JOURNAL OF BACTERIOLOGY, Jan. 1990, p. 193–203
0021-9193/90/010193-11$02.00/0

Vol. 172, No. 1

# Two Genes That Regulate Exopolysaccharide Production in *Rhizobium* sp. Strain NGR234: DNA Sequences and Resultant Phenotypes

JAMES X. GRAY, MICHAEL A. DJORDJEVIC, AND BARRY G. ROLFE*

*Plant Microbe Interactions Group, Research School of Biological Sciences, Australian National University, P.O. Box 475, Canberra, A.C.T., 2601, Australia*

Received 6 July 1989/Accepted 11 September 1989

**Two closely linked genes involved in the regulation of exopolysaccharide (EPS) production in *Rhizobium* sp. strain NGR234, *exoX* and *exoY*, were sequenced, and their corresponding phenotypes were investigated. Inhibition of EPS synthesis occurred in wild-type strains when extra copies of *exoX* were introduced, but only when *exoY* had been deleted or mutated or was present at a lower copy number. Normal EPS synthesis occurred in *Rhizobium* sp. when both *exoX* and *exoY* were introduced on the same replicon. Surprisingly, the presence of multiple copies of *exoY* in *exoY*::Tn*5* mutants of NGR234 adversely affected cellular growth. This was apparent when *exoY* was introduced into *exoY* mutants on IncP1 vectors, where the copy number was approximately 10, but was not apparent when present on much larger R-prime plasmids with lower copy numbers (approximately 3 per cell). Multiple copies of *exoX* did not adversely affect cellular growth of any strain. The *exoX* gene appeared analogous, in size and phenotype, to a previously described *Rhizobium leguminosarum* biovar *phaseoli* EPS gene, *psi* (D. Borthakur and A. W. B. Johnston, Mol. Gen. Genet. 207:149–154, 1987), and the deduced ExoX and Psi shared strikingly similar secondary structures. Despite this, ExoX and Psi showed little homology at the primary amino acid level, except for a central region of 18 amino acids. The interaction of ExoX and ExoY could form the basis of a sensitive regulatory system for EPS biosynthesis. The presence of a multicopy *exoX* in *Rhizobium meliloti* and *R. fredii* similarly abolished EPS biosynthesis in these species.**

A complex symbiotic interaction between rhizobia and specific legume plants, involving a wide array of genes, results in the formation of nitrogen-fixing root nodules (13, 28, 37). Overwhelming evidence indicates that genes responsible for exopolysaccharide (EPS) biosynthesis are involved in the infection process (28) and possibly host range specificity (35). EPSs are complex sugar polymers secreted from rhizobia that loosely encapsulate the cell surface. They are probably among the first molecules to come in contact with the root surface. Mutants of various *Rhizobium* species that lack EPS production ($Exo^-$) are often characterized by both poor infectivity and nodule formation (6, 9, 10, 17, 26). $Exo^-$ mutants cause little or no bacterial penetration or colonization of the root tissue, although the ability to induce a nodule meristem is retained (6, 9, 10, 17, 26).

*Rhizobium* sp. strain NGR234 is capable of infecting a broad range of legume species. $Exo^-$ mutants derived from this strain induce poor nodulation on most plants. Microscopic examination of the callus structures formed by $Exo^-$ mutants of strain NGR234 on the legume *Leucaena leucocephala* reveals little or no bacterial penetration or colonization (10). Further evidence for the involvement of EPS in infection was shown in mixed inoculation experiments in which $Exo^-$ mutants were inoculated together with a nonnodulating ($Nod^-$) NGR234 derivative (cured of the Sym plasmid) that produces normal EPS (14). Together, these strains induced normal nitrogen-fixing nodules and cohabited the nodule tissue. Furthermore, the addition of purified NGR234 EPS or oligosaccharide could substitute for the $Nod^-$ $Exo^+$ strain in these mixed inoculation experiments. Plants inoculated with some $Exo^-$ mutants together with EPS or oligosaccharide generated a proportion of normal nitrogen-fixing nodules (14). This indicated that purified EPS has "bioactivity" and may have acquired a direct and specific role in the symbiosis through an as yet unknown mechanism.

In *Rhizobium meliloti* a cluster of 12 genes affecting EPS biosynthesis occurs on the second megaplasmid, pRmeSU47b (18, 27). Transposon insertion mutations within these genes affect EPS production in *R. meliloti* and cause a variety of alfalfa infection deficiencies. Mutations in 9 of the 12 *exo* genes either abolish or severely reduce EPS synthesis; in most cases, non-nitrogen-fixing ($Fix^-$) nodules result (27).

In *Rhizobium leguminosarum* bv. *phaseoli*, two genes, termed *psi* and *psr*, involved in EPS biosynthesis are located on the Sym plasmid (6). Mutant strains defective in *psi* induce $Fix^-$ nodules. When multiple copies of *psi* were introduced on a plasmid into *R. leguminosarum* bv. *phaseoli*, both EPS synthesis and nodulation ability were inhibited. The presence in *R. leguminosarum* bv. *phaseoli* of equal copies of *psi* and another gene, *psr*, resulted in normal EPS production and nodule formation (7). Borthakur and Johnston (7) also demonstrated that Psr inhibits transcription of *psi*. The nucleotide sequence of *psi* showed that it specified a polypeptide of 86 amino acids with a hydrophobic N-terminal region spanning 41 amino acids. This suggested that the *psi* product was associated with the cell membrane (7). Psi is not the only regulatory protein involved in regulation of EPS synthesis. In *Rhizobium fredii*, another gene, *nodD2*, prevents EPS synthesis when carried on a multicopy plasmid (3). The mechanism by which the genes *psi* and *nodD2* inhibit EPS synthesis is unknown.

The timing and regulation of EPS synthesis may be critical

* Corresponding author.

TABLE 1. Bacterial strains and plasmids

| Strain or plasmid | Description | Reference or source |
|---|---|---|
| *Rhizobium* strains | | |
| NGR234 | Broad-host-range cowpea *Rhizobium* sp. | 44 |
| ANU280 | $Sm^r$ $Rif^r$ derivative of NGR234 | 10 |
| ANU2811 | ANU280 *exoY11*::Tn*5* | 10 |
| ANU2890 | ANU280 *exoY90*::Tn*5* | 10 |
| ANU2823 | ANU280 *exoY23*::Tn*5* | 10 |
| ANU2808 | ANU280 *exoY08*::Tn*5* | 10 |
| ANU2840 | ANU280 *exoY40*::Tn*5* | 10 |
| Rm1021 | $Sm^r$ derivative of wild-type *R. meliloti* | 30 |
| USDA191 | Wild-type *R. fredii* | 2 |
| *E. coli* strains | | |
| NM522 | Δ(*lac-pro*) F′ *lacZ*M15 *lacI*$^q$ *hsd-5* | 23 |
| HB101 | *leu proA2 rps120* ($Sm^r$) *hsdS20* | 5 |
| Plasmids | | |
| pD56 | Clone carrying wild-type *R. meliloti exo* DNA | 26 |
| R′3222 | Clone carrying wild-type NGR234 *exo* DNA | 11 |
| pBR322 | Cloning vector, $Ap^r$ $Tc^r$ | 5 |
| pHC11 | *EcoRI* fragment (*exoY11*::Tn*5*) in pBR322 | 11 |
| pRK2013 | Helper plasmid *tra oriT* ColE1 | 12 |
| pRK290 | Broad-host-range IncP vector, $Tc^r$ | 12 |
| pJJ016 | *Bam*HI-*Hin*dIII ($Km^r$) fragment of Tn*5* cloned into pRK290, $Tc^r$ $Km^r$ | Weinman[a] |
| pJG22 | Recombinant pJJ016 carrying wild-type NGR234 *exo* genes | This work |
| pJG22::Tn*5* | Same as pJG22, but with *exoY11*::Tn*5* | This work |
| pMP220 | Broad-host-range IncP promoterless *lacZ*, $Tc^r$ | 40 |
| pJG51[b] | 7.0-kb *Bgl*II subclone | This work |
| pJG52[b] | 5.0-kb *Pst*I subclone | This work |
| pJG53[b] | 1.9-kb *Cla*I-*Pst*I subclone | This work |
| pJG54[b] | 1,408-bp *Mlu*I-*Pst*I subclone *exoY′-lacZ*$^+$ | This work |
| pJG55[b] | 1,221-bp *Mlu*I-*Hin*dIII partial subclone | This work |
| pJG56[b] | 992-bp *Mlu*I-*Nru*I partial subclone | This work |
| pJG57[b] | 790-bp *Mlu*I-*Eco*RI subclone | This work |
| pJG58[b] | 1,121 *Nru*I-*Pst*I partial subclone | This work |
| pJG59[b] | 934-bp *Nru*I-*Hin*dIII partial subclone | This work |
| pJG60[b] | 705-bp *Nru*I subclone *ecoX′-lacZ*$^+$ | This work |
| pJG61[b] | 923-bp *Hin*dIII-*Pst*I partial subclone | This work |
| pJG62[b] | 736-bp *Hin*dIII subclone | This work |
| pJG63[b] | 630-bp *Eco*RI subclone | This work |
| pJG64[b] | 695-bp *Hin*dIII subclone | This work |
| pJG65[b] | 1,392-bp *Eco*RI subclone | This work |
| pJG66[b] | 887-bp *Nru*I-*Cla*I partial subclone | This work |
| pJG70[b] | The pJG60 insert in opposite orientation | This work |

