

# Isolation and Characterization of *ropA* Homologous Genes from *Rhizobium leguminosarum* Biovars *viciae* and *trifolii*

HENK P. ROEST,\* CEES-JAN P. BLOEMENDAAL, CAREL A. WIJFFELMAN,  
AND BEN J. J. LUGTENBERG

*Institute of Molecular Plant Sciences, Leiden University, Wassendarseweg 66, 2333 AL Leiden, The Netherlands*

Received 21 November 1994/Accepted 22 June 1995

*ropA* encodes a 36-kDa outer membrane protein of *Rhizobium leguminosarum* bv. *viciae* strain 248 which constitutes the low- $M_r$  part of antigen group III (R. A. de Maagd, I. H. M. Mulders, H. C. J. Canter Cremers, and B. J. J. Lugtenberg, *J. Bacteriol.* 174:214–221, 1992). We observed that genes homologous to *ropA* are present in strain 248 as well as in other *R. leguminosarum* strains, and we describe the cloning and characterization of two of these genes. Sequencing of a 2.2-kb *Bgl*III fragment from *R. leguminosarum* bv. *viciae* strain 248 that hybridizes with *ropA* revealed one large open reading frame of 1,074 bp encoding a mature protein of 38,096 kDa. Homology between this gene and *ropA* is 91.8% on the DNA level. Homology on the amino acid level is only 69.9% as a result of a frameshift. On the basis of homology and immunochemical characteristics, we conclude that this gene encodes the high- $M_r$  part of the outer membrane protein antigen group III that is repressed during symbiosis. We named this gene *ropA2*. The second gene that we cloned was the *ropA* homologous gene of *R. leguminosarum* bv. *trifolii* strain LPR5020. Except for amino acid 43, the N-terminal part of the corresponding protein appeared to be identical to the first 51 amino acids of RopA of strain 248. The transcription start sites of both genes were determined, and the promoter regions were compared with that of *ropA* of strain 248. No clear consensus sequence could be deduced. The relationship of *ropA* and *ropA2* of *R. leguminosarum* bv. *viciae* strain 248 with two similar genes from *Brucella abortus* is discussed.

Rhizobia induce the formation of root nodules on appropriate leguminous host plants (for a review, see reference 20). In alfalfa and pea, the host plants of *Rhizobium meliloti* and *Rhizobium leguminosarum* bv. *viciae*, respectively, the nodules are of the indeterminate type. Five well-defined regions can be distinguished in such nodules. This subdivision in regions is based on the developmental stages of bacteroid differentiation and can be helpful in understanding the regulation of both plant and bacterial genes involved in symbiosis (35).

The *Rhizobium* bacteroids, present in the cytoplasm of root nodule cells, show extensive differences with free-living bacteria at the molecular and biochemical levels. Differences in expression of *nod* genes have been studied by using antibodies, in situ hybridizations, and *gusA* fusions. Transcription of the inducible *nod* genes was still detectable in the invasion zone but not in the early symbiotic zone (30, 31). The synthesis of extracellular polysaccharide, a cell surface component, is also severely decreased in bacteroids. It was shown that *exo::TnphoA* fusions were mainly expressed in the invasion zone and not in the symbiotic zone of alfalfa nodules (26). This was also shown for the *R. leguminosarum* bv. *phaseoli* gene *pssA*, as detected by using a fusion with the *gusA* gene. The polysaccharide inhibition gene *psiA* was expressed throughout the nodule and is probably responsible for the decrease in *pssA* expression (19).

The bacterial cell envelope, whose role in symbiosis is becoming more and more significant, also shows some clear differences during bacteroid development. Down-regulation was observed for certain outer membrane proteins of *R. leguminosarum* bv. *viciae* strain 248 bacteroids (4). This phenomenon was also observed in bacteroids with a chromosomal back-

ground of other biovars (27). The outer membrane proteins of strain 248 can be divided into four prominent antigen groups, designated I through IV. Two of these, antigen groups II and III, are severely decreased in bacteroids from indeterminate-nodule-forming host plants (27). Recently the structural gene encoding the lower- $M_r$  part of antigen group III, called *ropA*, was isolated (5). The promoter of this gene contains some specific sequences, including a consensus binding site for the integration host factor (IHF) and inverted repeats (5). Transcription of this gene was, as was the case with the genes involved in extracellular polysaccharide synthesis and nodulation, detected only in the zones distal to the symbiotic zone (8). A fusion of this gene with *lacZ* did not provide an unambiguous answer to the question of which factor(s) is responsible for the decrease in expression. To test whether a regulatory consensus exists, we isolated two *ropA* homologous genes and compared the promoter regions of these apparently similarly regulated genes with the promoter from *ropA*.

## MATERIALS AND METHODS

**Strains and plasmids.** The strains and plasmids used in this study are listed in Table 1. Tetracycline was added to concentrations of 20 and 2  $\mu\text{g/ml}$  for *Escherichia coli* and *R. leguminosarum*, respectively. Rifampin, streptomycin, and spectinomycin were added to final concentrations of 20, 500, and 400  $\mu\text{g/ml}$ , respectively.

**Construction of the LPR5020 genomic library.** Genomic DNA was isolated as described by Meade et al. (24), partially digested with *Eco*RI, and fractionated over a 10 to 40% (wt/vol) continuous sucrose gradient for 22 h at 20,000 rpm at 16°C (22). Fragments ranging from 15 to 30 kbp were pooled, dialyzed, and ligated into cosmid pLAFR1 (13). Concatemeric DNA was then packed in phages by using the Gigapack Plus packaging system (Stratagene, La Jolla, Calif.). The library was subsequently transduced into cells of *E. coli* HB101. Transductants were selected on solidified LB medium supplemented with tetracycline.

**Colony hybridization.** Colonies of *E. coli* HB101 containing the genomic library were transferred to GeneScreen filters (New England Nuclear, Boston, Mass.). The filters were treated as described earlier (22) and hybridized with a nick-translated [ $\alpha$ - $^{32}\text{P}$ ]dCTP probe (Amersham International, Little Chalfont, England).

