The production of species-specific highly unsaturated fatty acyl-containing LCOs from Rhizobium leguminosarum bv. trifolii is stringently regulated by nodD and involves the nodRL genes

Schlaman, W.R.M.; Olsthoorn, M.M.A.; Harteveld, M.; Dorner, L.; Djordjevic, M.A.; Thomas-Oates, J.E.; Spaink, H.P.

MPMI Vol. 19, No. 3, 2006, pp. 215–226. DOI: 10.1094/MPMI-19-0215. © 2006 The American Phytopathological Society

e-Xtra*

The Production of Species-Specific Highly Unsaturated Fatty Acyl-Containing LCOs from *Rhizobium leguminosarum* bv. *trifolii* Is Stringently Regulated by *nodD* and Involves the *nodRL* Genes

Helmi R. M. Schlaman,[1] Maurien M. A. Olsthoorn,[2] Marga Harteveld,[1] Lucie Dörner,[1] Michael A. Djordjevic,[3] Jane E. Thomas-Oates,[2] and Herman P. Spaink[1]

[1]Institute of Biology Leiden, Leiden University, Clusius Laboratory, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands; [2]Department of Mass Spectrometry, Bijvoet Centre for Biomolecular Research, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands; [3]Australian Research Council Centre of Excellence for integrative Legume Research, GI Group, Research School of Biological Sciences, Australian National University, Sullivan's Creek Road, Acton, ACT 2601, Australia

Submitted 2 July 2005. Accepted 18 October 2005.

A proportion of the Nod factors of some *Rhizobium leguminosarum* bv. *trifolii* strains is characterized by the presence of highly unsaturated fatty acyl chains containing *trans* double bonds in conjugation with the carbonyl group of the glycan oligosaccharide backbone. These fatty acyl chains are $C_{18:3}$, $C_{20:3}$, $C_{18:4}$, or $C_{20:4}$ and have UV-absorption maxima at 303 and 330 nm. These Nod factors are presumed to be important for host-specific nodulation on clover species. However, in wild-type *R. leguminosarum* bv. *trifolii* ANU843, Nod factors with these characteristic acyl chains were not observed using standard growth conditions. They were observed only when *nod* genes were present in multiple copies or when transcription was artificially increased to higher levels by introduction of extra copies of the transcriptional regulator gene *nodD*. In a screen for the genetic requirements for production of the Nod factors with these characteristic structures, it was found that the region downstream of *nodF* and *nodE* is essential for the presence of highly unsaturated fatty acyl moieties. Mu-*lacZ* insertion in this region produced a mutant that did not produce detectable levels of the highly unsaturated fatty acyl-bearing Nod factors. The Mu-*lacZ* insertion was translationally fused to a putative new gene, designated *nodR*, in the *nodE-nodL* intergenic region; however, no predicted function for the putative NodR protein has been obtained from database homology searches. In a set of 12 wild-type strains of *R. leguminosarum* bv. *trifolii* originating from various geographical regions that were analyzed for the presence of a *nodR*-like gene, it was found that seven strains carry a homologous NodR open reading frame. Taken together, our results suggest a tightly controlled regulation of *nod* genes, in which we propose that it is the balance of transcriptional levels of *nodFE* and the *nodRL* genes that is critical for determining the presence of highly unsaturated fatty acyl moieties in the Nod factors produced by *R. leguminosarum* bv. *trifolii*.

Corresponding author: H. R. M. Schlaman;
E-mail: schlaman@rulbim.leidenuniv.nl

Current address of M. M. A. Olsthoorn: DSM Food Specialties, R & D, Department of Analysis, postal point 624-0305. P.O. Box 1, 2600 MA Delft, The Netherlands.

Current address of J. E. Thomas-Oates: Department of Chemistry, University of York, Heslington, York YO10 5DD, U.K.

Nucleotide sequence data is available in the NCBI GenBank database under accession numbers: AY924422 (ANU843^{RR}), AY926603 (0403), AY926604 (TA1), AY926597 (RBL52), AY926599 (RBL53), AY926600 (RBL54), AY926598 (LPR5020), AY926601 (CSF), and AY926602 (4S).

*The *e*-Xtra logo stands for "electronic extra" and indicates the HTML abstract available on-line contains supplemental material not included in the print edition, which presents alignments of the nucleotide sequences and the putative *nodR* ORFS as well as Southern hybridization data.

The gram-negative soil bacteria that collectively are called rhizobia are able to live in symbiosis with leguminous plants by the formation of root nodules in which the bacteria reside in a differentiated form, designated bacteroids. Within the root nodule, the bacteroids fix atmospheric nitrogen into ammonia used by the plant as a soluble nitrogen source. Root nodulation is a host-specific process in which one rhizobial species usually successfully nodulates only a limited number of leguminous plants. For instance, *Rhizobium leguminosarum* bv. *trifolii* nodulates clover plants (genus *Trifolium*), whereas its close relative, *R. leguminosarum* bv. *viciae,* nodulates pea (genus *Pisum*), lentil (genus *Lens*), vetch (genus *Vicia*), and sweet pea (genus *Lathyrus*).

Root nodule formation is a de novo process in which differentiated cells in the root cortex dedifferentiate and divide, resulting in a nodule primordium that ultimately enlarges due to rhizobia invading the plant primordium cells via endocytosis. In plants such as clover and vetch, nodules contain a persistent apical meristem and, therefore, are designated indeterminate nodules. Signals known as nodulation (Nod) factors, secreted from the rhizobial cell, trigger the formation of nodule primordia.

The Nod factors consist of a glycan backbone of two to six β-linked *N*-acetylglucosamine residues to which a fatty acyl chain of varying length and with varying degrees of unsaturation is attached. Therefore, Nod factors are designated lipochitin oligosaccharides (LCOs). Furthermore, other substituents, such as fucosyl, sulfate, methyl, carbonyl, or acetyl moieties, are attached to the glycan backbone (Spaink 2000). Nod factors are synthesized by enzymes encoded by bacterial nodu-

lation (*nod*, *nol,* and *noe*) genes that are, in most rhizobial species, present on a large, extra-chromosomal plasmid, designated the symbiosis plasmid (pSym). Some nodulation genes are present in all rhizobia species examined thus far (the so-called common nodulation genes; e.g., *nodABC* and *nodD*), whereas others are to be found in only a few or even a single species. The function of most *nod* gene products in the biosynthesis of the Nod factors is known or can be predicted on the basis of homology studies (Downie 1998). For instance, the NodE protein that belongs to a family of β-ketoacyl synthases, together with NodF, an acyl carrier protein, is involved in the biosynthesis of highly unsaturated fatty acyl chains that are transferred to the chitin oligosaccharide backbone by NodA. A few nodulation proteins have no enzymatic function in nodulation. For instance, NodD and NolR have been shown to act as a transcriptional activator and repressor, respectively, of nodulation gene transcription. NodD-dependent transcription activation requires the presence of a signal molecule from the plant, a flavonoid compound, such as naringenin or di-hydroxyflavone (DHF), as co inducer and is mediated by a consensus sequence in the promoter region, known as the *nod* box (Schlaman et al. 1998).

For structural analyses of Nod factors, *R. leguminosarum* bv. *trifolii* strains have been used that contain cloned nodulation genes (Olsthoorn et al. 2000; Spaink et al. 1995a; van der Drift et al. 1996) in order to increase Nod factor production levels. The Nod factor structures of *R. leguminosarum* bvs. *trifolii* and *viciae* are very much alike: a three- to five-residue-long glycan oligosaccharide backbone to which an acetyl moiety is attached on the nonreducing terminal residue. An additional acetyl group is attached to the reducing residue in *nodX*-bearing strains (Firmin et al. 1993; Olsthoorn et al. 2000; Ovtsyna et al. 1999). A subset of Nod factors from *R. leguminosarum* bv. *viciae* bears a NodE-dependent $C_{18:4}$ acyl chain with three *trans* double bonds at positions 2, 4, and 6 in conjugation with the carbonyl group, and a fourth *cis* double bond which is located at position 11 (Spaink et al. 1991). The presence of these Nod factors was shown to be essential for host-specific biological activity on vetch (Spaink et al. 1991; van Brussel et al. 1992). A subset of Nod factors produced by *R. leguminosarum* bv. *trifolii* bears, instead of $C_{18:4}$, a complex mixture of highly unsaturated fatty acyl chains, including $C_{18:3}$, $C_{20:3}$, and $C_{20:4}$, in which all the *trans* double bonds are conjugated with the carbonyl group and lack a *cis* double bond. These highly unsaturated fatty acids appear to be determined by the specialized characteristics of the NodE of *R. leguminosarum* bv. *trifolii* (Bloemberg et al. 1995). In addition, $C_{20:3}$- and $C_{20:4}$-bearing Nod factors are produced having only two and three double bonds in conjugation with the carbonyl group, respectively (Spaink et al. 1995a; van der Drift et al. 1996). Furthermore, both *R. leguminosarum* bv. *viciae* and *R. leguminosarum* bv. *trifolii* produce Nod factors containing the fatty acyl moieties which commonly occur as components of the phospholipids, such as $C_{16:0}$, $C_{16:1}$, $C_{18:0}$, and $C_{18:1}$ (Olsthoorn et al. 2000; Orgambide et al. 1995; Philip-Hollingsworth et al. 1995; Spaink et al. 1991, 1995a). The relative amount of the highly unsaturated fatty acid-bearing Nod factors produced by *R. leguminosarum* bv. *trifolii* is much lower than that of the $C_{18:4}$-containing Nod factors produced by *R. leguminosarum* bv. *viciae* (Spaink et al. 1991). It is noteworthy that Philip-Hollingsworth and associates did not detect Nod factors with highly unsaturated fatty acids from the *R. leguminosarum* bv. *trifolii* wild-type strain ANU843 (Philip-Hollingsworth et al. 1995). As a result of the longer carbon chain length, the lack of a *cis* double bond, or both, the novel *R. leguminosarum* bv. *trifolii* acyl groups are more hydrophobic than the $C_{18:4}$ fatty acyl chains of *R. leguminosarum* bv. *viciae* (Spaink et al. 1991). It has been shown that a mixture of Nod factors bearing $C_{18:2}$ and $C_{18:3}$ fatty acyl chains lacking the *cis* double bond is able to elicit nodule primordia on red clover, but not on pea, showing their host-specific character (Bloemberg et al. 1995). Thus, the major structural differences between the Nod factors of *R. leguminosarum* bv. *viciae* and *R. leguminosarum* bv. *trifolii* are found in the length and the degree of unsaturation of the fatty acyl chain. The difference in nodulation host range of these two biovars was proposed to be determined by the overall hydrophobicity of the highly unsaturated fatty acyl moieties of the Nod factors (Spaink et al. 1995a).

