MPMI Vol. 9, No. 3, 1996, pp. 151–163. Publication no. M-1996-0115-01R.

Characterization of *Rhizobium tropici* CIAT899 Nodulation Factors: The Role of *nodH* and *nodPQ* Genes in Their Sulfation

Jorge L. Folch-Mallol[1,2], Silvia Marroquí[1], Carolina Sousa[1], Hamid Manyani[1], Isabel M. López-Lara[4], Koen M. G. M. van der Drift[3], Johan Haverkamp[3], Carmen Quinto[2], Antonio Gil-Serrano[5], Jane Thomas-Oates[3], Herman P. Spaink[4], and Manuel Megías[1]

[1]Department of Microbiology and Parasitology, University of Seville, 41012 Seville, Spain; [2]Department of Plant Molecular Biology, Institute of Biotechnology, U.N.A.M., Cuernavaca, Morelos 62271, México; [3]Department of Mass Spectrometry, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands; [4]Institute of Molecular Plant Sciences, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands; and [5]Department of Organic Chemistry, University of Seville, 41012 Seville, Spain.
Received 2 June 1995. Accepted 18 December 1995.

We have purified and characterized the nodulation factors produced by *Rhizobium tropici* CIAT899. This strain produces a large variety of nodulation factors, these being a mixture of sulfated or nonsulfated penta- or tetra-chito-oligosaccharides to which any of six different fatty acyl moieties may be attached to nitrogen of the nonreducing terminal residue. In this mixture we have also found methylated or nonmethylated lipo-chitin oligosaccharides. Here we describe a novel lipo-chitin-oligosaccharide consisting of a linear backbone of 4 *N*-acetylglucosamine residues and one mannose that is the reducing-terminal residue and bearing a C18:1 fatty acyl moiety on the nonreducing terminal residue. In addition, we have identified, cloned, and sequenced *R. tropici nodH* and *nodPQ* genes, generated mutations in the *nodH* and *nodQ* genes, and tested the mutant strains for nodulation in *Phaseolus* and *Leucaena* plants. Our results indicate that the sulfate group present in wild-type Nod factors plays a major role in nodulation of *Leucaena* plants by strain CIAT899 of *R. tropici*.

Hemos purificado y caracterizado los factores de nodulación producidos por *Rhizobium tropici* CIAT899. Esta cepa produce una amplia variedad de factores de nodulación, siendo éstos una mezcla de tetra- o penta-quito-oligosacáridos sulfatados o no sulfatados, a los cuales se puede unir un ácido graso al nitrógeno del residuo terminal no reductor. En esta mezcla también hemos encontrado lipo-quitin-oligosacáridos metilados y no metilados. Aquí describimos un nuevo lipo-quitin-oligosacárido que consiste en un esqueleto lineal de 4 residuos de N-acetil-glucosamina y una manosa como residuo terminal reductor, y que porta un sustituyente ácido graso C18:1 en el residuo terminal no reductor. Además, hemos identificado, clonado y secuenciado los genes *nodH* y *nodPQ* de *R. tropici*, generado mutaciones en los genes *nodH* y *nodQ* y probado las cepas mutantes en cuanto a nodulación en plantas de *Phaseolus* y *Leucaena*. Nuestros resultados indican que el grupo sulfato presente en los factores de nodulación silvestres juegan un papel importante en la nodulación de plantas de *Leucaena* por la cepa CIAT899 de *R. tropici*.

Additional keywords: LCOs, mass spectrometry, *nodHPQ*, sulfation.

Corresponding author: Manuel Megías, Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Seville, Apdo. Postal 874, 41012 Seville, Spain; Tele.: 34-5-455-67-66; Fax: 34-5-462-81-62; E-mail: megias@cica.es

Nucleotide sequence data reported in this paper will appear in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under the accession number X87608.

Bacteria of the genus *Rhizobium* are able to fix nitrogen in symbiotic associations with leguminous plants. These associations are highly specific and depend on both partners, the bacterium and the plant, to be effective. *Rhizobium*-legume interaction is triggered by a signal exchange between the two symbiotic partners. It has been shown that bacteria secrete low molecular weight extracellular signals in response to plant regulators (Spaink and Lugtenberg 1994). The bacterial Nod factors (or nodulation factors) share a "core" structure consisting of three to five β-1,4-linked N-acetylated glucosamine (GlcNAc) residues to which a lipid moiety is added in amide linkage to the nonreducing terminal residue. This lipid moiety, as well as other functional groups such as sulfate or methyl groups attached to the lipo-chitin-oligosaccharide (LCO), vary between species (for a review see Dénarié and Cullimore 1993; Megías et al. 1993; and Carlson et al. 1994).

There is evidence suggesting that the different functional groups attached to the LCO determine host specificity. In this way, the structure of the lipid moiety on *R. leguminosarum* bvs. *viciae* and *trifolii* Nod factors is a major host determinant in this species (Spaink et al. 1995). Another clear example is the sulfate group present in *R. meliloti* LCOs, which is absolutely necessary for nodulation of its normal host alfalfa plants (Roche et al. 1991). In this species it has been shown that *nodPQ* and *nodH* genes are involved in the sulfation of the Nod factors and mutants in *nodH* are severely impaired in

their ability to nodulate alfalfa, since they produce Nod factors that lack the sulfate group present in the wild-type strain. However, single mutants in *nodP* or *nodQ* are still able to nodulate alfalfa plants, although with a certain delay (Roche et al. 1991).

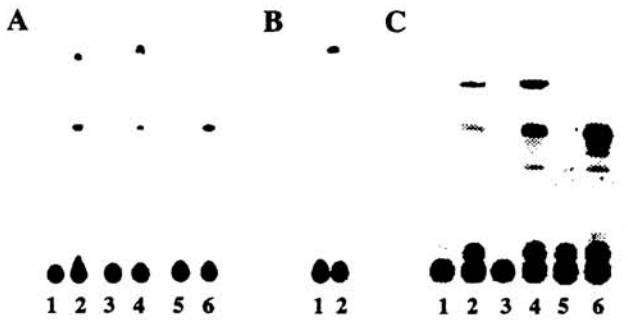

Fig. 1. Thin-layer chromatography analysis of Nod factors produced in the presence of **(A)** ^{14}C-acetate, **(B)** ^{35}S-sulphate, and **(C)** L-(methyl-^{14}C)-methionine of strains CIAT899 (lanes 1 and 2 in A, B, and C), RSP900/pCV38 (lanes 3 and 4 in A and C), and RSP900/pCV3804 (lanes 5 and 6 in A and C). Odd lanes without inducer, even lanes with 2 mM naringenin.

The biochemical function of NodP, NodQ, and NodH proteins in the biosynthetic pathway leading to the sulfation of *R. meliloti* LCOs has now been elucidated. NodP and NodQ are homologous to the *Escherichia coli* CysDNC system involved in the production of activated sulfate compounds. Thus, NodP and the N terminus of NodQ encode both subunits (CysD and CysN) of the ATP sulfurilase activity that produces APS (adenosine-5′-phosphosulfate), an activated sulfate compound, from ATP and sulfate (Schwedock and Long 1990). APS is then converted to PAPS (3′-phosphoadenosine-5′-phosphosulfate), a second activated sulfate compound, by means of the APS kinase activity (CysC) encoded by the carboxy terminus of NodQ (Schwedock and Long 1993). The NodH protein is probably the specific sulfotransferase that transfers sulfate from PAPS to the reducing end of the LCO (Schultze et al. 1995).

Rhizobium tropici is a broad host range species that can nodulate several nonrelated legumes such as *Leucaena*, *Phaseolus*, *Medicago*, and *Macroptilium* (Martínez et al. 1991). This species is tolerant to acidic soils and to high temperatures, and its symbiotic plasmids confer the ability to fix nitrogen when introduced into *Agrobacterium tumefaciens* (Martínez et al. 1985). In this paper we describe the structure

