Host specificity of Rhizobium leguminosarum is determined by the hydrophobicity of highly unsaturated fatty acyl moieties of the nodulation factors

Spaink, H.P.; Bloemberg, G.V.; Brussel, A.A.N. van; Lugtenberg, E.J.J.; Drift, K.M.G.M. van der; Haverkamp, J.; Thomas-Oates, J.E.

MPMI Vol. 8, No. 1, 1995, pp. 155–164.

Host Specificity of *Rhizobium leguminosarum* Is Determined by the Hydrophobicity of Highly Unsaturated Fatty Acyl Moieties of the Nodulation Factors

Herman P. Spaink,[1] Guido V. Bloemberg,[1] Anton A. N. van Brussel,[1] Ben J. J. Lugtenberg,[1] Koen M. G. M. van der Drift,[2] Johan Haverkamp,[2] and Jane E. Thomas-Oates[2]

[1]Institute of Molecular Plant Sciences, Clusius Laboratory, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands; [2]Bijvoet Center for Biomolecular Research, Department of Mass Spectrometry, Utrecht University, P.O.Box 80083, 3508 TB Utrecht, The Netherlands

Received 15 September 1994. Accepted 31 October 1994.

The *nodE* genes of *Rhizobium leguminosarum* bvs. *trifolii* and *viciae* strains are the major determinants of the host range of nodulation. Using mass spectrometry we have analyzed the structures of the major and minor lipo-oligosaccharides produced by two *R. leguminosarum* strains that differ only in the origin of the *nod* genes. The strain containing the *nod* genes of *R. l.* bv. *viciae* produces lipo-chitin oligosaccharide (LCO) molecules that contain the common *cis*-vaccenyl (C18:1) or a highly unsaturated acyl (C18:4) group that is determined by *nodE* (Spaink *et al.* Nature 354:125-130, 1991). Here we show that, in addition, minor quantities of analogous molecules that contain different common fatty acyl groups such as stearyl (C18:0), palmityl (C16:0), and palmitoyl (C16:1) groups are produced. The set of LCOs produced by the strain containing the *R. l.* bv. *trifolii nod* genes is similar, but major differences are found in the length and unsaturation of the fatty acyl groups of various LCOs. The *R. l.* bv. *trifolii* strain does not produce LCO molecules that contain the C18:4 fatty acyl group. Instead, a complex mixture of other unsaturated fatty acyl groups containing either two, three, or four double bonds is produced. These novel fatty acyl groups are more hydrophobic than the C18:4 fatty acyl group found in bv. *viciae*, as judged by high-performance liquid chromatography retention times, probably as a result of the longer carbon chain length, the lack of a *cis* double bond, or both. Another difference is that the relative amount of highly unsaturated fatty acid–containing LCOs produced by *R. l.* bv. *trifolii* is much smaller. An isogenic *R. l.* bv. *trifolii* strain containing a Tn5 insertion in the *nodE* gene only produces LCOs containing the common fatty acyl groups and not the highly unsaturated types. The results indicate that the difference in the host range of the *R. l.* bvs. *viciae* and *trifolii* is determined by the overall hydrophobicity of the highly unsaturated fatty acyl moieties of LCOs rather than by a specific structural feature.

Corresponding author: Herman P. Spaink.

H.P.S., G.V.B., K.M.G.M.V., and J.E.T.-O. have made equal contributions to this work.

Additional keywords: chitin, mass spectrometry, signal molecules.

The interaction between rhizobia, bacteria belonging to the genera *Rhizobium*, *Bradyrhizobium,* and *Azorhizobium*, and their leguminous host plants results in the formation of nitrogen-fixing root nodules in a species-specific way (see Brewin 1991). The specificities of various types of rhizobia for host plants are very different. Some strains are able to nodulate a wide range of mainly tropical plant genera, whereas other strains only nodulate plant species belonging to a few genera. Examples of the latter case are *Rhizobium leguminosarum* (*R. l.*) bv. *trifolii* which nodulates plants of the genus *Trifolium*, and the closely related *R. l.* bv. *viciae*, which nodulates plants of the genera *Lathyrus*, *Lens*, *Pisum,* and *Vicia*. Recently it has become apparent that the host specificity of rhizobia is based on a mutual recognition of signal molecules produced by plant and bacteria. Flavonoid signal molecules produced by the plant are recognized by the guest bacteria, a process that is mediated by the regulatory NodD protein (see Fisher and Long 1992; Spaink 1994). As a result of this recognition the NodD protein activates the transcription of the *nod* and *nol* genes and subsequently the bacteria produce lipo-chitin oligosaccharide (LCO) signal molecules that are specifically recognized by the host plant (see Dénarié and Cullimore 1993). The LCO molecules from many rhizobial strains have been identified and they all appear to consist of an acylated chitin fragment that can contain strain-specific modifications. It has been shown that several *nod* genes encode enzymes that are involved in the processes of fatty acid biosynthesis, chitin synthesis, and chitin modification (see Dénarié and Cullimore 1993; Spaink and Lugtenberg 1994). Through their essential role in the biosynthesis of the LCOs, the *nod* genes are major mediators of the host specificity of nodulation.

In the case of *R. l.* bvs. *trifolii* and *viciae*, the *nodE* gene has been shown to be the major determinant of the difference in their host specificities (Spaink *et al.* 1989). In *R. l.* bv. *viciae* the *nodE* gene is involved in the biosynthesis of a highly unsaturated C18:4 fatty acyl moiety which is present in a fraction of the LCOs. In *R. l.* bv. *trifolii* the biochemical role

of the *nodE* gene has not been shown. In this paper we describe the structural analysis of the LCOs that are produced by means of the *R. l.* bv. *trifolii nod* genes and the investigation of the role of the *nodE* gene in their production.

RESULTS

TLC analysis of lipo-chitin oligosaccharides.

We have constructed an isogenic set of *R. leguminosarum* strains which only differ in the origin of the *nod* genes. The results of the nodulation test (Fig. 1) show that these strains have normal nodulation characteristics either on *Vicia sativa* (in the case of plasmid pIJ1089) or *Trifolium pratense* (in the case of plasmids pRtRF101 and pRI4003). Recently we have shown that LCOs can be efficiently radiolabeled by growing the bacteria in the presence of [1-^{14}C]-glucosamine (Spaink *et al.* 1994). With this method, the two derivatives of strain LPR5045 harboring the plasmids pIJ1089 and pRtRF101, containing the *nod* genes of *R. l.* bv. *viciae* and *trifolii*, respectively, were grown in the presence or absence of the *nod* gene inducer naringenin and analyzed for the production of LCOs. TLC analysis of the *n*-butanol extracts of the cell pellet, the spent growth medium, and the cell-wash (Fig. 2A) showed several naringenin-inducible spots for each strain. Quantification of radioactivity using the ImageQuant software showed that the total radioactivity in the TLC profiles was similar for the two strains tested (Fig. 2C). As has been observed previously (Spaink *et al.* 1992), the major difference between the profiles of the two biovars is the lack of a major spot at a relative flow (Rf) of 60% (Fig. 2) in the *R. l.* bv. *trifolii* strain. In *R. l.* bv. *viciae* this spot, which is not observed in the case of a *nodE* mutant strain, has been identified as NodRlv-V(C18:4,Ac) (Spaink *et al.* 1991). In order to investigate the role of the *nodE* gene we have analyzed the LCOs produced by strain LPR5045.pMP2111, containing the *nodE11*::*Tn5* mutation (Fig. 2A). Surprisingly, the LCO profile of this strain was not different from the wild-type control. In addition, using another *R. l.* bv. *trifolii* strain (LPR5045.pRI4003; Fig. 2B, lanes 4 and 8), different culture conditions such as various induction times (2, 6, 12, and 16 hr; data not shown) and growth temperature (Fig. 2B), and using different extraction methods such as that of Bligh and Dyer (1959) (data not shown), we have been unable to detect a difference in the LCO profiles on TLC analysis. However, it was noted that at the lower growth temperature of 18° C the quantity of radiolabeled LCO molecules produced was increased (Fig. 2B). An effect of the growth temperature on the quantity of LCO produced by *R. l.* bv *trifolii* has also been reported by McKay and Djordjevic (1993).

