ELSEVIER

International Journal of Mass Spectrometry 219 (2002) 703–716

International Journal of Mass Spectrometry

www.elsevier.com/locate/ijms

The pivotal role of tandem mass spectrometry in structural determinations of Nod factors produced by Rhizobia Nod factors produced by wild-type strains of *Mesorhizobium huakii* and *Rhizobium* sp. *mus10*

Jean Claude Promé[a,*], Myriam Ferro[a,b], Frederic Debellé[c], Danielle Promé[a], Hari B. Krisnan[d]

[a] *Laboratoire de Synthèses et Physicochimie de Molécules d'Intérêt Biologique, UMR-CNRS 5068, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse, France*
[b] *Laboratoire de Chimie des Protéines, DBMS-CP, CEN-G, 38054 Grenoble, France*
[c] *Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, CNRS-INRA 31326, Castanet-Tolosan, France*
[d] *Department of Plant Pathology, University of Missouri, Columbia, MO 65211, USA*

Received 11 March 2002; accepted 13 May 2002

Dedicated to Dr. Yannik Hoppilliard on the occasion of her 60th birthday.

Abstract

Nod factors are signalling molecules secreted by bacteria of the Rhizobium family that plays an important role in initiating the early events of the nitrogen-fixing symbiosis between Rhizobia and legumes. They possess a lipochitooligomeric structure decorated by several functional groups. The structural details are essential for their specific recognition by plants and for switching on a genetic mechanism leading to the formation of nodules. Mass spectrometry is now able to fully characterise these molecules, even within mixtures and on minute amounts. It opens the possibility to study Nod factors from wild-type low-producing strains. It was demonstrated that Nod factors from the wild-type strain of *Mesorhizobium huakii* are almost the same as those produced by a genetically engineered over-producing strain. *Rhizobium* sp. *mus10* is an Indian strain that nodulates the plant *Sesbania*. It was found that several Nod factors produced by this strain are identical to those produced by African strains that nodulate *Sesbania* species. However, two other Nod factors possessed new structural features. This synthetic possibility was probably acquired by the *Rhizobium* sp. *mus10* strain by evolution in a different environment. (Int J Mass Spectrom 219 (2002) 703–716)

Keywords: Nod factors; Lipo-oligosaccharides; Chito-oligomers; Rhizobia; Electrospray; Fast atom bombardment; Tandem mass spectrometry

1. Introduction

Rhizobia are soil bacteria that are able to establish a nitrogen-fixing symbiosis with plants of the legume family. Infection of a legume by the appropriate bacterial symbiont leads to the formation of root or stem nodules in which reduction of atmospheric nitrogen into ammonium ions occurs. A key mechanism that pilots this interaction is bacterial synthesis and secretion in the external medium, close to the root plants,

* Corresponding author. E-mail: djcprom@wanadoo.fr

1387-3806/02/$ – see front matter
PII S1387-3806(02)00748-0

of water-soluble molecules, called Nod factors. The synthesis of Nod factors, is initiated by root plants secretions of flavonoids that activates the transcription of the rhizobial nodulation (*nod*) genes, whose translation products are enzymes involved in the de novo synthesis and in the secretion of Nod factors. It has been said that the main role of Nod factors is to initiate the rhizobium–legume symbiosis by "opening" the legume door for plant infections by their specific Rhizobium [12].

In 1990, Lerouge et al. [1] showed that the products of the *nod* genes of *Sinorhizobium meliloti* are sulphated tetramers of β1,4-linked *N*-acetyl-D-glucosamine (chitooligomers) in which the *N*-acetyl group of the non-reducing terminus is replaced by a variety of *N*-linked fatty acids. The main one is 2*E*,9*Z* hexadecadienoic acid. Later, tri-unsaturated and very long-chain ($\omega-1$)-hydroxylated fatty acids were also found [2,3]. These molecules, called Nod factors, were also named lipochitooligomers (LCOs). The preliminary work was done on an over-producing, genetically modified strain containing multicopies of the *nodABCIJ* genes. As very low amounts of compounds with an additional acetyl group were also detected, it was hypothesised that other *nod* genes play a role in Nod factors synthesis. An other modified strain was constructed by amplifying the number of copies of the regulatory *nod* genes. It was found that this new strain synthesises Nod factors with an additional acetyl group located on the *O*-6 position of the non-reducing end residue [4]. Same structures were detected in LCOs from the wild-type strain.