\* Corresponding author. Present address: Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Phone: 31 104087150. Fax: 31 104360225. Electronic mail address: roest@gen.fgg.eur.nl.

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Reference or source
<i>E. coli</i>		
JM101	$\Delta(lac-proAB) supE thi (F' traD36 proAB lacI^qZ\Delta M15)$	37
HB101	$F^- supE44 thi recA13 hsdS20 lacY1 rpsL20 (Str^r)$	1
<i>R. leguminosarum</i>		
248	Wild type bv. viciae	17
LPR5020	Wild type bv. trifolii, $Str^r$	16
RBL5523	bv. trifolii cured of pSym5, contains Sym plasmid pRL1JI	25
B434	bv. phaseoli cured of pRP2JI, contains pRL1JI. <i>rhiA</i> ::Tn5	9
ANU843	Wild type bv. trifolii	28
<i>R. meliloti</i> 1021	SU47, $Str^r$	24
Plasmids		
M13Tg130	M13 sequencing vector	18
pIC20H	IncColE1, multicopy cloning vector	23
pLAFR1	IncP, lambda <i>cos</i> $Mob^+$ $Tra^-$ cosmid vector	13
pRK2013	Helper plasmid for triparental matings	10
pGP1127	<i>E. coli himA</i> and <i>himD</i> cloned in pBR322	N. Goosen, Leiden, The Netherlands
pMP92	IncP cloning vector	32
pMP2202	2.05-kb <i>EcoRI</i> fragment containing <i>ropA</i> cloned in pMP92	5
pMP2204	2.05-kb <i>EcoRI</i> of pMP2202 cloned in pIC20H	This study
pMP2250	pLAFR1 clone with genomic DNA of strain 248	This study
pMP2251	6.4-kb <i>EcoRI</i> fragment of pMP2250 in pIC20H	This study
pMP2254	2.3-kb <i>BglII</i> fragment of pMP2251 in pIC20H	This study
pMP2256	pLAFR1 clone with genomic DNA of strain LPR5020	This study
pMP2264	2.1-kb <i>EcoRI</i> fragment of pMP2256 in pIC20H	This study
pMP2267	2.3-kb <i>BglII</i> insert of pMP2254 in pMP92	This study
pMP2276	<i>SalI</i> deletion of pMP2264, contains the LPR5020 <i>ropA</i> promoter	This study

**DNA manipulations and bacterial matings.** Recombinant DNA techniques were carried out essentially as described by Maniatis et al. (22). Plasmids were transferred from *E. coli* to *R. leguminosarum* by using triparental mating as described previously (10), with pRK2013 as the helper plasmid.

**DNA sequencing.** DNA sequencing was performed as described by Sanger et al. (29), using either M13 vector Tg130 or pIC20H plasmids with [ $\alpha$ - $^{32}$ S]dATP (Amersham). Primers were synthesized on a model 392 DNA synthesizer (Applied Biosystems, Maarsse, The Netherlands). Sequenase (U.S. Biochemical, Cleveland, Ohio) and T7 polymerase (Pharmacia LKB, Woerden, The Netherlands) were used according to the manufacturers' guidelines.

**Primer extension.** Total RNA from tryptose-yeast extract-grown bacteria was isolated as described previously (34). The primer was labeled with [ $\gamma$ - $^{32}$ P]dATP (Amersham), using polynucleotide kinase (Pharmacia) as described by Maniatis et al. (22). Primer extension was performed by using avian myeloblastosis virus reverse transcriptase (Pharmacia) annealed on approximately 20  $\mu$ g of RNA at 45°C. To prevent self-copying of labeled primer, actinomycin D was added to the reaction mixture in a final concentration of 50  $\mu$ g/ml.

**Preparation of protein samples, electrophoresis, and Western blotting (immunoblotting).** Cell lysates were prepared as described earlier (27). Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed on 11% gels as described previously (21). After separation on SDS-polyacrylamide gels, the cell lysate constituents were transferred to nitrocellulose sheets by electroblotting, using an LKB Multiphor 2117 (Pharmacia). Immunodetection with monoclonal antibodies (MAbs) was performed as described elsewhere (4).

**Southern blot analysis.** Digested genomic DNA was separated in 1% (wt/vol) agarose gels and transferred to GeneScreen filters (New England Nuclear) by standard methods (22). Hybridization was performed at 42°C in 50% formamide-5 $\times$  SSPE (1 $\times$  SSPE is 150 mM NaCl, 10 mM sodium phosphate, and 1 mM EDTA [pH 7.4])-5% (wt/vol) SDS. The blots were washed at 65°C for 15 min with 2 $\times$  SSPE-0.1% SDS, followed by 1 $\times$  SSPE-0.1% SDS and 0.1 $\times$  SSPE-0.1% SDS under the same conditions. Probes were nick translated with [ $\alpha$ - $^{32}$ P]dCTP (Amersham).

**Nucleotide sequence accession numbers.** The nucleotide sequence of the *R. leguminosarum* bv. viciae strain 248 *ropA2* gene has been deposited in the GenBank/EMBL library under accession number X80794. The GenBank/EMBL accession number of the promoter sequence and the 5' region of *ropA* of *R. leguminosarum* bv. trifolii strain LPR5020 is X80793.

## RESULTS

**Detection of the *ropA* homologous genes in strain 248 and other *Rhizobium* strains.** Antigen group III, one of the two

antigen groups whose levels are severely decreased during symbiosis, consists of two primary proteins, both of which are also present as derivatives with peptidoglycan residues of various sizes attached to them. This antigen group is detectable with MAb37 and MAb38. The MAb38 epitope is specific for *R. leguminosarum* bv. viciae strain 248. MAb 37 detects antigen group III in a variety of *Rhizobium* and *Agrobacterium* strains (3). As was shown earlier, *ropA* encodes only the lower- $M_r$  part of the group III antigens of *R. leguminosarum* bv. viciae strain 248 (5). This finding implies that there is another structural gene that encodes for the high- $M_r$  part of the proteins of the antigen group III. Because both gene products have the MAb 37 epitope, we tested whether we could detect this second gene on the basis of homology with *ropA*. Genomic DNA of *R. leguminosarum* bv. viciae strain 248 was digested with *EcoRI* and analyzed in a Southern blot experiment with the *ropA* gene as a probe. This analysis revealed two strong hybridizing *EcoRI* fragments approximately 2.1 and 6.4 kb in size (Fig. 1, lane 3). The size of the smallest fragment resembled the size of the *EcoRI* fragment that contained the *ropA* gene (5). The second fragment indicated that there is, in addition to the *ropA*-containing fragment, another fragment with strong homology to *ropA*. We investigated whether this second fragment contained the gene encoding the high- $M_r$  part of antigen group III.