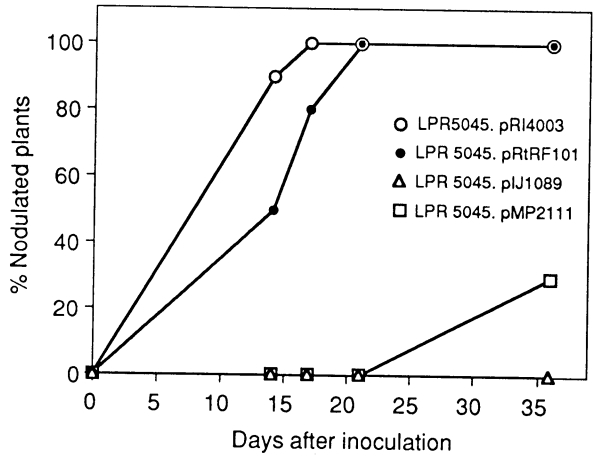

Fig. 1. Nodulation kinetics of strain LPR5045 containing various plasmids on *Trifolium pratense*. Plasmid pIJ1089 contains the *nod* genes of *Rhizobium leguminosarum* bv. *viciae*. Plasmids pRtRF101 and pRI4003 contain the *nod* genes of *R. l.* bv. *trifolii*. Plasmid pMP2111 is a derivative of plasmid pRtRF101 harboring a Tn*5* insertion in the *nodE* gene. For each strain 20 plants were inoculated under conditions which have been described previously (van Brussel *et al.* 1982).

HPLC analysis.

One-liter cultures of the four derivatives of strain LPR5045, containing the plasmids pIJ1089, pRtRF101, pRI4003, or pMP2111, were grown in the presence of the inducer naringenin, and LCOs were isolated using a revised purification method, as described in the experimental section. Before the final HPLC purification step, the samples were mixed with extracts of radiolabeled LCOs of the corresponding strain (Fig. 2B, lanes 5 to 7). One-minute fractions from the HPLC elution were collected. Part of these fractions was concentrated and analyzed using TLC and phosphorimaging. The results link the UV absorption profiles obtained from the photodiode array analyses (Fig. 3, panels A, B, and C) to the radioactive profiles observed on TLC (e.g., Fig. 3D). For fractions 1–10 of the HPLC elution the absorption profile of LCOs was undetectable due to the elution of the flavonoid inducer naringenin. The HPLC diode array profile of the sample derived from *R. l.* bv. *viciae* strain LPR5045.pIJ1089 (Fig. 3, panel A) contains two major peaks (fractions A11-12 and A15-16) with the typical absorption maximum at 303 nm, which has been shown previously to correspond to the compounds NodRlv-V(C18:4,Ac) and NodRlv-IV(C18:4,Ac), respectively (Spaink *et al.* 1991). In addition, two major peaks without specific absorption maxima were observed (fractions A18-19 and A24-25), which correspond to the compounds NodRlv-V(C18:1,Ac) and NodRlv-IV(C18:1, Ac) (Spaink *et al.* 1991). Very minor peaks were observed that elute at the position of the radiolabeled compounds that migrate on TLC at an Rf between 25 and 42% (Fig. 3D). These peaks had not been analyzed in previous studies.

The HPLC diode array profiles of the sample derived from *R. l.* bv. *trifolii* strains LPR5045.pRtRF101 (Fig. 3B) and LPR5045.pRI4003 (data not shown) were indistinguishable, except that in the latter strain the total amount of LCOs produced was less. The HPLC profiles contain only two major peaks that migrate at the same position as the compounds NodRlv-V(C18:1,Ac) and NodRlv-IV(C18:1,Ac) (Fig. 3D). Apparently, these strains do not produce the compounds NodRlv-V(C18:4,Ac) and NodRlv-IV(C18:4,Ac), which is consistent with the absence of radiolabeled material at an Rf of 60% on the thin-layer chromatograph (Fig. 2A). Instead, a complex pattern of minor peaks is observed (fractions B25 to B53) which have characteristic absorption maxima at 260 nm, 303 nm, or 330 nm, or combinations of these. The former two maxima would be consistent with fatty acyl groups that contain two or three *trans* double bonds in conjugation with the carbonyl, respectively, as has been observed for the LCOs of *R. meliloti* and *R. l.* bv. *viciae* (Spaink *et al.* 1991; Schultze *et al.* 1992; Demont *et al.* 1993). However, the compounds that

are responsible for these absorption maxima are more hydrophobic than the C18:4-containing LCOs of *R. l.* bv. *viciae*, since they all have much longer retention times than that of NodRlv-IV(C18:4,Ac). LCOs that have an absorption maximum at 330 nm have not been previously observed. This absorption maximum could be explained by the presence of four conjugated *trans* double bonds. Another major difference from the situation observed in *R. l.* bv. *viciae* is that only very minor quantities of the compounds having these characteristic absorption maxima are produced, as estimated by the UV absorbance profiles at 200 nm. After treatment of the sample with chitinase, followed by re-extraction, the intensity of the characteristic UV-absorbing peaks was greatly reduced and instead new, more hydrophobic compounds with the characteristic UV absorption maxima could be eluted with 70% acetonitrile/water (v/v) (data not shown). These results indicate that these specific peaks are chitinase sensitive and therefore related to chitin.

The HPLC elution profile of the sample derived from *R. l.* bv. *trifolii* strain LPR5045.pMP2111 (*nodE*::Tn5) (Fig. 3, panel C) monitored at 200 nm looks very similar to that of the wild-type control. This is consistent with the results of the radiolabeling experiments (Fig. 2A). However, this strain does not produce the compounds that are responsible for the characteristic absorption maxima observed in the sample from the wild type. In conclusion, the production of these compounds is *nodE*-dependent. The very low level of production of the *nodE*-dependent metabolites and their co-migration with *nodE*-independent metabolites (Fig. 3, panel D) explain why no effect of the *nodE* mutation was observed in the radiolabeling studies.

Structural analysis of common LCOs.

Fractions from the final HPLC purification step (Fig. 3) were analyzed using FAB-MS. In fractions 18 and 24 of the three strains analyzed, which all gave a strong UV absorption