The highly unsaturated fatty acyl chains in the LCOs from *R. leguminosarum* bv. *trifolii* are characterized by UV absorption maxima at either 303 or 330 nm and long reversed-phase high-performance liquid chromatography (RP-HPLC) retention times of the respective Nod factors due to their hydrophobicity (Spaink et al. 1995b; van der Drift et al. 1996). The characteristic UV absorption maxima are consistent with the presence of fatty acyl chains having a conjugated π system with three and four *trans* double bonds, in conjugation with the carbonyl group (van der Drift et al. 1996). Fast-atom-bombardment–collision-induced dissociation (FAB-CID)–mass spectrometry (MS)/MS analysis of the Nod factors prior to and following palladium-catalyzed reduction of the fatty acyl double bonds showed that up to four double bonds are present (van der Drift et al. 1996). The molecular masses and structures of the LCOs from *R. leguminosarum* bv. *trifolii* have been unambiguously determined using positive and negative ion mode FAB-MS, FAB-MS/MS, and electrospray quadrupole (ES-Q)-MS (Olsthoorn et al. 2000; van der Drift et al. 1996).

In this study, we have systematically screened strains of wild-type *R. leguminosarum* bv. *trifolii* grown under various conditions and a series of *R. leguminosarum* bv. *trifolii* strains and derivatives bearing various sets of cloned nodulation genes to reveal what is needed to produce Nod factors carrying highly unsaturated fatty acyl chains. Nod factor extracts from the strains screened were analyzed using HPLC with photodiode array detection. It was found that a higher level of *nod* gene products, due to either higher expression or copy number of *nod* genes, was an absolute requirement for observing these characteristic Nod factor structures under standard laboratory conditions. Moreover, the intergenic region between *nodE* and *nodL* in *R. leguminosarum* bv. *trifolii* ANU843, containing a putative gene designated *nodR*, was found to be essential. Our data suggest that a critical balance of *nodF*, *nodE,* and putative *nodR* product levels determines the final composition of the Nod factors that are produced by *R. leguminosarum* bv. *trifolii*.

RESULTS

We have systematically screened derivatives of wild-type *R. leguminosarum* bv. *trifolii* grown under various conditions and a series of *R. leguminosarum* bv. *trifolii* strains bearing various sets of cloned nodulation genes to reveal what is needed to produce Nod factors carrying highly unsaturated fatty acyl chains. Nod factor extracts from the strains screened were analyzed using HPLC with photodiode array detection (Table 1). Absorption peaks at 303 and 330 nm are indicative of the presence of highly unsaturated fatty acyl chains in the Nod factors, as was confirmed by nanoES tandem mass spectromic analysis of some representative Nod factor-containing HPLC fractions.

Effects of variations in chromosomal background and number of *nod* gene copies on structures of LCOs produced.

In previous studies, we described the structures of the LCOs produced by *R. leguminosarum* bv. *trifolii* strains containing

cloned nodulation genes (Olsthoorn et al. 2000; Spaink et al. 1995a). Either the pSym-cured strain LPR5045 complemented with plasmid pRtRF101 (copy number 5-10), harboring all *nod* genes except *nodX* from strain ANU843, or the wild-type strain ANU843^{RR}, harboring plasmid pRI4003 (copy number 2-3) that contains all *nod* and *nif* genes, was used. These two *R. leguminosarum* bv. *trifolii* strains differ in i) chromosomal background and ii) copy number of *nod* genes and thus, presumably, in the amount of Nod proteins present. Using mass spectrometric analysis, it was shown that the low-copy-number strain ANU843^{RR}.pRI4003 produces a large range of LCOs, including those bearing $C_{18:3}$, $C_{18:4}$, $C_{20:3}$, $C_{20:4}$, and $C_{20:5}$ fatty acyl chains (Olsthoorn et al. 2000), indicating that highly unsaturated fatty acyl-bearing LCOs also are produced when the *nod* genes from *R. leguminosarum* bv. *trifolii* are present on a low-copy-number plasmid.

Table 1. UV absorption maxima indicative of the number of conjugated double bonds present in the fatty acyl chain of lipo-chitin oligosaccharides of different derivative of *Rhizobium leguminosarum* bv. *trifolii* analyzed using high-performance liquid chromatography (HPLC) with UV photodiode array detection

Strain	Plasmid cn[a]	*nod* Genes present and comments	260 nm[b]	303 nm	330 nm	Percent[c]	MS[d]
ANU843^{RR}	0	All known *R. leguminosarum* bv. *trifolii nod* genes on symbiosis plasmid (pSym)	+	–	–	100	+[e]
idem		Grown at 12ºC	+	–	–	100	+
idem		Induced with seed exudate	+	–	–	100	–
ANU925	0	All known *R. leguminosarum* bv. *trifolii nod* genes with an insertion in putative *nodR* gene on pSym	+	–	–	100	–
LPR5045.pRtRF101	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes except *nodX* on plasmid	+	+	+	20	+[f]
ANU845.pRtRF101	5–10	*idem*	+	+	+	20	–[e]
ANU845.pRt032	5–10	*idem*	+	+	+	20	–
RBL1391.pRtRF101	5–10	*idem*	+	+	+	20	–
248.pRtRF101	5–10	All known *R. leguminosarum* bv. *viciae nod* genes on pSym and all known *R. leguminosarum* bv. *trifolii nod* genes except *nodX* on plasmid	+	+	+	20	–
Variation, cn[g]							
ANU843^{RR}.pRtRF101	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and additionally all known *R. leguminosarum* bv. *trifolii nod* genes except *nodX* on plasmid	+	+	+	20	–
ANU843^{RR}.pRI4003	2–5	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and additionally all known *R. leguminosarum* bv. *trifolii nod* and *nif* genes on plasmid	+	+	+	100	+[h]
RBL1391.pRI4003	2–5	All known *R. leguminosarum* bv. *trifolii nif* and *nod* genes on plasmid	+	+	+	20	–
248.pRI4003	2–5	All known *R. leguminosarum* bv. *viciae nod* genes on pSym with additionally all known *R. leguminosarum* bv. *trifolii nif* and *nod* genes on plasmid	+	+	+	25	–
Variation, source and cn[i]							
ANU843^{RR}.pMP283	5–10	*R. leguminosarum* bv. *trifolii nod* genes including *nodD* on pSym with extra plasmid-borne copies of *nodD* from *R. leguminosarum* bv. *trifolii*	+	+	+	30	–
ANU851.pMP283	5–10	*R. leguminosarum* bv. *trifolii nod* genes and a mutation in *nodD* on pSym with extra plasmid-borne copies of *nodD* from *R. leguminosarum* bv. *trifolii*	+	+	+	20	–
idem			+	+	+	100	–
RBL5584.pMP283	5–10	*R. leguminosarum* bv. *trifolii nod* genes and a mutation in *nodD* on pSym with extra plasmid-borne copies of *nodD* from *R. leguminosarum* bv. *trifolii*	+	+	+	100	–
ANU851.pMP604	5–10	*R. leguminosarum* bv. *trifolii nod* genes and a mutation in *nodD* on pSym with FITA *nodD* on plasmid. No flavonoid induction of bacterial culture	+	+	+	100	–
Variation, specific[j]							
ANU843^{RR}.pMP2121	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and additionally *nodFE* from *R. leguminosarum* bv. *trifolii* on plasmid	+	+	–	20	–
idem			+	+	–	100	+
ANU843^{RR}.pMP2091	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and additionally *nodR* from *R. leguminosarum* bv. *trifolii* on plasmid	+	–	–	100	–
ANU843^{RR}.pMP299	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and additionally *nod FERL* from *R. leguminosarum* bv. *trifolii* on plasmid	+	+	+	100	+
ANU843^{RR}.pMP3608	15–20	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and additionally *nod FERL* from *R. leguminosarum* bv. *trifolii* on plasmid	+	+	+	100	–
ANU925.pMP2091	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and with insertion in the putative *nodR* and additionally *nodR* from *R. leguminosarum* bv. *trifolii* on plasmid	+	–	–	100	–
ANU925.pMP3608	15–20	All known *R. leguminosarum* bv. *trifolii nod* genes on pSym and with insertion in the putative *nodR* and additionally *nod FERL* from *R. leguminosarum* bv. *trifolii* on plasmid	+	+	+	100	+
ANU845.pRt032:344	5–10	All known *R. leguminosarum* bv. *trifolii nod* genes except *nodX* and with insertion in putative *nodR* on plasmid	+	–	–	100	–

[a] Estimated copy number (cn) of introduced plasmid.
[b] + = detected, – = not detected.
[c] Percentage of extract analyzed using HPLC.
[d] HPLC fractions analyzed using mass spectrometry (MS).
[e.] Also analyzed by Philip-Hollingsworth and associates (1995) and Orgambide and associates (1995).
[f] Data have been described by Spaink and associates (1995a).
[g] Variation in cn of *nod* gene-bearing plasmid.
[h] Data have been described by Olsthoorn and associates (2000).
[i] Variation in source and copy number of *nodD.*
[j] Variation in specific *nod* genes.

To further delimit the variables that determine the production of LCOs with highly unsaturated acyl chains, we varied the chromosomal background. Plasmid pRtRF101 was introduced into strain ANU845 (which is ANU843 cured of its pSym) and RBL1391 (which is a pSym-cured derivative of *R. leguminosarum* bv. *viciae*). In both cases, similar results were obtained that were indistinguishable from the results obtained with strain LPR5045.pRtRF101. Indeed, in the HPLC isogram obtained from the LCOs produced by these strains, absorption maxima were observed at 260, 303, and 330 nm (Table 1), indicative of the presence of LCOs containing fatty acyl moieties with two, three, and four *trans* double bonds, respectively, in conjugation with the carbonyl group. These results indicate that the chromosomal background in which pRtRF101 is expressed does not influence the production of highly unsaturated fatty acyl-bearing LCOs. Plasmid pRt032, harboring the same set of *nod* genes as pRtRF101 but originating from a different laboratory (Table 2), then was introduced into ANU845.