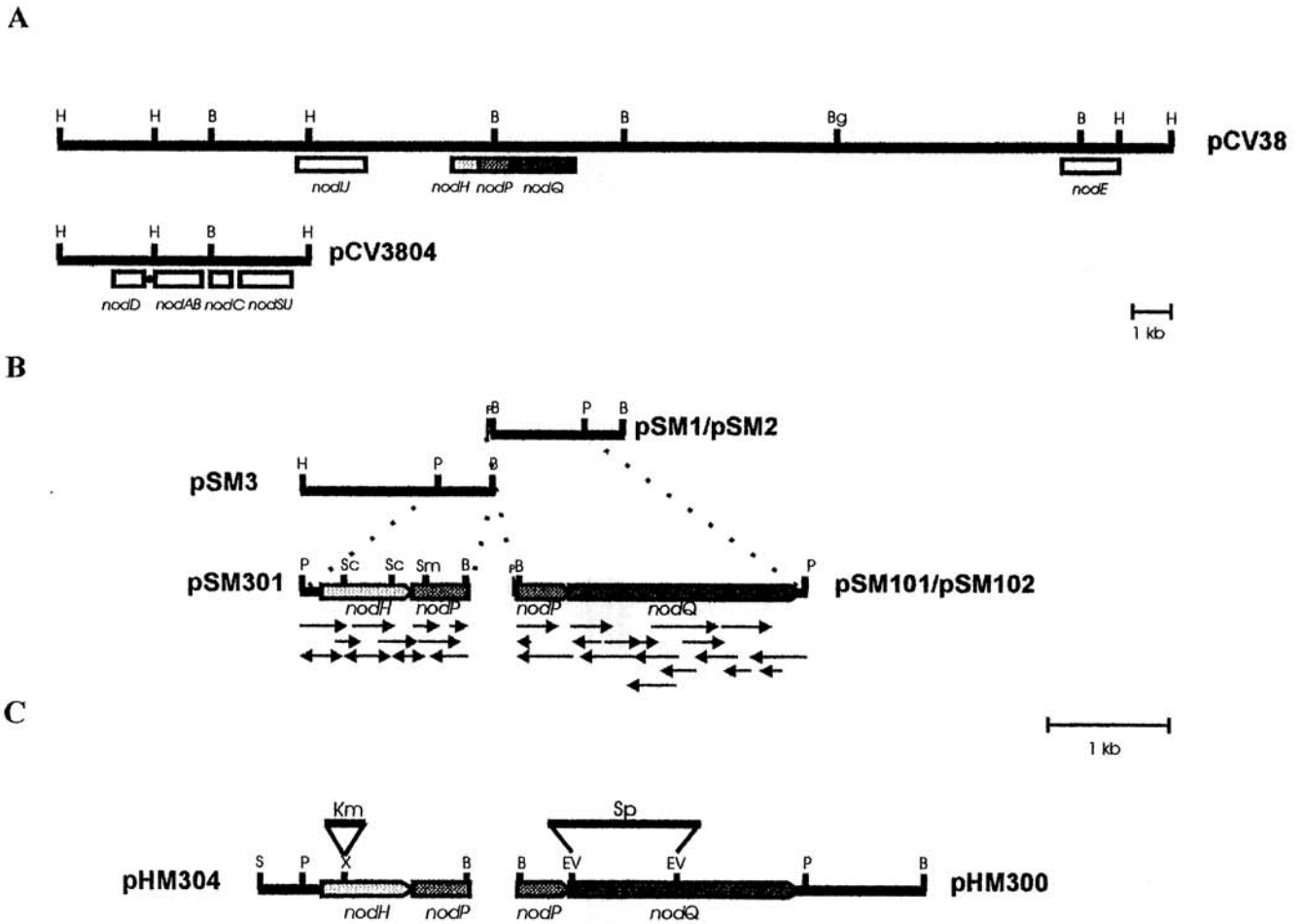

Fig. 2. **(A)** Location of *Rhizobium tropici* nodulation genes in cosmids pCV38 and pCV3804. Size and position of the genes are approximate except for *nodHPQ*. **(B)** Sequencing strategy of the *nodHPQ* region. Segments with single arrows indicate extent and direction of the reading in the different subclones obtained by exonuclease III digestion. Double-arrowed segments correspond to subclones constructed to complete the sequence. **(C)** Location of mutations generated in *nodH* and *nodQ* genes (plasmids pHM304 and pHM300 respectively). Km, kanamycin^R interposon; Sp, spectinomycin^R interposon; B, *Bam*HI; Bg, *Bgl*II; EV, *Eco*RV; H, *Hin*dIII; P, *Pst*I; S, *Sal*I; Sc, *Sac*I; Sm, *Sma*I; X, *Xho*I.

of the Nod factors produced by *R. tropici* CIAT899. We also present the identification, cloning, and sequencing of the *nodHPQ* region involved in the sulfation of some of the LCOs produced by this strain and the nodulation phenotypes of *nodH* and *nodQ* mutants on *Phaseolus* and *Leucaena* plants.

RESULTS

Purification and preliminary characterization of the Nod factors produced by *R. tropici*.

Using thin-layer chromatography (TLC) analysis of wild-type strain CIAT899, we observed two spots when induced by naringenin, whereas no spots are detected in the absence of the inducer (Fig. 1A lanes 1 and 2). These compounds run in TLC in a manner similar to other described Nod factors (data not shown, Spaink et al. 1992). These observations suggest that wild-type strain CIAT899 produces at least two different kinds of Nod factors. Of course these two spots seen on TLC could be composed of several different compounds, and indeed high-pressure liquid chromatography (HPLC) analysis showed that this is the case (see below). When apigenin, chrysin, or 7-hydroxy-flavone are used as inducers, the same spots can be observed (data not shown). We previously reported that genistein is not an inducer of *nod* gene expression in *R. tropici* (Sousa et al. 1993) and, as expected, no Nod factor production was observed when this flavonoid was used as an inducer of the wild-type strain CIAT899 (data not shown). One of the compounds observed shows very hydrophilic behavior (R_f = 0.8) when run on reversed phase TLC, suggesting the presence of a charged group on this molecule. Because of the similar behavior on TLC of this spot to that of *R. meliloti* NodRmIV(S) (data not shown) we decided to explore the possibility that this molecule was sulfated. When the same experiment is performed with radioactive sulfate, a single fast-moving band with the same R_f can be detected, indicating that this molecule is indeed sulfated, while the other is not (Fig. 1B). This correlates with the presence of *nodP*, *nodQ*, and *nodH* genes in this strain (see below), suggesting that they are functional and carry out the same reactions as in *R. meliloti*. Further evidence that the *nodHPQ* genes are functional in *R. tropici* arises from an experiment performed with cosmids pCV38 and pCV3804 (Fig. 2A, see below). These cosmids carry a region of the symbiotic plasmid of *R. tropici* in which several nodulation genes have been placed (Vargas et al. 1990). TLC analysis of the Nod factors produced by cosmid pCV38 in strain RSP900 (pSym cured derivative of wild-type strain CIAT899) shows a pattern identical to that of the Nod factors produced by wild-type strain CIAT899 (Fig. 1A, lanes 3 and 4), but cosmid pCV3804, which lacks the *nodHPQ* region (Fig. 2A), only produces the nonsulfated compounds (Fig. 1A, lanes 5 and 6), indicating that these genes are necessary for the sulfation of the Nod factors.

TLC analysis was also performed using L-(methyl-^{14}C) methionine as a radioactive label (Fig. 1C). The result of this experiment shows that both the fast-moving band and the slow-moving one become labeled, possibly indicating the presence of a methyl group in the Nod factors produced by wild-type strain CIAT899. This correlates with the presence of the *nodS* gene detected in this strain (M. Villalobos, unpublished results). When we analyzed the Nod factors produced by strain RSP900 containing cosmids pCV38 and pCV3804 in the presence of L-(methyl-^{14}C) methionine, we found that both cosmids produced labeled compounds (Fig. 1C, lanes 3, 4, 5, and 6), suggesting that the *nodS* gene is present and functional in cosmid pCV3804.

The Nod factors produced by *R. tropici* CIAT899 were then further purified. To achieve this, strain CIAT899/pMP604 was constructed in order to overproduce the Nod factors. Plasmid pMP604 contains a FITA (flavonoid independent transcription activation) *nodD* gene (Spaink et al. 1989), and it has been shown that the introduction of this plasmid into *R. leguminosarum* bv. *viciae* strains results in overproduction of Nod factors. The presence of this plasmid also avoids the need to add naringenin to the culture medium, which is important because this compound co-elutes with the sulfated Nod factors of *R. tropici*, making it difficult to detect and isolate them using HPLC (data not shown). TLC analysis of the Nod factors produced by strain CIAT899/pMP604 shows a pattern identical to that of those produced by wild-type strain CIAT899.

Strain CIAT899/pMP604 was cultivated in minimal media and the Nod factors were extracted by the addition of n-butanol; this extract was dried and resuspended in 40% acetonitrile/water (ACN/H_2O). A pre-purification step was performed by passing the ACN/H_2O extract over a Baker silica C18 column from which the Nod factors were eluted with 2 ml of 40% ACN/H_2O, 60% ACN/H_2O, and 80% ACN/H_2O. Each of these fractions was then analyzed using reversed phase HPLC and only in the 40% ACN/H_2O fraction were LCOs detected. Figure 3A shows a chromatogram of an extract of an uninduced culture of strain CIAT899/pMP281 (this

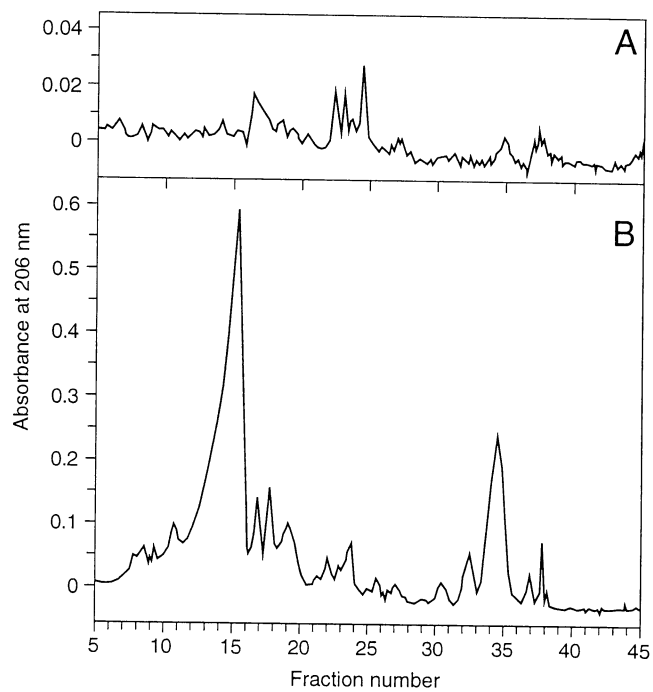

Fig. 3. High-pressure liquid chromatography chromatograms of Nod factors produced by strains **(A)** CIAT899/pMP281 and **(B)** CIAT899/pMP604. Amount of sample loaded was equivalent to that isolated from **(A)** 100 ml and **(B)** 200 ml of culture.

plasmid carries a wild-type *nodD* gene from *R. leguminosarum* bv. *viciae*) that was used as a control in order to be able to identify specific peaks in the chromatogram obtained from the CIAT899/pMP604 extract. Figure 3B shows the chromatogram of the extract obtained from strain CIAT899/pMP604 in which several specific peaks can be identified that elute in two different areas of the chromatogram. Fractions 9 to 22 contain more hydrophillic compounds than do fractions 31 to 37. These peaks were collected for mass spectrometric analysis.

Structure of *R. tropici* Nod factors.