[a] J. J. Weinman, Ph.D. thesis, Australian National University, 1986.
[b] These plasmids carry subcloned regions of the pJG22 insert DNA cloned into the vector pMP220 and are represented in Fig. 2.

for successful nodule formation. Borthakur and Johnston (7) have hypothesized that the apparent inhibition of EPS synthesis by *R. leguminosarum* bv. *phaseoli* in the bacteroid state is achieved by a repression of *psr*, which results in a derepressed *psi*. In *R. meliloti*, it appears that some EPS genes, which are actively transcribed during the free-living state, remain strongly expressed during symbiosis (24a). To date, it is not clear whether EPS is produced in the bacteroid state or indeed whether the regulation is the same for all *Rhizobium* species.

Molecular investigations have shown that five genetic loci involved in the synthesis of acidic exopolysaccharide in *Rhizobium* sp. strain NGR234 are clustered in a 15-kilobase (kb) region of DNA (11). Chen et al. (11) showed that the introduction into the wild-type ($Exo^+$) strain of R-prime plasmids containing Tn*5* insertions (mapped to genetic groups F and E) resulted in merodiploid transconjugants that were stably $Exo^-$. Conversely, when the corresponding wild-type allele was introduced into group F or E $Exo^-$ mutants, the episomally located *exo* allele was dominant, and stable $Exo^+$ transconjugants resulted. In this paper, we demonstrate that group F and E mutations are located within a single gene, *exoY*. In contrast to the hypothesis of Chen et al. (11) mutations in *exoY* do not act as dominant, negative mutations. Instead, we conclusively show that another *exo* gene less than 1 kb from the group F and E mutation sites is responsible for the $Exo^-$ phenotype of these merodiploid strains. This novel NGR234 *exo* gene, termed *exoX*, confers an $Exo^-$ phenotype only when it is present in a copy number above that of *exoY* or when *exoY* has been mutated or deleted.

## MATERIALS AND METHODS

**Strains, plasmids, and media.** Bacterial strains and plasmids used and constructed for this paper are listed in Table 1. All media used, BMM, TY, and LB, have been previously described (38).

**Recombinant DNA techniques.** DNA isolations, visualizations, and hybridizations were done by the methods of

Maniatis et al. (29). Hybond-N nylon membranes (Amersham, England) were used for DNA transfers. Restriction enzyme digests and ligations were performed according to the specifications of the manufacturers (Boehringer Mannheim Biochemicals, Indianapolis, Ind., and New England BioLabs, Inc., Beverly, Mass.). DNA probes were $^{32}P$ labeled with random primers (45).

Subcloned fragments of NGR234 *exo* DNA in vector pMP220 (Table 1; see Fig. 2) were constructed by using donor DNA from subclones used for DNA sequencing. Those fragments, which appear to be generated by partial restriction endonuclease digestion, were instead constructed by ligating the relevant fully digested fragments (agarose gel purified) and cloning the combined fragment into the vector Bluescript (Stratagene, San Diego, Calif.). To ensure that the original continuity of the DNA was preserved, the junctions of the ligated fragments were sequenced. These reconstructed partial fragments were then cloned into the vector pMP220 as single fragments, utilizing the wide choice of unique flanking restriction sites present in the Bluescript polylinker.

**Bacterial conjugation.** Broad-host-range recombinant plasmids were mobilized from *Escherichia coli* NM522 into *Rhizobium* spp. by a triparental patch mating technique with pRK2013 (12) as a helper plasmid. After a 24-h mating period, the patch was replica plated onto BMM selective medium supplemented with 30 μg of rifampin per ml to counterselect *E. coli* and with 4 μg of tetracycline per ml to select for transconjugants carrying pMP220 recombinant plasmids.

**Assays of *lacZ* expression from cloned *Rhizobium* promoters.** Gene expression from cloned *exo* gene promoters, fused to *E. coli lacZ* located on vector pMP220 (40), were determined by measuring β-galactosidase activity with *O*-nitrophenol-β-D-galactopyranoside as the substrate (32). The average values from at least nine repeat measurements are presented.

**Plant and acetylene reduction assays.** Seed sterilization, germination, inoculation, and growth of *L. leucocephala* (Lam.) Wit. var. Peru was described previously (10). Nodulated plants were tested for acetylene reduction by the method of Bender and Rolfe (4).

**Determination of copy numbers for R-prime and smaller IncP1 plasmids.** The copy number associated with very large R-prime plasmids (150 kb) was compared with that of smaller pRK290 (IncP1) recombinant plasmids (30 kb). This was achieved by preparing a Southern blot (1) of an agarose gel where equal amounts of *Eco*RI-digested genomic DNA isolated from Exo$^-$ mutant ANU2811 and transconjugants ANU2811(R'3222) and ANU2811(pJG22) were electrophoresed. The membrane was probed with the 10-kb *Rhizobium* DNA insert of pJG22, and the intensity of hybridizing DNA bands was measured by using an LKB Bromma 2202 Ultroscan Laser Densitometer. The ratio of the intensities for hybridizing bands between the various genomic DNA preparations reflects their respective copy numbers.

**DNA sequencing.** Plasmid vectors Bluescript SK$^+$ and SK$^-$ (Stratagene) and vectors M13mp18 and 19 (33) were used for cloning. Sequencing reactions were conducted using a Sequenase kit (U.S. Biochemical Corp., Cleveland, Ohio). The sequencing protocol provided with this kit is a modification of the dideoxy-chain termination technique (39), which uses a modified T7 DNA polymerase (42). The kit was used according to the manufacturer's instructions, except that a standard chase step of 1 μl of chase mix (1 mM each deoxynucleoside triphosphate) was included. Some sequencing reactions were also carried out by using the dideoxy-chain termination technique of Sanger et al. (39) with the Klenow fragment of *E. coli* DNA polymerase I.

Single-stranded DNA was isolated from Bluescript recombinants using the helper phage VCSM13 (Stratagene) by the following protocol. A single colony of the relevant clone and 1 μl of VCSM13 helper phage ($10^{13}$ pfu ml$^{-1}$) was inoculated into 2 ml of liquid LB supplemented with 50 μg of ampicillin per ml and shaken at 37°C. After 2 h, kanamycin was added to the culture to a final concentration of 70 μg ml$^{-1}$, and the culture was shaken at 37°C for 16 h. The phage particles were separated from the bacterial cells by centrifuging the culture for 5 min at 4°C in a 1.5-ml Eppendorf tube. The supernatant was centrifuged for 5 min at 4°C, transferred to a tube containing 300 μl of 25% polyethylene glycol and 2.5 M NaCl and incubated on ice for 30 min before centrifugation at 4°C for 5 min. The phage pellet was suspended in 100 μl of TE (10 mM Tris hydrochloride, 1 mM EDTA [pH 7.5]), extracted with 50 μl of phenol (equilibrated against TE), vortexed for 1 min, and centrifuged for 10 min at room temperature. An 80-μl volume of the aqueous phase was recovered, and the single-stranded DNA was precipitated with 2 volumes of 100% ethanol and suspended in 50 μl of TE.

Single-stranded DNA from M13 recombinant phages was isolated by the method of Heidecker et al. (24). The DNA sequence presented was sequenced entirely in both directions and was analyzed by using SEQ (a package of computer programs available at the Research School of Biological Sciences, Australian National University).

