exoX* in the symbiotic state, but it seems to be expressed in the nodule as was shown for the *exoX* mutant Rm124 (Keller *et al.* 1988). Negative regulation of EPS biosynthesis has also been reported for other bacteria. Kamoun *et al.* (1989) described the gene *psdA* which caused depression of EPS production in *Agrobacterium tumefaciens*, a bacterium closely related to *R. meliloti.* Analogous to *exoX,* the negative influence was shown to be dependent on the copy number. But *psdA* and *exoX* seem to be different since the *psdA* locus spans at least 2.8 kb, whereas *exoX* is smaller than 500 bp.

EPS production in *Rhizobium* is a complicated process. Gray and Rolfe (1990) presented a model of how EPS synthesis in *Rhizobium* can occur. They proposed that the gene products of *exoY* and ORF1 of *Rhizobium* sp. strain NGR234 (*exoY* and *exoF* of *R. meliloti*) are located in the inner membrane forming the processing complex for EPS synthesis. This complex is thought to be regulated by ExoX via protein-protein interaction with ExoY. Our results lead us to speculate that this complex consists of a much higher number of proteins (e.g., the membrane associated protein ExoQ and also the assumed periplasmic protein ExoF can be considered to be part of the complex). Further analysis of other genes of the *R. meliloti exo* gene cluster will help in an understanding of the process of EPS I production in *R. meliloti.*

MATERIALS AND METHODS

Bacterial strains and plasmids.

Strains and plasmids used in this work are listed in Tables 2 and 3.

Media and growth conditions.

E. coli strains were grown in Luria-Bertani (LB) medium (Maniatis *et al.* 1982) at 37° C. *R. meliloti* strains were grown either in tryptone-yeast (TY) medium (Beringer 1974) with 0.4 g/L $CaCl_2$, M98 medium (Keller *et al.* 1988), or LB medium supplemented with 2.5 mM $CaCl_2$ and 2.5 mM $MgSO_4$ at 30° C. Antibiotics were used at the following concentrations (per liter) for *E. coli*: kanamycin (Km), 50 mg; tetracyclin (Tc), 7.5 mg; gentamicin (Gm), 20 mg; ampicillin (Ap), 200 mg. For *R. meliloti*: neomycin (Nm), 120 mg; streptomycin (Sm), 600 mg; Tc, 10 mg; Gm, 50 mg.

Table 2. Strains used and constructed in this study

Strain	Relevant characteristics	Source or reference
E. coli		
JM83	*ara*, Δ(*lac*, *pro*), (φ80d*lacZ*ΔM15), *thi*, λ^-	Vieira and Messing 1982
DH5α	*rec*A1, *lac*U169, Φ80d*lacZ*ΔM15	Bethesda Research Laboratories[a]
S17-1	MM294, RP4-2-Tc::Mu-Km::Tn*7* chromosomally integrated	Simon *et al.* 1983
R. meliloti		
Rm2011	Wild-type, Nod^+, Fix^+, Inf^+, EPS^+, Cfw^+, Sm^R	Casse *et al.* 1979
Rm0540	*exo*Y::Tn*5*, EPS^-, Inf^-	Müller *et al.* 1988b
RmJQ13	*exo*Y::Tn*5*-Gm, EPS^-, Inf^-	Kapp *et al.* 1990
RmH43	*exo*V mutant with Gm^R cassette integrated in the *Bgl*II site	B. Enenkel, Bielefeld
Rm124	*exoX*::Tn*5*-B20, EPS^{++}, Inf^+, Fix^+	Keller *et al.* 1988
Rm2011-H3	Rm2011 carrying plasmid pMK47.H3 integrated into the genome, EPS^+, Inf^+, Fix^+	This work
Rm2011-H6	Rm2011 carrying plasmid pMK47.H6 integrated into the genome, EPS^-, Inf^-, Fix^-	This work
Rm2011-GK2	Rm2011 carrying plasmid pMK47.GK2 integrated into the genome, EPS^-, Inf^-, Fix^-	This work
Rm2011-GK4	Rm2011 carrying plasmid pMK47.GK4 integrated into the genome, EPS^-, Inf^-, Fix^-	This work
RmJQ13-H3	RmJQ13 carrying plasmid pMK47.H3 integrated into the genome, EPS^-, Inf^-, Fix^-	This work
RmH43-GK4	RmH43 carrying plasmid pMK18.GK4 integrated into the genome, EPS^-, Inf^-, Fix^-	This work
Rm2011-X2	Rm2011 carrying plasmid pMK202.X2 integrated into the genome, EPS^+, Inf^+, Fix^+	This work
Rm2011-X3	Rm2011 carrying plasmid pMK202.X3 integrated into the genome, EPS^+, Inf^+, Fix^+	This work

[a] Gaithersburg, MD.