Other LCOs were found in the culture medium of many Rhizobia. Their synthesis was progressively dissected by determination of the structural variations induced by inactivation or transfer of the *nod* genes. Mass spectrometry played a decisive role in these determinations. These results together with details on biological activities of Nod factors were presented in several reviews [5–12]. One of the main conclusions was that the "decorations" of the chitooligosaccharidic core define the host–plant range of the bacteria.

As said before, over-expression of the *nod* genes may change Nod factor structures by modifying the intracellular metabolism. Reliable structural determinations should be preferably done on wild-type strains even if they produce minute amounts of Nod factors. The sensitivity of mass spectrometry and its ability to work on mixtures permits to work on small amounts of partly purified material, thus avoiding the need of large-scale cultures.

This paper describes the structures of Nod factors from two weak-producing wild-type strains: Nod factors from an over-producing strain of *Mesorhizobium huakii* called *Ra5 (pA28)* were described previously [13]. Part of them exhibited an uncommon structure with glycolylated glucosamine instead of acetylated glucosamine at the reducing end. It may be asked whether Nod factor structures from the wild-type strain are identical. *Rhizobium* sp. *mus10* was isolated from a *Sesbania* plant grown in India. Strains isolated from other *Sesbania* growing in West Africa produced Nod factors with peculiar structures [14,16]. It was to be asked whether a geographically distant strains growing in a very different environment produced similar or different Nod factors.

2. Material and methods

The mass spectrometer was an Autospec 6F from Micromass (Manchester, UK). Different ion sources were used: electron impact and positive-ion ammonia chemical ionisation for sugar and fatty acid analysis, methane negative-ion chemical ionisation for dissociative electron capture analysis of pentafluorobenzyl derivatives of fatty acids, LSIMS and electrospray for Nod factor mixtures. The accelerating voltage was set to 8 kV except for electrospray for which it was reduced to half of its original value.

It was connected to a capillary gas chromatograph fitted with 0.32 i.d., 30-m length OV1 column for GC-MS and GC-MS–MS analysis. The carrier gas was helium. The MIKE scanning mode was used for dissociation studies of carboxylate anions of fatty acids.

It was fitted with a LSIMS-Cs gun source operating at 30 kV for LSIMS ionisation. The matrix was a 1:1 mixture of glycerol–metanitrobenzyl alcohol acidified with 10% trichloroacetic acid.

The electrospray source operated 4 kV higher than the accelerating voltage (4 kV). The cone voltage was set to 50 V. In the negative-ion mode, a diluted solution of Nod factors in 1:1 methanol–water was continuously infused at a 10 μL/min flow rate. In the positive-ion mode, the same solvent contained 1% acetic acid.

Collision-induced spectra were recorded by constant B/E-linked scans using the first half or the full instrument with a collision cell floating at half the accelerating voltage, depending on the intensity of the precursor ion.

Permethylation of $NaBD_4$-reduced Nod factors was performed according to Ciacanu and Derek [17] on few micrograms of material. The permethylated LCO was hydrolysed (2 M trifluoroacetic acid at 110 °C for 2 h), acetylated and analysed by GC-MS. The reducing end afforded a partly methylated and partly acetylated glucosaminitol whose identification was easy by EIMS.

Hydrogenation of LCOs was carried out in 80:20 ethanol–water mixture over platinum oxide for 2 h.

Nod factors were extracted from flavonoid-induced culture medium of rhizobia by butanol. For larger-scale cultures, solid-phase extraction on hydrophobic beads (Amberlite XAD-4) was preferred. Chromatography of the crude extract over C18-bound silica columns both removed most of the inorganic salts, insoluble polymeric material and very hydrophobic compounds that were co-extracted. Nod factors were generally eluted from these columns by 50:50 to 90:10 methanol–water mixtures. They were detected by monitoring the UV absorption of amide groups at 206 nm. Nod factors with carbonyl-conjugated double bonds in their acyl chains were characterised by diode-array detection.