**Extraction of RNA from *Rhizobium* sp.** All solutions were made up with sterile, deionized water and kept ice cold during use (unless otherwise indicated). A mid- to late-log-phase *Rhizobium* culture (optical density at 600 nm, 0.5 to 1) grown in 500 ml of BMM liquid medium at 30°C was harvested in Sorvall SS34 centrifuge tubes at 10,000 rpm for 5 min. The pellets were washed in TES (100 mM NaCl, 1 mM EDTA, 10 mM Tris hydrochloride [pH 7.4]) to aid in the removal of EPS from the cell surface, and the combined 15 ml of bacterial resuspension was transferred to Eppendorf tubes. After washing, the bacterial pellet was suspended in 500 μl of extraction buffer (10 mM Tris hydrochloride [pH 7.5], 5 mM EDTA, 5% sucrose, 300 mM $CH_3COONa$, 1% sodium dodecyl sulfate, 1% 2-β-mercaptoethanol). A 300-μl volume of phenol (equilibrated against TE [pH 8]) at 90°C was added and mixed, followed by the addition of 300 μl of $CHCl_3$-isoamyl alcohol (24:1). The tubes were vortexed and centrifuged for 10 min. The aqueous phase was recovered, and the nucleic acid was precipitated with 2.5 volumes of 100% ethanol. The supernatant was removed, and the pellet was suspended in 50 μl of TE. An equal volume of 8 M LiCl was added, and RNA was precipitated for 16 h at −20°C. Insoluble RNA was centrifuged for 15 min at 4°C and suspended in 50 μl of TE with vortexing. The LiCl precipitation was repeated to eliminate contaminating DNA. The final RNA pellet was suspended in 100 μl of TE and stored at −20°C. To determine the yield, a 5% sample was visualized on a morpholinepropanesulfonic acid-agarose gel.

**Mapping transcription initiation sites with S1 nuclease.** The method is a modification of that recently described by Calzone et al. (8). The DNA probe for the *exoY* transcript was end labeled at the *Sma*I site at nucleotide position 1700 (see Fig. 3), within the *exoY* coding region (see Fig. 2). The probe used to identify the intergenic transcribed region without any open reading frames (ORFs) was end labeled at the *Eco*RI site at nucleotide position 789 (see Fig. 2 and 3).

Restriction enzyme sites were 5′ end labeled with [γ-$^{32}$P]ATP by using polynucleotide kinase (New England BioLabs) after dephosphorylation by calf intestine alkaline phosphatase. The double-stranded probe fragment, now end labeled at its two 5′ ends, was restricted at an appropriate restriction site to free the probe fragment labeled only at the 5′ end of the antisense strand. This was recovered from an agarose gel slice by centrifugation in a TLS55 Beckman rotor at 55K rpm for 2 h at room temperature.

Hybridizations were conducted at 58°C for 18 h in a 10-μl volume of 80% formamide–80 mM sodium piperazine-*N*,*N*′-bis(2-ethanesulfonic acid (pH 6.8)–400 mM NaCl–10 mM disodium EDTA–100 μg of *Rhizobium* RNA–1 μg of $^{32}$P-end-labeled DNA and covered with a drop of paraffin oil to prevent evaporation. Digestion with S1 nuclease (Boehringer Mannheim) was at 37°C for 30 min and was started by adding to the hybridization solution a 200-μl volume of 30 mM $CH_3COONa$ (pH 4.5)–280 mM NaCl–1 mM $ZnSO_4$–300 U of S1 nuclease. Digestion with S1 nuclease was terminated by transferring the digestion solution (avoiding the paraffin) to a 20-μl solution containing 500 mM Tris hydrochloride (pH 9)–100 mM EDTA–50 μg of tRNA per ml. After ethanol precipitation and suspension of the samples in a 70% formamide gel loading buffer and electrophoresis through a 8 M urea–3% acrylamide gel, the protected probe DNA was visualized by autoradiography.

## RESULTS

**Analysis of the *Rhizobium* strains with multicopy wild-type *exo* genes.** A 10-kb *Bam*HI fragment known to contain *exo* genes (11) was cloned into the broad-host-range IncP1 vector pJJ016 (Table 1). This recombinant plasmid (pJG22) was transferred into ANU280 (a rifampin-resistant derivative of the wild-type strain NGR234) and into several Tn*5*-induced Exo$^-$ derivatives of ANU280 (ANU2811, ANU2890, ANU2823, ANU2808, and ANU2840). When pJG22 was transferred into ANU280, the transconjugants grew normally and the colony morphology remained Exo$^+$. Interestingly, when this plasmid was transferred into the Exo$^-$ mutants, transconjugants arose at a frequency of $3.4 \times 10^{-4}$ per recipient strain. The frequency of transfer was approximately 100-fold lower than for (i) the transfer of the vector (pJJ016) alone to these strains and (ii) the transfer of pJG22 or pJJ016 to ANU280. The Exo$^-$ mutants containing pJG22 were initially slow to appear, the growth rates of colonies were not uniform, and a mixture of Exo$^+$ and Exo$^-$ colony morphologies resulted (Fig. 1A). Both Exo$^+$ and Exo$^-$ transconjugants arose at approximately equal frequencies [52% Exo$^+$ and 48% Exo$^-$ in the case of ANU2811(pJG22) transconjugants]. Each of these colony types appeared clonal and was highly stable; no conversion from one colony morphology to another was observed when the different cell types were recultured on selective medium plates. Although these colony types were initially slow to appear, the subsequent growth rates of these transconjugant variants were normal.

To investigate the cause of the two colony morphologies for ANU2811(pJG22) transconjugants, plasmid DNA was recovered from single Exo$^+$ and Exo$^-$ isolates. The restriction digest profile of the plasmids recovered from several Exo$^+$ transconjugants was identical to that of the original plasmid pJG22 (Fig. 1B, lane 1a). In contrast, the profile of the plasmid DNA from several Exo$^-$ transconjugants was altered (Fig. 1B, lane 1b). These plasmids had lost the 0.6-kb *Eco*RI fragment, which contained the site of Tn*5* insertion in the mutant ANU2811, and now had a new *Eco*RI fragment, which was larger by 5.8 kb (the length of the Tn*5* sequences). Hybridization analysis (Fig. 1B) showed that the altered plasmids from the Exo$^-$ transconjugants now carried the Tn*5* sequence with no detection of the wild-type 0.6-kb *Eco*RI fragment. However, hybridization analysis of total genomic DNA isolated from these Exo$^-$ transconjugants (data not shown) revealed that the 0.6-kb and the 0.6-kb plus Tn*5* bands were both present. Thus, the altered plasmids probably resulted from double-reciprocal recombination events between sequences flanking the Tn*5* insertion in the background genome and the homologous sequences carried on the introduced plasmid.

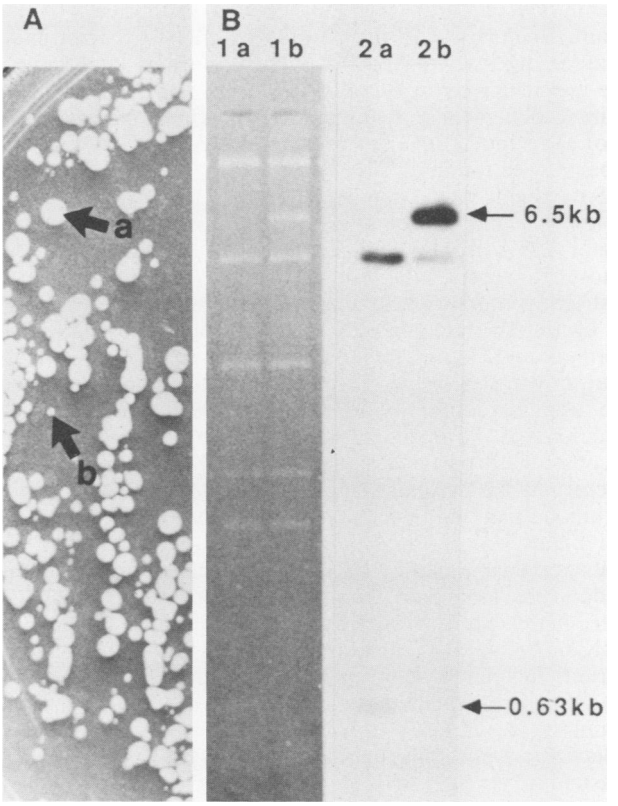


FIG. 1. Analysis of transconjugants resulting from the transfer of pJG22 into mutant ANU2811. (A) Exo$^+$ (a) and Exo$^-$ (b) colony types. (B) Analysis of plasmid DNA carried by Exo$^+$ and Exo$^-$ transconjugants. Lanes: 1a, *Eco*RI restriction profile of the plasmid from an Exo$^+$ colony; 1b, *Eco*RI restriction profile of the plasmid from an Exo$^-$ colony; 2a and 2b, autoradiograph of a Southern blot of lanes 1a and 1b, respectively, after hybridization with radioactively labeled plasmid pHC11. Plasmid pHC11 has a 6.5-kb *Eco*RI insert, cloned from ANU2811, that carries the Tn*5* insertion. The 0.63-kb band is the size of the wild-type *Eco*RI fragment, and the 6.5-kb band is this same *Eco*RI fragment carrying the Tn*5* insertion. The other hybridizing bands in both lanes 2a and 2b are due to cross-hybridization between kanamycin resistance gene sequences located on the vector pJJ016 and present within Tn*5*.