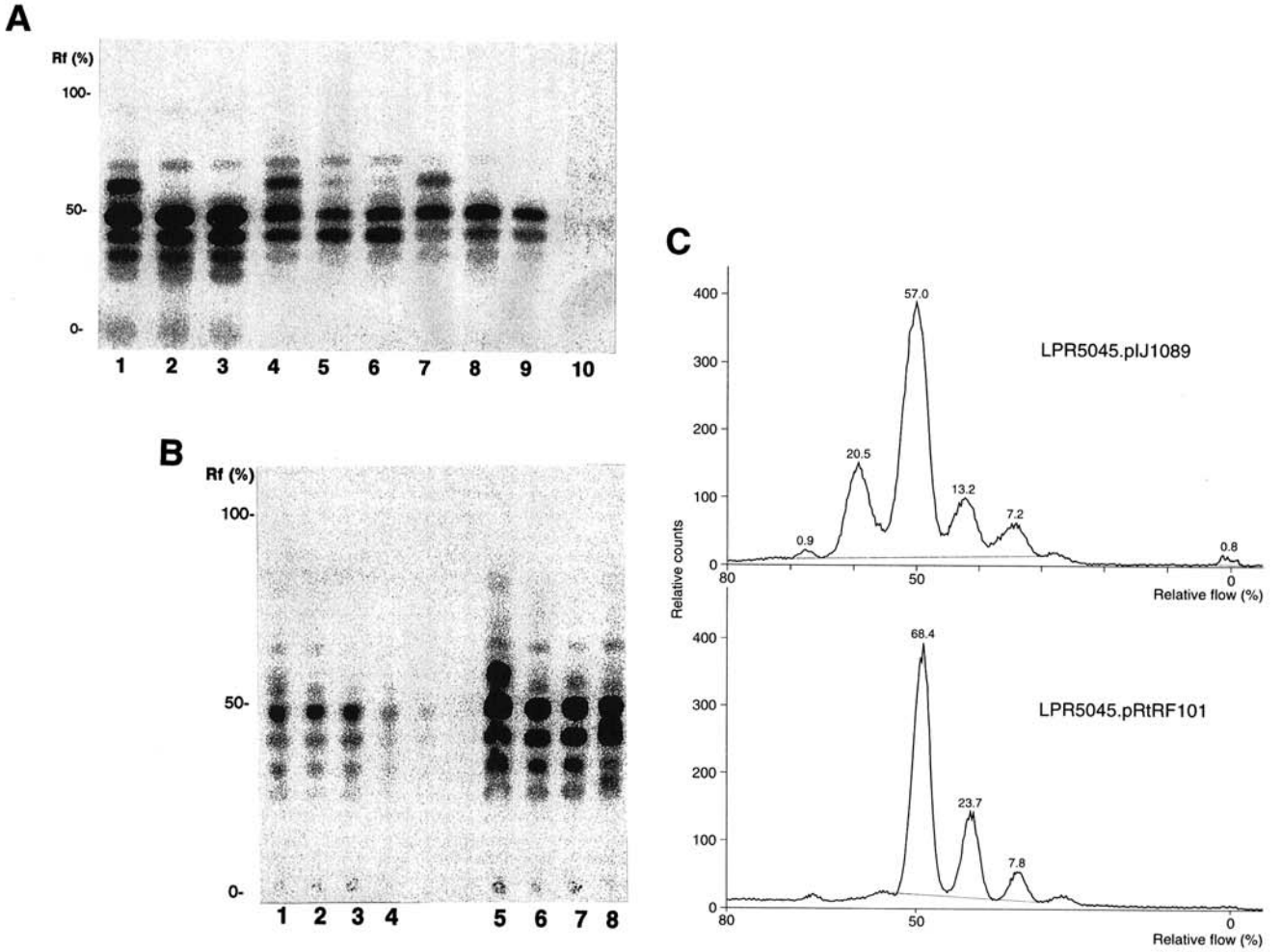

Fig. 2. Thin-layer chromatography analysis of lipo-chitin oligosaccharide molecules. **A,** Analysis of *n*-butanol extracts of cell pellets (lanes 1–3), spent growth medium (lanes 4–6) and wash of the cell pellets (lanes 7–9) of the *R. l.* bv. *viciae* strain LPR5045.pIJ1089 (lanes 1, 4, and 7), *R. l.* bv. *trifolii* strain LPR5045.pRtRF101 (lanes 2, 5, and 8) and *nodE*-mutant strain LPR5045.pMP2111 (lanes 3, 6, and 9) grown in the presence of the inducer naringenin. In lane 10 the *n*-butanol extract of the spent growth medium of strain LPR5045.pRtRF101 grown in the absence of naringenin was applied. **B,** Analysis of *n*-butanol extracts of total cell cultures of strains LPR5045.pIJ1089 (lanes 1 and 5), LPR5045.pRtRF101 (lanes 2 and 6), LPR5045.pMP2111 (lanes 3 and 7) and LPR5045.pRI4003 (lanes 4 and 8) grown at 28° C (lanes 1–4) or 18° C (lanes 5–8). **C,** Quantitation of the radioactivity present in lanes 5 (strain LPR5045.pIJ1089) and lane 6 (strain LPR5045.pRtRF101) of panel B. The values obtained by integration of the peak areas are indicated. The output was generated using the ImageQuant software.

at 200 nm (Fig. 3), $[M+H]^+$ pseudomolecular ions were observed at *m/z* 1,298 and 1,095, respectively. These ions correspond to LCOs containing 5 and 4 HexNAc residues, bearing one acetyl group and a C18:1 fatty acid moiety, as has been reported for the wild-type *R. l.* bv. *viciae* strain (Spaink *et al.* 1991). In fractions A11 and A16 (strain LPR5045.pIJ1089; Fig. 3, panel A) $[M+H]^+$ pseudomolecular ions were observed at *m/z* 1,292 and 1,089, respectively, corresponding to the previously described NodRlv-V(C18:4,Ac) and NodRlv-IV(C18:4,Ac) molecules (Spaink *et al.* 1991). These pseudomolecular ions were not present in mass spectra of the equivalent fractions from the other two strains (Fig. 3, panels B and C). Using fractions B18 and B24 (from strain LPR5045.pRtRF101; Fig. 3, panel B) we obtained CID mass spectra on collision with helium of the $[M+H]^+$ pseudomolecular ions at *m/z* 1,298 and 1,095, respectively. The spectra contain intense A^+-type fragment ions, formed by sequential cleavage of the glycosidic linkages, with charge retention on the nonreducing portion of the molecule (oxonium ions). The *m/z* values of these ions (468, 671, 874, and 1,077, from the *m/z* 1,298 precursor; 468, 671, and 874 from *m/z* 1,095) and the mass intervals between them allow the glycan sequence and site of fatty acid and acetyl substitution to be determined for each of the LCOs (Spaink *et al.* 1991). The data are consistent with a homologous set of LCOs consisting of a linear backbone of 5 and 4 HexNAc residues, and con-