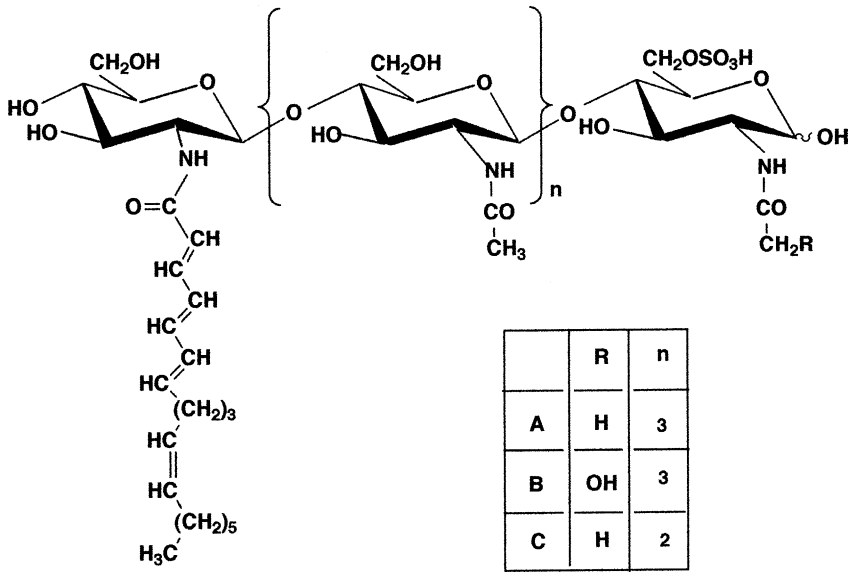
3. Results and discussion

3.1. Nod factors from wild-type strains of *Mesorhizobium huakii*

Astralagus sinicus is an important winter-growing green manure in the southern part of the People's Republic of China. Rhizobia have been found in the nodules of this legume grown in the People's Republic of China and Japan. These bacteria were assigned to a new Rhizobium species called *M. huakii*. In a previous paper [13], structures of Nod factors from the genetically modified *Ra5 (pA28)* strain of *M. huakii* were presented. The *pA28* plasmid carries the regulatory *nod D1* gene of *Rhizobium* sp. *NGR234*. It was introduced to by-pass the control of the gene expression by intrinsic regulatory genes. Mass spectrometry, 2D NMR and chemical analysis revealed the presence of two different structures A and B (Fig. 1) Both are chitopentamers, *O*-6 sulphated on the reducing end and *N*-acylated on the non-reducing terminus by 2*E*,4*E*,6*E*,11*Z* octadecatetraenoic acid. The main component (A) does not possess other modification. In the second one (B), the *N*-acetyl group of the reducing glucosaminyl end is replaced by an *N*-glycolyl one, probably induced by the oxidation of the methyl group of the *N*-acetyl moiety. This modification is very uncommon and we questioned whether it was introduced by the genetic modification or whether it represents a particularity of the *Ra5* strain.

The wild-type *Ra5* strain of *M. Huakii* was cultivated in the presence of naringenin as inducer of *nod* gene expression. The culture medium was extracted by butanol and the butanol extract was passed over a preparative C18 reversed-phase HPLC column, in an acetonitrile–water gradient, monitoring the UV absorption by diode-array detection. A wide UV absorbing peak was detected with maximum of absorption at 306 nm. Nod factors of the *Ra5 (pA28)* strain exhibited the same UV-adsorption maximum that was attributed to the three carbonyl-conjugated double bonds of the fatty acyl part. The corresponding fraction was collected and analysed by mass spectrometry.

Positive-ion LSIMS spectrum of the fraction exhibited a relatively complex pattern in the molecular range with a large background (Fig. 2). The most intense peaks correspond to different ionic species formed by compound A (mass 1329.5 Da). Indeed, MH^+ was at *m*/*z* 1330.5, MNa^+ and MK^+ ions were at *m*/*z* 1352.6 and 1368.6, respectively. The sulphate group induced the formation of several $(M–H + 2cat)^+$ ions such as



	R	n
A	H	3
B	OH	3
C	H	2

Fig. 1. Structures of LCOs produced by *Mesorhizobium huakii*.

$(M–H + 2Na)^+$ at *m/z* 1374.6, $(M–H + Na + K)^+$ at *m/z* 1390.6 and $(M–H + 2K)^+$ at *m/z* 1406.6. Loss of SO_3 (80 u) from MH^+, MNa^+ and MK^+ ions produced the ion series at *m/z* 1250.5, 1272.5 and 1288.5. No ions due to component B with a molecular mass shifted by 16 u can be detected.