A high frequency of recombination events was not observed for ANU2811 transconjugants carrying the same wild-type *exo* DNA cloned on the much larger (150-kb) R′3222 plasmid. The copy numbers associated with these two plasmids were determined (see Materials and Methods); R′3222 was present at approximately three copies per cell compared with approximately 10 copies per cell for pJG22. This result suggested that an elevated copy number of the

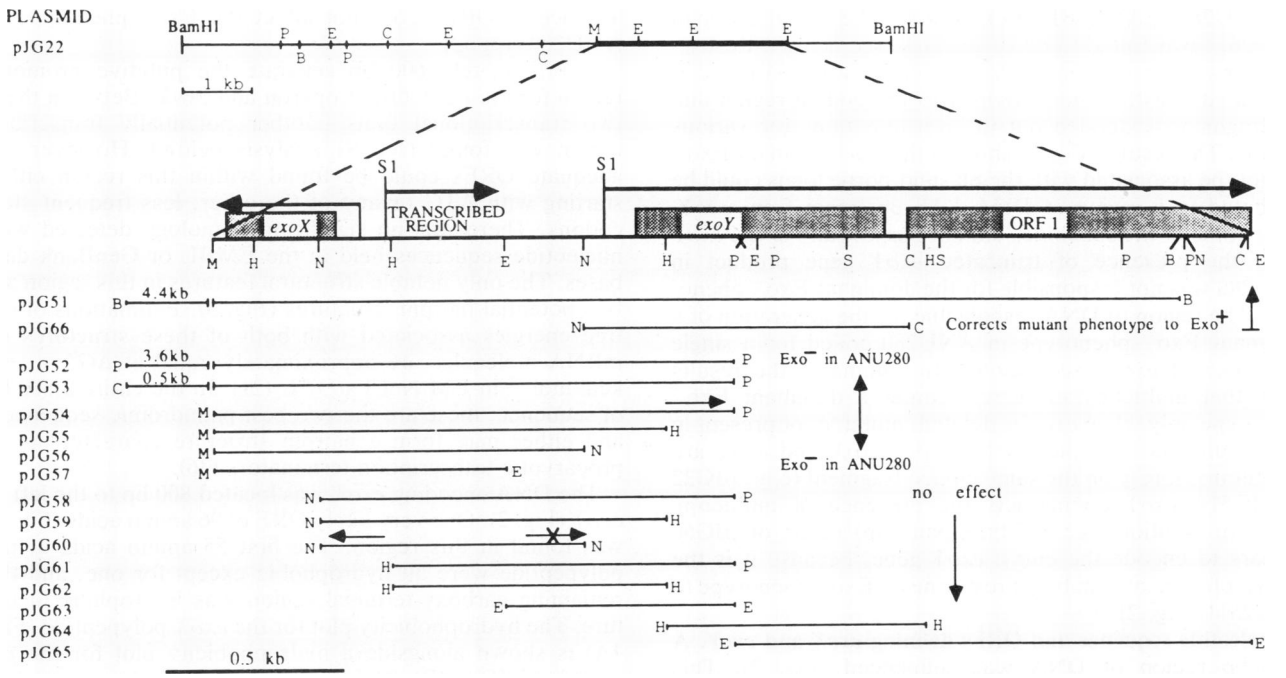


FIG. 2. Physical map of the 10-kb *Bam*HI fragment involved in EPS synthesis. The nucleotide sequence has been determined for the expanded region of the map. The extremities of the subcloned fragments and the phenotypes associated with these fragments, when used in complementation experiments, are also shown. Plasmids pJG51 and pJG66 complement the Exo$^-$ mutant ANU2811 to Exo$^+$. Plasmids pJG52 through to pJG57 confer an Exo$^-$ phenotype when present in ANU280. Plasmids pJG58 through to pJG65 do not alter the phenotypes of either ANU280 or ANU2811. Symbols: (—► and ⊣►) presence or absence, respectively, of β-galactosidase activity associated with these fragments when fused to *lacZ* in the direction of the arrow, ( ▩ ) putative coding regions of the sequenced genes (see Fig. 3), (—► and ◄—) direction of transcription. The sites of transcription initiation, as determined by S1 promoter mapping, are indicated. Restriction sites: B, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; M, *Mlu*I; N, *Nru*I; P, *Pst*I; S, *Sma*I.

wild-type allele is deleterious to the cell and consequently results in selection for normally rare recombination events.

We were thus intrigued as to why 52% of the ANU2811(pJG22) transconjugants remained stably Exo$^+$. To ensure that plasmids from the Exo$^+$ transconjugants had not suffered a small undetectable deletion or rearrangement of only a few critical nucleotides that rendered the putative deleterious gene(s) "harmless," plasmid DNA was recovered from Exo$^+$ and Exo$^-$ transconjugants and transformed into *E. coli*. These *E. coli* derivatives were used to transfer these plasmids into the Exo$^-$ mutant ANU2811 as well as into the original parental strain ANU280. The results for ANU2811 recipients were as follows: (i) unaltered plasmids originally recovered from the Exo$^+$ transconjugants again produced a mixture of Exo$^+$ and Exo$^-$ colonies, and (ii) pJG22 plasmids containing Tn*5* did not alter the Exo$^-$ phenotype of ANU2811 colonies. For ANU280 recipients the results were as follows: (i) unaltered pJG22 conferred no phenotypic change upon colony morphology or growth, and (ii) pJG22 containing Tn*5* conferred an Exo$^-$ phenotype upon 100% of the transconjugants. Therefore, pJG22 plasmids recovered from the original Exo$^+$ ANU2811(pJG22) colonies were indeed unaltered. In contrast, Tn*5* containing pJG22 plasmids conferred a dominant Exo$^-$ phenotype when present in ANU280.

**Symbiotic phenotypes of Exo$^+$ and Exo$^-$ transconjugants.** The wild-type strain ANU280 inoculated onto *L. leucocephala* forms between 5 and 23 cylindrical, pink pigmented, nitrogen-fixing nodules per plant after 4 weeks; in contrast, all Exo$^-$ mutant strains (e.g., ANU2811) form Fix$^-$ callus-like nodules (10). Exo$^+$ ANU2811(pJG22) transconjugants were able to form Fix$^+$ nodules on *L. leucocephala* that were indistinguishable from those formed by the wild-type strain. However, this was not due to a simple complementation of the mutant allele by the plasmid-borne wild-type allele. Examination of the bacteria recovered from the nodules demonstrated that, in the absence of antibiotic selection, more than 50% of the cells had lost both tetracycline resistance (vector marker) and kanamycin resistance (Tn*5* marker). This result indicated that these cells underwent recombination events and subsequent loss of plasmid during nodule passage. Similarly, the Exo$^-$ ANU280 (pJG22::Tn*5*) transconjugants induced Fix$^+$ nodules on *L. leucocephala*, but all bacteria isolated from nodules were Exo$^+$ and had neither tetracycline nor kanamycin resistance markers, which are both present on the plasmid. This result again indicated plasmid loss.