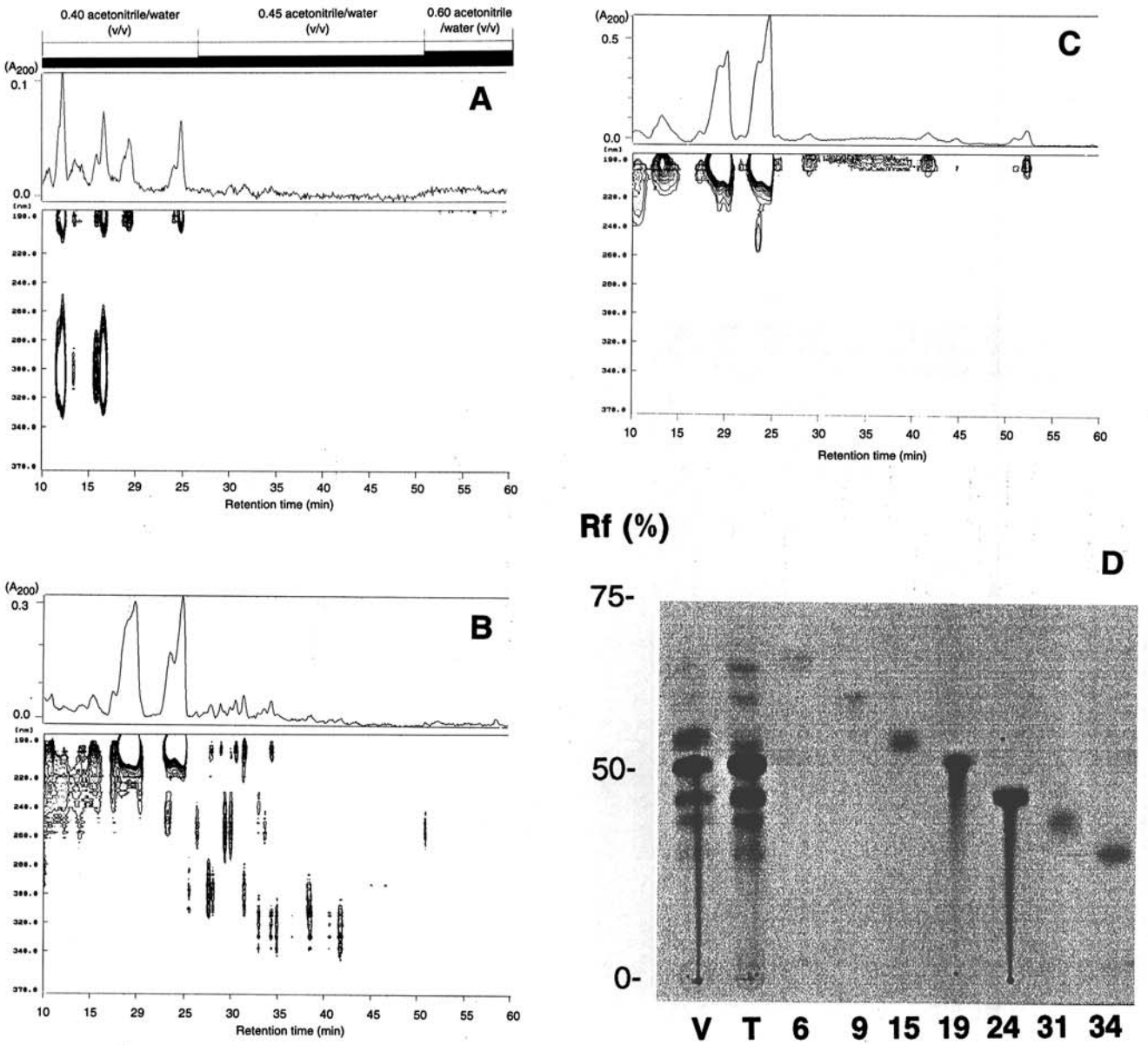

Fig. 3. High-performance liquid chromatography analysis of lipo-chitin oligosaccharide molecules. **A–C,** is shown the photodiode array detection of reverse-phase HPLC elutions with the gradient indicated on top of the figure. LCO extracts were obtained from strains LPR5045.pIJ1089 **(A)** (LPR5045.pRtRF101 **(B)** and LPR5045.pMP2111 **(C).** On top of each panel is shown the plot of the absorbance at 200 nm. HPLC fractions were taken at each minute. The applied LCO extracts were obtained from 25 ml (A) or 100 ml (B and C) culture volumes. In **D** is shown the TLC analysis of peak radioactivity fractions from the HPLC elution shown in panel B. One tenth of the fraction numbers indicated was analyzed. Also a fraction (1/100th) of the total extracts from strains LPR5045.pIJ1089 (lane V) and LPR5045.pRtRF101 (lane T) was analyzed.

taining an acetyl group and a C18:1 fatty acyl moiety substituted on the nonreducing terminal residue. The nature of the fatty acid and acetyl linkages was established following mild base treatment of the pentasaccharide species (Fig. 3, fraction B18), under conditions that hydrolyze ester but not amide bonds. FAB-MS and CID MS analyses of the product revealed a reduction of 42 amu in the *m/z* value of the pseudomolecular ion (*m/z* 1,256) and all the fragment ions (*m/z* 426, 629, 832, and 1,035), indicating removal of an ester-linked acetyl moiety from the nonreducing terminal residue with retention of the amide-bound fatty acyl chain. Similar analyses of the product of palladium-catalyzed reduction resulted in a mass increase of 2 amu in both the $[M+H]^+$ pseudomolecular ion (*m/z* 1,300) and the fragment ions generated from it (*m/z* 470, 673, 876, and 1,079), consistent with the reduction of a single double bond in the fatty acyl chain by incorporation of two atoms of hydrogen.

The position of the double bond in the fatty acid moiety from the LCO in fraction B18 was identified following conversion of the double bond via its epoxide, to a diol from which the di-TMS derivative was formed (Carlson *et al.* 1993). The electron impact mass spectrum contained diagnostic ions at *m/z* 187 and *m/z* 287, arising by fragmentation of the acyl chain between the carbon atoms bearing the TMS groups. These data correspond to a derivative of vaccenic acid (C18:1, Δ^{11}).

Samples derived from larger cultures of strain LPR5045.pRtRF101 that are equivalent to fractions B18 and B24 were analyzed by ^{1}H NMR. The 2D-COSY spectra were identical to those of the reference samples NodRlv-V(C18:1,Ac) and NodRlv-IV(C18:1,Ac), described previously (Spaink *et al.* 1991) (data not shown). These data confirm that the location of the *O*-acetyl moiety and the acyl group are identical to those in the reference compounds, and show that the sugar residues are joined in ß-1,4 -linkage.

In conclusion, the previously reported *nodE*-independent molecules NodRlv-V(C18:1,Ac) and NodRlv-IV(C18:1,Ac) appear to be common for both *R. l.* bvs. *viciae* and *trifolii*.

Structural analysis of minor *nodE*-independent LCOs.

Fractions A29-31, B29-31, and C29-31 (Fig. 3), which correspond to the spots at Rf 42% on TLC (Fig. 2), were analyzed using FAB-MS and CID MS. FAB mass spectra obtained for all three fractions contained an $[M+H]^+$ pseudomolecular ion at *m/z* 1,300. The CID mass spectrum obtained on collision of the *m/z* 1,300 species from fraction B31 contained A^+-type fragment ions at *m/z* 470, 673, 876, and 1,079, which correspond to a linear $HexNAc_5$-containing LCO bearing a C18:0 fatty acyl moiety and an acetyl group substituted on its nonreducing terminal residue. This absence of the *cis* Δ^{11} double bond found in vaccenic acid is consistent with the very different chromatographic mobility on ODS columns of this species compared with that of NodRlv-V(C18:1, Ac) (Carlson *et al.* 1993).

FAB-MS analyses of fractions A33, B33, and C33, corresponding to spots at Rf 30% on TLC, yielded spectra containing an $[M+H]^+$ pseudomolecular ion at *m/z* 1,097. CID mass spectrometric analysis of this ion in fraction B33 produced spectra containing A^+-type fragment ions at *m/z* 470, 673, and 876, indicating that these fractions contain a $HexNAc_4$-LCO that is substituted on its nonreducing terminal residue by a C18:0 fatty acyl group and an acetyl moiety.

Fraction pools A1-10, B1-10, and C1-10, corresponding to the Rf 70–80% region of the TLC, contained naringenin as a major component. To remove the naringenin these pooled fractions were further purified on a silica column as described (Spaink *et al.* 1991). FAB mass spectra obtained from these fractions contained $[M+H]^+$ pseudomolecular ions at *m/z* 1,270, 1,242, 1,067, and 1,039. These data, together with the chromatographic mobilities, suggest the presence of $GlcNAc_5$ and $GlcNAc_4$-containing LCOs substituted with an acetyl moiety and either a C16:1 or a C14:1 fatty acyl chain (see also Carlson *et al.* 1993). The FAB mass spectra of fractions A11 and B11 contained $[M+H]^+$ pseudomolecular ions at *m/z* 1,272, indicating the presence of the $HexNAc_5$ species substituted with a C16:0 acyl chain.

In addition to the $[M+H]^+$ ions described above, in several fractions of the *R. l.* bv. *trifolii* strains LPR5045.pRtRf101 and LPR5045.pMP2111 several other ions were observed in the mass spectra that were not observed from *R. l.* bv. *viciae* strain LPR5045.pIJ1089. It was possible to obtain CID mass spectra from some of these ions (Table 1). For example, the FAB mass spectra obtained from fractions B29 and C29 contained an $[M+H]^+$ pseudomolecular ion at *m/z* 892, which was not observed in fraction A29. The CID mass spectrum obtained on collision of this ion contained A^+-type fragment ions at *m/z* 468 and 671. These data indicate that fractions B29 and C29 contain a $HexNAc_3$-LCO bearing a C18:1 fatty acyl group and an acetyl moiety on its nonreducing terminal residue. The FAB mass spectra obtained from fractions B27 and C27 contained an $[M+H]^+$ pseudomolecular ion at *m/z* 1,113, which was not observed in fraction A27. The CID mass spectrum obtained on collision of this ion contained A^+-type fragment ions at *m/z* 892, 689, and 486. These data suggest the presence of a $HexNAc_4$-LCO bearing a C18:0 hydroxy fatty acid and an acetyl moiety. This assignment is supported by the observation that this compound was not susceptible to palladium-catalyzed reduction with hydrogen (data not shown). An alternative, but very unlikely, possibility would be an LCO molecule, the nonreducing saccharide of which contains two carbamoyl substituents and a C16:0 *N*-acyl group.

Structural analysis of *nodE*-dependent metabolites.