The same fraction was examined by negative-ion electrospray ionisation, performed at low cone voltage, after fifty times dilution to lower the inorganic salt concentration. The spectrum now revealed the presence of four different components. (Fig. 3).The intense peak at *m/z* 1328.5 is the $(M–H)^-$ ion of component A. The peak at *m/z* 1344.5 was attributed to the $(M–H)^-$ ion of component B, since its mass was 16 u higher than that of component A. The peak at *m/z* 1125.4, was produced by a third component (C). The 203 u mass difference between component A and C indicated one *N*-acetylglucosaminyl unit less in the oligochitin core. Finally, the peak at *m/z* 1227.3 was the $(M–H)^-$ ion of a fourth component (D) with its sodium salt at *m/z* 1249.3. It is likely that component D holds two sulphate groups as indicated by an abundant double-charged ion $(M–2H)^{2-}$ at *m/z* 613.1, since double-charged negative ions are not seen for the monosulphated species. The structure of compound D is not yet identified.

Negative-ions CID spectra of deprotonated sulphated LCOs did not give useful structural information. Indeed, the main routes are the detachment of the sulphate group that keeps the negative charge and, probably an electron detachment which explain the relatively low abundance of the ion products. The remaining ions are of very low intensity and the spectrum is complicated because of charge-remote fragmentations processes.

Because polyunsaturated fatty acyl groups are easily auto-oxidised by atmospheric oxygen, which may introduce artefactual +16 u shifts, the fraction was hydrogenated over platinum. Molecular masses were 8 Da shifted up, a result that fits with the presence of four double bonds in the acyl moiety. Fig. 4 presents the positive-ion electrospray–CID spectrum of the MH^+ ion from hydrogenated component B at *m/z* 1354. Loss of SO_3 gives *m/z* 1274. The mass interval between the $(MH–SO_3)^+$ ion and the B4 ion is 237 u instead of 221 u for component A. This indicates that the terminal reducing glucosamine holds the +16 mass increment. As glucosamine was the