**Complementation phenotype associated with subclones of the 10-kb *Bam*HI fragment.** As reported earlier (11), the introduction into ANU280 of large 150-kb R-prime plasmids (copy number, approximately three per cell) carrying the mutant alleles from either Exo$^-$ mutant ANU2808, ANU2811, ANU2823, ANU2840, or ANU2890 resulted in the transconjugants being Exo$^-$. Similarly, pJG22 with a Tn*5* insertion at the ANU2811 locus, also conferred an Exo$^-$ phenotype when transferred into ANU280. To determine whether the Exo$^-$ phenotype of ANU280 transconjugants was due to a dominant negative mutation or the presence of another element on the 10-kb *Bam*HI fragment, a series of subclones from this *Bam*HI fragment was cloned into the IncP1 vector pMP220 (40). Figure 2 is a summary of the fragments subcloned and their phenotypes when present in

the ANU280 or ANU2811 backgrounds. The Tn*5* insertion in ANU2811 was found to occur in *exoY* (see below). Restriction sites were chosen such that *exoY* was restricted at several sites within and proximal to the coding region and the fragments extended out in both directions for various lengths. The results (Fig. 2) showed that the dominant Exo$^-$ phenotype associated with the plasmid-borne locus could be attributed to a region of DNA 1 kb upstream from *exoY*. Since some of the subclones did not possess any of the *exoY* gene, the presence of truncated *exoY* gene product in ANU280 was not responsible for the dominant Exo$^-$ phenotype. The region of DNA responsible for the generation of a dominant Exo$^-$ phenotype in ANU280 coded for a single gene, named *exoX* (see below). In summary, the results show that multicopy *exoX* will confer a dominant Exo$^-$ phenotype only when *exoY* is deleted, mutated, or present in lower numbers of copies. When both *exoX* and *exoY* are present in entirety on the same cloned fragment (e.g., pJG22 or pJG51), *exoY* counteracts the presence of multicopy *exoX*. In addition, the 887-base-pair (bp) insert of pJG66 appears to encode the entire *exoY* gene, because it is the smallest fragment capable of restoring an Exo$^+$ phenotype to ANU2811 (Fig. 2).

**Nucleotide sequence and ORFs defining *exoX* and *exoY*.** A 2,800-bp region of DNA was sequenced (Fig. 3). This included DNA spanning the Tn*5* insertion sites for several Exo$^-$ mutants (ANU2808, ANU2811, ANU2823, and ANU2840) and extended sufficiently in both directions to include (i) the DNA sequences required to complement the Exo$^-$ phenotype of these mutants and (ii) the DNA sequences of the second gene, *exoX*. The DNA sequence and ORFs proposed to define *exoX* and *exoY* are shown in Fig. 3. Plots made by using a Positional Base Preference program (41) and the Fickett TESTCODE program (16) indicated that the three ORFs assigned to the 2,800 bp of sequence had a very high level of nonrandomness and are most likely coding regions (data not shown).

Chen et al. (11) showed that the Tn*5* insertion sites for 17 Exo$^-$ mutants mapped within two adjacent *Eco*RI fragments of 0.6 and 1.4 kb (Fig. 2). These *Eco*RI fragments were sequenced, and two very likely ORFs were found, designated *exoY* and ORF1. The Tn*5* insertion sites for 5 representatives of the 17 Exo$^-$ mutants in this region (ANU2808, ANU2811, ANU2823, ANU2840, and ANU2890) were sequenced and were found to occur within the coding region of *exoY* (Fig. 3). The ORF for *exoY* was 226 amino acids in length. This predicted polypeptide has a run of 24 entirely hydrophobic amino acids starting at residue 34, followed immediately by a lengthy 65-residue hydrophilic domain. The hydrophobic stretch is long enough to form a transmembrane region or to associate by hydrophobic interactions with other protein domains (34). No significant similarities were detected between this deduced ExoY protein and any of the protein sequences held in the NBRF protein data base.

The short 56-bp stretch of untranslated DNA between *exoY* and ORF1, coupled with the absence of promoterlike sequences, indicated that *exoY* and ORF1 form an operon. ORF1 extended beyond the region of DNA sequenced and the predicted protein for this reading frame was more than 312 amino acids in length. No Tn*5* insertions were obtained within this coding region, and hence its mutant phenotype is not known. None of the protein sequences held in the NBRF protein data base had any detectable homology with the predicted polypeptide from ORF1. The presence of an intact ORF1 is not required to complement ANU2811, and the absence of ORF1 does not affect the Exo$^+$ phenotype of ANU280 (Fig. 2).

Approximately 600 bp separate the putative promoter region for the *exoY*-ORF1 operon and *exoX*. Between these two transcriptional units another potentially transcribed region was found (see S1 analysis below). However, no adequate ORFs could be found within this region either starting with ATG or any of the other, less frequent start codons. There was no significant homology detected with nucleotide sequences held in the EMBL or GenBank data bases. The only notable structural features in this region are two potential hairpin structures (Fig. 3). Estimations of the free energies associated with both of these structures on mRNA molecules are approximately equal at $\Delta G^0 = -19$ kcal mol$^{-1}$ in 1 M NaCl at 37°C (21). In the entire 2,800 bp of sequence, these are the two best palindromic sequences, and either may form a hairpin structure characteristic of procaryotic transcription terminators (36).

The DNA encoding *exoX* was located 800 bp to the left of *exoY* (Fig. 2). One very likely ORF of 96 amino acids (Fig. 3) was found in this region. The first 55 amino acids of the polypeptide were all hydrophobic except for one, and the remaining carboxy-terminal region was hydrophilic in nature. The hydrophobicity plot for the *exoX* polypeptide (Fig. 4A) is shown alongside a hydrophobicity plot for the *psi* protein (7) (Fig. 4B). The similarity between these two plots is striking. Both proteins are very hydrophobic for the amino half and then rapidly make the transition to a hydrophilic nature for the remaining half. In addition, both proteins are similar in size (96 amino acids for ExoX and 86 amino acids for Psi) and predicted molecular weight (approximately 10,500 and 9,500, respectively). The similarity between these two proteins is less apparent at the primary amino acid level (Fig. 4C). There is, however, a conserved 18 amino acid region with 14 functionally similar amino acids (10 exact matches), just before the region in the protein where the transition from hydrophobic to hydrophilic amino acids occurs. This may be a conserved domain responsible for the similar phenotype associated with these two genes.

**Promoter mapping by S1 nuclease and *lacZ* fusions.** After analysis of the nucleotide sequence, no satisfactory ORF could be assigned to the DNA sequence occurring between *exoX* and *exoY*. It was not clear whether this intervening DNA sequence was an extensive 5′ untranslated region of either the putative *exoY*-ORF1 operon or the *exoX* operon or a separate transcriptional unit. Potential sites for transcription initiation in this region were mapped with S1 nuclease (Fig. 2, 3, and 5). The cluster of bands in lane 3 of Fig. 5 suggests that this intervening sequence is transcribed divergently from *exoX*. A transcriptional fusion to *lacZ*(pJG70) using the promoterless vector pMP220 showed no activity (Table 2), and this demonstrated that transcription of this region terminates upstream of the *Nru*I site at nucleotide position 989 (Fig. 3).

A clear band in lane 2 of Fig. 5 demonstrates that the transcription initiation site for *exoY* is downstream of the *Nru*I site at nucleotide position 989. An *exoY′-lacZ*$^+$ fusion (pJG54) at the *Pst*I site at nucleotide position 1408 (Fig. 3) showed significant *lacZ* activity (Table 2) and confirmed that the direction of transcription was from left to right (Fig. 2). Strong activity of an *exoX′-lacZ*$^+$ fusion (pJG60) (Table 2) at the *Nru*I site at nucleotide position 290 (Fig. 3) indicated that *exoX* was transcribed divergently from *exoY*.

**Identification of possible transcription and translation initiation signals.** The putative ribosome-binding site for *exoY* is 5′-TGGAGT-3′, this is identical to the presumptive ribo-