Fractions 9 to 37 from the HPLC run were submitted to positive mode collision induced dissociation tandem mass spectrometry (CID MS-MS). The results are summarized in Table 1, which contains the *m/z* values observed for the $[M - H]^-$ or $[M + H]^+$ pseudomolecular ions and the OXOH 14 M-type fragment ions observed on CID MS-MS together with the structures assigned to them. The spectra of fractions 9 to 22 (analyzed in the negative ion mode) contain $[M - H]^-$ pseudomolecular ions indicating the presence of LCOs mainly consisting of five GlcNAc residues, substituted with a sulfate group and some also with a methyl group. These LCOs are substituted with a range of common fatty acyl moieties. The LCO structures correspond broadly to those reported in *R. tropici* strain CFN299 by Poupot et al. (1993).

Table 1. Chemical composition of the lipo-chitin-oligosaccharides found in *Rhizobium tropici* CIAT899pMP604

Fraction	$[M - H]^-$ or $[M + H]^+$	TG-adduct	Oxonium type fragment ions (corresponding $[M + H]^+$ parent ion)	Designation
9 + 10	1306 m	1414	. . .	V(C16:1,S)
	1320 m	1428	. . .	V(C16:1,NMe,S)
	1322 *	. . .	1023, 820, 617, 414 (1324)	V(C16:0,NMe,S)
	1334	1442	1035, 832, 629, 426 (1336)	V(C18:1,S)
11 + 12	1306 m	1414	. . .	V(C16:1,S)
	1334	1442	. . .	V(C18:1,S)
	1348 *	1456	1049, 846, 643, 440 (1350)	V(C18:1,NMe,S)
13 – 16	1145	1253	846, 643, 440 (1147)	IV(C18:1,NMe,S)
	1306 m	1414	. . .	V(C16:1,S)
	1348 *	1456	. . .	V(C18:1,NMe,S)
17	1308 m	. . .	. . .	V(C16:0,S)
	1334 m	1442	. . .	V(C18:1,S)
	1348	1456	. . .	V(C18:1,NMe,S)
	1350 *	. . .	1051, 848, 645, 442 (1352)	V(C18:0,NMe,S)
18	1308 m	. . .	. . .	V(C16:0,S)
	1334 m	1442	. . .	V(C18:1,S)
	1348	1456	. . .	V(C18:1,NMe,S)
	1350 *	. . .	. . .	V(C18:0,NMe,S)
19 + 20	1147	. . .	848, 645, 442 (1149)	IV(C:18:0,NMe, S)
	1350	. . .	. . .	V(C18:0,NMe,S)
	1376 *	1484	1077, 874, 671, 468 (1378)	V(C20:1,NMe,S)
22	1378 *	. . .	1079, 876, 673, 470 (1380)	V(20:0,NMe,S)
31	1215 m	1323	1035, 832, 629, 426	IV-Man(C18:1)
	1256 *	1364	1035, 832, 629, 426	V(C18:1)
32	1270 *	1378	. . .	V(C18:1,NMe)
33	1270 *	1378	. . .	V(C18:1,NMe)
34 – 36	1053 m	1161	832, 629, 426	IV(C18:1)
	1067	1175	846, 643, 440	IV(C18:1,NMe)
	1270 *	1378	1049, 846, 643, 440	V(C18:1,NMe)
37	1067 m	1175	. . .	IV(C18:1,NMe)
	1270 *	1378	. . .	V(C18:1,NMe)

[a] Data given for fractions 31 – 37 were obtained in the positive ion mode. MS-MS data were recorded in the positive ion mode for all fractions. Asterisk indicates the major component in each fraction; "m" indicates minor components.

Fractions 31 to 37 contain neutral LCO species (see Table 1), the most interesting of which is found in fraction 31; in the positive ionization mode an $[M + H]^+$ pseudomolecular ion is observed at *m/z* 1,215, indicating the presence of an LCO consisting of four GlcNAc residues substituted with a C18:1 fatty acyl chain and a hexose residue. The CID spectrum (Fig. 4) of this pseudomolecular ion contains A^+-type fragment ions formed by sequential cleavage of each glycosidic linkage, with charge retention on the nonreducing portion of the molecule (oxonium ions). The *m/z* values of these ions (426, 629, 832, 1,035, and 1,215) are consistent with a entirely novel LCO consisting of a linear backbone of four GlcNAc residues and a hexose as the reducing-terminal residue and bearing a C18:1 fatty acyl moiety on the nonreducing terminal residue (see Fig. 4). To identify the hexose present in the LCO found in HPLC fraction 31, monosaccharide composition analysis was carried out. Monosaccharide methyl glycosides were released on methanolysis and converted to their trimethylsilyl (TMS) ethers prior to gas chromatography–mass spectrometry (GC-MS) analysis. Comparison of the retention times and the mass spectra of the TMS methyl glycosides obtained from fraction 31 with those for authentic hexose standards allowed identification of mannose as the hexose residue.

Localization of *nodPQ* and *nodH* homologues in *R. tropici* CIAT899.

The isolation and preliminary characterization of the nodulation region of *R. tropici* CIAT899 is described elsewhere (Vargas et al. 1990). Using heterologous probes, we have localized by hybridization several nodulation genes in cosmid pCV38, which carries a region of the symbiotic plasmid of *R. tropici* CIAT899 (Fig. 2A). This strain contains at least four copies of the regulatory gene *nodD*, one of them, which proved to be functional, being located in cosmid pCV38 (Sousa et al. 1993). However, no *syrM* homologue was detected when hybridizing against an *R. meliloti syrM* probe (data not shown). We also looked for the presence of *nodL*, but no hybridization was detected in cosmid pCV38 nor in total DNA from the wild-type strain CIAT899 (data not shown). This result is in agreement with the lack of an acetyl group in *R. tropici* Nod factors (see Table 1). Several nodulation genes described for other species such as *nodABC*, *nodE*, *nodIJ* (Vargas et al. 1990), *nodS*, and *nodU* (M. Villalobos, unpublished results) are also present in this strain (Fig. 2A). Homologues of *nodP* and *nodQ* have also been found when hybridizing against internal fragments of *R. meliloti nodP* and *nodQ* genes (Fig. 5A and B, respectively). It can be seen that a common 3.4-kb *Bam*HI fragment hybridizes against *R. meliloti nodP* and *nodQ* probes both in cosmid pCV38 (Fig. 5A and B, lane 1) and in total DNA from the wild-type strain (Fig. 5A and B, lane 2). In the case of *nodP*, there is an additional 7.5-kb *Bam*HI fragment (Fig. 5A, lanes 1 and 2) that

corresponds to the left-hand contiguous region of the 3.4-kb fragment, indicating that the *nodP* gene has an internal *Bam*HI restriction site. In both cases (*nodP* and *nodQ*) a common extra band of 2.9 kb can be detected in total DNA from strains CIAT899 (Fig. 5A and B, lane 2) and RSP900 (Fig. 5A and B, lane 3). This may account for the presence of at least another copy of *nodP* and *nodQ* genes not located in the symbiotic plasmid, as was previously reported for *R. meliloti* (Schwedock and Long 1992). Finally, Figure 5C shows a hybridization assay against an *R. tropici* CIAT899 *nodH* probe once this gene had been localized in cosmid pCV38 by sequencing techniques. This experiment confirms the existence of *nodH* in single copy in this strain, since only the 7.5-kb *Bam*HI fragment of cosmid pCV38 (lane 1) and of total DNA from the wild-type strain (lane 2) hybridize against the homologous probe, while no bands were detected for strain RSP900 (lane 3).

Sequence analysis of *nodH*, *nodP*, and *nodQ* genes of *R. tropici* CIAT899.

Based on the results obtained in the hybridization assays, we proceeded to subclone two fragments of cosmid pCV38 in a pUC-derivative vector. These fragments were the hybridizing 3.4-kb *Bam*HI fragment common to *nodP* and *nodQ* genes, and the contiguous 4.8-kb *Hin*dIII-*Bam*HI fragment on the left-hand side of this, yielding plasmids pSM1/pSM2 (depending on the orientation) and pSM3, respectively (Fig. 2B). Once the physical restriction map of both fragments was