DNA manipulations.

Plasmid DNA was isolated from *E. coli* as described by Priefer (1984). DNA restriction, ligation, and agarose gel electrophoresis were conducted essentially as described by Maniatis *et al.* (1982). Transformation of *E. coli* cells was performed according to Morrison (1977).

DNA sequencing.

Some clones were obtained by inserting defined restriction fragments of plasmid pRmPM157.8 into the sequencing vectors pSVB23, pSVB30, pSVB31, and pK18 (Arnold and Pühler 1988; Pridmore 1987). Appropriate subclones for DNA sequencing were constructed by creating a set of nested deletions (Henikoff 1984). The DNA sequence was obtained for both strands by the chain termination method (Sanger *et al.* 1977) using double-stranded DNA (Arnold *et al.* 1988). For some subclones, the sequence was determined by the chemical degradation method as described by Maxam and Gilbert (1980) with some modifications (Arnold *et al.* 1988). The Tn*5*-B20 insertion junctions were sequenced using an oligodeoxynucleotide primer complementary to the first nucleotides of the *lacZ* gene integrated in IS50L (Simon *et al.* 1989).

Analysis of the nucleotide and amino acid sequences.

The nucleotide and amino acid sequences were analyzed using the sequence analysis programs (ANALYSEQ) of Staden (1986). The coding probability was calculated by means of the codon-usage method (Staden and McLachlen 1982) employing a codon usage table as described by Buendia *et al.* (1991). For homologous DNA and amino acid sequences the databanks at the EMBL, Heidelberg, Germany, were screened using the FASTA programs (Pearson and Lipman 1988). The predicted gene products were analyzed for their hydrophobicity following the procedure of Eisenberg *et al.* (1984) and for potential β-turns using the program PC/Gene (Genofit) according to the method of Chou and Fasman (1978). Possible procaryotic signal peptides were searched for using the matrix published by von Heijne (1986).

Plasmid integration mutagenesis and cloning of subfragments.

DNA fragments used for plasmid integration mutagenesis were cloned in several steps in both orientations into mobilizable vectors using appropriate restriction sites for digestion of plasmid pRmPM157.8. Two subclones were obtained by using deletion subclones of p135a and subsequent digestion with *Stu*I. The resulting hybrid plasmids were transferred from the broad host range mobilizing strain *E. coli* S17-1 (Simon *et al.* 1983) to *R. meliloti* Rm2011 or *R. meliloti* mutants according to Simon (1984). Hybrid plasmids not able to replicate in *R. meliloti* were maintained by single crossing over, selecting for the vector-encoded antibiotic resistance. Transconjugants were assayed for the symbiotic phenotype and their ability to produce EPS I.

Exopolysaccharide (EPS I) production.

EPS I production by *R. meliloti* strains was detected by the Calcofluor white (Cfw) method as previously described by Hynes *et al.* (1986). For quantitative analysis of EPS I production methods according to Keller *et al.* (1988) were used.

Plant nodulation assay.

R. meliloti strains were assayed for their symbiotic phenotype on different plants. *Medicago sativa* cv. "Du-Puits" (Saatgutveredelung Lippstadt, Germany) was used for nodulation assays. Seed surface sterilization, inoculation of the seedlings and growth of the plants on nitrogen-free medium were performed as described by Müller *et al.* (1988b). After 4 wk the plants were inspected for their symbiotic phenotype.

ACKNOWLEDGMENTS

We thank B. Enenkel for strain RmH43, A. Schäfer for plasmid pK18mob, and U. Schramm for the *lacZ* sequencing primer. In addition, we thank J. Kalinowski, R. Krey, M. Schmehl, and F. Vorhölter for computing help. We also acknowledge the excellent technical assistence of A. Kleikmann. This work was supported by a grant of the Deutsche Forschungsgemeinschaft (Pu28/13-2).