Because all *Rhizobium* and *Agrobacterium* strains tested contain antigen group III, we also tested for the presence of *ropA* homology in the genomes of some other strains. Three *R. leguminosarum* strains, bv. trifolii strain ANU843, bv. viciae strain RBL5523, and bv. phaseoli strain B434 (Fig. 1, lanes 1, 2, and 4, respectively) also showed strong hybridization signals with the *ropA* probe, despite the fact that antigen group III could not be detected in these strains when MAb 38 was used (3). The degree of hybridization of *ropA* with *R. meliloti* strain 1021 was moderate under the conditions used (Fig. 1, lane 5).

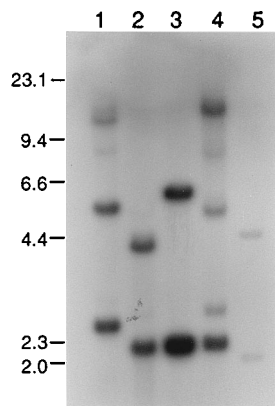


FIG. 1. Southern blot hybridization of *Eco*RI-digested total DNA probed with *ropA* of *R. leguminosarum* bv. *viciae* strain 248. Lanes: 1, *R. leguminosarum* bv. trifolii strain ANU843; 2, *R. leguminosarum* bv. *viciae* strain RBL5523; 3, *R. leguminosarum* bv. *viciae* strain 248; 4, *R. leguminosarum* bv. phaseoli strain B434; 5, *R. meliloti* 1021. *Hind*III-digested lambda DNA fragment sizes are indicated on the left.

**Isolation and subcloning of a cosmid clone isolated from a genomic library of strain 248 that hybridizes with *ropA*.** To obtain the *ropA* homologous region of strain 248, approximately 1,600 colonies of a cosmid library of strain 248 were transferred to GeneScreen filters, lysed, and tested in a Southern blot assay for hybridization with the probe containing the *ropA* gene. From these 1,600 colonies, seven colonies that hybridized with the probe were selected. The corresponding colonies on the master plates were selected and grown in LB supplemented with tetracycline. The cosmids of these seven colonies were isolated, digested with *Eco*RI, and again analyzed in a Southern blot assay using the same *ropA* probe as was used in the colony blot assay. Of these seven cosmid clones, two contained the 2.1-kb *Eco*RI fragment on which *ropA* is located, and five contained the 6.4-kb *Eco*RI fragment. We selected one of the latter cosmids, pMP2250, for further studies.

**Cosmid pMP2250 contains a MAb 38 epitope-expressing gene.** While all *Rhizobium* strains tested express the MAb37 epitope of antigen group III, only antigen group III from *R. leguminosarum* bv. *viciae* strain 248 contains the MAb38 epitope (3). This characteristic was used to detect expression of the MAb38 epitope in *R. leguminosarum*. Cosmid pMP2250 was cut with various restriction enzymes, and fragments hybridizing with a *ropA* probe were subcloned into pMP92, a broad-host-range cloning vector. Plasmids were transconjugated to *R. leguminosarum* bv. *viciae* strain RBL5523, and transconjugants were tested for expression of the MAb38 epitope (data not shown). Plasmid pMP2267, containing a 2.2-kb *Bgl*II fragment, is sufficient to express the MAb38 epitope. Figure 2 shows the detection of the MAb38 epitope in antigen group III in *R. leguminosarum* bv. *viciae* strain 248 cell lysates (lane A) and the expression of this epitope in isolated cell lysates of *R. leguminosarum* bv. *viciae* strain RBL5523 containing pMP2267 (lane D). Only the largest of the two major protein constituents of antigen group III of strain 248, comprising the approximately 40-kDa primary protein and its derivatives substituted with peptidoglycan residues, is produced. Lane C shows detection of the MAb38 epitope in cell lysates of *R. leguminosarum* bv. *viciae* strain RBL5523 containing pMP2202. This strain produces a number of protein bands with molecular masses of between 35 and 38 kDa. We conclude from these results that plasmids pMP2202 and pMP2267 are

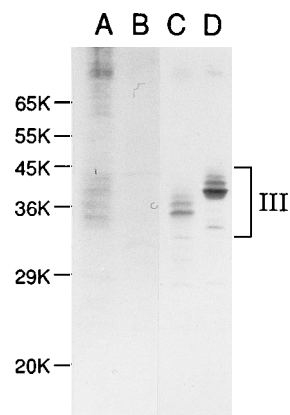


FIG. 2. Immunological detection of group III antigens in cell lysates by using MAb38. Lanes: A, *R. leguminosarum* bv. *viciae* strain 248; B to D, *R. leguminosarum* bv. *viciae* strain RBL5523 containing pMP92, pMP2202, and pMP2267, respectively. Positions of molecular weight markers are indicated on the left in kilodaltons, and the antigen group III region is indicated on the right.

responsible for expression of the *R. leguminosarum* bv. *viciae* strain 248 outer membrane antigen group III primary proteins of 35 and 40 kDa, respectively. As was observed previously for pMP2202 (5), expression of the MAb38 epitope from plasmid pMP2267 could not be detected in *E. coli* cells.