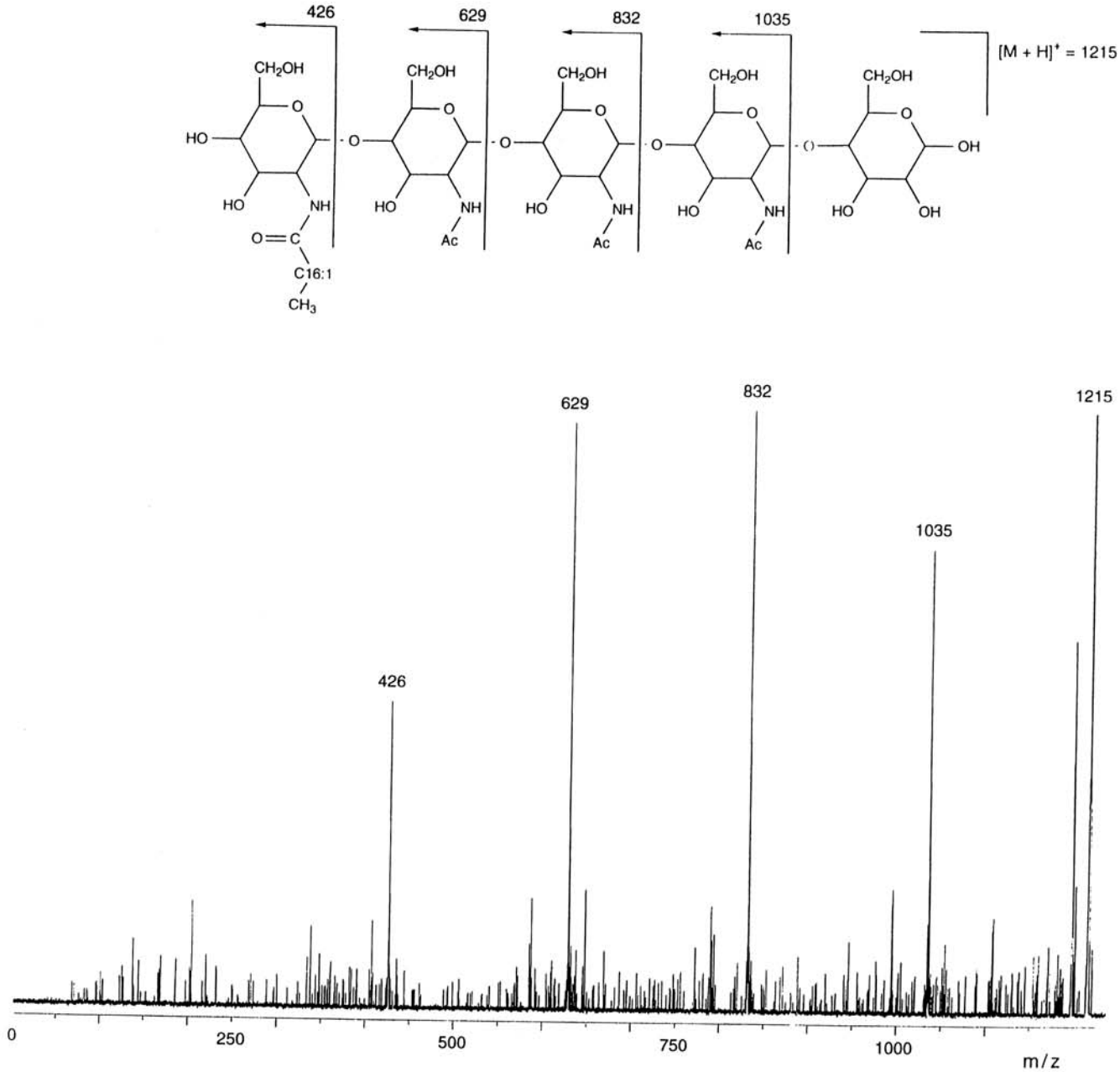

Fig. 4. Collision induced dissociation mass spectrum of high-pressure liquid chromatography fraction 7 (Fig. 3B) from *Rhizobium tropici* strain CIAT899; *m/z* values are quoted as nominal masses.

obtained, three new subclones were constructed in plasmid pBluescript SK II+ (as described in Materials and Methods) in order to determine the nucleotide sequence of part of the region that was homologous to *R. meliloti nodPQ* probes. Figure 2B shows the sequencing strategy followed, the extent and direction of the reading indicated by arrows.

The resulting 3,809 nucleotide (nt) sequence has been submitted to the EMBL Nucleotide Sequence Data Library and it will appear under the accession number X87608. A number of open reading frames (ORFs) defined by an ATG and a stop codon were found. Among those ORFs larger than 300 bp, three ORFs, designated 1 to 3, were found to be homologous to the *R. meliloti nodH*, *nodP*, and *nodQ* genes, respectively. The three of them are preceded by putative ribosome binding sites (rbs) (Shine and Dalgarno 1974) and their codon usage suggests that they are translated.

ORF 1 (*nodH*).

The first ORF lies between the ATG at position 181 and the TAA stop codon at position 930, encoding a polypeptide of 250 amino acids (aa) with a predicted M_r of 28,492, almost identical to the M_r of *R. meliloti* NodH protein (28,548) (Horvath et al. 1986). This ORF is preceded by the putative rbs GGAAA at positions –13/–9, which coincides both in nucleotide sequence and spacing with the corresponding rbs found upstream of its *R. meliloti* counterpart (Debellé and Sharma 1986).

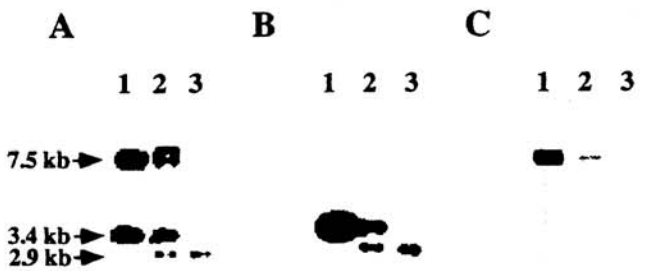

Fig. 5. Southern blot hybridizations against **(A)** *Rhizobium meliloti nodP* and **(B)** *nodQ* and **(C)** *R. tropici nodH* probes of *Bam*HI restriction digested DNA of cosmid pCV38 (lane 1 in A, B, and C), total *Bam*HI-restricted DNA of strain CIAT899 (lane 2 in A, B, and C) and total *Bam*HI-restricted DNA of strain RSP900 (lane 3 in A, B, and C).

Table 2. Sequence comparison of *Rhizobium tropici* CIAT899 *nodH*, *nodP*, and *nodQ* genes

Open reading frame (ORF)	Homologue sequence	Homology		
			Protein	
		Nucleotide (%)	Similarity (%)	Identity (%)
ORF 1 (*nodH*)	*R. meliloti nodH*[a]	70	77	68
ORF 2 (*nodP*)	*R. meliloti nodP*[b]	75	92	68
	Azospirillum brasilense nodP[c]	66	81	66
	Escherichia coli K-12 *cysD*[d]	64	81	62
ORF 3 (*nodQ*)	*R. meliloti nodQ*[b]	74	87	77
	A. brasilense nodQ[c]	60	74	57
	E. coli K-12 *cysN*[d]	53	69	47
	E. coli K-12 *cysC*[d]	55	68	56

[a] Debellé and Sharma (1986)
[b] Schwedock and Long (1989) and Cervantes et al. (1989).
[c] Vieille and Elmerich (1990)
[d] Leyh et al. (1992)

The sequence homology detected between ORF 1 and the *R. meliloti nodH* gene is very strong (Table 2; Fig. 6), thus suggesting that ORF 1 encodes the *R. tropici nodH* gene.

The hydropathic character of the protein deduced from ORF 1 was studied according to the Kyte-Doolittle algorithm (Kyte and Doolittle 1982) and we identified a putative membrane-spanning segment next to the carboxy-terminus, since the hydropathy value of this segment longer than 19 aa averages greater than +1.6.

ORF 2 (*nodP*).

ORF 2 starts 5 nt downstream of ORF 1 at position 936 and extends through 900 bp to the TGA stop codon at position 1,835. This ORF encodes a protein of 300 aa, with a calculated M_r of 34,756, very similar to the predicted M_r of the *nodP* gene products of *R. meliloti* (34,763) (Cervantes et al. 1989; Schwedock and Long 1989) and to that of the free-living diazotroph *Azospirillum brasilense* (34,098) (Vieille and Elmerich 1990). The initiation ATG codon is preceded by the putative rbs AAGGA at positions –14/–10 and is followed by the CAAT motif complementary to the 5′ end of the bacterial 16S rRNA present 12 nt downstream of this ATG position (Petersen et al. 1988).

Computer analysis shows that the polypeptide deduced from ORF 2 is highly homologous to the NodP proteins of *R. meliloti* and *A. brasilense* as well as to the *Escherichia coli* CysD ATP-sulfurilase subunit (Leyh et al. 1992) (Table 2; Fig. 6).

No membrane-spanning sequences could be detected when applying the Kyte-Doolittle algorithm. Indeed, the putative NodP protein exhibits a relatively polar character.

ORF 3 (*nodQ*).

From the ATG start codon at position 1,835 to the TAG stop codon at position 3,733, ORF 3 encodes a protein of 633 residues, with a deduced M_r of 69,712, intermediate between the calculated M_r of the *nodQ* gene products of *R. meliloti* (70,613) and *A. brasilense* (66,965). The putative rbs GGAAGG is found 9 nt upstream of the first ATG. Interestingly, in all three species the ATG initiation codon of *nodQ* overlaps the TGA stop codon of *nodP* by 1 nt. The presence of the motif TGATG suggests that in *R. tropici* the translation of *nodP* and *nodQ* genes is coupled as it is in *R. meliloti* (Cervantes et al. 1989).

Sequence comparison of the protein deduced from ORF 3 with the NodQ proteins of *R. meliloti* and *A. brasilense* revealed high homology in both cases (Table 2; Fig. 6). The amino-terminal region of *R. tropici* NodQ shows important homology with the second *E. coli* ATP-sulfurilase subunit, CysN, whereas its carboxy-terminal region is clearly homologous to the *E. coli* CysC protein, which has APS-kinase activity (Fig. 6), as has been described for *R. meliloti* NodQ protein (Schwedock et al. 1994). In contrast to *E. coli*, in *Rhizobium* both activities involved in the biosynthesis of activated forms of sulfate are probably encoded by a single polypeptide, NodQ.