Table 3. Plasmids used and constructed in this study

Plasmid	Relevant characteristics	Source or reference
pSUP102	pACYC184-*mob*; Tc^r, Cm^r	Simon *et al.* 1983
pSUP104	Broad host range plasmid pACYC184 derivative; *mob*, Tc^r, Cm^r	Priefer *et al.* 1985
pSUP202	pBR325-*mob*; Tc^r, Cm^r, Ap^r	Simon *et al.* 1983
pSVB23, 30, 31	pUC8-derivative, Ap^r	Arnold and Pühler 1988
pK18	pUC18-derivative, Km^r	Pridmore 1987
pJQ47	pK18-*mob*, Km^r	J. Quandt, Bielefeld
pK18mob	pK18-*mob*, Km^r	A. Schäfer, Bielefeld
pRmPM157.8	pSUP102 with a 7.8-kb *Eco*RI fragment of megaplasmid 2	Müller *et al.* 1988b
p135a	pSVB23 with a 2.5-kb *Xho*I-*Hin*dIII sub-fragment of pRmPM157.8	This work
pRmMK124	pRmPM157.8-derivative, *exoX*::Tn*5*-B20	Keller *et al.* 1988
pRmMK129	pRmPM157.8-derivative, *exoY*::Tn*5*-B20	Keller *et al.* 1988
pMK47.H3	pJQ47 with a 0.7-kb *Hin*dIII fragment (Fig. 6)	This work
pMK47.H6	The same as pMK47.H3, but the *Hin*dIII fragment integrated in opposite direction	This work
pMK47.GK2	pJQ47 with a 931-bp *Bgl*II-*Kpn*I fragment (Fig. 6).	This work
pMK18.GK4	pK18mob with the 931-bp *Bgl*II-*Kpn*I fragment in opposite direction.	This work
pMK104.X2	pSUP104 with a 718-bp fragment of p135a carrying *exoX* (Fig. 1)	This work
pMK104.X3	pSUP104 with a 441-bp fragment of p135a carrying *exoX* (Fig. 1)	This work
pMK202.X2	pSUP202 with a 718-bp fragment of p135a carrying *exoX* (Fig. 1)	This work
pMK202.X3	pSUP202 with a 441-bp fragment of p135a carrying *exoX* (Fig. 1)	This work

LITERATURE CITED

Arnold, W., and Pühler, A. 1988. A family of high-copy-number vectors with single end-label sites for rapid nucleotide sequencing. Gene 70:171-179.

Arnold, W., Rump, A., Klipp, W., Priefer, U., and Pühler, A. 1988. Nucleotide sequence of a 24206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. J. Mol. Biol. 203:715-738.

Beringer, J. E. 1974. R factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. 84:188-198.

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.

Borthakur, D., Barker, R. F., 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 of *Rhizobium leguminosarum* required for exopolysaccharide synthesis and nodulation of peas: Their primary structure and their interaction with 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 which inhibits exopolysaccharide production and which is required for symbiotic nitrogen fixation. Mol. Gen. Genet. 200:278-282.

Buendia, A. M., Enenkel, B., Koplin, R., Niehaus, K., Arnold, W., and Pühler, A. 1991. The *Rhizobium meliloti exoZ/exoB* fragment of megaplasmid 2: 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.

Casse, F., Boucher, C., Julliot, J. S., Michel, M., and Dénarié, J. 1979. Identification and characterization of large plasmids in *Rhizobium meliloti* using agarose gel electrophoresis. J. Gen. Microbiol. 113:229-242.

Chakravorty, A. K., Zurkowski, W., Shine, J., and Rolfe, B.G. 1982. Symbiotic nitrogen fixation: Molecular cloning of *Rhizobium* genes involved in exopolysaccharide synthesis and effective nodulation. J. Mol. Appl. Genet. 1:585-596.

Chen, H., Batley, M., Redmond, J., and Rolfe, B. G. 1985. Alteration of the effective nodulation properties of a fast-growing broad range *Rhizobium* due to changes in exopolysaccharide synthesis. J. Plant. Physiol. 120:331-349.

Chou, P. Y., and Fasman, G. D. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. Relat. Areas Mol. Biol. 47:45-148.

Coplin, D. L., and Cook, D. 1990. Molecular genetics of extracellular polysaccharide biosynthesis in vascular phytopathogenic bacteria. Mol. Plant-Microbe Interact. 3:271-279.

Djordjevic, M. A., Gabriel, D. W., and Rolfe, B. G. 1987. *Rhizobium*—the refined parasite of legumes. Annu. Rev. Phytopathol. 25:145-168.

Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. 179:125-142.

Finan, T. M., Hirsch, A. M., Leigh, J. A., Johansen, E., Kuldau, G.A., Deegan, S., Walker, G. C, and Signer, E. R. 1985. Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell 40:869-877.

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

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 phenotype. J. Bacteriol. 172:193-203.

Gray, J. X., Zhan, H., Levery, S. B., Battisti, L., Rolfe, B. G., and Leigh, J. A. 1991. Heterologous exopolysaccharide production in *Rhizobium* sp. strain NGR234 and consequences for nodule development. J. Bacteriol. 173:3066-3077.

Gray, J. X., and Rolfe, B. G. 1990. Exopolysaccharide production in *Rhizobium* and its role in invasion. Mol. Microbiol. 4:1425-1431.

Halverson, L. J., and Stacey, G. 1986. Signal exchange in plant-microbe interactions. Microbiol. Rev. 4:193-225.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351-359.

Hynes, M. F., Simon, R., Müller, P., Niehaus, K., Labes, M., and Pühler, A. 1986. The two megaplasmids of *Rhizobium meliloti* are involved in the effective nodulation of alfalfa. Mol. Gen. Genet. 202:356-362.

Jähnig, F. 1990. Structure predictions of membrane proteins are not that bad. Trends Biochem. Sci. 15:93-95.

Jiang, X.-M., Neal, B., Santiago, F., Lee, S. J., Romana, L. K., and Reeves, P. R. 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar *typhimurium* (strain LT2). Mol. Microbiol. 5:695-713.

Kamoun, S., Cooley, M. B., Rogowsky, P. M., and Kado, C. I. 1989. Two chromosomal loci involved in production of exopolysaccharide in *Agrobacterium tumefaciens*. J. Bacteriol. 171:1755-1759.

Kapp, D., Niehaus, K., Quandt, J., Müller, P., and Pühler, A. 1990. Cooperative action of *Rhizobium meliloti* nodulation and infection mutants during the process of forming mixed infected alfalfa nodules. Plant Cell 2:139-151.

Keller, M., Arnold, W., Kapp, D., Müller, P., Niehaus, K., Schmidt., M., Quandt, J., Weng, W. M., and Pühler, A. 1990. *Rhizobium meliloti* genes involved in exopolysaccharide production and infection of alfalfa nodules, Pages 91-97 in: *Pseudomonas:* Biotransformations, Pathogenesis, and Biotechnology. S. Silver, A. M. Chakrabarty, B. Iglewski, and S. Kaplan, eds. American Society for Microbiology, Washington, DC.

Keller, M., Müller, P., Simon, R., and Pühler, A. 1988. *Rhizobium meliloti* genes for exopolysaccharide synthesis and nodule infection located on megaplasmid 2 are actively transcribed during symbiosis. Mol. Plant-Microbe Interact. 1:267-274.

Labes, M., Pühler, A., and Simon, R. 1990. A new family of RSF1010-derived expression and *lac*-fusion broad-host-range vectors for gram-negative bacteria. Gene 89:37-46.

Latchford, J. W., Borthakur, D., and Johnston, A. W. B. 1991. The products of *Rhizobium* genes, *psi* and *pss*, which affect exopolysaccharide production, are associated with bacterial cell surface. Mol. Microbiol. 5:2107-2114.

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.

Leigh, J. A., Signer, E. R., and Walker, G. C. 1985. Exopolysaccharide deficient mutants of *Rhizobium meliloti* that form ineffective nodules. Proc. Natl. Acad. Sci. USA 82:6231-6235.

Long, S. R. 1989. *Rhizobium*-legume nodulation: Life together in the underground. Cell 56:203-214.

Long, S., Reed, J., Himawan, J., and Walker, G. C. 1988. Genetic analysis of a cluster of genes required for synthesis of the calcofluor-binding exopolysaccharide of *Rhizobium meliloti*. J. Bacteriol. 170:4239-4248.

Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. Molecular Cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maxam, A., and Gilbert, W. 1980. Sequencing end-labeled DNA with bas-specific chemical cleavages. Methods Enzymol. 65:499-531.

McClure, W. R. 1985. Mechanisms and control of transcription initiation in prokaryotes. Annu Rev. Biochem. 54:171-202.

Morrison, D. A. 1977. Transformation in *E. coli*: Cryogenic preservation of competent cells. J. Bacteriol. 132:349-351.