In the FAB mass spectra obtained from many fractions of the HPLC eluate of the *R. l.* bv. *trifolii* strain LPR5045.pRtRF101 (Fig. 3) $[M+H]^+$ pseudomolecular ions were observed that were not present in the spectra of equivalent fractions from the *nodE*-mutant strain LPR5045.pMP2111 and the *R. l.* bv. *viciae* strain LPR5045.pIJ1089. These ions have *m/z* values of 1,091, 1,093, 1,117, 1,119, 1,121, 1,320, and 1,322, which are presumed to correspond to *nodE*-dependent LCOs (see below). When sample quantities allowed it, CID MS analyses of these ions were carried out, resulting in the identification of several novel LCOs (Table 1). In all cases the pseudomolecular ions were shown to correspond to LCOs that varied only in the nature of the fatty acyl moiety, the length of the glycan backbone, or both. Based on their CID MS fragmentation patterns, the fatty acyl moieties of the *nodE*-dependent LCOs listed in Table 1 are assigned as C18:3, C20:2, C20:3, and C20:4. Several ions (e.g., the ion at *m/z* 1,117, see Table 1) that yielded identical fragmentation patterns on CID MS

appeared at several very different retention times on HPLC; this suggests that these species differ only in bearing isomeric fatty acyl moieties.

The positive ion mass spectrum (data not shown) of fraction B27 contained $[M+H]^+$ pseudomolecular ions at *m/z* 1,113 and 1,091. Palladium-catalyzed reduction of the LCOs in this fraction resulted in the disappearance of the ion at *m/z* 1,091 and the appearence of an ion at *m/z* 1,097. The ion at *m/z* 1,113, which was also detected in fraction C27 of strain LPR5045.pMP2111, was not shifted on palladium-catalyzed reduction. These data indicate that fraction B27 contains the LCOs IV(C18:3, Ac) and IV(C18:0-OH, Ac). The former is an obvious candidate for causing the characteristic absorption maximum at 303 nm (Fig. 3B). The lack of a Δ^{11} *cis* double bond in the fatty acyl moiety of this IV(C18:3, Ac) LCO would explain its much longer retention time on reversed phase HPLC compared with that of the IV(C18:4, Ac) LCO of *R. l.* bv. *viciae*.

LCOs derived from a large-scale 200-L culture, and corresponding approximately to fractions 40–44 of Fig. 3 (panel B) were isolated and subjected to detailed structural analysis. Positive ion FAB-MS analysis (data not shown) of this sample yielded pseudomolecular ions at *m/z* 1,117 ($[M+H]^+$ for IV(C20:4, Ac)), *m/z* 1,225 (corresponding to its thioglycerol adduct), *m/z* 1,119 ($[M+H]^+$ for IV(C20:3, Ac) and *m/z* 1,257 (corresponding to its thioglycerol adduct). The CID mass spectrum of parent ion *m/z* 1,117 is shown in Figure 4A. The fragment ions at *m/z* 490, 693, and 896 arise as shown in the fragmentation scheme. The CID mass spectrum of parent ion *m/z* 1,119 (data not shown) contained analogous fragment ions at *m/z* 492, 695, and 898. These data are consistent with LCOs composed of a linear HexNAc tetramer, the nonreducing residue of which is substituted with a C20:4, or a C20:3, fatty acid moiety and an acetyl group. An aliquot of the sample was submitted to palladium-catalyzed reduction with hydrogen to confirm the degree of unsaturation of the fatty acid moieties. The mass spectrum (data not shown) of the product contained a single $[M+H]^+$ pseudomolecular ion at *m/z* 1,125. This indicates a mass increase of 8 amu for the ion at *m/z*

Table 1. Summary of lipo-chitin oligosaccharide molecules from *Rhizobium leguminosarum* biovar *trifolii* characterized by CID MS

HPLC fraction	$[M+H]^+$	A^+-Type fragment ions	Assignment
R. l. biovar *trifolii*[a]			
B18	1298	468, 671, 874, 1077	V(C18:1,Ac)
B24	1095	468, 671, 874	IV(C18:1,Ac)
B27	1113	486, 689, 892	IV(C18:0-OH, Ac)
	1091	464, 667, 870	IV(C18:3, Ac)
B29	892	468, 671	III(C18:1,Ac)
B31	1300	470, 673, 876, 1079	V(C18:0,Ac)
B33	1097	470, 673, 876	IV(C18:0,Ac)
B35	**1121**	494, 697, 900	IV(C20:2,Ac)
	1320	490, 693, 896, 1099	V(C20:4,Ac)
B38	1123	496, 699, 902	IV(C20:1, Ac)
	1322	492, 695, 898, 1101	V(C20:3,Ac)
	1117	490, 693, 896	IV(C20:4,Ac)
B41	894	470, 673	III(C18:0,Ac)
B42	**1117**	490, 693, 896	IV(C20:4,Ac)
B43	**1117**	490, 693, 896	IV(C20:4,Ac)
	1119	492, 695, 898	IV(C20:3,Ac)

[a] Ions indicated in boldface were not found in equivalent fractions of LPR5045.pMP2111.

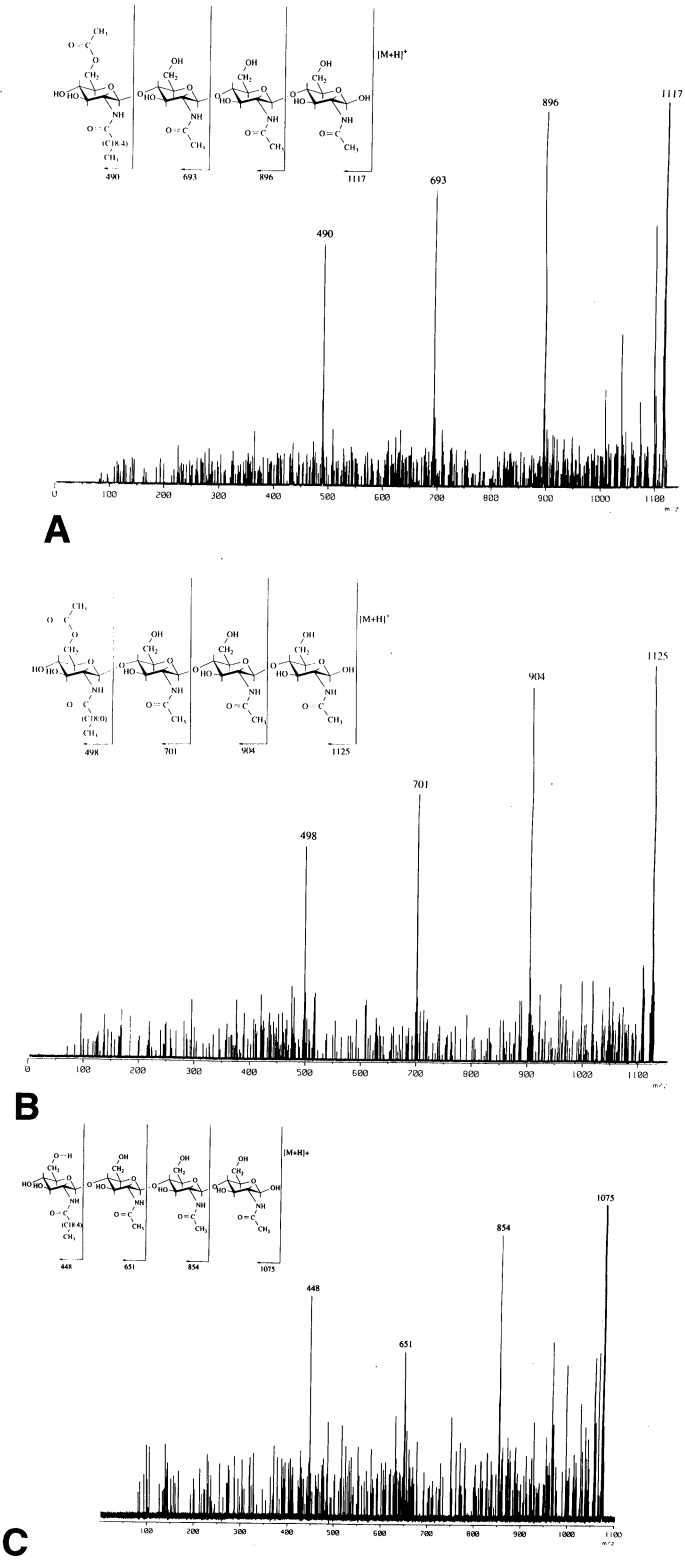

Fig. 4. Collision-induced dissociation tandem mass spectrometry of a *nodE*-dependent lipo-chitin oligosaccharide and its derivatives from strain *Rhizobium leguminosarum* bv. *trifolii* strain LPR5045.pRtRF101. **A,** Spectrum and fragmentation scheme for the $[M+H]^+$ ion at *m/z* 1,117. **B,** Spectrum and fragmentation scheme for the $[M+H]^+$ ion at *m/z* 1,125 following palladium-catalyzed reduction. **C,** Spectrum and fragmentation scheme for the $[M+H]^+$ ion at *m/z* 1,075 following mild base treatment. The poor signal to noise ratio in the latter spectrum is a result of the very small amount of sample available for this experiment.