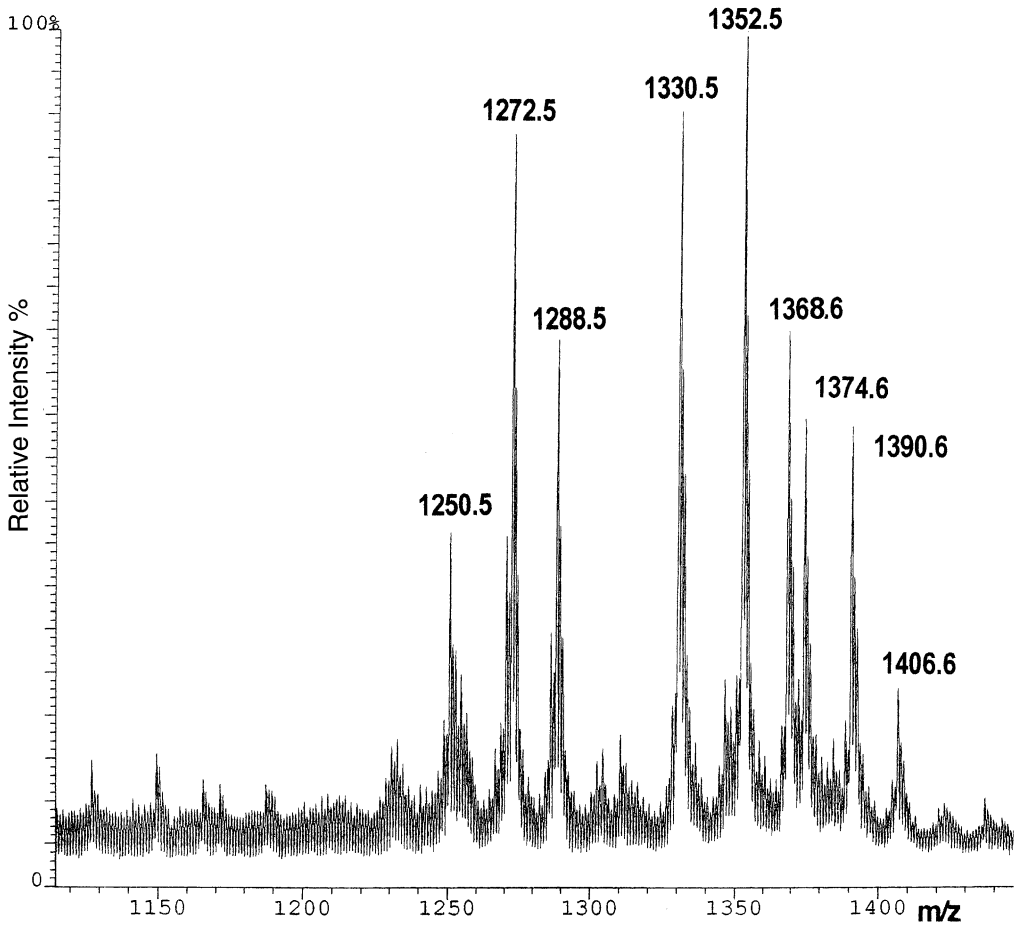


Fig. 2. Molecular mass range of the LSIMS positive-ion spectrum of Nod factors produced by the *Ra5* wild-type sprain of *Mesorhizobium huakii*.

only sugar liberated by hydrolysis, the +16 increment is likely due to the replacement of the *N*-acetyl group by a glycolyl one on the reducing glucosaminyl residue, as in the *Ra5 (pA28)* strain. Mass intervals of 203 u between other ions of the B-series characterise a non-substituted inner core. The glucosaminyl non-reducing end is *N*-substituted by octadecanoic acid, as indicated by the mass of the B-1 ion at *m*/*z* 428.

This structural analysis was not continued further as the above results indicated that the A and B Nod factors are both produced by the *Ra5* strain and the genetically modified *Ra5 (pA28)* strain of *M. huakii*. The wild-type strain produces additionally a tetrameric component (C) of structure close to that of component A, and a disulphated component D, which can be only detected by negative-ion electrospray, a technique that was not used in the previous work.

Similar results were obtained from Nod factors isolated from the wild-type strain *7635R* of *M. Huakii*. Compound A was the main component, components B and C were present but in lesser proportion than in the *Ra5* strain.

3.2. Structural determination of Nod factors produced by Rhizobium sp. mus10

Sesbania rostrata is an aquatic legume growing in Africa. It establishes highly specific interactions with