Table 2. Rhizobial strains and plasmids

Strains, plasmids	Characteristics[a]	Reference
Bacterial strains		
0403	Wild-type strain *Rhizobium leguminosarum* bv. *trifolii*	Gardiol and Dazzo 1986
248	Wild-type strain *R. leguminosarum* bv. *viciae* carrying pSym pRLJI	Josey et al. 1979
4S	Wild-type strain *R. leguminosarum* bv. *trifolii*	M. W. Breedveld
ANU843^{RR}	Spontaneous rifampicin-resistance derivative of wild type strain *R. leguminosarum* bv. *trifolii* ANU843, which was isolated in Australia, reisolated from nitrogen-fixing nodules on red clover, carrying pSym843, Rif^r	Olsthoorn et al. 2000
ANU845	PSym-cured derivative of *R. leguminosarum* bv. *trifolii* ANU843, Spc^r	Schofield et al., 1983
ANU851	Derivative of ANU843 carrying a Tn*5* insertion in *nodD* gene on pSym843, Km^r, Rif^r	Djordjevic et al. 1985
ANU925	Derivative of ANU843 carrying the Mu dII1743 insertion 344 in the putative nodR gene on pSym843, Km^r	Lewis-Henderson and Djordjevic 1991
K11	Derivative of ANU843 carrying Tn*5* insertion in *nodE* gene on pSym843, Km^r, Rifr^R	Djordjevic et al. 1985
LPR5020	Wild-type strain *R. leguminosarum* bv. *trifolii*, carrying pSym5, Str^r; isolated in U.K.	Hooykaas et al. 1981
LPR5045	pSym-cured derivative of *R. leguminosarum* bv. *trifolii* RCR5, Rif^r	Hooykaas et al. 1982
RBL1391	pSym-cured derivative of *R. leguminosarum* bv. *viciae* 248, Rif^r	Schlaman et al. 1992
RBL5584	LPR5045 carrying pSym843 *nodD*::Tn*5*	Josey et al. 1979
RBL51	Wild-type strain *R. leguminosarum* bv. *trifolii* isolated from red nodules on clover roots obtained in a very cold winter; isolated in the Netherlands	This study
RBL52	Wild-type strain *R. leguminosarum* bv. *trifolii* isolated from clover nodules obtained along the high way; isolated in the Netherlands	This study
RBL53	Wild-type strain *R. leguminosarum* bv. *trifolii*, isolated in the Netherlands	Spaink et al. 1989b
RBL54	Wild-type strain *R. leguminosarum* bv. *trifolii*, isolated in Italy	Canter Cremers et al. 1988
RBL56	Wild-type strain *R. leguminosarum* bv. *trifolii*, isolated in the Netherlands	Canter Cremers et al. 1988
TA1	Wild-type strain *R. leguminosarum* bv. *trifolii*, Nod$^+$ on clover species but cold sensitive nodulation on *Trifolium subterraneum* cv. Woogenellup, isolated in Tasmania, Australia	Gibson 1968
TOM	Wild-type *R. leguminosarum* bv. *viciae* strain able to nodulate pea cv. Afghanistan.	Lie 1978
Tromsø 20-15	Wild-type strain *R. leguminosarum* bv. *trifolii* isolated from white clover nodules obtained in the arctic zone	M. Svenning
Plasmids[b]		
pCR4-TOPO	T-tailed vector for cloning of Taq polymerase-generated DNA fragments, Km^r	Invitrogen life technologies
pMP283	IncP plasmid containing *nodD* from *R. leguminosarum* bv. *trifolii* ANU843, Tc^r	Spaink et al. 1987b
pMP299	IncP plasmid pMP220[2] additionally containing *nodFERL* from ANU843, Tc^r	This study
pMP604	IncP plasmid containing a hybrid *nodD* conferring flavonoid-independent transcription activation (FITA), Tc^r	Spaink et al. 1989c
pMP2009	PCR-amplified *nodE-nodL* intergenic region from CSF in pCR4-TOPO vector, Km^r	This study
pMP2016	PCR-amplified *nodE-nodL* intergenic region from 4S in pCR4-TOPO vector, Km^r	This study
pMP2017	PCR-amplified *nodE-nodL* intergenic region from 0403 in pCR4-TOPO vector, Km^r	This study
pMP2019	PCR-amplified *nodE-nodL* intergenic region from RBL52 in pCR4-TOPO vector, Km^r	This study
pMP2039	PCR-amplified *nodE-nodL* intergenic region from RBL53 in pCR4-TOPO vector, Km^r	This study
pMP2091	IncP plasmid containing *nodR*-open reading frame of ANU843 downstream of *nodA* promoter, Tc^r	This study
pMP2092	PCR-amplified *nodE-nodL* intergenic region from LPR5020 in pUC21 vector, Cb^r	This study
pMP2094	PCR-amplified *nodE-nodL* intergenic region from RBL54 in pCR4-TOPO vector, Km^r	This study
pMP2095	PCR-amplified *nodE-nodL* intergenic region from TA1 in pCR4-TOPO vector, Km^r	This study
pMP2121	IncP plasmid pMP92 containing *nodFE* from ANU843, Tc^r	This study
pMP3402	IncP plasmid containing MCS downstream of the inducible *nodA* promoter from *R. leguminosarum* bv. *viciae* 248, Tc^r	G. Bloemberg
pMP3608	IncQ plasmid pMP190 containing *nodFERL* from ANU843, Cm^r Str^r	This study
pRt032	Self-transmissible plasmid containing 14-kb *Hin*dIII fragment from pSym843 with all known *nod* genes except *nodX*, Cb^r; confers on strain LPR5045 the ability to nodulate efficiently on red clover (Spaink et al. 1995b)	Schofield et al. 1984
pRt032:344	Plasmid pRt032 with a Mu dII1734 insertion in the putative *nodR* gene, Km^r, Cb^r	This study
pRtRF101	IncP plasmid containing 14-kb *Hin*dIII fragment from pSym843 with all known *nod* genes except *nodX*, Tc^r; confers on strain LPR5045 the ability to nodulate efficiently on red clover (Spaink et al. 1995b)	Fisher et al. 1985
pRI4003	IncW plasmid containing all known *nod* and *nif* genes from pSym843, Spc^r; confers on strain LPR5045 the ability to nodulate efficiently on red clover (Spaink et al. 1995b)	Innes et al. 1988
pUC21	General cloning vector, Cb^r	Vieira and Messing 1991

[a] Antibiotic resistance (^{r}): Cb, carbenicillin; Cm, chloramphelicol; Km, kanamycin; Rif, rifampicin; Spc, spectinomycin; Str, streptomycin; Tc, tetracyclin. pSym = symbiosis plasmid, PCR = polymerase chain reaction, MCS = multiple cloning site, and Nod$^+$ = forming root nodules.
[b] Plasmids pMP92, pMP190, and pMP220 have been described (Spaink et al. 1987a).

The HPLC isogram obtained from its Nod factors was indistinguishable from that obtained from ANU845.pRtRF101 (Table 1), confirming that production of highly unsaturated fatty acyl moieties in the LCOs is independent of the chromosomal background but requires at least some of the *R. leguminosarum* bv. *trifolii nod* genes localized on the Sym plasmid.

In addition, plasmid pRI4003 was introduced into *R. leguminosarum* bv. *viciae* strains RBL1391 and 248. The isograms obtained on HPLC separation with photodiode array detection of the LCOs produced by these two strains were similar to those obtained from strains ANU843^{RR}.pRtRF101 and ANU843.pRI4003 (Table 1). Based on the absorption maxima of 260, 303, and 330 nm that were observed, we conclude that the strains with different chromosomal backgrounds harboring pRI4003 all produce LCOs containing fatty acyl chains with one, two, three, and four double bonds in conjugation with the carbonyl group, indicating that the nodulation genes on the Sym plasmid of *R. leguminosarum* bv. *trifolii* ANU843 are sufficient for the synthesis of these LCOs.

Effect of variation in growth and induction conditions on LCO structures produced.

In the HPLC isogram obtained on fractionation of the LCOs produced by wild-type strain ANU843^{RR}, an absorption maximum is observed at 260 nm, in addition to those at 206 and 220 nm, which is indicative of the presence of LCOs having fatty acyl chains with two double bonds in conjugation with the carbonyl group. Absorption maxima at 303 and 330 nm were not be observed, indicating the absence of detectable levels of LCOs carrying highly unsaturated acyl chains having three and four double bonds in conjugation with the carbonyl group, respectively (Table 1). Differences in growth and induction conditions can change the relative quantities of different types of LCOs produced (Demont et al. 1994; McKay and Djordjevic 1993; Olsthoorn et al. 2000); therefore, the growth temperature of the culture was lowered from 28 to 12ºC. The results were essentially indistinguishable, although some differences in the type and relative amounts of LCO structures were observed. FAB-MS analysis (data not shown) of all HPLC fractions of the LCOs from strain ANU843^{RR} grown at 28 and 12ºC revealed the presence of a range of LCO structures, including $C_{18:0}$-, $C_{18:1}$-, $C_{18:2}$-, $C_{20:0}$-, $C_{20:1}$-, and $C_{20:2}$-bearing species, but did not provide any evidence for the presence of $C_{18:3}$-, $C_{20:3}$-, or $C_{20:4}$-bearing LCOs. Thus, it was not possible to detect LCOs with the characteristic highly unsaturated fatty acyl chains observed in cultures of strains harboring pRtRF101 or pRI4003 from the wild-type strain.

When strain ANU843^{RR} was induced with seed exudate from red clover instead of naringenin in order to represent more natural conditions, we also failed to observe peaks with absorption maxima at 303 or 330 nm (Table 1). This result suggests that the failure to observe these absorption maxima in isograms of LCO extracts of wild-type strain ANU843^{RR} is not due to inappropriate induction conditions.

Effects of variations in *nod* gene expression level on LCOs produced.

The amount of Nod proteins can be influenced by both *nod* gene copy number and level of *nod* gene expression. Alteration of transcription levels can be achieved by introducing additional copies of the transcriptional regulator NodD (Spaink et al. 1987a and b). To determine whether a higher copy number of isogenic *nodD* is able to induce production of highly unsaturated fatty acyl-containing LCOs, plasmid pMP283, containing *R. leguminosarum* bv. *trifolii* ANU843 *nodD*, was introduced into ANU843^{RR} (wild type) and ANU851 (*nodD* mutant). By analyzing approximately 25% of the LCO extracts by HPLC with photodiode array detection, peaks with absorption maxima at 260, 303, and 330 nm were observed from the LCO extracts of both transconjungants (Table 1), indicating that *R. leguminosarum* bv. *trifolii* ANU843^{RR} is able to produce LCOs with highly unsaturated fatty acyl chains in the presence of an increased number of copies of *nodD*. As a control, pMP283 also was crossed into strain RBL5584, which is a derivative of LPR5045 containing pSym843 with a Tn*5* mutation in *nodD*. As expected, the LCOs produced by this strain also were shown to carry the highly unsaturated fatty acyl chains (Table 1).