Finally, it is interesting to point out the significant similarities (ranging from 52.5 to 58.4%) identified on computer

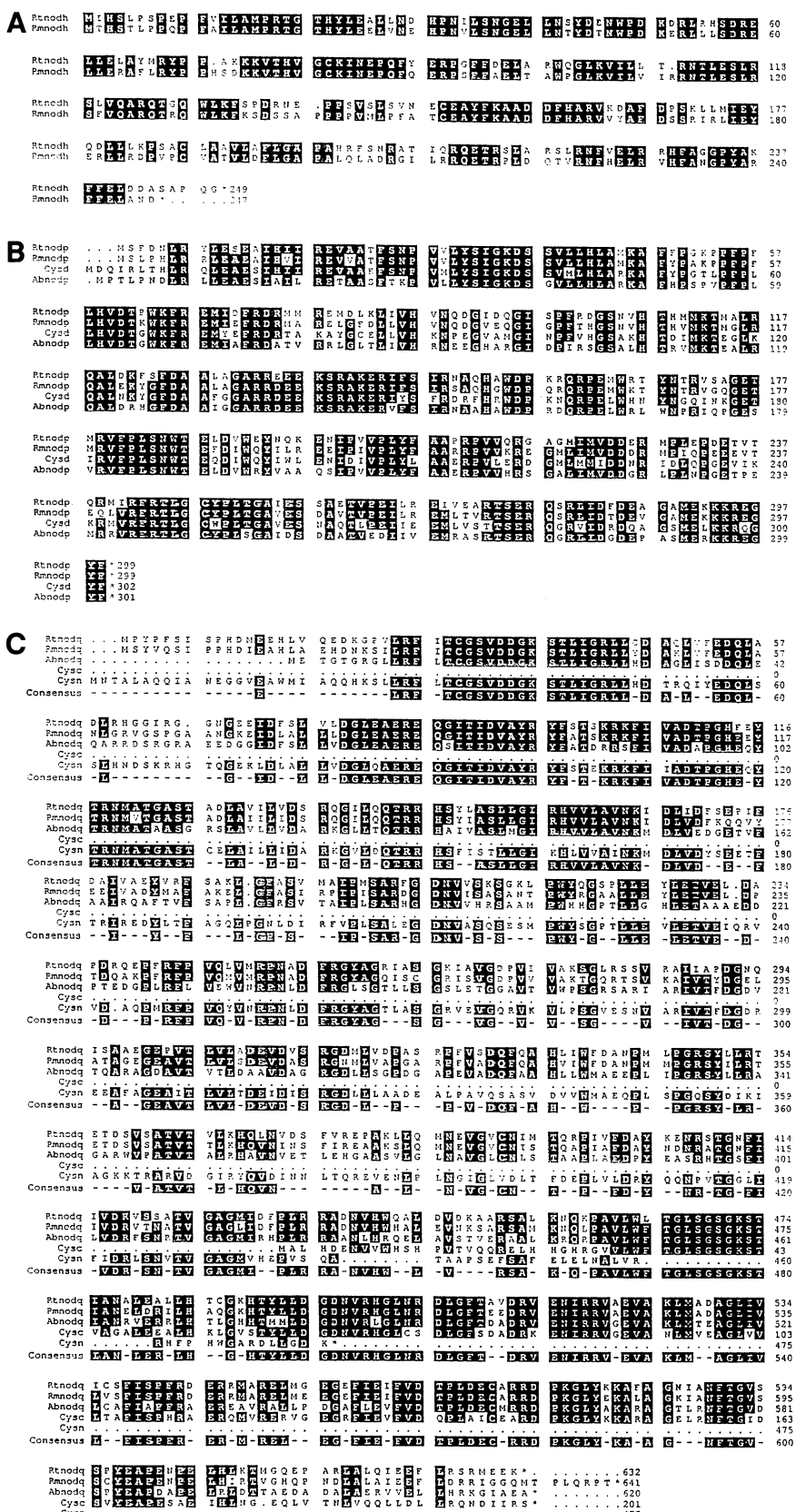

Fig. 6. Comparison of **(A)** *nodH*-encoded, **(B)** *nodP*-encoded, and **(C)** *nodQ*-encoded amino acid sequences of *Rhizobium tropici* (Rtnodh, Rtnodp, and RtnodQ), *R. meliloti* (Rmnodh, Rmnodp, and Rmnodq), *Azospirillum brasilense* (Abnodp and AbnodQ), *Escherichia coli* CysD ATP sulphurilase subunit (Cysd) and *E. coli* CysN ATP sulphurilase subunit (Cysn) and APS kinase (Cysc). Identical residues are enclosed in black boxes. Conservative substitutions are indicated by shaded boxes.

analysis between ORF 3 and some translation elongation factors from both procaryotic (*Halobacterium halobium, Thermococcus celer, Thermus termophilus*) and eucaryotic organisms (*Trypanosoma cruzi, Porphyra purpurea, Homo sapiens, Drosophila melanogaster, Glycine max*) (data not shown). The explanation for this homology, previously reported for *R. meliloti* NodQ (Cervantes et al. 1989) is found in a common GTP-binding domain in all these proteins, whose function in *Rhizobium* NodQ is not yet clear (Schwedock et al. 1994).

As in the case of NodP protein, NodQ seems not to exhibit any transmembrane domain.

Nodulation phenotypes of wild-type strain and mutants in *nodH* and *nodQ*.

Table 3 summarizes the nodulation phenotype and Nod factor synthesis detected on TLC analysis of strains CIAT899, RSP900/pCV38, and RSP900/pCV3804, in addition to the mutant strains in *nodH* and *nodQ* genes (see Materials and Methods). These five strains are able to nodulate two nonrelated host plants of *R. tropici*, namely *Phaseolus* and *Leucaena*, although strains RSP900/pCV38 and RSP900/pCV3804 are unable to fix nitrogen, since the *nif-fix* gene cluster is not present in these cosmids.

The nodulation phenotype on *Phaseolus vulgaris* plants is similar among all the strains studied, although a slight delay in nodulation is observed for strain RSP900/pCV3804. However, the final number of nodules and their size is similar. It is worth noting that wild-type strain CIAT899 and *nodH*$^-$ and *nodQ*$^-$ mutants induce nodule formation only in the upper part of the root system, while strains RSP900/pCV38 and RSP900/pCV3804 produce nodules all along the root, maybe in response to nitrogen starvation by the plant.

In *Leucaena leucocephala* plants, besides the reduced nodule number, a lengthy delay in nodulation was observed when inoculated with strain RSP900/pCV3804 (10% decrease) and the *nodH*$^-$ mutant strain (59% decrease).

TLC analysis of the Nod factors produced by all five strains in the presence of naringenin and ^{14}C-glucosamine or ^{35}S-sulfate shows that the fast-moving hydrophilic band corresponding to the sulfated LCO family only disappears in those strains that lack (RSP900/pCV3804) or carry a mutation in the *nodH* gene.

These results indicate that the sulfated Nod factors synthesized by *R. tropici* CIAT899 are important for nodulation of *Leucaena* plants.

Table 3. Nodulation phenotype and Nod factor production of different *Rhizobium tropici* strains

	Nodulation phenotype[a]		Nod factor synthesis[b]	
Strain	***Phaseolus***	***Leucaena***	**^{14}C-glucosamine**	**^{35}S-sulphate**
CIAT899	100	100	2	1
RSP900/pCV38	100[c]	88	2	1
RSP900/pCV3804	100[c]	10	1	0
CIAT899 *nodH*$^-$	100	59	1	0
CIAT899 *nodQ*$^-$	100	100	2	1

[a] Nodule formation 20 days (*Phaseolus vulgaris*) or 35 days (*Leucaena leucocephala*) after inoculation with the different strains.
[b] Nod factor production in the presence of naringenin and ^{14}C-glucosamine or ^{35}S-sulphate, as detected by thin-layer chromatography analysis. The number indicates the number of bands detected.
[c] Nodules are distributed all along the root.

DISCUSSION

One of the main processes involved in the specificity of the *Rhizobium*-legume interaction is the synthesis of low molecular weight LCOs (Lerouge et al. 1990; Spaink et al. 1991). *Rhizobium tropici* is a broad host range species that can nodulate several nonrelated legumes, so it offers a convenient system to determine some of the mechanisms of the specificity of the *Rhizobium*-legume interaction. In this work we have purified and characterized the nodulation factors of *R. tropici* strain CIAT899 and related their structures with the *nodHPQ* genes found in this strain. We also have determined the role of the sulfate group present in *R. tropici* Nod factors and the nodulation ability of this strain on two different hosts, namely *Leucaena* and *Phaseolus*.

TLC analysis of the Nod factors produced by wild-type strain CIAT899 induced by naringenin reveals that *R. tropici* synthesizes two families of Nod factors, a hydrophilic family containing sulfated compounds, and a neutral family comprising nonsulfated LCOs (Fig. 1; Table 1).

One possible explanation for the broad host range exhibited by *R. tropici* is that this species could synthesize different Nod factors in response to different plant regulators, so we decided to explore the production of Nod factors using different flavonoids as inducers. The result of this experiment shows no difference in the pattern of Nod factor production when chrysin, 7-hydroxy-flavone, or apigenin is used as inducer of wild-type strain CIAT899 (data not shown). The possibility that root exudates from different hosts trigger the production of different sets of Nod factors still remains to be explored, although it is also possible that the great variety of Nod factors produced by this strain (see below) is sufficient to account for its broad host range.

Upon further characterization, we have shown that the presence of the *nodHPQ* genes is essential for the sulfation of the Nod factors produced by strain CIAT899. Cosmid pCV3804 carries a region of the pSym of wild-type strain CIAT899. This region contains the common nodulation genes *nodABC*, the host specific gene *nodS,* and the regulatory gene *nodD*, but lacks the *nodHPQ* region (Fig. 2A). Nod factor production by this cosmid in the pSym cured derivative RSP900 was explored and only the nonsulfated LCOs are synthesized (Fig. 1A), indicating the necessity of the *nodHPQ* genes for the production of sulfated Nod factors.

NodS has been proposed to carry out the methylation of Nod factors in other rhizobia (Geelen et al. 1993), so we decided to explore whether this gene is functional in *R. tropici*. For this purpose, Nod factor production was induced by naringenin in the presence of L-(methyl-^{14}C) methionine in strains CIAT899, RSP900/pCV38, and RSP900/pCV3804 (Fig. 1C). All three strains are able to produce methylated Nod factors, indicating that the *nodS* gene is functional in *R. tropici*.

Following this preliminary characterization, we purified and characterized the Nod factors produced by *R. tropici*.

Strain CIAT899 produces a wide variety of Nod factors, and we were able to detect as many as 16 different LCOs. The general structure of these compounds is similar to that reported by Poupot et al. (1993) for *R. tropici* strain CFN299, although these authors only describe two LCOs present in this strain. In contrast to their results, we have identified a wide variety of LCOs and describe a novel LCO produced by *R. tropici* strain CIAT899. This molecule has a pentasaccharide backbone consisting of four GlcNAc residues and a hexose as the reducing-terminal residue and bears a C18:1 fatty acyl moiety on the nonreducing terminal residue. In other broad host range species, such as *Rhizobium* sp. NGR234, the presence of additional monosaccharide residues (although not as part of the backbone) has been reported (Price et al. 1992) and it has been suggested that these residues could play a role in the ability to nodulate a wide variety of legumes. Experiments are under way to explore this possibility in strain CIAT899.