**Nucleotide sequencing and characterization of the fragment involved in expression of the high- $M_r$  part of antigen group III.**

The 2.2-kb *Bgl*II fragment was the smallest fragment that we could obtain that still resulted in expression of the upper part of antigen group III. This fragment was sequenced, and the result is shown in Fig. 3. Analysis of this fragment with the codon preference program of the University of Wisconsin Genetics Computer Group software package revealed one open reading frame large enough to encode the protein of antigen group III. This open reading frame extends from nucleotides 664 to 1830. An additional start codon, however, was observed at position 757 (Fig. 3), in frame with the first start codon. The open reading frame starting with the initiation codon at position 757, which we refer to as ORF38b, contains a consensus ribosome binding site starting 11 nucleotides upstream of the start of translation. This consensus, GGAGG, indicated in boldtype face in Fig. 3 at positions 746 to 750, was not found upstream of the initiation codon starting at position 664. We therefore assume that ORF38b is responsible for the production of the 40- to 45-kDa proteins containing the MAb38 epitope. ORF38b encodes a putative protein of approximately 38,096 kDa. By applying the standard weight-matrix analysis of von Heijne (36) to the amino acid sequence of ORF38b, a putative processing site, characteristic for exported proteins, was calculated. This processing site is located between residues 22 and 23 of the immature protein product. After processing, the calculated molecular mass of the mature protein encoded by ORF38b is 36.028 kDa, which is in reasonable agreement with the value of 40 kDa observed in SDS-PAGE. The putative signal sequence of this protein has the same length as the signal sequence of RopA. The deduced amino acid sequence of this protein is almost 70% identical to the amino acid sequence of RopA (Fig. 4) and shows a C-terminal phenylalanine residue often observed in outer membrane proteins, which is thought to be important for incorporation into the outer membrane (33). We therefore conclude that ORF38b encodes the high- $M_r$  part of the outer membrane protein group III antigens of strain 248. Because the gene products of *ropA*

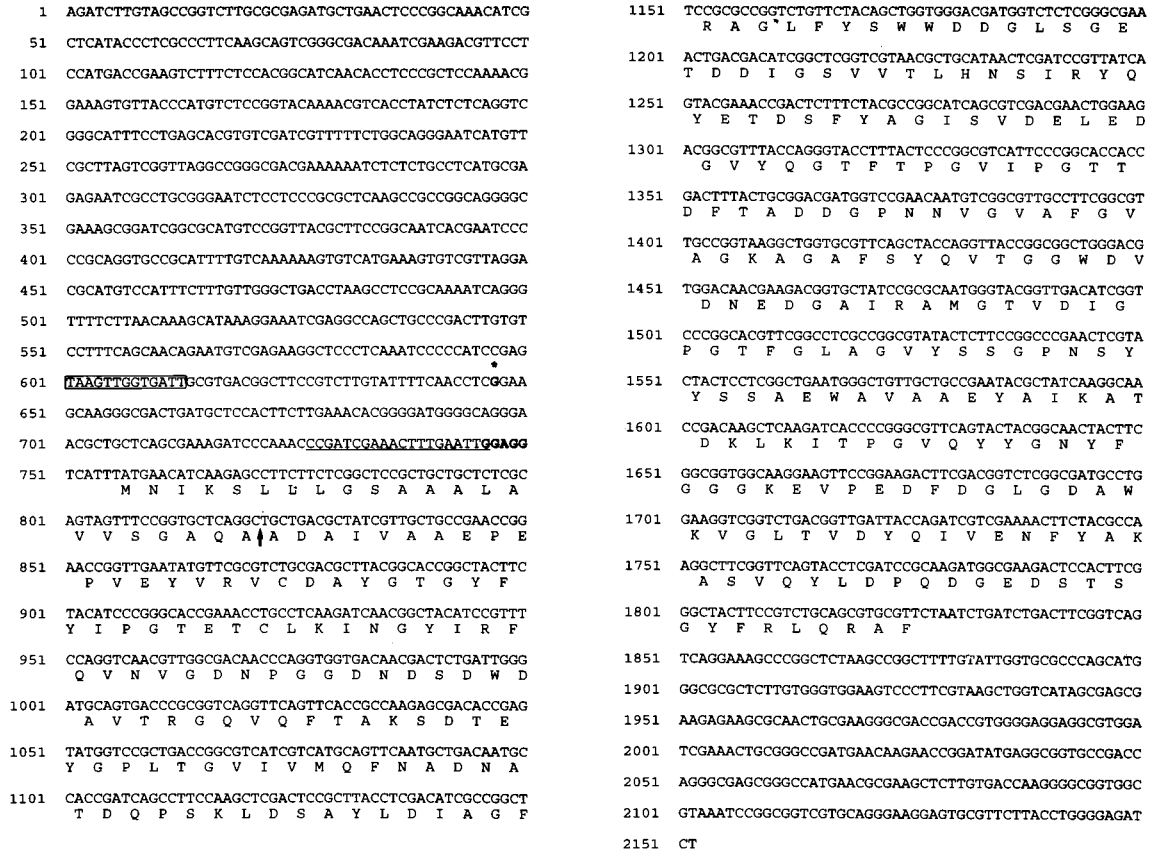


FIG. 3. DNA sequence of the 2.2-kb *Bgl*II fragment containing the *R. leguminosarum* bv. *viciae* strain 248 *ropA2* gene. The predicted amino acid sequence of ORF38b is indicated. The putative signal sequence cleavage site is indicated by an arrow. The underlined nucleotides indicate the primer used in the determination of the transcription start (marked by an asterisk). The IHF consensus binding site is boxed. The putative ribosome binding site is in boldface.

and ORF38b are responsible for the complete antigen group III of *R. leguminosarum* bv. *viciae* strain 248 and are very homologous (Fig. 4), we propose the name *ropA2*, in analogy with the presence of multiple *nodD* gene copies in the genomes of various *Rhizobium* species.