1,117 and 6 amu for that at *m/z* 1,119, corresponding to the reduction of four or three double bonds, respectively. The CID mass spectrum (Fig. 4B) of *m/z* 1,125 contained fragment ions at *m/z* 498, 701, and 904, demonstrating that the reduced fatty acid moiety is attached to the nonreducing sugar residue. The mass spectrum (data not shown) of the product of mild base treatment contained $[M+H]^+$ pseudomolecular ions at *m/z* 1,075 and 1,077. The CID mass spectrum of the former (Fig. 4C) contained fragment ions at *m/z* 448, 651, and 854, while that of the latter gave ions at *m/z* 450, 653, and 856. Since the parent ions and all fragment ions were shifted by 42 amu, this indicates that the only ester-linked group was an acetyl group that was substituted on the nonreducing residue in both compounds. The proposed structures were confirmed by the analysis of the products of permethylation. The mass spectrum obtained (data not shown) contained $[M+H]^+$ pseudomolecular ions at *m/z* 1,271 and *m/z* 1,273 and CID mass spectra (data not shown) contained fragment ions at *m/z* 504, 749, and 994, from the parent at *m/z* 1271, and at *m/z* 506, 751, and 996 from the parent at *m/z* 1,273, indicating the incorporation of the expected number of methyl groups.

The identity of the HexNAc residues was determined following methanolysis and TMS-derivatization. The resulting monosaccharide derivatives were analyzed using GC-MS and were shown to co-elute with authentic GlcNAc standards. Together with the sensitivity to chitinase, these results demonstrate that the backbone of the lipo-oligosaccharides is composed of β-1→4 linked GlcNAc residues. Taken together, these data demonstrate the presence of NodRlt-IV(C20:4, Ac) and NodRlt-IV(C20:3, Ac) LCOs in this sample. The IV(C20:4, Ac) LCO present in fraction B42 is the only candidate for causing the characteristic absorption maximum at 330 nm at retention time 42 min in Figure 3B. The most likely explanation for the 330 nm absorption maximum is that the fatty acyl moiety of the IV(C20:4, Ac) LCO contains four *trans*-conjugated double bonds. A further confirmation of the fatty acyl structures by NMR analysis was not possible due to the very low yield of the *nodE*-dependent metabolites. Two large-scale fermentations (200-L) did not yield sufficient material for 2D-proton NMR analysis.

DISCUSSION

The *nod* genes of *R. l.* bvs. *trifolii* and *viciae* strains determine the host range of nodulation (Fig. 1; Spaink *et al.* 1989). We have analyzed the structures of the major and minor LCOs produced by two *R. leguminosarum* (*R. l.*) strains differing only in the origin of the *nod* genes. The results show that both strains produce complex mixtures of LCOs. The complexity of the mixtures and the physical similarity of the components make complete purification in several cases technically impossible. However, using CID tandem mass spectrometry together with specific chemical modifications, the structures of a number of the components in the mixtures could be determined without the need for complete fractionation. In this way all of the major and most of the minor LCOs from both strains could be identified. The results from these analyses are summarized in Table 1.

The mixtures of LCOs produced by the two biovars are similar, with the components differing only in their glycan chain length and the nature of the fatty acyl moieties. Without exception, an *O*-acetyl substituent is present on the nonreducing terminal GlcNAc residue. Recently, it has been shown for *R. l.* bv. *viciae* that the presence of this *O*-acetyl group is determined by the NodL protein, which functions as an acetyl transferase (Bloemberg *et al.* 1994). Indeed, *R. l.* bv. *trifolii* also contains a *nodL* gene, located in the *nodFERL* operon (Djordjevic and Weinman 1991), which can produce a functional NodL protein (G. V. Bloemberg, unpublished results).

In *R. l.* bv. *viciae* the length of the glycan backbone is confined to five or four GlcNAc residues. In addition in the *R. l.* bv. *trifolii* strain two minor LCOs are produced that consist of three GlcNAc residues (Table 1). An LCO containing three GlcNac residues has also been described for *R. meliloti* (Schultze *et al.* 1992).

The two *R. l.* bvs. *viciae* and *trifolii* resemble each other in that, independently of the presence of the *nodE* gene, LCOs are produced that contain fatty acyl moieties that commonly occur as components of the phospholipids, such as vaccenoyl (C18:1), stearyl (C18:0), palmitoyl (C16:1), and palmityl (C16:0) moieties. In addition, *nodE*-dependent fatty acyl moieties are found that have multiple double bonds. The two biovars appear to differ in the nature of these fatty acyl groups and also in the relative quantities of LCOs containing them. Whereas in *R. l.* bv. *viciae* a single *nodE*-dependent C18:4 fatty acyl group is present in a major class of LCOs, *R. l.* bv. *trifolii* produces a very large variety of minor classes of *nodE*-dependent LCOs. The fatty acyl moieties of the latter LCOs have in common with the C18:4 of *R. l.* bv. *viciae* that they contain conjugated double bonds, as judged by their UV-absorption spectra. However, the number of these conjugated double bonds appears to be variable in the case of *R. l.* bv. *trifolii,* resulting in UV absorbance maxima at 260, 303, and 330 nm. Furthermore, comparison of the mass spectrometric data and the UV absorption maxima suggests that the *cis* double bond, present in the C18:4 of *R. l.* bv. *viciae,* is absent in several of the *nodE*-dependent LCOs of *R. l.* bv. *trifolii.* The major distinguishing feature of the *nodE*-dependent LCOs produced by the two biovars is that the fatty acyl substituents always differ in their hydrophobicity as judged by their affinity for the hydrophobic matrices of TLC plates and HPLC columns. The *nodE*-dependent fatty acyl moieties of *R. l.* bv. *trifolii* seem to be more hydrophobic, a result of either a longer carbon chain or the lack of a *cis* double bond. Interestingly, the observation that *R. l.* bv. *trifolii,* in contrast to bv. *viciae,* produces LCOs containing only three GlcNAc residues, could indicate that the former biovar uses this as a second route to producing more hydrophobic LCOs.

Since LCOs seem to be the major mediators of the host specificity of nodulation (Lerouge *et al.* 1990; Spaink *et al.* 1991; Truchet *et al.* 1991) and considering the crucial role of the *nodE* gene in the nodulation of various clover species (Spaink *et al.* 1989; Fig. 1), it is to be expected that the *nodE*-dependent LCOs play an important role in the nodulation process. Our data demonstrate that the major difference between the LCOs from *R. l.* bvs. *viciae* and *trifolii* lies in the hydrophobicity of the *nodE*-dependent fatty acyl moieties, thus indicating that the host specificity of these two biovars is determined by the hydrophobicity of the highly unsaturated fatty acyl moieties of the LCOs. However, the biological significance of the *nodE*-dependent LCOs of *R. l.* bv. *trifolii* is not yet clear. Many attempts have been made to develop a test

system to measure host specific effects of LCOs on clover plants. However, until now, it has not been possible to show the induction of nodule primordium formation by *R. l.* bv. *trifolii* LCOs (unpublished results). This could be due to the fact that nodule primordia-like structures can also be formed on various clover species in the absence of LCOs (McIver *et al.* 1993). Due to the limited production of *nodE*-dependent LCOs, it is also possible that the amounts of the specific signal molecules in the bioassays could have been too low to cause an observable effect. Our identification of various *nodE*-dependent LCOs makes it possible to undertake the chemical synthesis of such molecules in large quantities as has been published for the LCOs of *R. meliloti* (Ikeshita *et al.* 1994; Nicolaou *et al.* 1992; Wang *et al.* 1994), which should then allow these questions to be answered.