Müller, P., Enenkel, B., Hillemann, A., Kapp, D., Keller, M., Quandt, J., and Pühler, A. 1988a. Genetic analysis of two DNA regions of the *Rhizobium meliloti* genome involved in the infection process of alfalfa nodules. Pages 26-32 in: Molecular Genetics of Plant-Microbe Interactions. R. Palacios and D. P. S. Verma, eds. American Phytopathological Society, St. Paul, MN.

Müller, P., Hynes, M. F., Kapp., D., Niehaus, K., and Pühler, A. 1988b. The two classes of *Rhizobium meliloti* infection mutants

differ in exopolysaccharide production and in coinoculation properties with nodulation mutants. Mol. Gen. Genet. 211:17-26.
Pearson, W. R., and Lipman, D. J. 1988. Improved tools for biological sequence analysis. Proc. Natl. Acad. Sci. USA 85:2444-2448.
Priefer, U. B. 1984. Isolation of plasmid DNA. Pages 14-25 in: Advanced Molecular Genetics. A. Pühler and K. N. Timmis, eds. Springer-Verlag, Berlin.
Priefer, U. B., Simon, R., and Pühler, A. 1985. Extension of the host range of *Escherichia coli* vectors by incorporation of RSF1010 replication and mobilization functions. J. Bacteriol. 163:324-330.
Pridmore, R. D. 1987. New and versatile cloning vectors with kanmycin-resistance marker. Gene 56:309-312.
Pühler, A., Enenkel, B., Hillemann, A., Kapp, D., Keller, M., Müller, P., Niehaus, K., Priefer, U. B., Quandt, J., and Schmidt, C. 1988. *R. meliloti* and *R. leguminosarum* mutants defective in surface polysaccharide synthesis and root nodule development. Pages 423-430 in: Nitrogen Fixation: Hundred Years After. H. Bothe, F. DeBruijn, and W. Newton, eds. Fischer, Stuttgart.
Reed, J. W., Capage, M., and Walker, G. C. 1991. *Rhizobium meliloti exoG* and *exoI* mutations affect the ExoX/ExoY sytem for modulation of exopolysaccharide production. J. Bacteriol. 173:3776-3788.
Reuber, T. L., Long, S., and Walker, G. C. 1991. Regulation of *Rhizobium meliloti exo* genes in free-living cells and *in planta* using Tn*phoA* fusions. J. Bacteriol. 173:426-434.
Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
Simon, R. 1984. High frequency mobilisation of gram negative bacterial replicons by the *in vitro* constructed Tn*5*-Mob transposon. Mol. Gen. Genet. 196:413-420.
Simon, R., Priefer, U. B., and Pühler, A. 1983. A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in Gram negative bacteria. Biotechnology 1:784-791.
Simon, R., Quandt, J., and Klipp, W. 1989. New derivatives of transposon Tn*5* suitable for mobilization of replicons, generation of operon fusions and induction of genes in Gram-negative bacteria. Gene 80:161-169.
Staden, R. 1986. The current status and portability of our sequence handling software. Nucleic Acids Res. 14:217-231.
Staden, R., and McLachlan, A. D. 1982. Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucleic Acids Res. 10:141-156.
Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. 1973. Improved estimation of secondary structure in ribonucleic acids. Nature 246:40-41.
Tolmasky, M. E., Staneloni, R. J., and Leloir, L. F. 1982. Lipid-bound saccharides in *Rhizobium meliloti*. J. Biol. Chem. 257:6751-6757.
Ugalde, R. A., Handelsman, J., and Brill, W. J. 1986. Role of galactosyltransferase activity in phage sensitivity and nodulation competitiveness of *Rhizobium meliloti*. J. Bacteriol. 166:148-154.
Vieira, J., and Messing, J. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259-268.
von Heijne, G. 1985. Signal sequences: The limits of variation. J. Mol. Biol. 184: 99-105.
von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683-4690.
Zhan, H., and Leigh, J. A. 1990. Two genes that regulate exopolysaccharide production in *Rhizobium meliloti*. J. Bacteriol. 172:5254-5259.
Zhan, H., Gray, J. X., Levery, S. B., Rolfe, B. G., and Leigh, J. A. 1990. Functional and evolutional relatedness of genes for exopolysaccharide synthesis in *Rhizobium meliloti* and *Rhizobium* sp. strain NGR234. J. Bacteriol. 172:5245-5253.