Another difference between the structures reported by Poupot et al. (1993) and our results is the presence of six different kinds of fatty acyl moieties in strain CIAT899 (C16:0, C16:1, C18:0, C18:1, C18:2, C20:0, and C20:1, Table 1). The structure of the fatty acyl moiety on the Nod factors has been shown to be a major host determinant for *R. leguminosarum* (Spaink et al. 1995), so it is possible that this variety of fatty acyl chains present in *R. tropici* Nod factors enables them to be recognized by a wide range of host plants, and we are exploring this possibility.

Extensive study has been focused on *R. meliloti nodHPQ* genes, since these three host-specific genes are the main determinants of host-specificity in this species, mediating the sulfation of Nod factors. It has been shown that *nodH* mutants lose their ability to nodulate alfalfa (*Medicago sativa*) but acquire this ability for the nonhomologous host common vetch (*Vicia sativa*) (Debellé and Sharma 1986). In contrast, mutants in *nodQ* extend the host range of *R. meliloti*, including both alfalfa and vetch plants (Cervantes et al. 1989). As for *R. tropici*, the introduction of *R. meliloti nodHPQ* genes into strains CFN299 and CIAT899 increases twofold the ability of both receptor strains to nodulate *Phaseolus vulgaris* plants, due to an increase in the production of sulfated Nod factors (Martínez et al. 1993). More recently, the necessity of the sulfate moiety in *R. tropici* Nod factors in overcoming an effective nodulation on *Phaseolus vulgaris* was demonstrated (Martínez et al. 1994). In this work, we show that this need can also be extensive in *Leucaena leucocephala* plants. Thus, *R. tropici* CIAT899 *nodHPQ* genes are relevant to the host range exhibited by this strain.

Until recently, the presence of the *nodH* gene was exclusive to *R. meliloti*, since no copies of this gene had been described for any other *Rhizobium* strain. However, a gene highly homologous to the *R. meliloti nodH* gene (70% identity) has been sequenced in *R. "hedysari,"* a bacterium isolated from the root nodules of *Hedysarum coronarium*, a drought- and alkali-tolerant legume (Tola et al. 1994).

It is striking that *nodH* exhibits such a different location in the three species mentioned above. In *R. meliloti* this gene is located next to *nodFEG* genes, although its transcription is in the opposite direction. In *R. tropici*, *nodH* is situated immediately upstream of *nodPQ*, it being very probable that the three together form a transcriptional/translational unit, whereas in *R. "hedysari"* the *nodH* copy is located downstream of the common nodulation genes *nodABC*. This different structural organization of nodulation genes in several *Rhizobium* species could be the result of genomic recombination mechanisms, for instance those mediated by insertion sequences (IS). For this reason, it is important to point out the finding of an 800-nt ORF approximately 400 bp downstream of the *R. meliloti nodPQ* operon, which exhibits 63% homology with IS66 of *Agrobacterium tumefaciens* (Schwedock and Long 1993). However, we could not find any IS downstream of the *R. tropici nodQ* gene (data not shown).

The transcriptional organization of *R. tropici nodPQ* genes is not yet known, but both genes are probably cotranscribed as one mRNA, as it has been previously reported for their *R. meliloti* counterparts (Cervantes et al. 1989), since both species share the common motif TGATG, in which the termination codon of *nodP* overlaps the initiation codon of *nodQ*. This system is characteristic of heterodymeric proteins in order to provide a stoichiometric proportion of each polypeptide chain in the dimer (Kozac 1983), which correlates with NodP and the N terminus of NodQ encoding both subunits of ATP-sulfurilase activity.

The possibility cannot be excluded that the *nodH* gene, located 5 nt upstream of *nodPQ*, belongs to the same transcriptional unit. Transcriptional/translational fusions (in study) will confirm if there is a translational coupling.

As described above, the *R. meliloti* NodPQ proteins are homologous to the *E. coli* CysDNC system involved in the production of the activated sulfate compounds APS and PAPS; the sulfate group is transferred from this last compound to the reducing terminal residue of the LCO by means of the sulfotransferase activity of NodH. The *R. tropici* CIAT899 NodPQ and NodH proteins probably carry out the same biochemical functions in the synthesis of the sulfate moiety present in the sulfated LCOs produced by this strain, since their protein sequences are very similar to their *R. meliloti* counterparts as well as to the *E. coli* CysDNC system (Table 2).

TLC experiments with wild-type strain CIAT899 and strains RSP900/pCV38, RSP900/pCV3804, CIAT899 *nodH*$^-$, and CIAT899 *nodQ*$^-$ (Table 3) show that, as expected, strains RSP900/pCV3804 (which lacks the complete *nodHPQ* region) and CIAT899 *nodH*$^-$ do not produce sulfated Nod factors. However, the mutant strain CIAT899 *nodQ*$^-$ still produces sulfated LCOs, as do single *nodQ*$^-$ mutants of *R. meliloti* (Roche et al. 1991), probably due to the presence of extra copies of *nodQ* (and also of *nodP*) in this strain, as can be deduced from heterologous Southern hybridizations (Fig. 5A and B). In *R. meliloti*, the existence of two copies of the *nodPQ* operon, $nodP_1Q_1$ and $nodP_2Q_2$, has been shown, each located on one of the symbiotic plasmids of this species, $pSym_a$ and $pSym_b$. In addition, there is evidence for a third chromosomal sulfate-activating locus, *saa*, which corresponds to the housekeeping methionine and cysteine biosynthesis system, unable to undertake sulfation of LCOs (Schwedock and Long 1992). The extra allele(s) of *nodQ* in CIAT899 could account for the sulfation of LCOs in the absence of the *nodQ* copy present in cosmid pCV38. In contrast, homologous hybridization shows that *nodH* is present in single copy in *R. tropici* CIAT899 (Fig. 5C), similarly to its *R. meliloti* counterpart, and thus mutants in *nodH* are unable to produce sulfated LCOs.

The presence or absence of a sulfate group in *R. tropici* LCOs also correlates with the nodulation phenotypes found for the aforementioned strains on *Phaseolus* and *Leucaena* plants (Table 3). All five strains studied are able to nodulate *Phaseolus vulgaris*, while strains RSP900/pCV3804 and CIAT899 *nodH*$^-$ (which do not produce sulfated LCOs), show very low nodulation efficiency on *Leucaena leucocephala*, indicating that the sulfated Nod factors probably play a crucial role in nodulation of this host plant by *R. tropici*. Studies of the Nod factors of *R. fredii* show that an N-methyl group and a fucose or methyl fucose are important for nodulation of *Leucaena*. The combined results indicate that for nodulation of *Leucaena* an N-methyl group and either a sulfate or a (methyl)fucose residue is needed.

MATERIALS AND METHODS

Bacterial strains and plasmids.

Bacterial strains and plasmids used in this study are listed in Table 4.

Media and bacterial growth conditions.

Rhizobium strains were grown at 28°C in TY (Beringer 1974) or in minimal B$^-$ medium as described by Spaink et al. (1992) for the production of Nod factors. *Escherichia coli* cultures were grown in Luria broth medium (Miller 1972) at 37°C. Antibiotics were added to the following final concentrations (in μg ml $^{-1}$): rifampicin (Rif), 50; tetracycline (Tc), 10; gentamycin (Gm), 35; spectinomycin (Sp), 80; ampicillin (Ap), 100; kanamycin (Km), 30; streptomycin (Sm), 100. Flavonoids were diluted in ethanol and used at a final concentration of 2 μM.

Genetic techniques.

In order to mobilize pVK102 derivatives from *E. coli* to *Rhizobium*, triparental matings were carried out as described by Kondorosi et al. (1982) using pRK2073 as a helper plasmid. Selection was on TY with Tc and Rif. The plasmids carrying mutations in *nodH* and *nodQ* were introduced into *Rhizobium* strains by biparental mating with *E. coli* S17-1. Marker exchange experiments were performed according to Ruvkun and Ausubel (1981).

Total genomic DNA and large- and mini-scale plasmid and cosmid DNA preparations were carried out as described by Maniatis et al. (1982). DNA manipulations, including restriction digests, ligations, transformations, electrophoresis, and bacterial transformation were performed according to Maniatis et al. (1982).

The *R. meliloti* ^{32}P-labeled *nodP* and *nodQ* probes used in the heterologous Southern experiments were the 1.15-kb *Sal*I and 1.2-kb *Sal*I-*Kpn*I fragments of plasmid pPP346, respectively. The homologous *R. tropici* digoxigenin-labeled (Boehringer-Mannheim, Germany) *nodH* probe was the 0.4-kb *Sac*I fragment of plasmid pSM3. Hybridization conditions were 50% formamide at 42°C.

DNA sequencing and computer methods.