The *nodD* gene product is one of the determinants of host specificity of nodulation, a feature that, at least in part, is due to its selectivity toward flavonoid co-inducers (Spaink et al. 1987b). To analyze whether this feature might influence the structures or amounts of LCOs being produced, a plasmid bearing a *nodD* hybrid gene that activates transcription of *nod* genes independently of the presence of flavonoids (FITA *nodD*, pMP604) was introduced into ANU851, a strain lacking its own *nodD*. Interestingly, the HPLC isogram obtained on HPLC fractionation of LCOs from strain ANU851.pMP604 (Fig. 1A) was similar to that obtained from the LCOs produced

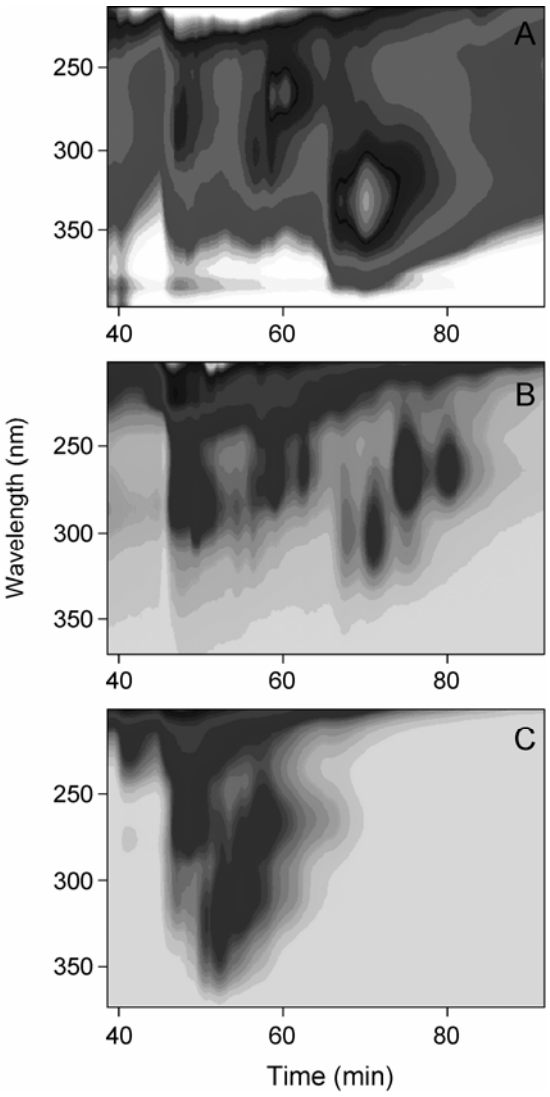

Fig. 1. High-performance liquid chromatography analysis of lipo-chitin oligosaccharides (LCOs) using photodiode array detection. Isograms of LCOs from *Rhizobium leguminosarum* bv. *trifolii* strains **A,** ANU851.pMP604, **B,** ANU843^{RR}.pMP2121, and **C,** ANU843^{RR}.pMP299.

by LPR5045.pRtRF101 and ANU845.pRt032, indicating that strain ANU851.pMP604 is able to produce LCOs with *R. leguminosarum* bv. *trifolii*-specific highly unsaturated fatty acyl chains (Table 1).

In conclusion, our results indicate that wild-type *R. leguminosarum* bv. *trifolii* ANU843 contains the entire genetic complement necessary to produce LCOs bearing highly unsaturated fatty acyl moieties but these can be detected only when the copy number of *nod* genes is increased, when FITA *nodD* or a higher copy number of *nodD* is introduced.

Effect of varying the copy number of specific *nod* genes on LCOs produced.

To determine which specific *nod* genes present in the 14-kb *Hin*dIII fragment cloned in pRtRF101 and pRt032 are responsible for the production of LCOs with highly unsaturated fatty acyl chains, plasmids carrying fragments of these plasmids were introduced into *R. leguminosarum* bv. *trifolii* ANU843^{RR}. First, the effect of plasmid pMP2121 (Fig. 2) containing nodFE, the genes shown to be responsible for the production of the highly unsaturated fatty acyl chains, was examined. Surprisingly, the HPLC isogram of LCOs from ANU843^{RR}.pMP2121 showed peaks with absorption maxima at 260 and 303 nm but not at 330 nm (Table 1; Fig. 1C). The same result was obtained repeatedly, even when the whole LCO extract obtained from the 3-liter culture was loaded onto the HPLC column (Table 1), indicating that LCOs are being produced having fatty acyl chains with three but not four double bonds in conjugation with the carbonyl group. To confirm this surprising observation, relevant HPLC fractions were analyzed using nanoES-MS in combination with CID-MS/MS. In the tandem mass spectra, evidence was obtained for the presence of LCOs carrying fatty acyl moieties including $C_{18:0}$, $C_{18:1}$, $C_{18:2}$, $C_{20:0}$, $C_{20:1}$, $C_{20:2}$, and $C_{20:3}$, but not $C_{18:3}$, $C_{18:4}$, and $C_{20:4}$ (data not shown). Thus, in addition to the LCOs that are produced by wild-type strain ANU843^{RR}, strain ANU843^{RR}.pMP2121 also is able to produce LCOs containing a $C_{20:3}$ moiety. These observations indicate that increased copy numbers of the *nodE* and *nodF* genes are not sufficient for the production of detectable amounts of 330 nm-absorbing fatty acyl chains and that one or more additional genes are necessary that are specifically involved in the production of the LCOs with these fatty acyl chains.

In order to identify these genes, the incP plasmid pMP299 carrying the entire host-specific *nod* region from *nodF* to the 5′ end of *nodM* (Fig. 2) was introduced into strain ANU843^{RR}. In the HPLC isogram obtained from the LCOs produced by ANU843^{RR}.pMP299, absorption maxima are observed at 260 and 303 nm; however, most importantly, also at 330 nm (Fig. 1D; Table 1), indicating the presence of the *R. leguminosarum* bv. *trifolii*-specific LCOs bearing fatty acyl chains that have four double bonds in conjugation with the carbonyl group. To analyze the effect of the copy number, the IncQ plasmid pMP3608 (copy number 15-20) carrying the same set of *nod* genes as pMP299 (copy number 5-10) was crossed into ANU843^{RR} and the LCOs produced by this strain were analyzed by HPLC with photodiode array detection. Again, absorption peaks at 260, 303, and 330 nm were observed in the isogram (Table 1). Mass spectrometric analysis of the HPLC fractions obtained from ANU843^{RR}.pMP299 showing the characteristic 303- and 330-nm absorbance maxima in the isogram was carried out in order to identify these LCOs. The results of CID-MS/MS using nanoES-MS carried out on low-intensity sodium-cationized molecules at *m/z* 1,111.6, 1,113.6, 1,139.7, 1,141.1, 1,155.7, 1,157.7, 1,314.7, 1,316.7, 1,342.8, 1,344.8, 1,384.8, and 1,386.8 revealed the presence of IV($C_{18:4}$, Ac), IV($C_{18:3}$, Ac), IV($C_{20:4}$, Ac), IV($C_{20:3}$, Ac), IV($C_{20:4}$-OH, Ac), IV($C_{20:3}$-OH, Ac), V($C_{18:4}$, Ac), V($C_{18:3}$, Ac), V($C_{20:4}$, Ac), V($C_{20:3}$, Ac), IV($C_{20:4}$, Ac,Ac), and V($C_{20:3}$, Ac,Ac), respectively. The nanoES tandem mass spectrum of the precursor ion at *m/z* 1,342.8, corresponding to V($C_{20:4}$,Ac), is shown in Figure 3. A range of different fragment ions is present, corresponding to A-, B-, C-, Y-, AY-, and BY-type ions, allowing the structure to be identified (see fragmentation scheme). The series of B_n, C_n, and Y_n ions observed at *m/z* 512, 715, 918, and 1,121; *m/z* 530, 733, 936, and 1,139; and *m/z* 244, 447, 650, and 853, respectively, show that the acetyl group is located on the nonreducing-terminal HexNAc residue. The *m/z* value of the B_1 ion indicates that, in addition to this acetyl substituent, a $C_{20:4}$ fatty acyl chain is attached to the nonreducing-terminal HexNAc residue.

Our data indicate that extra copies of genes downstream of *nodE* and upstream of *nodM* are essential and sufficient for the production of the 330-nm-absorbing highly unsaturated fatty acyl chains in *R. leguminosarum* bv. *trifolii* LCOs. The *nodL* gene located in this region has a well-established function as an acetyl transferase (Bloemberg et al. 1994); its activity may exert an influence on whether the special fatty acids are transferred to the growing LCO. In between the *nodE* and *nodL* genes on plasmid pRt032, a Mu dII1734 insertion, designated mutant 344, was roughly mapped (Lewis-Henderson and Djordjevic 1991). This mutant plasmid, as present in the

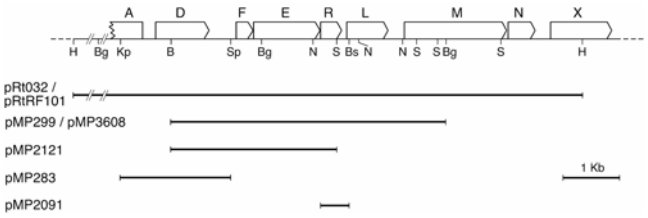

Fig. 2. Overview of *Rhizobium leguminosarum* bv. *trifolii* ANU843 *nod* genes present in the different plasmids used. Top: physical map of part of the *nod* gene region of pSym843. Names of the *nod* genes are given in capitals and the direction of transcription is indicated by arrows. The plasmids pMP299 and pMP3608 contain identical inserts in incP and incQ plasmids, respectively. Bottom: inserts present in the different plasmids. In mutant 344, the insertion Mu dII1734 in *nodR* is flanked by the following sequences: 5′ GAAAGAGCGCCCGCTCCGG↓GTTTGGCACTCGGAG CGGCT 3′ which is 270 bp downstream of the *Nco*I site in *nodE*. Restriction sites: B, *Bam*HI; Bg, *Bgl*II; Bs, *Bsr*GI; H, *Hin*dIII; Kp, *Kpn*I; N, *Nco*I; S, *Sal*I.

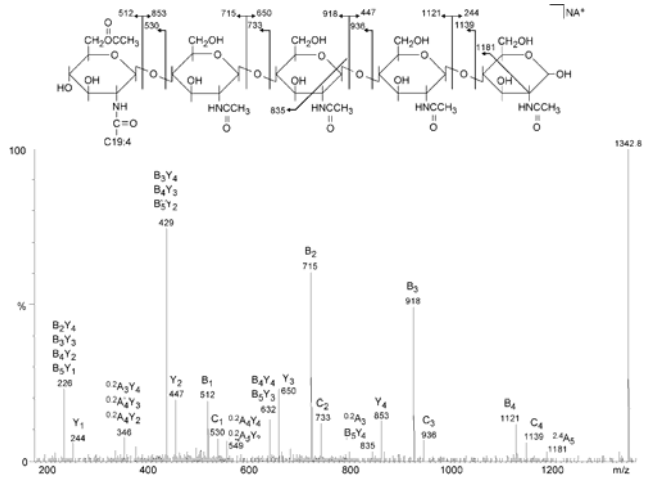

Fig. 3. Nano electrospray quadrupole tandem mass spectrum of lipo-chitin oligosaccharide component V($C_{20:4}$, Ac) produced by *Rhizobium leguminosarum* bv. *trifolii* ANU843^{RR}.pMP299 on collision of *m/z* 1,342.8 from High-performance liquid chromatography fraction having 330-nm absorbance upon photodiode array detection. The *m/z* values are quoted as nominal masses.

pSym-cured strain ANU845, displays β-galactosidase activity, but only upon induction with the flavonoid 7,4′-dihydroxyflavone (to relatively low levels), indicating that a translational fusion is established with the promoter less-truncated *lacZ*, present on the transposon, and some unknown gene present in this region (data not shown). Therefore, the LCOs that were produced by ANU845.pRt032:344 were analyzed by HPLC with photodiode array detection. Indeed, from this strain, no 330-nm-absorbing peaks were observed, indicating that no highly unsaturated fatty acyl chains were produced (Table 1). This result strongly suggests that the intergenic region between *nodE* and *nodL* is responsible for the production of the characteristic *R. leguminosarum* bv. *trifolii* LCOs.