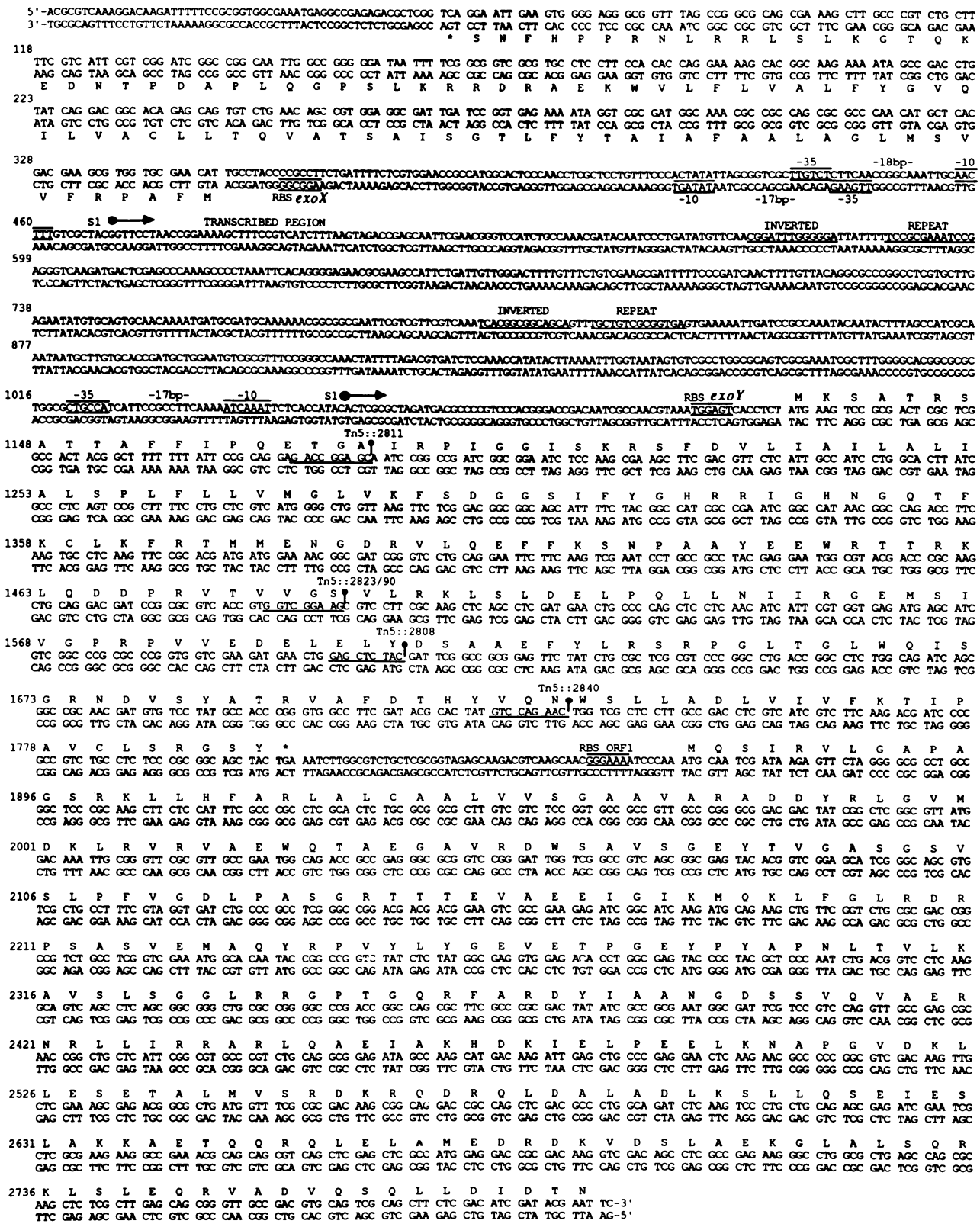


FIG. 3. Uninterrupted DNA sequence of the region represented diagrammatically in the expanded portion of Fig. 2. The translated regions are indicated with the amino acids being represented by the standard single-letter code. Note that the sense strand for *exoX* is the bottom strand, and the sense strand for *exoY* and ORF1 is the top strand. The proposed −35 and −10 transcription signals and ribosome-binding sites for each gene are indicated by lines above and below the relevant sequences. Also indicated are the Tn*5* insertion sites within *exoY* and the transcription initiation sites as determined by S1 promoter mapping. The discussed inverted repeat sequences are underlined. These data have been submitted to GenBank under accession no. M28454.

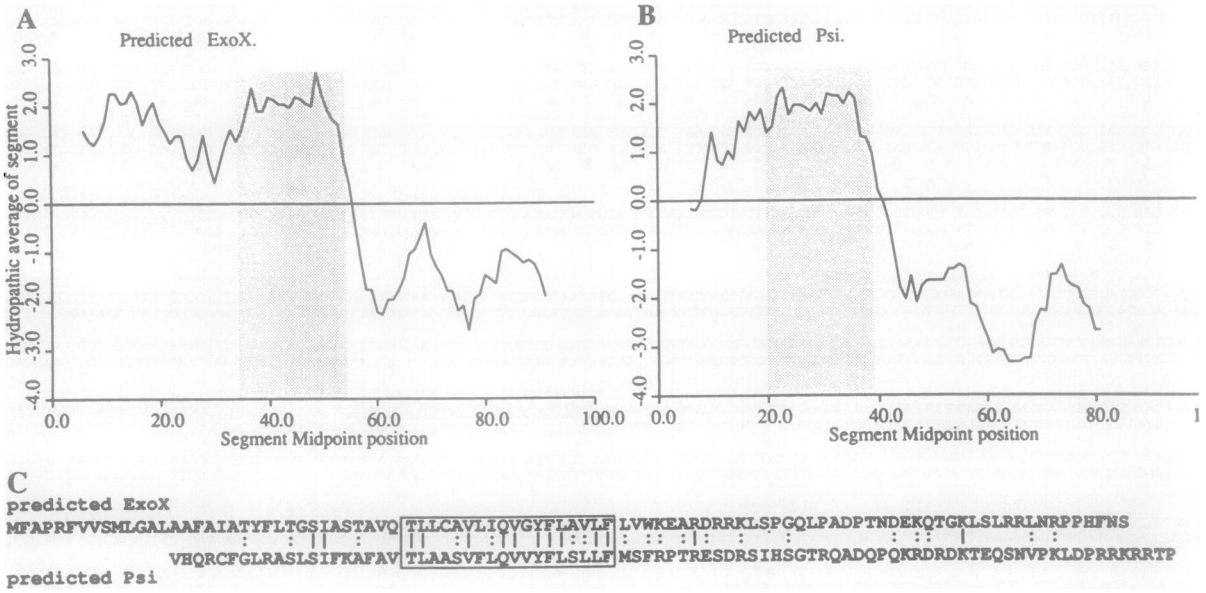


FIG. 4. Comparisons between deduced polypeptide sequences of *exoX* and *psi*. The hydrophobicity plots of ExoX (A) and Psi (B) are shown. Both plots were generated by using a hydropathy program with the values of Kyte and Doolittle (25) and written by Weinman (Ph.D. thesis). The values on the horizontal axis represent the amino acid residue positions, and the values on the vertical axis represent the hydropathic averages of an 11-residue segment of the polypeptide. (C) Best alignment between the deduced protein sequences of ExoX and Psi, without the introduction of any gaps. Exact amino acid matches are indicated by an unbroken line, and functionally similar amino acids are indicated by a broken line. The region of highest homology is enclosed within a box, corresponding to the shaded regions of A and B. Amino acids have been grouped into the following families: acidic and amidic (D, E, N, Q), basic (H, K, R), polar (A, G, P, S, T), nonpolar (I, L, M, V), aromatic (F, W, Y), and cysteine (C).

some-binding site for *R. meliloti nodA* (43) and is similar (four out of six nucleotides match) to those of *nodD* and *nodH* (15, 20). S1 promoter mapping experiments indicated that the start of transcription of *exoY* was approximately 70 bp upstream of the putative start codon. Upstream of this position, corresponding with the −35 position, was the sequence 5′-CTGCCA-3′; this had four out of six matches with the possible *R. meliloti nodA*, *nodF*, and *nodH* −35 sequences (19, 20). An optimal 17-bp spacing existed between this −35 sequence and a downstream potential −10 sequence, with four out of six matches to the procaryote consensus sequence (31) or five out of six matches to a presumptive *nodD* −10 sequence (20). The putative *exoX* ribosome-binding site, 5′-AGGCGG-3′, had five out of six matches with the procaryote consensus sequence (22). Virtually consensus-like (31) promoter sequences [5′-TTGAag-(17-bp space)-TATAgT-3′] exist 50 bp upstream of the proposed *exoX* start codon.

**EPS synthesis in other *Rhizobium* species is affected by *exoX*.** EPS synthesis in two other *Rhizobium* strains, *R. meliloti* Rm1021 and *R. fredii* USDA 191, was also inhibited by the introduction of multiple copies of *exoX* without extra copies of *exoY*. When pJG22::Tn*5* was transferred into strains Rm1021 and into USDA 191, and the transconjugants were cultured on solid BMM medium, the colonies were Exo$^-$ and appeared identical to those of ANU280(pJG22::Tn*5*) transconjugants. Strains Rm1021 and USDA 191 have a visibly Exo$^+$ colony morphology when cultured on BMM. Introduction of pJG22 into these two *Rhizobium* strains had no effect on EPS production. Although we have observed extensive DNA homology between the cloned NGR234 *exo* genes (11) and the cloned *R. meliloti* pRmeSU47b *exo* genes (27) (unpublished results), a probe from within the proposed coding region of *exoX* did not hybridize to the cloned *R. meliloti exo* genes encoded on pD56 under low-stringency conditions. The intragenic *exoX* probe was a 190-bp fragment from the *Hin*dIII site at position 100 to the *Nru*I site at position 290. In contrast, an intragenic *exoY* probe, from the *Hin*dIII site at position 1216 to the *Cla*I site at position 1867, did show strong homology to *R. meliloti* sequences cloned on cosmid pD56 (27) but did not appear to flank transposon insertions to *exo* loci in this region (data not shown).