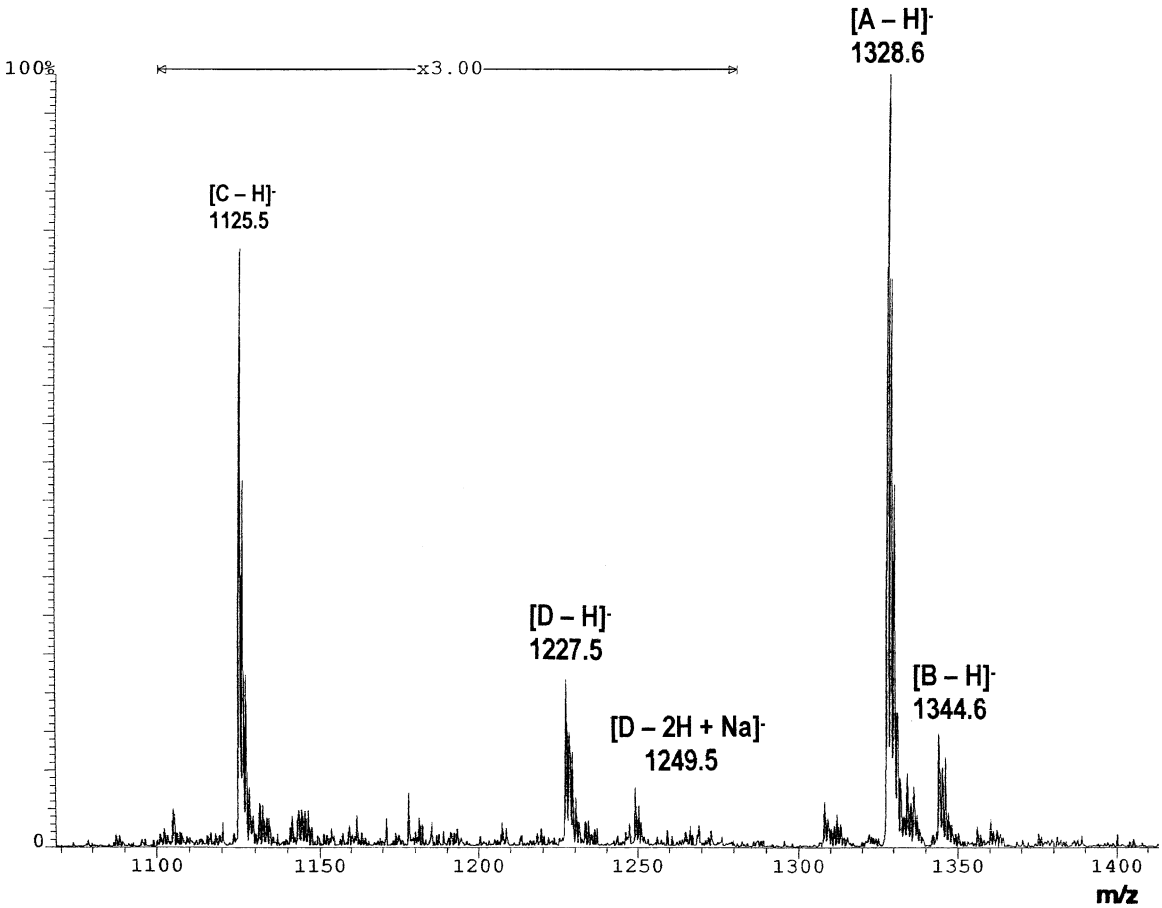


Fig. 3. Upper mass range of the negative-ion electrospray spectrum of Nod factors produced by the *Ra5* wild-type strain of *Mesorhizobium huakki*.

some bacterial strains leading to the formation of root and stems nodules. *Azorhizobium caulinodans* was isolated from nodules of *S. rostrata* grown in the Sahel region of West Africa. The structures of the secreted Nod factors were unusual. The chitopentameric backbone was substituted by two different neutral sugars. The main component holds arabinose (pentose) and fucose (6-deoxyhexose) at the reducing *N*-acetyl glucosaminyl end. Fucose was located on position 6 while arabinose occupies position 3 [15]. The remaining substitutions were more common. The non-reducing glucosaminyl terminus is *N*-methylated and *N*-acylated by acyl groups from membrane fatty acids (octadecanoic and 11-octadecenoic acids). An additional carbamoyl group is on position *O*-6 of the same sugar residue.

Sinorhizobium saheli and *Sinorhizobium teranga* bv. *Sesbaniae* were isolated in Senegal from *Sesbania grandiflora* and *Sesbania aculeata*, respectively. These African strains are also able to nodulate *S. rostrata*. Nod factors are chitopentamers substituted at the reducing end by arabinose and fucose, *N*-methylated and *N*-acylated by long chain “common” fatty acids at the non-reducing end and carbamoylated on the same residue [16]. Neither the carbamoyl group nor fucose and arabinose were precisely located because of too low amounts of material. These molecules are very similar to those from *A. caulinodans*. The Nod factor