To further examine the role of the *nodE-nodL* intergenic region in the production of LCOs with highly unsaturated fatty acids, the 344 mutation was transferred into the Sym plasmid, generating strain ANU925. As expected, no 303- or 330-nm-absorbing peaks were detected in LCO preparations from this strain (Table 1). Complementation of strain ANU925 with wild-type sequences from the *nodE-nodL* intergenic region on plasmid pMP2091 (see below) did not result in LCOs with absorption maxima at 303 or 330 nm, in contrast to the result obtained with ANU925 containing plasmid pMP3608, on which *nodF* and *nodE* additionally are present (Table 1). Together, the results show that an increased copy number of only the DNA region in between *nodE* and *nodL* is not sufficient for the production of LCOs with highly unsaturated fatty acyl chains, but that extra copies of *nodF* and *nodE* are required as well.

Sequence of the *nodR* open reading frame of *R. leguminosarum* bv. *trifolii* ANU843^{RR}.

The region between *nodE* and *nodL* is rather large in *R. leguminosarum* bv. *trifolii* ANU843^{RR} (approximately 0.5 kb), as deduced from Southern hybridization data (Lewis-Henderson and Djordjevic 1991), and similar in size to the comparable region in the closely related *R. leguminosarum* bv. *viciae* 248 (643 bp). In the latter species, no open reading frame (ORF) is present in this region (Canter Cremers et al. 1989). However, in *R. leguminosarum* bv. *trifolii* ANU843, of which the nucleotide sequence downstream of *nodE* has been only partly published (GenBank accession number X16620), a putative gene called *nodR* has been postulated to be present (Lewis-Henderson and Djordjevic 1991). The putative *nodR* gene is expected to be part of one transcriptional unit together with *nodFE* and *nodL* because the *nodE-nodL* intergenic region lacks an inducible promoter (Canter Cremers et al. 1989). To determine the nucleotide sequence of the entire *nodE-nodL* intergenic region of *R. leguminosarum* bv. *trifolii* ANU843^{RR}, polymerase chain reaction (PCR) was used to amplify the region between the 3′ end of *nodE* and the 5′ end of *nodL*. Using Pfu polymerase, which has proofreading activity, several independent reactions were performed, from which the material was pooled before cloning, resulting in a plasmid series called pMP2091. The nucleotide sequence of the insert of three independent clones of pMP2091 was determined. In all three cases, an identical nucleotide sequence was obtained, showing the reliability of the PCR reaction. Our sequence extended 264 nucleotides beyond the sequence that is present in GenBank (accession number X16620) and it revealed two nucleotide differences (at positions 22 and 61) in the overlapping region. One ORF (called ORF1) of 129 amino acids could be deduced from the sequence, predicting a protein with a molecular weight of 14,959 Da and a *p*I of 11.01. It should be noted that 96 nucleotides downstream of the ATG start codon of this ORF1, a second potential ATG translational start codon is present in the same frame, giving rise to ORF2. ORF2 encodes a protein of 97 amino acids with a predicted molecular weight of 10,948 Da and a *p*I of 9.92. The deduced amino acid sequence of the ORFs did not reveal any significant homology with sequences in the NCBI database. The highest scores obtained using TBLASTN was an identity, maximally, of 30 to 40% with translated sequences of zebrafish (BX004832 and the overlapping AL92899); to which, however, no function could be assigned.

Using the cloned *nodR* as a probe, it was possible to determine whether the putative *nodR* mutant strain ANU845.pRt032:344 indeed contains an insertion in the predicted *nodR* ORF. Southern hybridization showed that, in wild-type strain ANU843^{RR}, an *Nco*I fragment of approximately 0.95 kb hybridized with the *nodR* probe that is not digested with *Hin*dIII, in accordance with the nucleotide sequence. In the mutant, two *Nco*I fragments of 6.6 and 2.7 kb hybridized with the *nodR* probe and the smaller fragment was digested once with *Hin*dIII, resulting in a 1.5-kb band hybridizing with *nodR*. This is in accordance with expectation because the Mu dII1734 insertion contains an internal *Hin*dIII site approximately 1 kb from the end of the transposon (Castilho et al. 1984). These data indicate that insertion 344 is located in the putative *nodR* gene of plasmid pRt032. The exact location of the insertion was determined by nucleotide sequencing and showed that the insertion is present 270 bp downstream of the *Nco*I site in *nodE* and in frame with the predicted ORF1 and ORF2 of the putative *nodR* gene.

Conservation of a *nodR* ORF in *R. leguminosarum* bv. *trifolii* strains.

To determine whether sequences homologous to the putative *nodR* gene are present in other strains of *R. leguminosarum* bv. *trifolii*, Southern hybridization was performed with total DNA from wild-type strains LPR5020 and Tromsø 20-15. These strains were chosen because they have a geographic origin distant from that of ANU843^{RR}. A strongly hybridizing signal (0.9-kb *Nco*I-*Hin*dIII fragment) was observed with DNA from LPR5020 but no signal was obtained with DNA from strain Tromsø 20-15.

To exclude failure of detection in strain Tromsø 20-15 due to low sensitivity of the method, it was decided to use the more sensitive method of PCR to amplify the DNA region between *nodE* and *nodL,* using appropriate primers. The analysis was extended in that a larger collection of various wild-type strains of *R. leguminosarum* bv. *trifolii* was assayed to establish whether the presence of a *nodR* ORF is a general feature among *R. leguminosarum* bv. *trifolii* strains. In addition to strains LPR5020 and Tromsø 20-15, the following strains were analyzed: 0403, 4S, CSF, RBL51, RBL52, RBL53, RBL54, RBL56, and TA1. PCR reactions were performed using a gradient cycler using a 10-step range of annealing temperatures from 47.2 to 58.7ºC to favor annealing of the primers to the template DNA. A PCR product was amplified from strains LPR5020, 0403, 4S, CSF, RBL52, RBL53, RBL54, and TA1 of approximately 600 bp (data not shown). Optimal annealing temperatures giving the highest production of amplified DNA were 48.1 (strain 4S), 56.6 (strains CSF, 0403 and RBL53), 57.1 (RBL54), and 57.9ºC (TA1 and RBL54) (data not shown). For strain 0403, an additional, much smaller PCR product was observed of approximately 300 bp (data not shown). For strains Tromsø 20-15, RBL51, and RBL56, no PCR product was amplified despite several trials with variations to the experimental conditions. A control experiment, in which primers specific for *nodE* were used, showed that a *nodE* PCR fragment was amplified from these strains, confirming the suitability of the method used (data not shown).

The 600-bp PCR products of the various strains obtained in three independent reactions were pooled and cloned into the vector pCR4-TOPO, resulting in the following plasmid series:

pMP2009 (strain CSF), pMP2016 (strain 4S), pMP2017 (strain 0403), pMP2019 (strain RBL52), pMP2039 (strain RBL53), pMP2094 (strain RBL54), and pMP2095 (strain TA1). The PCR product obtained from strain LPR5020 was cloned into the vector pUC21, giving pMP2092. The nucleotide sequence of the plasmid insert was determined, and the DNA sequence of all strains, including ANU843^{RR}, showed a very high degree of similarity. The most striking difference is that strains ANU843^{RR}, 0403, and TA1 have two short additional stretches of five and six nucleotides, respectively. When the various sequenced regions were analyzed for the longest ORF, it was found that, for strains LPR5020, RBL52, RBL53, RBL54, 4S, and CSF, an ORF was predicted encoding a protein of 119 (LPR5020) or 101 (other strains) amino acids that are very similar in the different strains. The corresponding ORF in strain ANU843^{RR}, designated ORF2 above, encodes a protein consisting of 97 amino acids. The longer ORF1 (129 amino acids) that overlaps ORF2 was observed only in strain ANU843^{RR} and not in any of the other *R. leguminosarum* bv. *trifolii* strains due to an upstream TGA stop codon.

Strikingly, in strains TA1 and 0403, a short frame of 10 amino acids similar to the beginning of the NodR ORF is present. This ORF is so short due to the presence of a TGA stop codon. Similarity at the protein level extends for another 12 amino acids followed by another stop codon although, further downstream, strong conservation of DNA sequence remains.

Induction of nodulation by *R. leguminosarum* bv. *trifolii* strains.

To establish a possible phenotype for the putative *R. leguminosarum* bv. *trifolii nodR* mutant, clover plants were inoculated with strains ANU925 and ANU845.pRt032:344 and compared with plants inoculated with the wild-type strain ANU843^{RR}. For strain ANU925, the onset of white clover nodulation was slightly delayed (3 to 4 days), but no phenotype was detected on red clover. No nodulation phenotype on either white or red clover plants was observed for strain ANU845.pRt032:344 (data not shown). These data confirm published data on clover nodulation (Djordjevic et al. 1985). In contrast, mutants in the *nodL* gene completely abolish nodulation on the natural host plant red clover (Canter Cremers et al. 1989), showing no polar effect of the 344 mutation on *nodL* function.

It has been reported that a *nodE* mutant strain of *R. leguminosarum* bv. *trifolii* ANU843, strain K11, is able to nodulate pea cv. Afghanistan in contrast to the wild-type strain (Djordjevic et al. 1985). This very unusual feature may be explained by the fact that strain K11 does not produce LCOs with highly unsaturated fatty acyl chains, because it lacks the keto-acyl reductase activity of the NodE protein. Such LCOs are assumed to inhibit successful nodulation of pea because *R. leguminosarum* bv. *viciae*, the natural microsymbiont of pea, does not synthesize such fatty acyl chains. Strain K11 contains a Tn*5* transposon insertion in the *nodE* gene, which may have polar effects on downstream genes. Taking into consideration our data suggesting a *nodR* gene with a likely function in the synthesis of highly unsaturated fatty acyl chains, located downstream of *nodE*, it was hypothesized that pea nodulation by K11 was not (solely) due to absence of its NodE function but to the additional absence of the NodR protein. To test this hypothesis, pea plants in Afghanistan were inoculated with the *R. leguminosarum* bv. *trifolii* wild-type strains LPR5020, RBL52, RBL53, RBL54, CSF, and 4S (containing a putative *nodR* gene); Tromsø 20-15, 0403, TA1, RBL51, and RBL56 (lacking a putative *nodR* gene); and the *nodR* mutant strains ANU925 and ANU845.pRt032:344. *R. leguminosarum* bv. *trifolii* ANU843^{RR} and *R. leguminosarum* bv. *viciae* 248 served as negative controls. *R. leguminosarum* bv. *trifolii* K11 and *R. leguminosarum* bv. *viciae* TOM were the positive control bacteria. The experiment was repeated three times and nodulation was scored four weeks after inoculation. The results showed that both positive control strains *R. leguminosarum* bv. *viciae* TOM and *R. leguminosarum* bv. *trifolii* K11 nodulated all of the pea plants successfully, whereas none of the other strains tested formed nodules (data not shown). The result also shows that the Mu insertion causing the *nodR* mutants does not destabilize the *nodFERL* transcript because, if this were the case, nodulation of Afghanistan pea would be expected for the *nodR* mutants as it is for the *nodE* mutant. It had to be concluded that the inhibition of pea nodulation by the *nodE* gene of *R. leguminosarum* bv. *trifolii* is independent of the presence of the *nodR* ORF.

