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::Tn5* 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 melliloti* 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 non-nodulating (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 melliloti* 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. melliloti* 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

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhizobium strains</td>
<td><strong>NGR234</strong> Broad-host-range cowpea Rhizobium sp.</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td><strong>ANU280</strong> Sm' Rif' derivative of NGR234</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td><strong>ANU2811</strong> ANU280 exoY11::Tn5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td><strong>ANU2890</strong> ANU280 exoY90::Tn5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td><strong>ANU2823</strong> ANU280 exoY23::Tn5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td><strong>ANU2808</strong> ANU280 exoY08::Tn5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td><strong>ANU2840</strong> ANU280 exoY40::Tn5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td><strong>Rm1021</strong> Sm' derivative of wild-type <em>R. meliloti</em></td>
<td>30</td>
</tr>
<tr>
<td></td>
<td><strong>USDA191</strong> Wild-type <em>R. fredii</em></td>
<td>2</td>
</tr>
<tr>
<td><em>E. coli</em> strains</td>
<td><strong>NM522</strong> Δ(lac-pro) F' lacZM15 lacY1 hsdS20</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td><strong>HB101</strong> leu proA2 rpsL20 (Sm') hsdS20</td>
<td>5</td>
</tr>
<tr>
<td>Plasmids</td>
<td><strong>pDS6</strong> Clone carrying wild-type <em>R. meliloti</em> exo DNA</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td><strong>R3222</strong> Clone carrying wild-type NGR234 exo DNA</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td><strong>pBR322</strong> Cloning vector, Ap' Tc'</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td><strong>pHC11</strong> EcoRI fragment (exoY11::Tn5) in pBR322</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td><strong>pRK2013</strong> Helper plasmid tra orfT ColE1</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td><strong>pRK290</strong> Broad-host-range IncP vector, Tc'</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td><strong>pJG106</strong> BamHI-HindIII (Km') fragment of Tn5 cloned into pRK290, Tc' Km'</td>
<td>Weinman*</td>
</tr>
<tr>
<td></td>
<td><strong>pJG22</strong> Recombinant pJG106 carrying wild-type NGR234 exo genes</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pJG22::Tn5</strong> Same as pJG22, but with exoY11::Tn5</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pMP220</strong> Broad-host-range IncP promoterless lacZ, Tc'</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td><strong>pG51a</strong> 7.0-kb BglII subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG52b</strong> 5.0-kb PstI subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG53b</strong> 1.9-kb Clal-PstI subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG54b</strong> 1,408-bp MluI-PstI subclone exoY-lacZ'</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG55a</strong> 1,221-bp MluI-HindIII partial subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG56a</strong> 992-bp MluI-NruI partial subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG57a</strong> 790-bp MluI-EcoRI subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG58a</strong> 1,121 NruI-PstI partial subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG59b</strong> 934-bp NruI-HindIII partial subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG60b</strong> 705-bp NruI subclone ecoX-lacZ'</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG61a</strong> 923-bp HindIII-PstI partial subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG62b</strong> 736-bp HindIII subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG63b</strong> 630-bp EcoRI subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG64b</strong> 695-bp HindIII subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG65b</strong> 1,392-bp EcoRI subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG66b</strong> 887-bp NruI-ClaI partial subclone</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td><strong>pG70b</strong> The pG60 insert in opposite orientation</td>
<td>This work</td>
</tr>
</tbody>
</table>


† 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. *phasesi* 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 stage, 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 Tn5 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 (Amer-
sham, England) were used for DNA transfers. Restriction
enzyme digests and ligations were performed according to
the specifications of the manufacturers (Boehringer Mann-
heim Biochemicals, Indianapolis, Ind., and New England
BioLabs, Inc., Beverly, Mass.). DNA probes were 32P
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
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 plas-
mids 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 promot-
ers.** Gene expression from cloned *exo* gene promoters, fused
to *E. coli lacZ* located on vector pMP220 (40), were deter-
mined by measuring β-galactosidase activity with O-nitro-
phenol-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). Nod-
ulated plants were tested for acetylene reduction by the
method of Heidecker et al. (24). The DNA sequence pre-
sented was sequenced entirely in both direc-
tions and was analyzed by using SEQ (a package of
computer programs available at the Research School of Biolog-
ical Sciences, Australian National University).