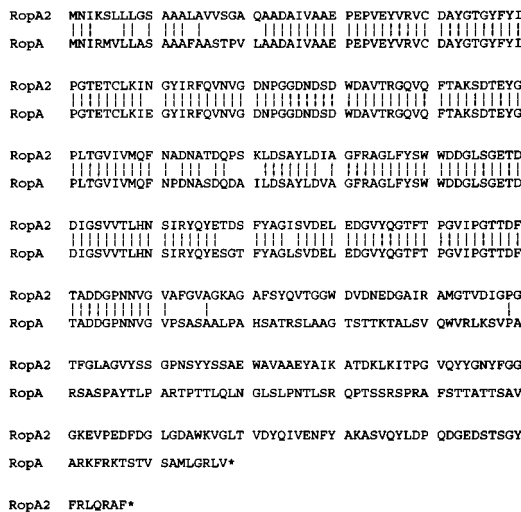


FIG. 4. Alignment of the amino acid sequences of RopA and RopA2 of *R. leguminosarum* bv. *viciae* strain 248.

**Isolation and sequencing of the untranslated upstream and 5' region of *R. leguminosarum* bv. trifolii strain LPR5020 *ropA*.** To deduce a possible consensus sequence in the promoter region of outer membrane protein antigen group III responsible for down-regulation during symbiosis, we wanted to isolate the *ropA* homologous genes of *R. leguminosarum* bv. trifolii strain LPR5020 and compare the upstream promoter region with that of the *ropA* and *ropA2* genes of *R. leguminosarum* bv. *viciae* strain 248. As mentioned earlier, *ropA* appears to be very well conserved among various *R. leguminosarum* strains (Fig. 1). *R. leguminosarum* bv. *viciae* RBL523, a strain derived from *R. leguminosarum* bv. trifolii strain LPR5020, contains a *ropA* homologous *Eco*RI fragment of the same size as the *Eco*RI fragment on which *ropA* of strain 248 is located (5). To isolate this fragment, we made a cosmid library of strain LPR5020 and tested approximately 1,500 colonies in the colony hybridization assay with a *ropA* probe. Five colonies containing *ropA* homologous DNA were isolated. They were further analyzed by restriction enzyme analysis and Southern hybridization. One cosmid contained the *Eco*RI fragment of 2.1 kb that hybridized with the *ropA* probe (Fig. 1, lane 2). This is probably the fragment containing the *ropA* gene of *R. leguminosarum* bv. trifolii strain LPR5020. The other four contained the 4.3-kb *Eco*RI fragment (Fig. 1, lane 2). The latter fragment is probably responsible for the high-*M<sub>r</sub>* part of the antigen group III of *R. leguminosarum* bv. trifolii strain LPR5020. The 2.1-kb *Eco*RI fragment was isolated, and restriction sites were mapped. The restriction patterns of LPR5020 *ropA* and strain

```

1 GAATTGAGAACACCATGTTCTAAACTGGCGCCTTGTGTTGCTTGTAGC
51 CGGCGGTAGAGTTTATGCGGCTTACCAGTATGTTTCGACCGGGTTCGGAGC
101 GTGGAGAATTGACAAATGGGGTGTAACTGTCGGCTTGGAGTGGAGGAT
151 TTCGTCGCGGGGGCCGCTTTGCCTGTTTTTCGTCTAGAAATTTGAATTTG
201 TTGTATAAAAATCACGCTGCCGGGAAAGTAGCCCGTGAGAACCCATCTG
251 GCGTATATCGAGAATTTGCTTGAAGAGCCTTGTATGTGATGACGGTTTGG
301 CGGTTGATCGATCCGCACGCGGAGATGGCTGTTCAAACCTGGGCCGGCTT
351 CAAATCGACTACCCGATCTGAAGCCAGCCCGATCGACAAAAGGAATGGAT
401 TGGTAAATGAACATCAGAATGGTTTTGCTTGCATCAGCAGCAGCATTTGC
451 TGCATCGACCGGTTCTTGCAGCTGACGCTATTGTTGCTGCTGAGCCGG
501 AACCGTTGAGTATGTTCCGCTCTGCGAGCTGACGGCACCGCTACTTC
551 TACATCCCGGG
    Y I P
  
```

FIG. 5. Nucleotide sequence of the 5' end of the open reading frame (positions 407 to 561), the untranslated 5' region (positions 300 to 406), and the upstream promoter region (positions 1 to 299) of the *R. leguminosarum* bv. trifolii strain LPR5020 *ropA* gene. The sequence for the primer used to determine the start of transcription is underlined. The asterisk is positioned at the start of transcription; the arrow indicates the putative signal sequence cleavage site. The putative ribosome binding site is shown in boldface. The boxed regions indicate the two putative IHF binding sites.

248 *ropA* were very similar, and we sequenced the part of the fragment that, as judged from the restriction pattern, contained the start of transcription and translation and the upstream promoter region of the *ropA* gene of LPR5020 (Fig. 5). The sequenced part of the fragment is 95.2% identical to the same region in *ropA* of strain 248. Comparing the *ropA* genes of strain 248 and LPR5020, we identified the putative start of the open reading frame of RopA of *R. leguminosarum* bv. *viciae* strain LPR5020. With one exception, the 51 N-terminal amino acids are identical to the 51 N-terminal amino acids of the RopA protein of strain 248. A putative ribosome binding site was identified at positions 388 to 394, a location identical to that of the putative ribosome binding site of the *ropA* gene of strain 248 (5).

**Determination of the transcription start sites of *ropA* of strain LPR5020 and *ropA2* of strain 248 and characterization of the untranslated upstream regions.** Down-regulation of *ropA* was shown to take place at the level of transcription (8). To identify the regions upstream of *ropA2* of strain 248 and of *ropA* of strain LPR5020, we determined the transcription start sites of these genes and compared them with the *ropA* promoter of strain 248. To determine the transcriptional start site of LPR5020 *ropA*, we used a synthetic primer antisense to nucleotides 379 to 398 (Fig. 4) for primer extension. The result is shown in Fig. 6A. This experiment revealed a transcript starting at position 300, 107 nucleotides upstream of the initiation codon (see also Fig. 4). No clear -10 and -35 sequences could be identified in the upstream promoter region. Two possible integration host factor (IHF) binding sites, indicated in Fig. 4, were observed at positions 185 to 197 and 262 to 274. These sites are known to play a role in regulation of transcription (14). No further homologies to recognition sites for regulatory proteins could be identified. The transcription start site of *ropA2* was determined in a similar way. A synthetic primer antisense to the nucleotides at positions 727 to 746 (Fig. 3) was used, and primer extension identified the start of transcription at position 647 (Fig. 3). Also, no consensus -10 and -35 promoter sequences could be identified. In addition, no clear homologies with any recognition site for regulatory proteins was detected except for an upstream activator sequence (2) at