Cloning of the 3.4-kb *Bam*HI and 4.8-kb *Hin*dIII-*Bam*HI fragments of cosmid pCV38 yielded plasmids pSM1/pSM2 (depending on the orientation) and pSM3, respectively (Fig. 2B). The 2.4-kb *Pst*I fragment of pSM1 (consisting of the *Pst*I restriction site of pUC19 polylinker) and the 1.4-kb *Pst*I-*Bam*HI fragment of pSM3 were subcloned in pBluescript SK II+ (pSKII+) vector (Fig. 4B). The resulting plasmids pSM101/pSM102 (opposite orientation) and pSM301 were

Table 4. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Reference or source
Rhizobium tropici		
CIAT899	wild-type, Rif^R, Cm^R, Ap^R	Martínez et al. (1985)
RSP900	pSym cured-derivative of CIAT899	Vargas et al. (1990)
Escherichia coli		
HB101	restriction-minus; *recA* background	Boyer and Roulland-Dussoix (1969)
DH5α	α–complementation	Bethesda Research Laboratory
S17-1	*tra* genes for plasmid mobilization	Simon et al. (1983)
Plasmids		
pCV38	28.8-kb *Hin*dIII fragment of the pSym of CIAT899 cloned in cosmid pVK102	Vargas et al. (1990)
pCV3804	6.4-kb *Hin*dIII fragment (partial digestion) of pCV38 subcloned in pRK404	Vargas et al. (1990)
pRK2073	helper plasmid, Sp^R	Ditta et al. (1986)
pPP346	plasmid carrying the *R. meliloti* 41 *nod* region	Putnoky and Kondorosi (1986)
pUC19	Mp13 cloning vector; Ap^R, Lac$^+$	Yanisch-Perron et al. (1985)
pBluescript II SK+	pUC-derivative, F1 replication origin, Ap^R	Stratagene Cloning Systems (La Jolla, CA)
pSP70	cloning vector, Ap^R	Promega (Madison, WI)
pSM1	3.4-kb *Bam*HI fragment of pCV38 in pUC19	This work
pSM2	same as pSM1, opposite orientation	This work
pSM3	4.8-kb *Hin*dIII-*Bam*HI fragment of pCV38 in pUC19	This work
pSM101	2.4-kb *Pst*I fragment of pSM1 in pSKII +	This work
pSM102	same as pSM101, opposite orientation	This work
pSM301	1.4-kb *Pst*I-*Bam*HI fragment of pSM3 in pSKII +	This work
Kan^R Genblock^R (*Eco*RI)	Km^R interposon	Pharmacia (Uppsala, Sweden)
pHP45Ω	Sm^R/Sp^R interposon in pHP45 vector, Ap^R	Prentki and Krisch (1984)
pJQ200SK	*sacB*, Lac$^+$, Gm^R	Quandt and Hynes (1993)
pHM146	Sp^R interposon cloned in *Eco*RV in pSM1	This work
pHM300	4.6-kb *Bam*HI fragment of pHM146 in pJQ200SK	This work
pHM303	1.7-kb *Sal*I-*Bam*HI fragment of pSM3 cloned in pJQ200SK	This work
pHM304	Km^R interposon insertion in *Xho*I of pHM303	This work

used to generate a series of nested deletions according to the procedure of Henikoff (1987) using the Double Stranded Nested Deletion Kit from Pharmacia (Uppsala, Sweden). Clones with overlapping deletions covering both DNA strands were sequenced by the dideoxy chain termination method (Sanger et al. 1977) in an A.L.F. automatic sequencer from Pharmacia LKB with the universal M13(-20) and reverse primers.

To complete the DNA sequence, the 400-pb *Pst*I-*Sac*I and *Sac*I-*Sac*I and the 300-pb *Sac*I-*Sma*I fragments of plasmid pSM301 were cloned in pSKII+ and sequenced with the reverse and universal primers. The global sequencing strategy is shown in Figure 2B.

Sequence analysis and comparisons with sequences contained in the data bases were accomplished with the programs Brujene (J. Vara, Centro Nacional de Biotecnología, Madrid, Spain) and DNA Strider (C. Marck, Dept. Biologie, Centre d'Études Nucleaires de Saclay, 9119 Gif-sur-Yvette, France) and the GCG Sequence Analysis software package version 8.0 (Devereux et al. 1984).

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under the accession number X87608.

Construction of mutants in *nodQ* and *nodH* genes.

To construct the *nodQ* mutant, the 0.8-kb *Eco*RV fragment of plasmid pSM1 carrying the *nodQ* gene was substituted for the spectinomycin^R interposon (purified from plasmid pHP45Ω and then cloned in pBluescript and pSP70 plasmids in order to generate *Eco*RV ends) raising plasmid pHM146. The resulting 4.6-kb *Bam*HI insert was cloned in plasmid pJQ200SK, which carries the *sacB* gene responsible for plasmid suicide in gram negative bacteria when grown in 10% saccharose. After introduction of plasmid pHM300 into *R. tropici* CIAT899, Km^{R}Gm^S colonies were selected among the recombinant clones isolated in TYKmSac. Hybridization assays against both the Sp^R interposon and *nodQ* probes confirmed the insertion site (data not shown).

Construction of the *nodH* mutant was carried out by insertion of the kanamycin^R interposon in the single *Xho*I restriction site of plasmid pHM303 (1.7-kb *Sal*I-*Bam*HI fragment of pSM3 carrying *nodH*, cloned in pJQ200SK). After introduction of the resulting plasmid pHM304 into *R. tropici* CIAT899, Km^{R}Gm^S colonies were selected from the Km^{R}sac recombinant clones. Hybridization against both the Km^R interposon and an internal fragment of *nodH* confirmed the insertion site (data not shown).

TLC analysis of Nod factors.

TLC analysis was performed according to Spaink et al. (1992). Briefly, *Rhizobium* strains were grown on minimal B$^-$ medium supplemented with the appropriate antibiotics and flavonoids, where indicated. Cultures were grown to saturation and the supernatant was extracted with 0.5 ml of water-saturated n-butanol. The n-butanol was evaporated to dryness and the resulting powder resuspended in 40 µl of n-butanol. Five microliters of this solution was applied to the TLC plate (ODS:100% octadecyl silanization, from Sigma Co., St. Louis, MO), where the Nod factors were separated using 50% ACN/H_2O (vol/vol) as the mobile phase.

For radiolabeling of LCOs the following compounds and quantities were used: 25 µCi of ^{14}C-acetate (specific activity 45 to 60 mCi/mmol, from New England Nuclear, Boston, MA) or 100 µCi of $MgSO_4$-^{35}S (specific activity 602 mCi/mmol, from Amersham, Int.) or 1 µCi of L-(methyl-^{14}C) methionine (specific activity 55 mCi/mmol, from Amersham, Int.) or 0.5 µCi of glucosamine hydrochloride ^{14}C (specific activity 52 mCi/mmol, from Amersham, Int.).

TLC plates were exposed to Kodak X-Omat R film for 15 days and the film was developed with Kodak reagents according to the manufacturer's instructions.

HPLC analysis of the Nod factors.

Nod factors from strain CIAT899 derivatives were obtained from a 2-liter culture. A pre-purification step was performed by passing the crude extract through a Baker silica C18 column from which Nod factors were eluted using different ACN/H_2O ratios (40, 60, and 80%). These samples were analyzed by HPLC with a reversed phase superpac pep-s column (5 µm, 4 × 250 mm, Pharmacia) at a rate of 0.7 ml min^{-1}. The 40% fraction contained the Nod factors and was reinjected onto the HPLC for separation of the different peaks using isocratic elutions in 20 (5 min), 30 (30 min), 40 (30 min), and 60% (15 min) ACN/H_2O and then a linear gradient in 10 min from 60 to 100% acetonitrile. Elution of Nod factors was followed by absorbance at 206 nm. Fractions were collected each 2 min.

Mass spectrometry.

FAB-mass spectra were obtained in the negative or positive mode using MS1 of a JEOL JMS-SX/SX102A tandem mass spectrometer using −10 or 10 kV accelerating voltage. The FAB gun was operated at 6 kV with an emission current of 10 mA using xenon as the bombarding gas. Spectra were scanned at a speed of 30 s for the full mass range specified by the accelerating voltage used and recorded and processed on a Hewlett Packard HP9000 series data system using the JEOL Complement software. Tandem mass spectra (MS/MS) were obtained on the same instrument using helium as the collision gas (in the third field free region) at a pressure sufficient to reduce the parent ion to one-third of its original intensity. One to three microliters of the sample solutions was loaded into a matrix of thioglycerol.

Composition analysis.

Fifty percent of HPLC fraction 7 was dried and 200 µl of methanolic HCl (1 N HCl in dry MeOH) was added. The mixture was incubated at 80°C for 4 h. After cooling, 50 µl of pyridine was added and the sample was dried under vacuum. The residue was redissolved in 250 µl of MeOH and the solvent evaporated. Trimethylsilylation was carried out by adding 50 µl of a TMS reagent (pyridine/BSTFA/TMCS; 10/5/1; vol/vol/vol) to the dried sample and allowing the mixture to react for 10 min at room temperature. Standard solutions of glucose, mannose, and galactose were prepared in the same way with a final concentration of 1 ng µl^{-1} TMS reagent. Finally, 1-µl aliquots of the silylated sample and standards were directly injected onto the GC column.

GC/MS analysis.

Monosaccharide TMS methyl glycosides were analyzed by means of gas chromatography–mass spectrometry on a Fisons MD 800 mass spectrometer fitted with an Interscience GC

8000 series gas chromatograph with an on-column injector and helium as the carrier gas at a flow rate of 2 ml min^{-1}. The compounds were separated on a DB 5MS column (0.32 × 20 m, J&W Scientific) with the following temperature program: 110°C for 2 min, a gradient of 30°C min^{-1} to 140°C, holding for 2 min at 140°C, then a gradient of 4°C min^{-1} to 240°C and finally holding the temperature at 280°C for 10 min. Mass spectra were obtained under conditions of electron impact and were scanned from *m/z* 50 to 350.

Plant nodulation tests.

The nodulation tests were performed as described by Vargas et al. (1990). Reisolation of bacteria from nodules was done according to Vargas et al. (1990).

ACKNOWLEDGMENTS

This work was supported, in part, by grants from CICYT, Spain (BIO 93-0716-C04-01) and from the Instituto de Cooperación Iberoamericano (Ministerio de Educación y Ciencia, Spain). C. S. and J. F. were supported by fellowships from Conserjería de Educación y Ciencia, Junta de Andalucía, Ministerio de Asuntos Exteriores and General Direction of Academic Staff Affairs (D.G.A.P.), México. J. F. was supported by an EMBO fellowship for carrying out research in the Leiden group. H. P. S. is supported by the Netherlands Organisation for Scientific Research (NWO). J. M. L. L. was supported by a fellowship from the European Union.