## DISCUSSION

In this report we have presented the nucleotide sequence for 2,800 bp of DNA involved in the synthesis of EPS for *Rhizobium* sp. strain NGR234. To assist in assigning transcriptional units and genes to the ORFs, a combination of S1 promoter mapping, *lacZ* transcriptional fusion experiments, and analysis of the phenotypes associated with subcloned regions was employed to complement the DNA sequencing and its computer analysis.

The Exo$^-$ mutants used in this study resulted from single Tn*5* insertions into the wild-type genome of ANU280, and the locations of the mutations were mapped to specific *Eco*RI fragments (11). R-prime plasmids carrying mutated *exo* genes were used previously to define two types of ANU280 transconjugants with repressed EPS synthesis. The introduction to ANU280 of R-prime plasmids carrying Tn*5* insertions corresponding to the ANU2811 or ANU2890 allele resulted in Exo$^-$ colonies (persistent dominant phenotype). R-prime plasmids carrying Tn*5* insertions corresponding to the ANU2808 or ANU2840 allele conferred an Exo$^-$ phenotype on ANU280, which upon prolonged incubation resulted in Exo$^+$ colonies (leaky dominant phenotype). We have shown that all four Tn*5* insertion sites were located within *exoY*. The two insertion sites for the persistent dominant

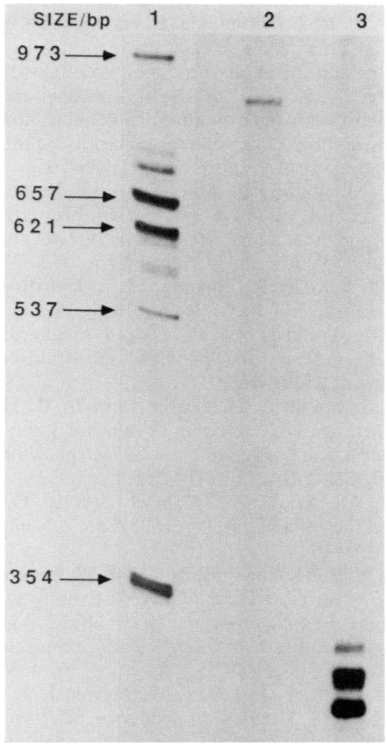


FIG. 5. Autoradiograph showing the protected probe DNA after S1 digestion. Lane 1 is *Cla*I-digested λ DNA with the sizes of the authentic bands indicated by arrows. Lane 2 shows the length of single-stranded probe DNA from the 5′ $^{32}$P-labeled *Sma*I site (at nucleotide position 1700, within the *exoY* coding region), which was protected by mRNA. Lane 3 shows the lengths of probe DNA from the 5′ $^{32}$P-labeled *Eco*RI site (at nucleotide position 789, within the untranslated, transcribed region), which was protected by mRNA. The two probes used in lanes 2 and 3 had a common 3′ unlabeled *Cla*I site 2.2 kb from the *Sma*I site (lane 2) and 1.3 kb from the *Eco*RI site (lane 3).

mutant alleles were at amino acid residue positions 20 (ANU2811) and 124 (ANU2890), and the two insertion sites for the leaky dominant alleles were at positions 161 (ANU2808) and 202 (ANU2840). Whatever the function of ExoY may be, it is evident that it has some wild-type activity when at least 161 of the total 226 amino acids have been translated.

The production of EPS by NGR234 appears to be strictly controlled by the products of *exoY* and *exoX*. The evidence suggests that the gene product of *exoX* is a repressor of EPS synthesis, because an elevated copy number of this gene results in the inhibition of EPS synthesis by *Rhizobium* sp. When *exoX* is carried on an IncP1 plasmid in the absence of *exoY*, the phenotype of the ANU280 merodiploid transconjugants is Exo$^-$. Normal EPS production occurs when the copy number of *exoY* is increased to a level equal to that of *exoX*; cloned fragments carrying both of these genes do not confer the Exo$^-$ phenotype on ANU280 transconjugants. Since normal EPS synthesis is sensitive to slight elevations of the copy number of *exoX* relative to that of *exoY*, this suggests the possibility of an interaction either between the products of the two genes or the product of *exoY* and the promoter of *exoX*. Conditions favoring repressed expression of *exoY* or enhanced expression of *exoX* in the wild-type strain might lead to repression of EPS production.

TABLE 2. β-Galactosidase activity of *lacZ* fusions in wild-type background

| Plasmid | Construction | Activity[a] | SD |
|---|---|---|---|
| pMP220 | Vector | 139 | 26 |
| pJG54 | *exoY′-lacZ*$^+$ | 1,881 | 302 |
| pJG60 | *exoX′-lacZ*$^+$ | 1,575 | 84 |
| pJG70 | 990-bp *Nru*I-*lacZ*$^+$ | 28 | 7 |

[a] Activity units are as defined by Miller (32).

Transposon Tn*5* insertions into the genomic copy of *exoY* abolish the ability of ANU280 to synthesize acidic EPS (10). This Exo$^-$ phenotype is probably due to the presence of *exoX* unchecked by *exoY* rather than to the mutation of *exoY* as a structural gene. Normal EPS production is restored by the introduction of fragments carrying the wild-type allele for *exoY*. Recombinant R-prime plasmids carrying wild-type *exoY* (with very low copy numbers of approximately three per cell) will correct the EPS phenotype of *exoY*::Tn*5* mutants in 100% of the cases (11). When the copy number of this wild-type *exoY* allele is increased on a recombinant IncP1 plasmid (copy number of approximately 10 per cell) and transferred into *exoY*::Tn*5* mutants, the frequency of correction to Exo$^+$ in the transconjugants is no longer 100%, but is now 52%. The other 48% of the transconjugants remain Exo$^-$. In the Exo$^-$ transconjugants, a normally rare double-reciprocal recombination event occurs between the DNA flanking the genomic Tn*5* insertion and homologous *Rhizobium* DNA cloned on the plasmid. Therefore, an Exo$^-$ transconjugant has several copies of the nearby *exoX* on the plasmid and only a single copy of the wild-type *exoY* allele in the genome. This imbalance in favor of *exoX* results in the inhibition of EPS biosynthesis by these *Rhizobium* cells. The Exo$^+$ transconjugants of ANU2811 carrying pJG22, on the other hand, appear to have been complemented by the introduced fragment, because the plasmid has not been altered in any way. However, since these transconjugants were slow to appear, another explanation is that these cells have undergone a suppressor mutation elsewhere in the genome to compensate for the presence of extra copies of *exoY*. We conclude that the presence of *exoY* at approximately 10 copies per cell is deleterious to the cell growth of *exoY*::Tn*5* mutants. This is supported by the strong selection for wild-type *Rhizobium* cells when merodiploid strains are passaged through *L. leucocephala* nodules. The reason why ANU280 is not affected in the same way by many copies of *exoY* has eluded us thus far. One explanation is that the Tn*5* insertion into *exoY* is polar to a downstream gene that is not present on the cloned *Rhizobium* DNA of pJG22. Possibly the absence of this putative gene creates an intolerance to high levels of *exoY*. The much larger cloned inserts of the R-prime plasmids would contain the whole *exoY*-ORF1 operon.

EPS regulatory systems similar to that of NGR234 may also occur in other *Rhizobium* species. Plasmids carrying *exoX* in the absence of *exoY* inhibited EPS synthesis in *R. meliloti* and *R. fredii* (*R. leguminosarum* bv. *phaseoli* was not tested). Although an *exoY*-specific hybridization probe strongly hybridized to cloned *R. meliloti* DNA, there was no detection of an *exoX* homolog by hybridization even under low-stringency conditions. This is not so surprising, when it is noted that the amount of DNA homology between *exoX* and the likely *R. leguminosarum* bv. *phaseoli* equivalent gene, *psi*, is very low (33%).