MATERIALS AND METHODS

Bacterial strains and plasmids.

The *Rhizobium* strain LPR5045, which was used as the chromosomal background in this study, is a Sym plasmid-cured derivative of *R. l.* bv. *trifolii* RBL5020 (Hooykaas *et al.* 1982). Plasmids pIJ1089 and pRtRF101, IncP derivatives that contain all the known *nod* genes from *R. l.* bv. *viciae* strain 248 and *R. l.* bv. *trifolii* strain ANU843 (Rolfe *et al.* 1981), respectively, have been described by Downie *et al.* (1985) and Fisher *et al.* (1985). Plasmid pRI4003 is a derivative of the low copy number IncW plasmid pRI40, which contains all the known *nod* and *nif* genes from *R. l.* bv. *trifolii* strain ANU843 (Innes *et al.* 1988). Both plasmids pRtRF101 and pRI4003 confer on strain LPR5045 the ability to efficiently nodulate on *T. pratense* (Fig. 1). pMP2111, a derivative of pRtRF101 that contains a *Tn5* insertion in the *nodE* gene, was constructed by homogenotization of the transposon mutant K11 (Djordjevic *et al.* 1985) and backcrossing to *Escherichia coli* strain JM109 as described previously (Canter Cremers *et al.* 1989). pMP2111 was isolated and restriction analysis showed that the plasmid contains the *nodE11::Tn5* mutation. The kinetics of nodulation on *T. pratense* by strain LPR5045.pMP2111 were similar to those observed with strain K11 (Fig. 1). Nodulation tests were performed as described previously (van Brussel *et al.* 1982).

Bacterial growth conditions and purification of lipo-chitin oligosaccharides.

Bacteria were grown in B^- medium containing tetracycline (2 mg L^{-1}) as described previously (Spaink *et al.* 1989). For induction, naringenin (Sigma Chemical Co.. St. Louis, MO) was added to a concentration of 2 μM. LCO molecules were extracted from the cell cultures with 0.2 volume of *n*-butanol and dried under vacuum. For large-scale purification, samples were redissolved overnight in acetonitrile/water (60:40, v/v) by vigorous shaking. In contrast to the method we have previously used for purification of LCOs (Spaink *et al.* 1991) no chromatography on silica columns was performed, unless indicated otherwise, to allow the isolation of more hydrophobic LCO species. Instead, samples were passed through octadecyl silica (ODS) columns (1 mg of sorbent; J. T. Baker n.v., Deventer, The Netherlands) and subsequently used for high-performance liquid chromatography (HPLC) using a Pep-S column (5 μm, 5 × 250 mm; Pharmacia LKB Biotechn. Inc.). Columns were eluted with a stepwise gradient of mixtures of acetonitrile and water (highest grade, both obtained from Baker) at a flow of 1 ml/min, and monitored with an RSD 2140 photodiode array detector (Pharmacia LKB Biotechn. Inc.). Samples were stored under argon. Care was taken to avoid contact of the samples with plastics for which LCO molecules have a strong affinity. Chitinase treatment was performed by incubating LCO samples for 24 hr in aqueous solution (pH = 6.0) with chitinase (0.01 U) from *Streptomyces griseus* (Sigma Chemical Co., St. Lous, MO.). After digestion LCOs were extracted with *n*-butanol, dried under vacuum and dissolved in acetonitrile/water (60:40, v/v).

Radioactive labeling and TLC analysis.

For radioactive labeling of LCOs, 1 ml bacterial cultures were grown for 8 hr in the presence of 0.4 μCi 1-^{14}C-labeled glucosamine (specific activity 50 mCi.mmol^{-1}, Amersham). After centrifugation for 30 min at 15,000 × *g* cells were washed with 1 ml of B^- medium and again centrifuged. The pellet was resuspended in 1 ml of B^- medium. The three resulting fractions were extracted with 200 μl of *n*-butanol and, after drying, redissolved in 40 μl of water-saturated *n*-butanol. TLC analysis was performed as described previously (Spaink *et al.* 1991) using ODS plates (Sigma) and acetonitrile/water (50:50, v/v) as the solvent. Radioactivity was detected using a phosphorimaging system of Molecular Dynamics Comp. using the ImageQuant software.

NMR analysis.

^{1}H NMR spectra were recorded on a Bruker MSL.400 spectrometer (equipped with an Aspect 3000 data-system). The thoroughly dried samples were dissolved in D_2O. A description of the 2D-COSY experiments is given by Aue *et al.* (1976).

Reduction.

Samples were dissolved in 500 μl of 2-butanol saturated with water and nitrogen was bubbled through the solution for 30 min. Ten milligrams of 10% palladium/carbon catalyst was added and hydrogen was bubbled through the solution, which was gently stirred for 4 hr. The organic solvent was removed after centrifugation and the catalyst was washed three times with 100 μl of dimethyl sulfoxide (DMSO). The solvent and washings were pooled and brought onto a mini ODS-column (Alltech) primed with 5 ml each of methanol, acetonitrile/water (30:70, v/v), 2-butanol, and DMSO, and the reduced LCO eluted with 3 ml of acetonitrile/water (75:25, v/v). The LCO-containing fraction was dried under vacuum and the residue was redissolved in 20 μl of DMSO prior to FAB-MS analysis.

De-*O*-acetylation.

Samples were de-*O*-acetylated in 100 μl of 1:1 methanol/25% NH_4OH (18 hr, room temperature) and subsequently the mixture was dried under nitrogen. The residue was redissolved in 10 μl of DMSO prior to FAB-MS analysis.

Permethylation.

Three hundred milligrams of freshly ground NaOH was added to samples, which were dissolved in 200 μl of DMSO. 250 μl of methyl iodide was added, after 10 min a further 250

µl was added, and after a further 20 min, 500 µl was added. The reaction was stopped after 20 min by addition of 1 ml of sodium thiosulphate solution (≈100 mg ml^{-1}) and 1 ml of chloroform. The two-phase system was mixed and after centrifugation the aqueous layer was removed. The organic phase was washed three times with 1 ml of water, and evaporated to dryness under nitrogen. The residue was redissolved in 20 µl of DMSO prior to FAB-MS analysis.

Composition analysis.

The de-*O*-acetylated sample was dried after the removal of 3–5 µl for mass spectrometric analysis. Two hundred microliters of 1M HCl in dry MeOH was added (4 hr, 80° C) and, after cooling, 50 µl of pyridine and 50 µl of acetic anhydride were added. After 30 min the sample was dried under vacuum, the residue was redissolved in 250 µl of MeOH, and the solvent was evaporated. The residue was dissolved in 40 µl of trimethylsilyl (TMS) reagent (pyridine/bis(trimethylsilyl) trifluoroacetamide/trimethylchlorosilane, 10:5:1, v/v/v) and was allowed to react for 10–15 min at room temperature when an aliquot of 1 µl was injected onto the GC column. Standards containing approximately 1 mg of *N*-acetylglucosamine (GlcNAc), *N*-acetylmannosamine, and *N*-acetylgalactosamine were prepared in the same way, and were dissolved in 500 µl of TMS reagent.

Fatty acid analysis.