**Extraction of RNA from Rhizobium sp.** All solutions were
made up with sterile, deionized water and kept on ice
during use (unless otherwise indicated). A mid-
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 CH3COONa, 1% sodium
dodecyl sulfate, 1% 2-mercaptoethanol). A 300-μl volume of
phenol (equilibrated against TE) at 90°C was added and mixed,
followed by the addition of 300 μl of CHCl3-isooamyl 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 precipi-
tation 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 *Smal* 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 *EcoRI* site at nucleotide position 789 (see Figs. 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
garose gel slice by centrifugation in a TLS55 Beckman rotor
at 35K 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 piperezine-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 (Boehr-
ingren 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$_3$COONa (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% form-
amide 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 BamHI fragment known to contain exo
genes (11) was cloned into the broad-host-range IncP1 vector
pJJO16 (Table 1). This recombinant plasmid (pJG22) was trans-
ferred into ANU280 (a rifampin-resistant derivative of the
wild-type strain NGR234) and into several Tn5-induced Exo–
derivatives of ANU280 (ANU2811, ANU2890, ANU2823, ANU2808, and ANU2840). When pJG22 was trans-
ferred 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 × 10$^{-4}$
per recipient strain. The frequency of transfer was approxi-
mately 100-fold lower than for (i) the transfer of the vector
(pJJO16) alone to these strains and (ii) the transfer of pJG22
or pJJO16 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 sub-
sequent 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 restric-
tion 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 6.6-kb
EcoRI fragment, which contained the site of Tn5 insertion in
the mutant ANU2811, and now had a new EcoRI fragment,
which was larger by 5.8 kb (the length of the Tn5 sequences).

Hybridization analysis (Fig. 1B) showed that the altered
plasmids from the Exo+ transconjugants now carried the
Tn5 sequence with no detection of the wild-type 0.6-kb
EcoRI fragment. However, hybridization analysis of total
genomic DNA isolated from these Exo+ transconjugants
(data not shown) revealed that the 0.6-kb and the 6.6-kb plus
Tn5 bands were both present. Thus, the altered plasmids
probably resulted from double-reciprocal recombination
events between sequences flanking the Tn5 insertion in the
background genome and the homologous sequences carried
on the introduced plasmid.

A high frequency of recombination events was not ob-
erved 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 BamHI 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. (EZ) 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, BglII; C, ClaI; E, EcoRI; H, HindIII; M, MluI; N, NruI; P, PstI; S, SmaI.

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 genes (i.e., 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 Tn5 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 Tn5 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, Tn5 containing pJG22 plasmids conferred a dominant EXO⁻ phenotype when present in ANU280.

Symplastic 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 (Tn5 marker). This result indicated that these cells underwent recombination events and subsequent loss of plasmid during nodule passage. Similarly, the EXO⁻ ANU280 (pJG22::Tn5) 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 BamHI 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 Tn5 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 BamHI fragment, a series of subclones from this BamHI 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 Tn5 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 counters 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 Tn5 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 Tn5 insertion sites for 17 Exo" mutants mapped within two adjacent EcoRI fragments of 0.6 and 1.4 kb (Fig. 2). These EcoRI fragments were sequenced, and two very likely ORFs were found, designated exoY and ORF1. The Tn5 insertion sites for 5 representative 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 hydrophobic 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 Tn5 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 ΔG = -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 (Fig. 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 nuclelease 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 NruI 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 NruI site at nucleotide position 989. An exoY'-lacZ⁺ fusion (pJG54) at the PsrI site at nucleotide position 1408 (Fig. 5) 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 NruI 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-
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′-AGGCCG-3′, had five out of six matches with the procaryote consensus sequence (22). Virtually consensus-like (31) promoter sequences [5′-TTGAgg-(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::Tn5 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::Tn5) 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 HindIII site at position 100 to the NruI site at position 290. In contrast, an intragenic _exoY_ probe, from the HindIII site at position 1216 to the ClaI 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 Tn5 insertions into the wild-type genome of ANU280, and the locations of the mutations were mapped to specific _EcoRI_ 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 Tn5 insertions corresponding to the ANU2811 or ANU2890 allele resulted in Exo− colonies (persistent dominant phenotype). R-prime plasmids carrying Tn5 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 Tn5 insertion sites were located within _exoY_. The two insertion sites for the persistent dominant
Transposon Tn5 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_:Tn5 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_:Tn5 mutants, the frequency of revertants to _exo^+_ in the transconjugants is no longer 100%, but is now 52%. The other 48% of the transconjugants remain _exo^-_. In the _exo^-_ transconjugant, a normally rare double-reciprocal recombination event occurs between the DNA flanking the genomic Tn5 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_:Tn5 mutants. This is supported by the strong selection for wild-type _Rhizobium_ cells when merodiplod 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 Tn5 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 hydrophilic domains of other proteins. One protein that may be a candidate for this type of multimeric association is that encoded by exyY, which has an internal hydrophobic region spanning 24 amino acids. Currently we are investigating whether the product of exyY represses the transcription of exoX or, alternatively, whether there is a posttranslational association between the products of genes exyY 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 Martie 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 exoX to avoid the duplicative naming of homologous loci.

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

LITERATURE CITED