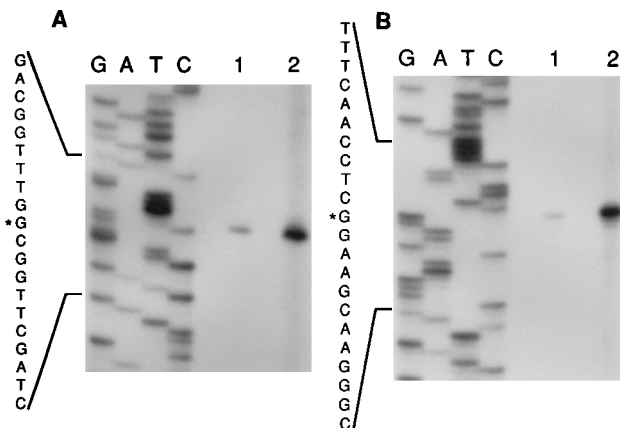


FIG. 6. Determination of the transcriptional start site by primer extension of *R. leguminosarum* bv. trifolii strain LPR5020 *ropA* (A) and *R. leguminosarum* bv. *viciae* 248 *ropA2* (B). A sequence ladder (lanes G, A, T, and C) was run simultaneously with the extension product. Lanes: 1, approximately 5  $\mu$ g of mRNA; 2, approximately 20  $\mu$ g of mRNA. The transcription start sites of both genes are indicated on the left with asterisks.

positions 548 to 563 and an IHF binding site at positions 601 to 613; the latter is indicated in Fig. 3.

**Alignment of the *ropA* and *ropA2* promoter regions.** The genes *ropA* and *ropA2* are responsible for the expression of antigen group III in *R. leguminosarum* bv. *viciae* strain 248. This antigen group is down-regulated during bacteroid development. It has been shown that the down-regulation of *ropA* takes place on the level of transcription (8). Because we are interested in identifying a putative regulatory promoter sequence responsible for this phenomenon, we compared the consensus obtained by aligning the two *ropA* promoter regions with the promoter region of *ropA2*. The assumption was made that any region of homology should be at approximately the same distance from the start of transcription for both *ropA* and *ropA2*. No clear regions of homology were observed (data not shown). Whether this means that *ropA* and *ropA2* are regulated differently remains to be established.

## DISCUSSION

In this study, we describe the isolation of two genes on the basis of their homology with the *ropA* gene of *R. leguminosarum* bv. *viciae* strain 248. The first gene was isolated from the same strain as *ropA*, namely, *R. leguminosarum* bv. *viciae* strain 248, whereas the second was the *ropA* gene of *R. leguminosarum* bv. trifolii strain LPR5020.

In addition to the restriction fragment containing the *ropA* sequence, a second restriction fragment containing a *ropA* homologous sequence was detected in strain 248. The latter sequence, which is 91.8% identical to *ropA*, encodes a protein of approximately 40 kDa that, like RopA, reacts with both MAb 37 and MAb38. MAb38 detects an epitope that is present only in antigen group III outer membrane proteins of strain 248, and we concluded that this sequence is the gene that encodes the 40-kDa major protein of antigen group III. The gene that encodes the high- $M_r$  part of antigen group III, designated *ropA2*, is 1,074 nucleotides in size and encodes a protein with a calculated molecular mass of 38.096 kDa, whereas the *ropA* open reading frame (ORF38) is 957 nucleotides long and encodes a protein with a calculated mass of 33.686 kDa. Since the determined (*ropA*) and calculated (*ropA2*) signal sequence cleavage sites are both located between amino acids 22 and 23,

the resulting mature proteins are 36.028 and 31.486 kDa for *ropA2* and *ropA*, respectively, in agreement with their difference in mobility in SDS-polyacrylamide gels and with the molecular weights of the primary proteins of this antigen group III. Despite the 91.8% homology between *ropA* and *ropA2* on the DNA level, the deduced amino acid sequences show only 69.9% homology overall. As the result of an insertion at position 1390 (Fig. 2), the subsequent frameshift results in the lack of homology beyond amino acid 211 (Fig. 4).

As was shown in earlier studies, *ropA* and *ropA2* are down-regulated during symbiosis (4). In situ hybridization experiments with an antisense *ropA* probe revealed a sudden decrease in mRNA level at the transition from prefixation zone II to interzone II-III (8). Because *ropA* and *ropA2* are very homologous and hybridize under stringent conditions (Fig. 1), it is likely that in this in situ experiment, mRNAs of both genes were detected. Because no signal could be detected in situ with an antisense *ropA* probe, it is also likely that neither of these genes is transcribed in the symbiotic zone of the nodule and that both are regulated at the level of transcription. Whether this regulation occurs for both genes in the same region of the nodule cannot be concluded from these data.

It was shown for isolated cell envelope profiles that differences between various *Rhizobium* strains can be substantial with respect to antigen group III (3). For example, *R. leguminosarum* bv. trifolii strain ANU843 may, as judged from the protein profile of isolated cell envelopes, contain only one active gene (6). However, we detected two *ropA* homologous fragments (Fig. 1, lane 1). Whether one of these genes is inactive or whether, for instance, both genes encode the same protein remains to be established. In this report, we show that RBL5523, a strain derived from LPR5020, also contains two fragments that hybridize with *ropA* (Fig. 1, lane 2). This finding suggests a subdivision of their antigen group III into two parts, similar to the situation in strain 248: one fragment that contains the *ropA* gene, and another fragment that contains a homologous sequence presumably encoding the higher- $M_r$  part of antigen group III of this strain. The isolation of the *ropA* gene from *R. leguminosarum* bv. trifolii strain LPR5020 was possible only on the basis of homology with *ropA* of strain 248. Expression could not be checked by using MAb38, since MAb 38 reacts only with RopA or RopA2 of strain 248, and the other MAb against antigen group III, MAb37, reacts with all tested *Rhizobium* strains. Because RopA was not expressed in *E. coli*, evidence for expression of this gene was obtained by using a primer extension experiment. We used a specific primer hybridizing to a region in the 5' untranslated region of the *ropA* gene of strain LPR5020, tested isolated RNA for the presence of *ropA* transcripts, and determined the transcription start site. It is striking that the 5' untranslated region is of the same length and 100% conserved between *ropA* genes of strains 248 and LPR5020.