The phenotype associated with multiple copies of *exoX* is the same as that already reported for the *R. leguminosarum*

bv. *phaseoli* gene *psi* (6). The proposed *psi* polypeptide is comprised of 86 amino acids (7) and this is similar to the proposed 96-amino-acid *exoX* polypeptide. In addition, the hydrophobicity plots for these proteins are strikingly similar. At the primary sequence level, however, there is less similarity between the proteins encoded by *exoX* and *psi*, except for an 18-amino-acid domain in which 14 of the residues are functionally similar (10 exact matches). It is possible that *exoX* and *psi* are related genes, where the only evolutionary constraints have been within the 18-amino-acid domain and the overall tertiary structure of the protein (i.e., maintaining a hydrophobic amino-terminal half and a hydrophilic carboxy-terminal half). It is possible that the hydrophobic amino-terminal region is inserted into the membrane, as already suggested for Psi (7). Alternatively, this hydrophobic region may associate, by hydrophobic interactions, with other protein subunits to form a multimeric complex. The length of the hydrophobic region (55 amino acids for the ExoX protein) suggests that both possibilities are plausible. The homologous domain between proteins encoded by *exoX* and *psi* occurs in the hydrophobic region just before the polypeptide makes a rapid hydrophilic transition. Therefore, the most amino-terminal 20 or so amino acids could form a transmembrane signal peptide, which would still leave the conserved hydrophobic region available for association with hydrophobic domains of other proteins. One protein that may be a candidate for this type of multimeric association is that encoded by *exoY*, which has an internal hydrophobic region spanning 24 amino acids. Currently we are investigating whether the product of *exoY* represses the transcription of *exoX* or, alternatively, whether there is a posttranslational association between the products of genes *exoY* and *exoX*.

## ACKNOWLEDGMENTS

We sincerely thank Jeremy Weinman for his many critical readings of the manuscript and for his discussions involving data interpretation. We thank Marie Oaks for correcting the grammar. Thanks also to Shou-wei Ding for trouble shooting tips in DNA sequencing, to Paul Keese for advice on RNA isolation, and also to Brett Tyler for a working S1 promoter mapping protocol. Finally, Graham C. Walker is acknowledged for providing cloned *R. meliloti exo* genes to avoid the duplicative naming of homologous loci.

J.X.G. is the recipient of a Commonwealth Postgraduate Research Award.

## LITERATURE CITED

1. **Amersham Corp.** 1985. Amersham handbook. Transfer of DNA to Hybond membranes. Membrane transfer and detection methods. Amersham Corp., Amersham, England.
2. **Appelbaum, E., T. McLoughlin, M. O'Connell, and N. Chartrain.** 1985. Expression of symbiotic genes of *Rhizobium japonicum* USDA 191 in other rhizobia. J. Bacteriol. **163**:385–388.
3. **Appelbaum, E. R., D. V. Thompson, K. Idler, and N. Chartrain.** 1988. *Rhizobium japonicum* USDA 191 has two *nodD* genes that differ in primary structure and function. J. Bacteriol. **170**:12–20.
4. **Bender, G. L., and B. G. Rolfe.** 1985. A rapid plant assay for the *Parasponia-Rhizobium* symbiosis. Plant Sci. Lett. **19**:277–284.
5. **Bolivar, F., R. Rodriguez, P. J. Greene, M. Betlach, H. L. Heynecker, H. W. Boyer, J. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
6. **Borthakur, D., J. A. Downie, A. W. B. Johnston, and J. W. Lamb.** 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.
7. **Borthakur, D., and A. W. B. Johnston.** 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.
8. **Calzone, F. J., R. J. Britten, and E. H. Davidson.** 1987. Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. Methods Enzymol. **152**:611–632.
9. **Chakravorty, A. K., W. Zurkowski, J. Shine, and B. G. Rolfe.** 1982. Symbiotic nitrogen fixation: molecular cloning of *Rhizobium* genes involved in exopolysaccharide synthesis and effective nodulation. J. Mol. Appl. Gen. **1**:585–596.
10. **Chen, H., M. Batley, J. Redmond, and B. G. Rolfe.** 1985. Alteration of the effective nodulation properties of a fast-growing broad host range *Rhizobium* due to changes in exopolysaccharide synthesis. J. Plant Physiol. **120**:331–349.
11. **Chen, H., J. X. Gray, M. Nayudu, M. A. Djordjevic, M. Batley, J. W. Redmond, and B. G. Rolfe.** 1988. Five genetic loci involved in the synthesis of acidic exopolysaccharides are closely linked in the genome of *Rhizobium* sp. strain NGR234. Mol. Gen. Genet. **212**:310–316.
12. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 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.
13. **Djordjevic, M. A., D. W. Gabriel, and B. G. Rolfe.** 1987. *Rhizobium*: the refined parasite of legumes. Annu. Rev. Phytopathol. **25**:145–168.
14. **Djordjevic, S. P., H. Chen, M. Batley, J. W. Redmond, and B. G. Rolfe.** 1987. Nitrogen-fixing ability of exopolysaccharide synthesis mutants of *Rhizobium* sp. strain NGR234 and *Rhizobium trifolii* by the addition of homologous exopolysaccharide. J. Bacteriol. **169**:53–60.
15. **Egelhoff, T. T., R. F. Fisher, T. W. Jacobs, J. T. Mulligan, and S. R. Long.** 1985. Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. DNA **4**:241–248.
16. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10**:5303–5318.
17. **Finan, T. M., A. M. Hirsch, J. A. Leigh, E. Johansen, G. A. Kuldau, S. Deegan, G. C. Walker, and E. R. Signer.** 1985. Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell **40**:869–877.
18. **Finan, T. M., B. Kunkul, G. F. De Vos, and E. R. Signer.** 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. J. Bacteriol. **167**:66–72.
19. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, and S. R. Long.** 1988. Specific binding proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. Genes Dev. **2**:282–293.
20. **Fisher, R. F., J. A. Swanson, J. T. Mulligan, and S. R. Long.** 1987. Extended region of nodulation genes in *Rhizobium meliloti* 1021. II. Nucleotide sequence, transcription start sites and protein products. Genetics **117**:191–201.
21. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.
22. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1981. Translation initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
23. **Gough, J. A., and N. E. Murray.** 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. J. Mol. Biol. **166**:1–19.
24. **Heidecker, G., J. Messing, and B. Gronenborn.** 1980. A versatile primer for DNA sequencing in the M13mp2 cloning system. Gene **10**:69–73.

24a. **Keller, M., P. Müller, R. Simon, and A. Pühler.** 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.

25. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
26. **Leigh, J. A., E. R. Signer, and G. C. Walker.** 1985. Exopolysaccharide deficient mutants of *Rhizobium meliloti* that form inef-

fective nodules. Proc. Natl. Acad. Sci. USA **82**:6231–6235.
27. **Long, S., J. W. Reed, J. Himawan, and G. C. Walker.** 1988. Genetic analysis of a cluster of genes required for the synthesis of the calcofluor-binding exopolysaccharide of *Rhizobium meliloti*. J. Bacteriol. **170**:4239–4248.
28. **Long, S. R.** 1989. *Rhizobium*-legume nodulation: life together in the underground. Cell **56**:203–214.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. **149**:114–122.
31. **McClure, W. R.** 1985. Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem. **54**:171–204.
32. **Miller, J. H.** 1972. Experiments in molecular genetics, 1st ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
34. **Oliver, D. B.** 1987. Periplasm and protein secretion, p. 56–69. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
35. **Philip-Hollingsworth, S., R. I. Hollingsworth, F. B. Dazzo, M. A. Djordjevic, and B. G. Rolfe.** 1989. 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.
36. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
37. **Rolfe, B. G., and P. M. Gresshoff.** 1988. Genetic analysis of legume nodule initiation. Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**:297–319.
38. **Rolfe, B. G., P. M. Gresshoff, and J. Shine.** 1980. Rapid screening for symbiotic mutants of *Rhizobium* and white clover. Plant Sci. Lett. **19**:277–284.
39. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
40. **Spaink, H. P., R. J. H. Okker, C. A. Wijffelman, E. Pees, and B. J. J. Lugtenberg.** 1987. Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. Plant Mol. Biol. **9**:27–39.
41. **Staden, R.** 1984. Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes. Nucleic Acids Res. **12**:551–567.
42. **Tabor, S., and C. C. Richardson.** 1987. Sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84**:4767–4771.
43. **Török, I., E. Kondorosi, T. Stepkowski, J. Pósfai, and A. Kondorosi.** 1984. Nucleotide sequence of *Rhizobium meliloti* nodulation genes. Nucleic Acids Res. **12**:9509–9524.
44. **Trinick, M. J.** 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucocephala*, *Mimosa* spp, *Acacia farnesiana* and *Sesbania grandiflora* and their affinities with other rhizobia groups. J. Appl. Bacteriol. **49**: 39–53.
45. **Whitfeld, P. L., P. H. Seeburg, and J. Shine.** 1982. The human pro-opiomelanocortin gene: organization, sequence, and interspersion with repetitive DNA. DNA **1**:133–143.