DISCUSSION

The results described in this article make it possible to define the mechanism underlying the differential production of LCOs carrying highly unsaturated fatty acyl chains by *R. leguminosarum* bv. *trifolii* ANU843. Our data resolve the apparent inconsistency between the results of our research groups (Olsthoorn et al. 2000; Spaink et al. 1995a; van der Drift et al. 1996) and of Orgambide and associates (1995) and Philip-Hollingsworth and associates (1995), who failed to identify production of LCOs with highly unsaturated fatty acids by *R. leguminosarum* bv. *trifolii*.

The LCOs carrying highly unsaturated fatty acyl chains were not detected in *R. leguminosarum* bv. *trifolii* wild-type strain ANU843^{RR} under any of the experimental conditions we have tested. They were observed only in strain ANU843 and derivatives thereof when the *nod* genes were either i) present in a system with a higher copy number (more than one copy) or ii) expressed at higher levels due to the introduction of additional copies of the *nodD* gene. We show, using HPLC analysis with photodiode array detection, that, among others, minor amounts of LCOs absorbing at 303 and 330 nm then were produced. The structures were analyzed using MS, which indicated the presence of LCOs with C_{20} and C_{18} fatty acyl chains containing three and four double bonds. Due to the high sensitivity of the nanoES-Q time-of-flight mass spectrometer, especially in the MS/MS mode, it was possible to identify these structures clearly.

Absolute requirement for higher levels of *R. leguminosarum* bv. *trifolii* Nod proteins.

Bacterial strains harboring *R. leguminosarum* bv. *trifolii nod* gene-containing plasmids with different copy number and from different sources (pRI4003, pRtRF101, or pRt032) all clearly showed the production of the highly unsaturated fatty acyl chain-bearing LCOs (Table 1). The plasmid pRI4003 is reported to have a very low copy number of approximately 2 to 3 (Innes et al. 1988); therefore, it can be concluded that the system regulating the production of the highly unsaturated fatty acids is extremely sensitively controlled by protein expression levels. The chromosomal background did not affect the results, indicating that only the nodulation genes of the Sym plasmid of *R. leguminosarum* bv. *trifolii* ANU843 are responsible for the synthesis of these LCOs.

Higher expression of *nod* genes can be achieved by introduction of a plasmid harboring the transcriptional regulator *nodD* into ANU843. It was found that introduction of the isogenic *R. leguminosarum* bv. *trifolii nodD*, present on pMP283, resulted in detectable levels of highly unsaturated fatty acyl-carrying LCOs. Introduction of the FITA *nodD* gene on plasmid pMP604 also led to the production of LCOs with highly

unsaturated fatty acyl chains. This result shows that the production of LCOs with highly unsaturated fatty acids is independent of the used induction conditions.

The failure to produce detectable amounts of LCOs with highly unsaturated fatty acyl chains by the wild-type strain ANU843 under standard laboratory conditions might be explained by reduced expression of the *nod* genes required for their synthesis, due to binding of a repressor protein to the *nodFERL* promoter region. A postulated repressor protein may compete with NodD for binding to the promoter and, thus, may block transcription of these host-specific *nod* genes. Such postulated repression would be overcome when extra copies of *nod* genes are present (titration of the postulated repressor by binding sites) or when more NodD protein is produced (competition for binding). An NolR-like *nod* gene repressor might be present in *R. leguminosarum* bv. *trifolii* ANU843, because DNA homology to *Sinorhizobium meliloti nolR* has been shown (Kiss et al. 1998). However, it has been suggested by the same authors that a NolR-like protein probably would not bind to the *nodF* promoter region because it lacks a NolR-target sequence (Kiss et al. 1998). For *R. leguminosarum* bv. *viciae*, it has been shown that, in addition to NodD, an alternative protein binds to the *nodF* promoter that does not bind to any of the other *nod* promoters, although its identity and function are still unknown (Okker et al. 2001; Schlaman et al. 1992).

Role of *nodF*, *nodE*, and *nodR*.

We have shown that extra copies of *nodF* and *nodE*, borne on pMP2121 in *R. leguminosarum* bv. *trifolii* ANU843^{RR}, resulted in production of LCO preparations with 303-nm-absorbing peaks, but not 330-nm peaks. Our data indicate that the highly unsaturated fatty acyl moieties, including $C_{18:4}$ fatty acids (Bloemberg et al. 1995), might be sufficient for blocking nodulation of Afghanistan pea. Western blot analysis showed that pMP2121-bearing strains produce clearly detectable levels of NodE protein (Olsthoorn 2000) and, therefore, it might be concluded that higher levels of NodF and NodE and normal levels of the other Nod proteins are not sufficient for the synthesis of the whole spectrum of *R. leguminosarum* bv. *trifolii* LCOs under laboratory conditions. For this, the *nodRL* region was found to be essential. Importantly, the insertion mutant 344, in which a Mu dII*lac* insertion is located downstream of *nodE* (plasmid pRt032:344), did not produce 330-nm-absorbing LCOs. The nucleotide sequence of the *nodE-nodL* intergenic region of 490 bp predicted two overlapping ORFs, ORF1 (129 amino acids) and ORF2 (97 amino acids), designated *nodR*. Evidence for translation of the postulated *nodR* gene is based on β-galactosidase expression from plasmid pRt032:344 after addition of inducing flavonoid compounds.

Unfortunately, we were unable to show whether a polar effect of the insertion in *nodR* is responsible for the lack of production of highly unsaturated fatty acylated Nod factors because strain ANU843^{RR} carrying additional copies of only the *nodR* region on plasmid pMP2091 did not produce LCOs with highly unsaturated fatty acyl moieties. The most likely explanation for this absence of complementation is that the balance of transcriptional levels of *nodFE* and the *nodR* gene is critical for determining the presence of highly unsaturated fatty acyl moieties in the Nod factors. Although it cannot be excluded that a high level of NodL protein (acetyl transferase) also is required for highly unsaturated fatty acids to be transferred to the growing LCO chain, the following arguments make this extremely unlikely. i) Transposon insertions in the upstream-located *nodF* or *nodE* genes are not polar for the function of *nodL* gene, as shown by normal acetylation patterns of the Nod factors in these mutants (Spaink et al. 1991, 1992). This indicates that, for functional activity of the *nodL* gene, there is no need for inducible expression from the promoter upstream of *nodF*. ii) Supporting this notion is the observation that a *nodL* mutant rhizobial strain in which the *nodL* gene was expressed on a plasmid with extremely low constitutive expression was able to restore red clover nodulation of an *R. leguminosarum* bv. *trifolii nodL* mutant (Canter Cremers et al. 1989). Therefore, although formal proof is still lacking, we propose that the *nodR* gene is essential for the production of Nod factors with $C_{20:4}$ highly unsaturated fatty acyl moieties.

Function of the putative NodR protein.

The putative *nodR* gene is very likely to be functional because a *nodR* ORF is present in 7 of 12 wild-type *R. leguminosarum* bv. *trifolii* strains. In all these cases, the predicted protein showed a high degree of similarity with strain ANU843^{RR} ORF2, encoding a 97-amino-acid protein, that, therefore, is proposed to represent the minimal *nodR* gene. In another two of the 12 strains analyzed, strong DNA homology existed in the *nodE-nodL* intergenic region and a short ORF corresponding to the first 10 amino acids of the *nodR* ORF was present which did not extend due to the presence of a stop codon. No information about the *nodE-nodL* intergenic region was obtained in three of the 12 wild-type strains analyzed. Such evolutionary conservation suggests functionality of the *nodR* ORF in *R. leguminosarum* bv. *trifolii*. This assumption is supported by sequence comparison with the *nodE-nodL* intergenic region in the closely related *R. leguminosarum* bv. *viciae* 248, in which no conservation was detectable at all. Unfortunately, it is not possible to link our sequence data from this large set of *R. leguminosarum* bv. *trifolii* strains to LCO structures because, to date, structural data are unavailable from these strains.

At present, it is only possible to speculate as to the function of the putative NodR protein because, to date, no sequence homology has been found. This is surprising considering the fact that the complete genomes of many organisms have been deposited, including a huge collection of random sequence data from an unknown number of oceanic microorganisms (Venter et al. 2004). However, it can be noted that, for the presumed acyl transferase *nodA,* no homologous sequences are to be found in the NCBI database, although this is an essential gene for Nod factor production. Therefore, it is tempting to speculate that NodR is an alternative acyl transferase to NodA that is specifically required for linking the very hydrophobic 330-nm-absorbing fatty acyl chains to the glycan backbone. Alternatively, NodR might function in the biosynthesis of highly unsaturated fatty acids as a specific β-keto-acyl reductase or hydroxy-acyl dehydratase required to transform $C_{20:3}$ into $C_{20:4}$. Our data indicate that stringent control of the levels of the Nod proteins involved in LCO-specific acyl chain production is essential. Such interplay between Nod proteins has precedent, as exemplified by the interference of the functions of NodL and NodS (López-Lara et al. 2001), and so also may occur between the (putative) acyl transferases NodA and NodR.

Role of LCOs in nodulation.

The absence of a clear nodulation phenotype in the *nodR* mutant strain leaves open the possibility that the putative *nodR* may function in cultivar-specific nodulation. In this regard, it might be relevant to note that *R. leguminosarum* bv. *trifolii* TA1 is unable to nodulate the subterranean clover cv. Woogenellup, whereas strain ANU843 is capable of a normal symbiosis (Lewis-Henderson and Djordjevic 1991). It should be noted that the nodulation host range of *R. leguminosarum* bvs. *trifolii*

and *viciae* overlaps, as is illustrated by the subterranean clover that can be nodulated by both biovars, whereas white clover cannot be nodulated by *R. leguminosarum* bv. *viciae*. This promiscuous behavior complicates the gene-phenotype relationship.