In our effort to identify a common feature in the promoter regions of the two *ropA* genes and *ropA2*, we compared the promoter regions with each other. We found that all three promoters contain one or two IHF binding sites (14). In the promoter of the *ropA* gene of strain LPR5020, two IHF binding sites are observed. One (IHF1; Fig. 4) is situated at a similar position as the IHF binding site of strain 248 *ropA*; the other is situated more upstream (IHF2; Fig. 4). The *ropA2* promoter contains only one IHF consensus sequence, located at positions 601 to 613. Because IHF can act both positively and negatively on transcription (14), this site may be involved in regulation of these genes. A retardation complex was observed in a bandshift experiment in which a fragment containing the *ropA* promoter was incubated with cell extracts of *E.*

TABLE 2. Homology between *R. leguminosarum* bv. viciae strain 248 RopA and RopA2 and *B. abortus* Omp2a and Omp2b<sup>a</sup>

	RopA	RopA2	Omp2a	Omp2b
RopA		69.9	32.2	33.4
RopA2	75.6		40.3	35.9
Omp2a	54.4	62.0		86.0
Omp2b	54.0	58.6	90.7	

<sup>a</sup> Percentage identity is given above the diagonal, and percentage similarity is given below the diagonal.

*coli* and *R. leguminosarum* bv. viciae 248. This complex was absent when a cell extract of an IHF-negative mutant of *E. coli* was used (29a). These results suggest a function of this type of regulatory sequence in the regulation of *ropA*. The IHF binding sites of *ropA2* and *ropA* of strain LPR5020 are not situated in a very A+T-rich region. Moreover, no signal was detected in a Southern blot analysis on total *EcoRI*-digested genomic DNA of various *Rhizobium* strains with plasmid pGP1127, which contains the two genes that encode the proteins that make up the IHF (data not shown). If there are IHF-encoding genes in *Rhizobium*, their nucleotide sequences are not very homologous to those of the IHF genes of *E. coli*.

Another striking similarity between the promoter regions of the *rop* genes is the absence of a clear consensus sequence for the binding of RNA polymerase. The recognition sequence of the RNA polymerase is determined by the type of  $\sigma$  factor used. A number of  $\sigma$  factors and their consensus sequences are known (15), but no consensus sequence for any of the  $\sigma$  factors could be discerned. It might well be that transcription of outer membrane protein-encoding genes in *Rhizobium* species requires a yet unknown  $\sigma$  factor that is inactive or not present in bacteroids. No other (possible) regulatory consensus sequence could be observed.

A recent sequence database search with the sequence of *ropA* revealed a striking similarity with two genes of a facultative intracellular gram-negative bacterium, *Brucella abortus* (11). These genes, designated *omp2a* and *omp2b* (12), are believed to encode porins. The homologies of the Omp2a and Omp2b proteins with RopA and RopA2 are listed in Table 2. It was shown that only one of the genes (*omp2b*) was actively transcribed in *B. abortus* and *E. coli* (12). This does not rule out expression of the *omp2a* gene under different environmental conditions or only in *B. abortus*. Additional experiments indicate that after infection of the host cells, the expression of *omp2b* is also severely depleted (10a), a phenomenon similar to the down-regulation of both *ropA* and *ropA2*. We could locate an IHF consensus binding site approximately 50 nucleotides upstream of the translational start site of *omp2b*, but because the start of transcription was not determined, any significance of this observation remains to be established.

The data presented here indicate that *ropA* and *ropA2* produce the outer membrane proteins of antigen group III. A possible function in phage attachment might be concluded from data presented earlier (7). However, although they are likely to represent porins, their exact function is still not clear; we hope, after isolating mutants in the *ropA* and/or *ropA2* genes, to get an indication of their function and regulation both in free-living bacteria and in bacteroids.

#### REFERENCES

- Boyer, H. W., and D. Roulland-Dussoix. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* **41**:459-472.
- Buck, M., S. Miller, M. Drummond, and R. Dixon. 1986. Upstream activator