The fatty acid was released from the *nodE*-independent Nod metabolite from *R. l.* bv. *trifolii* in 100 µl of 5% KOH/MeOH (18 hr, 70° C). Twenty microliters of 3.5 M acetic acid was added followed by 200 µl of diethyl ether and the solution was washed twice with 200 µl of water. After evaporating the solvent under nitrogen, 0.5 ml of diazomethane in diethyl ether was added and the reaction was allowed to proceed at room temperature for 20 min. The solvent was evaporated and 10 µl of *m*-chloroperbenzoic acid (0.26 M in CH_2Cl_2) and 60 µl of CH_2Cl_2 was added. After 10 min of sonication, the mixture was allowed to react in the dark (18 hr, room temperature). The solvent was evaporated and the residue was redissolved in 50 µl of toluene. Five to ten beads of washed and activated Amberlyst 15 resin were added and the mixture was sonicated for 30 min. The solvent was removed and the resin was washed three times with 300 µl of *n*-hexane and three times with 300 µl of diethyl ether, each with 1 min sonication. The pooled organic fractions were dried under vacuum and 50 µl of TMS reagent was added. The mixture was allowed to react for 10–15 min at room temperature when an aliquot of 1 µl was injected onto the GC column.

FAB-MS and CID MS analyses.

FAB mass spectra were obtained in the positive ion mode using MS1 of a JEOL JMS-SX/SX102A tandem mass spectrometer using 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 sec 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. Collision induced dissociation tandem mass spectra (CID 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 were loaded into a matrix of thioglycerol (1–2 µl).

GC-MS analysis.

The samples were analyzed using a JEOL JMS-AX505W mass spectrometer fitted with a Hewlett Packard 5890 gas chromatograph using an on-column injector and helium as the carrier gas at a flow rate of 2 ml min^{-1}. The fatty acid derivatives were separated on a BP-1 column (0.32 mm × 25 m, SGE) using the following temperature program: 120° C for 2 min, a gradient of 40 C min^{-1} to 180° C, holding for 4 min at 180° C, then a gradient of 4° C min^{-1} to 280° C and finally holding the temperature at 280° C for 10 min. For the composition analysis the following temperature program was used: 110° C for 2 min, a gradient of 40° C min^{-1} to 160° C, holding for 2 min at 160° C, then a gradient of 4° C min^{-1} to 240° C and finally holding the temperature at 240° C for 10 min. Mass spectra were obtained under conditions of electron impact and were recorded using linear scanning from *m/z* 35–600 at an accelerating voltage of 3 kV.

Lipo-chitin oligosaccharide nomenclature.

Nomenclature of the LCOs is according to Spaink *et al.* (1991). Briefly, a Roman numeral following the species indication (e.g., NodRlv for *R. l.* bv. *viciae*) refers to the number of glucosamine units and a term in parentheses indicates fatty acyl carbon number and degree of unsaturation and *O*-acetylation.

ACKNOWLEDGMENTS

We are grateful to A. Aarts and P. van Dillewijn for their help in the initial stages of this research and K. Erkelens (Leiden University) for his assistance with the NMR analysis. This work was supported in part by The Netherlands Foundation for Chemical Research (with financial aid from the Netherlands Organization for Scientific Research), the Royal Netherlands Academy of Arts and Sciences, and a contract from the European Community (BIO2-CT93-0400, DG12 SSMA).

LITERATURE CITED

Aue, W. P., Bartholdi, E., and Ernst, R. R. 1976. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. J. Chem. Phys. 64:2229-2247.

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. Mol. Microbiol. 11:793-804.

Brewin, N. J. 1991. Development of the legume root nodule. Annu. Rev. Cell Biol. 7:191-226.

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 nodulation genes on the Sym plasmid of *Rhizobium leguminosarum* biovar *viciae*. Plant Mol. Biol. 13:163-174.

Carlson, R. W., Sanjuan, J., Ramadas Bhat, U., Glushka, J., Spaink, H. P., Wijfjes, A. H. M., van Brussel, A. A. N., Stokkermans, T. J. W., Peters, K., and Stacey, G. 1993. The structures and biological activities of the lipo-oligosaccharide nodulation signals produced by Type I and Type II strains of *Bradyrhizobium japonicum*. J. Biol. Chem. 268:18372-18381.

Demont, N., Debellé, F., Aurelle, H., Dénarié, J., and Promé, J. C. 1993. Role of the *Rhizobium meliloti nodF* and *nodE* genes in the biosynthesis of lipo-oligosaccharidic nodulation factors. J. Biol. Chem. 268:20134-20142.

Dénarié, J., and Cullimore, J. 1993. Lipo-oligosaccharides nodulation factors: A new class of signaling molecules mediating recognition and morphogenesis. Cell 74:951-954.

Djordjevic, M. A., Schofield, P. R. and Rolfe, B. G. 1985. Tn5 mutagenesis of *R. trifolii* host-specific nodulation genes result in mutants with altered host-range ability. Mol. Gen. Genet. 200:463-471.

Djordjevic, M. A., and Weinman, J. J. 1991. Factors determining host recognition in the clover-*Rhizobium* symbiosis. Aust. J. Plant Physiol. 18:543-557.

Downie, J. A., Knight, C. D., Johnston, A. W. B., and Rossen, L. 1985. Identification of genes and gene products involved in the nodulation of peas by *Rhizobium leguminosarum*. Mol. Gen. Genet. 198:255-262.

Fisher, R. F., and Long, S. R. 1992. *Rhizobium*-plant signal exchange. Nature (London) 357:655-660

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.

Hooykaas, P. J. J., Schijndewindt, F. G. M., and Schilperoort, R. A. 1982. Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to and expression in other Rhizobia and *Agrobacterium tumefaciens*. Plasmid 8:73-82.

Ikeshita, S., Sakamoto, A., Nakahara, Y., Nakahara, Y., and Ogawa, T. 1994. Synthesis of the root nodule-inducing factor NodRm-IV(C16:2,S) of *Rhizobium meliloti* and related compounds. Tetrahedron Lett. 35:3123-3126.

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.

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

McIver, J., Djordjevic, M. A., Weinman, J. J., and Rolfe, B. G. 1993. Influence of *Rhizobium leguminosarum* biovar *trifolii* host specific nodulation genes on the ontogeny of clover nodulation. Protoplasma 172:166-179.

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.

Nicolaou, K. C., Bockovich, N. J., Carcanague, D. R., Hummel, C. W., and Even, L. F. 1992. Total synthesis of the NodRm-IV factors, the Rhizobium nodulation signals. J. Am. Chem. Soc. 114:8701-8702.

Rolfe, B. G., Gresshoff, P. M., and Shine J. 1980. Rapid screening for symbiotic mutants of *Rhizobium* and white clover. Plant Sci. Lett. 19:277-284.

Schlaman, H. R. M., Okker, R. J. H., and Lugtenberg, B. J. J. 1992. Regulation of nodulation gene expression by NodD in rhizobia. J. Bacteriol. 174:5177-5182.

Schultze, M., Quiclet-Sire, B., Kondorosi, E., Virelizier, H., Glushka, J. N., Endre, G., Géro, S. D., and Kondorosi, A. 1992. *Rhizobium meliloti* produces a family of sulfated lipo-oligosaccharides exhibiting different degrees of plant host specificity. Proc. Natl. Acad. Sci. U.S.A. 89:192-196.

Spaink, H. P. 1994. The molecular basis of the host specificity of the *Rhizobium* bacteria. Ant. van Leeuwenhoek 65:81-98.

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., 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 (London) 354:125-130.

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

Spaink, H. P., Wijfjes, A. H. M., van der Drift, K. M. G. M., Haverkamp, J., Thomas-Oates, J. E., and Lugtenberg, B. J. J. 1994. Structural identification of metabolites produced by the NodB and NodC proteins of *Rhizobium leguminosarum*. Mol. Microbiol. 13:821-831.

Truchet, G., Roche, P., Lerouge, P., Vasse, J., Camut, S., De Billy, F., Promé, J.-C., and Dénarié, J. 1991. Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. Nature 351:670-673.

van Brussel, A. A. N., Tak, T., Wetselaar, A., Pees, E., Wijffelman, C. A. 1982. Small leguminosae as test plants for nodulation of *Rhizobium leguminosarum* and other *Rhizobia* and *Agrobacteria* harbouring a leguminosarum plasmid. Plant Sci. Lett. 27:317-325.

Wang, L.-X., Li C., Wang, Q.-W., and Hui, Y.-Z. 1994. Chemical synthesis of NodRm-1:the nodulation factor involved in *Rhizobium meliloti*-legume symbiosis. J. Chem. Soc. Perkin Trans. 1:621-629.