In conclusion, the regulation of the production of the highly unsaturated fatty acid-bearing LCOs in *R. leguminosarum* bv. *trifolii* is complex and, therefore, still not fully understood. In this work, it has been demonstrated that regulation of the transcription level downstream of *nodFE* genes plays an essential role in this process. The question remains as to the function of this regulatory mechanism in the *Rhizobium*–plant symbiosis. Although these LCOs are not detectable in wild-type strains, they have been shown to have an important role in host-specific nodulation (Bloemberg et al. 1995). Schlaman and associates (1992) indicated that it is very likely that there is a different pattern of *nod* gene expression during root nodulation than under laboratory conditions.

The function of decorations of the LCOs during the later stages of the symbiotic process has barely been studied in any system. It might be expected that there is a function of the unsaturated fatty acyl chains in LCOs during the later stages of the infection, not only in *R. leguminosarum* bv. *trifolii* but also in *R. leguminosarum* bv. *viciae* The latter is strongly supported by recent data obtained from *Vicia sativa* inoculation experiments that suggest that unsaturated fatty acyl chain-bearing LCOs are produced mainly at the stage of nodule initiation (Tak et al. 2004). These data, together with those presented in this study, suggest that production of LCOs with highly unsaturated fatty acids is differentially controlled during progression of indeterminate nodulation and, thus, underlines the importance of studying the role of LCOs in planta. Therefore, it is very challenging to study the expression of the *nodFE* and *nodR* genes in planta; for instance, by performing in situ hybridization studies.

MATERIALS AND METHODS

Bacterial strains and plasmids.

Rhizobial strains and plasmids are listed in Table 2. The mini-Mu transposon dII1734 containing the promoterless *Escherichia coli lacZ* structural gene without the first seven codons (Castilho et al. 1984) was inserted into plasmid pRt032 (harboring the carbenicillin resistance gene) and selected by kanamycin resistance, yielding pRt032:344. Strain ANU925 has been made by screening for double reciprocal crossover of the Mu dII1734 mutation 344 into pSym843. Therefore, strain ANU843 containing plasmid pRt032:344 was generated and cells that spontaneously lost their carbenicillin resistance but retained kanamycin resistance and flavonoid-inducible *lacZ* activity were isolated. An overview is given in Figure 2 of the plasmids harboring *nod* genes of *R. leguminosarum* bv. *trifolii* ANU843 aligned to part of the Sym-plasmid pSym843.

Plasmid pMP2091 was constructed by pooling the PCR-amplified DNA of ANU843^{RR} (discussed below) from three independent reactions, purifying it using a Qiagen PCR purification kit, and digesting it with *Eco*RI and *Bsr*GI, followed by ligation into pMP3402 digested with *Eco*RI and *Asp*718I. Similarly, plasmid pMP2092 was constructed using PCR-amplified DNA from strain LPR5020 and pUC21 as a vector. Plasmids pMP2009, pMP2016, pMP2017, pMP2019, pMP2039, pMP2094, and pMP2095 were obtained by thymidine/adenine cloning of PCR-amplified DNA from *R. leguminosarum* bv. *trifolii* strains CSF, 4S, 0403, RBL52, RBL53, RBL54, and TA1, respectively, using vector pCR4-TOPO following the manufacturer's instructions (Invitrogen Life Technologies, San Diego, CA, U.S.A.).

Plasmids were crossed from *E. coli* into rhizobial strains using triparental mating with pRK2013 as the mobilizing plasmid (Ditta et al. 1980).

Molecular genetic analyses.

Template DNA from *R. leguminosarum* bv. *trifolii* strains for PCR reactions was obtained by boiling a freshly grown colony for 10 min in water. After centrifugation for 1 min, the supernatant was used as template DNA in subsequent reactions. The *nodE-nodL* intergenic region was amplified by PCR using primers oMP558 (5′ CTG*GAATTC*AGACAAGTGTGAAGAGGC 3′) and oMP559 (5′ CGAGA*TGTACA*A*CCCGGG*GGA 3′) containing *Eco*RI, *Bsr*GI, and *Sma*I restriction sites, respectively, for cloning purposes (indicated in italics). For strains ANU843^{RR} and LPR5020, the PCR reaction was performed in a final volume of 50 µl in the presence of RNaseA (0.1 µg/µl) and using 5 µl of template and Pfu polymerase (Stratagene, La Jolla, CA, U.S.A.) according to the manufacturer's instructions. DNA amplification of ANU843^{RR} was carried out using the following profile: 1 cycle of 2 min at 94ºC; 35 cycles of: 1 min at 94ºC, 1 min at 59ºC, and 2 min at 72ºC; and 1 cycle of 10 min at 72ºC in a Genius DNA thermo cycler (Techne, Cambridge). Essentially the same profile was used for strain LPR5020 but with a gradient of annealing temperatures ranging from 42.5 to 62.6ºC using a gradient cycler (Biometra T gradient; Biometra, Göttingen, Germany). For other strains, the PCR reactions were performed in a final volume of 25 µl containing 2.5 µl of template, 50 pg of each primer, 0.5 mM dNTPs, and 2.5 U of Super Taq DNA polymerase (HT Biotechnology, Cambridge) in the buffer supplied with the enzyme. DNA amplification was carried out in a gradient cycler using the following profile: 1 cycle of 2 min at 94ºC; 35 cycles of: 1 min at 94ºC, 1 min of annealing (10 steps in the range from 47.2 to 58.7ºC), and 2 min at 72ºC; and 1 cycle of 10 min 72ºC.

Nucleotide sequence determination was performed at BaseClear (Leiden, The Netherlands) using primers oMP12 (5′ GA TCCAACCAATCAATTTTACCAAT 3′) for plasmid pMP2091 and standard M13 reverse primer for plasmid pMP2092 and the pCR4-TOPO-derived plasmids. Sequences were analyzed using Vector NTI software.

For Southern hybridizations, total DNA was isolated from rhizobial strains using Qiagen genomic tips according to the manufacturer's instructions. Total DNA (2 µg) was digested with *Nco*I or *Nco*I and *Hin*dIII, purified using phenol extraction followed by precipitation, separated by gel electrophoresis, blotted onto nitrocellulose membranes, and detected using a digoxigenin (DIG)-labeled probe following the protocol given by the manufacturer (Roche Molecular Biochemicals, Mannheim, Germany). The probe was synthesized by PCR using primers oMP558 and oMP559, 0.5 µl of purified PCR product from ANU843^{RR} as template DNA, DIG-11-UTP, and Super Taq DNA polymerase (HT Biotechnology) according to described methods (Roche protocol). Hybridizing bands were visualized by exposing X-ray films.

Isolation of LCOs and HPLC analysis.

Production and isolation of Nod factors was performed as described (Olsthoorn et al. 2000). Nod factor production was induced with naringenin (3.5 µM) unless indicated otherwise. For strains harboring plasmids, the growth medium was supplemented with appropriate antibiotics (tetracycline at 2 mg/liter or spectinomycin at 10 mg/liter was used). Purification and separation of LCOs by HPLC was carried out as described, using photodiode array detection (Olsthoorn et al. 2000). Either the whole extract or approximately 25% of it was injected onto the column. When the whole extract was injected, 1-min fractions were collected and dried for mass spectrometric analysis.

MS.

Positive ion mode FAB-MS and nanoES-Q time-of-flight MS were carried out as described (Olsthoorn et al. 2000). The HPLC fractions from *R. leguminosarum* bv. *trifolii* strains ANU843^{RR}.pMP2121 and ANU843^{RR}.pMP299, which show the characteristic 303- or 330-nm absorption maxima in the isogram, were pooled and redissolved in 10 to 1,000 µl in acetonitrile/water/formic acid (50:50:0.1, vol/vol/vol) prior to nanoES-MS and nanoES-MS/MS analysis, and 5 mM NaCl was added in order to displace potassium.

Plant nodulation assays and preparation of seed exudate.

Seed were sterilized and germinated as described (van Brussel et al. 1986). For nodulation assays, clover seedlings were put on agar slopes of Jensen medium (Vincent 1970) in sterile tubes, each infected with 100 µl of a fresh bacterial solution in Jensen medium (absorption at 660 nm of 0.1 to 0.2), and the roots were shielded from the light. Germinated seed of pea cultivars were inoculated by rolling them through a layer of freshly grown bacteria, then placed on lids of dark-brown-colored glass pots filled with liquid Jensen medium. Plants were incubated at 20ºC with a 16-h daytime.

For seed exudates, seed of *Trifolium pratense* were incubated in Jensen medium (Vincent 1970) for a week at 4ºC, within which time the seed germinated. After removal of the seedlings, the exudate was used for induction of LCO production.

ACKNOWLEDGMENTS

This work was supported in part by The Netherlands Organization for Earth and Life Sciences (with financial aid from the Netherlands Organization for Scientific Research). We would like to thank A. Aarts, P. van Dillewijn, and G. Bloemberg (Leiden University, The Netherlands) for the construction of pMP3608, pMP2121, and pMP3402, respectively; M. Svenning (Tromsø University, Norway) and M. Breedveld (Agricultural University Wageningen, The Netherlands) for the gift of strains Tromsø 20-15 and 4S, respectively; and T. Tak for his help in performing the nodulation experiments on Afghanistan pea. We greatly acknowledge J. Weinman (Australian National University, Canberra, Australia) for sequencing the site of insertion of Mu dII1734 in the mutant 344.

LITERATURE CITED

Bloemberg, G. V., Kamst, E., Harteveld, M., van der Drift, K. M. G. M., Haverkamp, J., Thomas-Oates, J. E., Lugtenberg, B. J. J., and Spaink, H. P. 1995. A central domain of *Rhizobium* NodE protein mediates host specificity by determining the hydrophobicity of fatty acyl moieties of nodulation factors. Mol. Microbiol. 16:1123-1136.

Bloemberg, G. V., Thomas-Oates, J. E., Lugtenberg, B. J. J., and Spaink, H. P. 1994. Nodulation protein NodL of *Rhizobium leguminosarum* O-acetylates lipo-oligosaccharides, chitin fragments and N-acetylglucosamine in vitro. Mol. Microbiol. 11:793-804.

Canter Cremers, H. C. J., Wijffelman, C. A., Pees, E., Rolfe, B. G., Djordjevic, M. A., and Lugtenberg, B. J. J. 1988. Host specific nodulation of plants of the pea cross-inoculation group is influenced by genes in fast growing Rhizobium downstream nodC. J. Plant Physiol. 132:398-404.

Canter Cremers, H. C. J., Spaink, H. P., Wijfjes, A. H. M., Pees, E., Wijffelman, C. A., Okker, R. J. H., and Lugtenberg, B. J. J. 1989. Additional genes on the symplasmid of *Rhizobium leguminosarum* biovar *viciae*. Plant Mol. Biol. 13:163-174.

Castilho, B. A., Olfson, P., and Casadaban, M. J. 1984. Plasmid insertion mutagenesis and lac gene fusion with mini-mu bacteriophage transposons. J. Bacteriol. 158:488-495.