- sequences are present in the promoters of nitrogen fixation genes. *Nature* (London) **320**:374–378.
3. **de Maagd, R. A.** 1989. Studies on the cell surface of the root-nodulating bacterium *Rhizobium leguminosarum*. Ph.D. thesis. Leiden University, Leiden, The Netherlands.
  4. **de Maagd, R. A., R. de Rijk, I. H. M. Mulders, and B. J. J. Lugtenberg.** 1989. Immunological characterization of *Rhizobium leguminosarum* outer membrane antigens using polyclonal and monoclonal antibodies. *J. Bacteriol.* **171**:1136–1142.
  5. **de Maagd, R. A., I. H. M. Mulders, H. C. J. Canter Cremers, and B. J. J. Lugtenberg.** 1992. Cloning, nucleotide sequencing, and expression in *Escherichia coli* of a *Rhizobium leguminosarum* gene encoding a symbiotically repressed outer membrane protein. *J. Bacteriol.* **174**:214–222.
  6. **de Maagd, R. A., C. van Rossum, and B. J. J. Lugtenberg.** 1988. Recognition of individual strains of fast-growing rhizobia by using profiles of membrane proteins and lipopolysaccharides. *J. Bacteriol.* **170**:3782–3785.
  7. **de Maagd, R. A., F. B. Wientjes, and B. J. J. Lugtenberg.** 1989. Evidence for divalent cation (Ca<sup>2+</sup>)-stabilized oligomeric proteins and covalently bound protein-peptidoglycan complexes in the outer membrane of *Rhizobium leguminosarum*. *J. Bacteriol.* **171**:3989–3995.
  8. **de Maagd, R. A., W.-C. Yang, L. Goosen-de Roo, I. H. M. Mulders, H. P. Roest, H. P. Spaink, T. Bisseling, and B. J. J. Lugtenberg.** 1994. Down-regulation of expression of the *Rhizobium leguminosarum* outer membrane protein gene *ropA* occurs abruptly in interzone II-III of pea nodules and can be uncoupled from *nif* gene expression. *Mol. Plant-Microbe Interact.* **7**:276–281.
  9. **Dibb, N. J., J. A. Downie, and N. J. Brewin.** 1984. Identification of a rhizosphere protein encoded by the symbiotic plasmid of *Rhizobium leguminosarum*. *J. Bacteriol.* **158**:621–627.
  10. **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.
  - 10a. **Ficht, T.** Personal communication.
  11. **Ficht, T. A., S. W. Bearden, B. A. Sowa, and L. G. Adams.** 1988. A 36-kilodalton *Brucella abortus* cell envelope protein is encoded by repeated sequences closely linked in the genomic DNA. *Infect. Immun.* **56**:2036–2046.
  12. **Ficht, T. A., S. W. Bearden, B. A. Sowa, and L. G. Adams.** 1989. DNA sequence and expression of the 36-kilodalton outer membrane protein gene of *Brucella abortus*. *Infect. Immun.* **57**:3281–3291.
  13. **Friedman, A. M., S. R. Long, S. E. Brown, W. J. Buikema, and F. M. Ausubel.** 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene* **18**:289–296.
  14. **Friedman, D. I.** 1988. Integration host factor: a protein for all reasons. *Cell* **55**:545–554.
  15. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**:839–872.
  16. **Hooykaas, P. J. J., A. A. N. van Brussel, H. den Dulk-Ras, G. M. S. van Slogteren, and R. A. Schilperoort.** 1981. Sym-plasmid of *Rhizobium trifolii* expressed in different Rhizobial species and *Agrobacterium tumefaciens*. *Nature* (London) **291**:351–353.
  17. **Josey, D. P., J. L. Beynon, A. W. B. Johnston, and J. E. Beringer.** 1979. Strain identification in *Rhizobium* using intrinsic antibiotic resistance. *J. Appl. Bacteriol.* **46**:343–350.
  18. **Kieny, M. P., R. Lathe, and J. P. Lecocq.** 1983. New versatile cloning and sequencing vectors based on bacteriophage M13. *Gene* **26**:91–99.
  19. **Latchford, J. W., D. Borthakur, and A. W. B. Johnston.** 1991. The products of *Rhizobium* genes, *psi* and *pss*, which affect exopolysaccharide production, are associated with the bacterial cell surface. *Mol. Microbiol.* **5**:2107–2114.
  20. **Long, S. R.** 1989. Rhizobium-legume nodulation: life together in the underground. *Cell* **56**:203–214.
  21. **Lugtenberg, B., J. Meyers, R. Peters, P. van der Hoek, and L. van Alphen.** 1975. Electrophoretic resolution of the major outer membrane protein of *Escherichia coli* K12 into four bands. *FEBS Lett.* **58**:254–258.
  22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  23. **Marsh, J. L., M. Erfle, and E. J. Wykes.** 1984. The PIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. *Gene* **32**:481–485.
  24. **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 Tn5 mutagenesis. *J. Bacteriol.* **149**:114–122.
  25. **Priem, W. J. E., and C. A. Wijffelman.** 1984. Selection of strains cured of the *Rhizobium leguminosarum* Sym plasmid by using small bacteriocin. *FEMS Microbiol. Lett.* **25**:247–251.
  26. **Reuber, T. L., S. Long, and G. C. Walker.** 1991. Regulation of *Rhizobium meliloti* *exo* genes in free-living cells and in planta examined by using TnphoA fusions. *J. Bacteriol.* **173**:426–434.
  27. **Roest, H. P., L. Goosen-de Roo, C. A. Wijffelman, R. A. de Maagd, and B. J. J. Lugtenberg.** Outer membrane protein changes during bacteroid development are independent of nitrogen fixation and differ between indeterminate and determinate nodulating host plants of *Rhizobium leguminosarum*. *Mol. Plant-Microbe Interact.* **8**:14–22.
  28. **Rolfe, B. G., P. M. Gresshoff, and J. Shine.** 1980. Rapid screening method for symbiotic mutants of *Rhizobium leguminosarum* biovar trifolii and white clover plants. *Plant Sci. Lett.* **19**:277–284.
  29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
  - 29a. **Schlaman, W. R. M.** Personal communication.
  30. **Schlaman, W. R. M., B. Horvath, E. Vijgenboom, R. J. H. Okker, and B. J. J. Lugtenberg.** 1991. Suppression of nodulation gene expression in bacteroids of *Rhizobium leguminosarum* biovar *viciae*. *J. Bacteriol.* **173**:4277–4287.
  31. **Sharma, S. B., and E. R. Signer.** 1990. Temporal and spatial regulation of the symbiotic genes of *Rhizobium meliloti* in planta revealed by transposon Tn5-*gusA*. *Genes Dev.* **4**:344–356.
  32. **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 pRL1J1. *Plant Mol. Biol.* **9**:27–39.
  33. **Struyvé, M., M. Moons, and J. Tommassen.** 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* **218**:141–148.
  34. **van Slogteren, G. M. S., J. H. C. Hoge, P. J. J. Hooykaas, and R. A. Schilperoort.** 1983. Clonal analysis of heterogenous crown gall tumour tissues induced by wild-type and shooter mutant strains of *Agrobacterium tumefaciens*: expression of T-DNA genes. *Plant Mol. Biol.* **2**:321–333.
  35. **Vasse, J., F. de Billy, S. Camut, and G. Truchet.** 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* **172**:4295–4306.
  36. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14**:4683–4690.
  37. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.