LITERATURE CITED

Beringer, J. E. 1974. R-factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. 120:421-429.

Boyer, H. S., and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *E. coli*. J. Mol. Biol. 41:459-472.

Carlson, R. W., Price, N. P. J., and Stacey, G. 1994. The biosynthesis of rhizobial lipo-oligosaccharide nodulation signal molecules. Mol. Plant-Microbe Interact. 7:684-695.

Cervantes, E., Sharma, S. B., Maillet, F., Vasse, J., Truchet, G., and Rosenberg, C. 1989. The *Rhizobium meliloti* host range *nodQ* gene encodes a protein which shares homology with translation elongation and initiation factors. Mol. Microbiol. 3:745-755.

Debellé, F., and Sharma, S. B. 1986. Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. Nucleic Acids Res. 14:7453-7472.

Dénarié, J., and Cullimore, J. 1993. Lipo-oligosaccharide Nodulation factors: Minireview. A new class of signalling molecules mediating recognition and morphogenesis. Cell 74:951-954.

Devereux, J., Haeberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12:387-395.

Ditta, G. S. 1986. Tn5 mapping of *Rhizobium* nitrogen fixation genes. Methods Enzymol. 118:519-528.

Geelen D., Mergaert, P., Geremia, R. A., Goormachtig, S., Van Montagu, M., and Holsters, M. 1993. Identification of *nodSUIJ* in Nod locus *1* of *Azorhizobium caulinodans*: Evidence that *nodS* encodes a methyltransferase involved in Nod factor modification. Mol. Microbiol. 9:145-154.

Henikoff, S. 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods. Enzymol. 155:156-165.

Horvath, B., Kondorosi, E., John, M., Schmidt, J., Török, I., Barabas, I., Wieneke, U., Schell, J., and Kondorosi, A. 1986. Organization, structure and symbiotic function of *Rhizobium meliloti* nodulation genes determining host specificity for alfalfa. Cell 46:335-343.

Kondorosi, A., Kondorosi E., Pankhurst, C. E., Broughton, W. J., and Banfalvi, Z. 1982. Mobilization of a *Rhizobium meliloti* megaplasmid carrying nodulation and nitrogen fixation genes into other rhizobia and *Agrobacterium*. Mol. & Gen. Genet. 188:433-439.

Kozac, M. 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. Microbiol. Rev. 47:1-45.

Kyte, J., and Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J.-C., and Dénarié, J. 1990. Symbiotic host specificity of *R. meliloti* is determined by a sulphated and acylated glucosamine oligo-saccharide signal. Nature 344:781-784.

Leyh, T. S., Suo, Y., and Vogt, T. F. 1992. The DNA sequence of the sulfate activation locus from *Escherichia coli* K-12. J. Biol. Chem. 267:10405-10410.

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

Martínez, E., Pardo, M. A., Palacios, R., and Ceballos, M. A. 1985. Reiteration of nitrogen fixation gene sequences and specificity of *Rhizobium* in nodulation and nitrogen fixation in *Phaseolus vulgaris*. J. Gen. Microbiol. 131:1779-1786.

Martínez, E., Poupot, R., Promé, J.-C., Pardo, M. A., Segovia, L., Truchet, G., and Dénarié, J. 1993. Chemical signalling of *Rhizobium* nodulating bean. Pages 171-175 in: New Horizons in Nitrogen Fixation. R. Palacios, J. Mora, and W. E. Newton, eds. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Martínez, E., Poupot, R., Rogel, M. A., Laeremans, T., Vanderleyden, J., and Hernández-Lucas, I. 1994. Sulfated Nod metabolites from *Rhizobium* species. IS MPMI, 7th, Edinburgh.

Martínez, E., Segovia, L., Martín-Mercante, F., Franco, A. A., Graham, P., and Pardo, M. A. 1991. *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. Int. J. Syst. Bacteriol. 41:417-426.

Megías, M., Folch, J. L., and Sousa, C. 1993. Control of the expression of bacterial genes involved in symbiotic nitrogen fixation. World J. Microbiol. Biotechnol. 9:444-454.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Petersen, G. B., Stockwell, P. A., and Hill, D. F. 1988. Messenger RNA recognition in *E. coli*: A possible second site of interaction with 16S ribosomal RNA. EMBO J. 7:3957-3962.

Poupot, R., Martinez, E., and Promé, J.-C. 1993. Nodulation factors from *R. tropici* CIAT899 are sulfated or nonsulfated chitopentasaccharides containing a *N*-methyl-*N*-acyl glucosaminyl terminus. Biochemistry 32:10430-10435.

Prentki, P., and Krisch, H. M. 1984. In vitro insertional mutagenesis with a selectable DNA fragment. Gene 29:303-313.

Price, N. P. J., Relic, B., Talmont, F., Lewin, A., Promé, D., Pueppke, S. G., Maillet, F., Dénarié, J., Promé, J.-C., and Broughton, W. J. 1992. Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated, and fucosilated, nodulation signals that are O-acetylated or sulphated. Mol. Microbiol. 6:3575-3584.

Putnoky, P., and Kondorosi, A. 1986. Two gene clusters of *Rhizobium meliloti* code for early essential nodulation function and a third influences nodulation efficiency. J. Bacteriol. 169:102-106.

Quandt, J., and Hynes, M. F. 1993. Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. Gene 127:15-21.

Roche, P., Debellé, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Dénarié, J., and Promé, J-C. 1991. Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. Cell 67:1131-1143.

Ruvkun, G. B., and Ausubel, F. M. 1981. A general method for site directed mutagenesis in procaryotes. Nature 289:85-88.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

Schultze, M., Staehelin, C., Röhrig, H., John, M., Schmidt, J., Kondorosi, E., Schell, J., and Kondorosi, A. 1995. *In vitro* sulfotransferase activity of *Rhizobium meliloti* NodH protein: Lipochitooligosaccharide nodulation signals are sulfated after synthesis of the core structure. Proc. Natl. Acad. Sci. USA 92:2706-2709.

Schwedock, J., and Long, S. R. 1989. Nucleotide sequence and protein products of two new nodulation genes in *Rhizobium meliloti*, *nodP* and *nodQ*. Mol. Plant-Microbe Interact. 2:181-194.

Schwedock, J., and Long, S. R. 1993. An open reading frame downstream of *Rhizobium meliloti* *nodQ1* shows nucleotide sequence similarity to an *Agrobacterium tumefaciens* insertion sequence. Mol. Plant-Microbe Interact. 7:151-153.

Schwedock, J. S., Liu, C., Leyh, T. S., and Long, S. R. 1994. *Rhizobium*

meliloti NodP and NodQ form a multifunctional sulfate-activating complex requiring GTP for activity. J. Bacteriol. 176:7055-7064.

Schwedock, J. S., and Long, S. R. 1990. ATP sulphurylase activity of the *nodP* and *nodQ* gene products of *Rhizobium meliloti*. Nature 348:644-647.

Schwedock, J. S., and Long, S. R. 1992. *Rhizobium meliloti* genes involved in sulfate activation: the two copies of *nodPQ* and a new locus, *saa*. Genetics 132:899-909.

Shine, J., and Dalgarno, L. 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA 71: 1342-1346.

Simon, R., Priefer, U., and Pühler, A. 1983. A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technology 1:784-791.

Sousa, C., Folch, J. L., Boloix, P., Megías, M., Nava, N., and Quinto, C. 1993. A *Rhizobium tropici* DNA region carrying the amino-terminal half of a *nodD* gene and a *nod*-box like sequence confers host range extension. Mol. Microbiol. 9:1157-1168.

Spaink, H. P., Aarts, A., Stacey, G., Bloemberg, G. V., Lugtenberg, B. J. J., and Kennedy, E. P. 1992. Detection and separation of *Rhizobium* and *Bradyrhizobium* Nod metabolites using thin-layer chromatography. Mol. Plant-Microbe Interact. 5:72-80.

Spaink, H. P., Bloemberg, G. V., van Brussel, A. A. N., Lugtenberg, B. J. J., van der Drift, K. M. G. M., Haverkamp, J., and Thomas-Oates, J. E. 1995. Host specificity of *Rhizobium leguminosarum* is determined by the hydrophobicity of highly unsaturated fatty acyl moieties of the nodulation factors. Mol. Plant-Microbe Interact. 8:155-164.

Spaink, H. P., and Lugtenberg, B. J. J. 1994. Role of rhizobial lipo-chitin oligosaccharide signal molecules in root nodule organogenesis. Plant Mol. Biol. 26:1413-1422.

Spaink, H. P., Okker, R. J. H., Wijffelmann, C. A., Tak, T., Goosen de Roo, L., Pees, E., van Brussel, A. A. N., and Lugtenberg, B. J. J. 1989. Symbiotic properties of Rhizobia containing a flavonoid independent hybrid NodD product. J. Bacteriol. 171:4045-4053.

Spaink, H. P., Sheeley, D. M., van Brussel, A. A. N., Glushka, J., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N., and Lugtenberg, B. J. J. 1991. A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. Nature 354:125-130.

Tola, E., Giacomini, A., Ollero, F. J., Squartini, A., and Nuti, M. P. 1994. *Rhizobium "hedysari,"* host-specific symbiont of *Hedysarum coronarium*, carries a gene highly homologous to *nodH* of *Rhizobium meliloti*. IS MPMI, 7th, Edinburgh.

Vargas, C., Martinez, L. J., Megías, M., and Quinto, C. 1990. Identification and cloning of nodulation genes and host specificity determinants of the broad host range *Rhizobium leguminosarum* bv. *phaseoli* strain CIAT899. Mol. Microbiol. 4:1899-1910.

Vieille, C., and Elmerich, C. 1990. Characterization of two *Azospirillum brasilense* Sp7 plasmid genes homologous to *Rhizobium meliloti nodPQ*. Mol. Plant-Microbe Interact. 3:389-400.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13phage cloning vectors and hosts strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33:103-119.