Demont, N., Ardourel, M., Maillet, F., Promé, D., Ferro, M., Promé, J.-C., and Dénarié, J. 1994. The *Rhizobium meliloti* regulatory *nodD3* and *syrM* genes control the synthesis of a particular class of nodulation factors *N*-acetylated by (w-1)-hydroxylated fatty acids. EMBO (Eur. Mol. Biol. Organ.) J. 13:2139-2149.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria: Construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. U.S.A. 77:7347-7351.

Djordjevic, M. A., Schofield, P. R., and Rolfe, B. G. 1985. Tn-*5* mutagenesis of *Rhizobium trifolii* host specific nodulation genes result in mutants with altered host range ability. Mol. Gen. Genet. 200:463-471.

Downie, J. A. 1998. Function of Rhizobial nodulation genes. Pages 387-402 in: The *Rhizobiaceae*. H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas, eds. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Firmin, J. L., Wilson, K. E., Carlson, R. W., Davies, A. E., and Downie, J. A. 1993. Resistance to nodulation of cv. Afghanistan peas is overcome by *nodX*., which mediates an O-acetylation of the *Rhizobium leguminosarum* lipo-oligosaccharide nodulation factor. Mol. Plant-Microbe Interact. 10:351-360.

Fisher, R. F., Tu, J. K., and Long, S. R. 1985. Conserved nodulation genes in *Rhizobium meliloti* and *Rhizobium trifolii*. Appl. Environ. Microbiol. 49:1432-1435.

Gardiol, A. E., and Dazzo, F. B.. 1986. Biosynthesis of *Rhizobium trifolii* capsular polysaccharide: Enzymatic transfer of pyruvate substitutions into lipid-bound saccharide intermediates. J. Bacteriol. 168:1459-1462.

Gibson, A. H. 1968. Nodulation failure in *Trifolium subterraneum* L. cv. Woogenellup (Syn. Marrar). Austr. J. Agric. Sci. 19:907-918.

Hooykaas, P. J. J., Peerbolte, R., Regensburg-Tuink, A. J. G., de Vries, P., and Schilperoort, R. A. 1982. A chromosomal linkage map of *Agrobacterium tumefaciens* and a comparison with the maps of *Rhizobium* spp. Mol. Gen. Genet. 188:12-17.

Hooykaas, P. J. J., van Brussel, A. A. N., den Dulk-Ras, A., van Slogteren, G. M. S., and Schilperoort, R. A.. 1981. Symplasmid of *Rhizobium trifolii* expressed in different rhizobial species and *Agrobacterium tumefaciens*. Nature 291:351-353.

Innes, R. W., Hirose, M. A., and Kuempel, P. L. 1988. Induction of nitrogen-fixing nodules on clover requires only 32 kilobase pairs of DNA from the *Rhizobium trifolii* symbiosis plasmid. J. Bacteriol. 170:3793-3802.

Josey, D. P., Beynon, J. L., Johnston, A. W. B., and Beringer, J. B. 1979. Strain identification in Rhizobium using intrinsic antibiotic resistance. J. Appl. Microbiol. 46:343-350.

Kiss, E., Mergaert, P., Olàh, B., Kereszt, A., Staehelin, C., Davies, A. E., Downie, J. A., Kondorosi, A., and Kondorosi, E. 1998. Conservation of *nolR* in the *Sinorhizobium* and *Rhizobium* genera of the *Rhizobiaceae* family. Mol. Plant-Microbe Interact. 11:1186-1195.

Lewis-Henderson, W. R., and Djordjevic, M. A. 1991. *nodT*, a positively-acting cultivar specificity determinant controlling nodulation of *Trifolium subterraneum* by *Rhizobium leguminosarum* biovar *trifolii*. Plant Mol. Biol. 16:515-526.

Lie, T. A. 1978. Symbiotic specializiation in pea plants: The requirements of specific *Rhizobium* strains for peas from Afghanistan. Proc. Assoc. Appl. Biol. 88:462-465.

López-Lara, I. M., Kafetzopoulos, D., Spaink, H. P., and Thomas-Oates, J. E. 2001. Rhizobial NodL O-acetyl transferase and NodS N-methyl transferase functionally interfere in production of modified Nod factors. J. Bacteriol. 183:3408-3416.

McKay, I. A., and Djordjevic, M. A. 1993. Production and excretion of nod metabolites by *Rhizobium leguminosarum* bv. *trifolii* are disrupted by the same environmental factors that reduce nodulation in the field. Appl. Environ. Microbiol. 59:3385-3392.

Okker, R. J., Spaink, H. P., Lugtenberg, B. J., and Schlaman, H. R. M. 2001. Mutants in the *nodFEL* promoter of *Rhizobium leguminosarum* bv. *viciae* reveal a role of individual nucleotides in transcriptional activation and protein binding. Arch. Microbiol. 175:152-160.

Olsthoorn, M. M. A. 2000. Screening of bacterial glycoconjugate structures. Pages 45-122 in: Screening and detailed structural analysis of bacterial glycoconjugates: A multi-technique approach. Ph.D. thesis, Utrecht University, The Netherlands.

Olsthoorn, M. M. A., Stokvis, E., Haverkamp, J., Spaink, H. P., and Thomas-Oates, J. E. 2000. Growth temperature regulation of host-specific modifications of rhizobial lipo-chitin oligosaccharides: The function of *nodX* is temperature regulated. Mol. Plant-Microbe Interact. 13:808-820.

Orgambide, G. G., Lee, J., Hollingsworth, R. I., and Dazzo, F. B. 1995. Structurally diverse chitolipooligosaccharide nod factors accumulate primarily in membranes of wild type *Rhizobium leguminosarum* biovar *trifolii*. Biochemistry 34:3832-3840.

Ovtsyna, A. O., Rademaker, G. J., Esser, E., Weinman, J., Rolfe, B. G., Tikhonovich, I. A., Lugtenberg, B. J., Thomas-Oates, J. E., and Spaink, H. P. 1999. Comparison of characteristics of the *nodX* genes from various *Rhizobium leguminosarum* strains. Mol. Plant-Microbe Interact. 12:252-258.

Philip-Hollingsworth, S., Orgambide, G. G., Bradford, J. J., Smith, D. K., Hollingsworth, R. I., and Dazzo, F. B. 1995. Mutation or increased copy number of *nodE* has no effect on the spectrum of chitolipooligosaccha-

ride nod factors made by *Rhizobium leguminosarum* bv. *trifolii*. J. Biol Chem. 270:20968-20977.

Schlaman, H. R. M., Lugtenberg, B. J. J., and Okker, R. J. H. 1992. The NodD protein does not bind to the promoters of inducible nodulation genes in bacteroids of *Rhizobium leguminosarum* biovar *viciae*. J. Bacteriol. 174:6109-6116.

Schlaman, H. R. M., Phillips, D. A., and Kondorosi, E. 1998. Genetic organisation and regulation of the nodulation genes. Pages 361-386 in: The *Rhizobiaceae*. H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas, eds. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Schofield, P. R., Djordjevic, M. A., Rolfe, B. G., Shine, J., and Watson, J. M. 1983. A molecular linkage map of nitrogenase and nodulation genes in *Rhizobium trifolii*. Mol. Gen. Genet. 192:459-465.

Schofield, P. R., Ridge, R. W., Rolfe, B. G., Shine, J., and Watson, J. M. 1984. Host specific nodulation is encoded on a 14 kb DNA fragment in *Rhizobium trifolii*. Plant Mol. Biol. 3:3-11.

Spaink, H. P. 2000. Root nodulation and infection factors produced by rhizobial bacteria. Annu. Rev. Microbiol. 54:257-288.

Spaink, H. P., Okker, R. J. H., Wijffelman, C. A., Pees, E., and Lugtenberg, B. J. J. 1987a. Promoters in the nodulation region of *Rhizobium leguminosarum* Sym plasmid pRL1JI. Plant Mol. Biol. 9:29-37.

Spaink, H. P., Wijffelman, C. A., Pees, E., Okker, R. J. H., and Lugtenberg, B. J. J. 1987b. *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. Nature 328:337-339.

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

Spaink, H. P., Weinman, J., Djordjevic, M. A., Wijffelman, C. A., Okker, R. J. H., and Lugtenberg, B. J. J. 1989b. Genetic analysis and cellular localization of the *Rhizobium* host specificity-determining NodE protein. EMBO (Eur. Mol. Biol. Organ.) J. 8:2811-2818.

Spaink, H. P., Wijffelman, C. A., Okker, R. J. H., and Lugtenberg, B. E. J. 1989c. Localization of functional regions of the *Rhizobium nodD* product using hybrid *nodD* genes. Plant Mol. Biol. 12:59-73.

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.

Spaink, H. P., Aarts, A., Stacey, G., Bloemberg, G. V.; Lugtenberg, B. 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. 1995a. 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., Wijfjes, A. H. M., and Lugtenberg, B. J. J. 1995b. *Rhizobium* NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligosaccharides. J. Bacteriol. 177:6276-6281.

Tak, T., van Spronsen, P. C., Kijne, J. W., van Brussel, A. A. N., and Boot, K. J. M. 2004. Accumulation of lipochitin oligosachharides and NodD-activating compounds in an efficient plant-*Rhizobium* nodulation assay. Mol. Plant-Microbe Interact. 17:816-823.

van Brussel, A. A. N., Bakhuizen, R., van Spronsen, P. C., Spaink, H. P., Tak, T., Lugtenberg, B. J. J., and Kijne, J. W. 1992. Induction of pre-infection thread structures in the leguminous host plant by mitogenic lipo-oligosaccharides of *Rhizobium.* Science 257:70-72.

van Brussel, A. A. N., Zaat, S. A. J., Canter Cremers, H. C. J., Wijffelman, C. A., Pees, E., Tak, T., and Lugtenberg, B. J. J. 1986. Role of plant root exudate and symplasmid-localized nodulation genes in the synthesis by *Rhizobium leguminosarum* of Tsr factor, which causes thick and short roots on common vetch. J. Bacteriol. 165:517-522.

van der Drift, K. M., Spaink, H. P., Bloemberg, G. V., van Brussel, A. A. N., Lugtenberg, B. J., Haverkamp, J., and Thomas-Oates, J. E. 1996. *Rhizobium leguminosarum* bv. *trifolii* produces lipo-chitin oligosaccharides with nodE-dependent highly unsaturated fatty acyl moieties. An electrospray ionization and collision-induced dissociation tandem mass spectrometric study. J. Biol. Chem. 271:22563-22569.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., Fouts, D. E., Levy, S., Knap, A. H., Lomas, M. W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y. H., and Smith, H. O. 2004. Environmental genome shotgun sequencing of the Sargasso sea. Science 304:66-74.

Vieira, J., and Messing, J. 1991. New pUC-derived cloning vectors with different selectable markers and DNA replication origins. Gene 100:189-194.

Vincent, J. M. 1970. Pages 75-76 in: A Manual for the Practical Study of the Root Nodule Bacteria. International Biological Program handbook no. 15. Blackwell Scientific Publications, Ltd., Oxford, U.K.