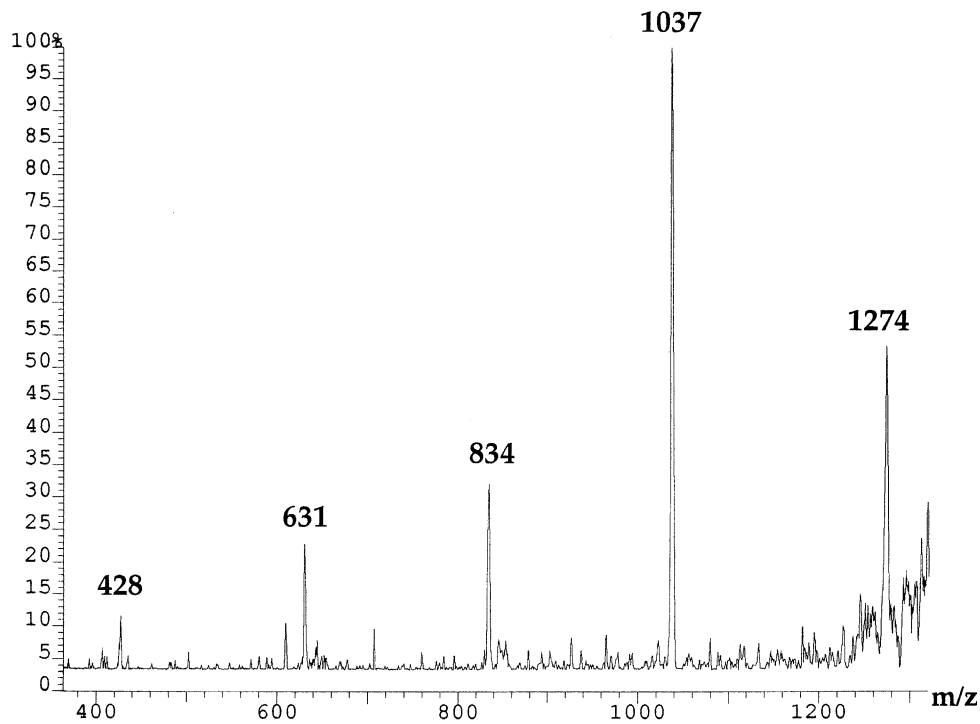


Fig. 4. CID mass spectrum of protonated component B at *m*/*z* 1354. This ion was produced by the electrospray source and its CID spectrum was recorded by B/E scans using the first half part of the instrument and an electrically grounded collision cell.

double glycosylation by arabinose and fucose appears unique for African *Sesbania* symbionts and probably reflects a high selection pressure from the *Sesbania* host plants.

Rhizobium sp. *mus10* is a particular bacterial strain isolated in India from nodules of *Sesbania macrocarpa*. It was asked whether Nod factors produced by this strain isolated from plant growing in a geographically distant area and in a different environment are similar to those produced by the African strains.

Rhizobium sp. *mus10* was cultivated in the presence of naringenin as *nod* genes inducer. The growth medium (30 L) was filtered and passed through an Amberlite XAD-4 column that retained hydrophobic materials. Nod factors eluted from the column by methanol–water mixtures. The crude fractions still contained some polymeric material (exopolysaccharides) that was eliminated by passing through an open C18 reverse-phase column. Nod factors were eluted by acetonitrile–water mixtures and purified by C18 reversed-phase HPLC, monitoring UV absorption at 206 nm. Many peaks were detected and collected. Nod factor-containing fractions were detected by positive-ion LSIMS, pointing out the characteristic pattern of their backbone fragmentation (series of peaks separated by 203 u). Five fractions were collected and the MH^+ ions of six different Nod factors (named from I to VI) were characterised by LSIMS (Table 1). CID of each MH^+ indicated that all of them are chitopentamers substituted on both ends. It can be noted that components III and IV had identical spectra. As they were eluted at different chromatographic times, we believe that they are isomeric compounds. The intensity of molecular ion attributed to component V was much smaller than that from component IV eluted in the same fraction.

An aliquot of each fraction was hydrolysed in acidic conditions. The liberated monosaccharides were

Table 1
Nod factors from *Rhizobium* sp. *mus10*: LSIM spectra of fractions from HPLC separations

Fraction number	Nod factor	$(M + H)^+$	B-1 ion	B-2 ion	B-3 ion	B-4 ion
1	I	1477	501	704	907	1110
2	II	1433	457	660	863	1066
3	III	1459	483	686	889	1092
4	IV	1459	483	686	889	1092
	V	1591	483	686	889	1092
5	VI	1461	485	688	891	1094

The *m*/*z* values of the B-type ions were measured by CID spectra of MH^+ ions recorded at constant B/E-linked scanning using the first half of the tandem instrument. They are numbered from the non-reducing end of the chitooligosaccharidic backbone. B-type ions are due to cleavages of glycosidic bonds with charge retention on the non-reducing end (oxonium ions)

identified by ammonia-CI GC-MS of peracetyl derivatives. All fractions contained glucosamine, *N*-methyl glucosamine and fucose (6-deoxygalactose). One of them (fraction 4) liberated also a small amount of arabinose. The chromatographic peaks due to fucose (four peaks) and arabinose (four peaks) overlap and relative amounts (about 2% arabinose) were estimated by monitoring the MNH_4 ions at *m*/*z* 350 for fucose and 336 for arabinose.

The overall mass of the substitutions on the reducing glucosaminyl moiety was deduced from the mass difference between B-4 and MH^+ ions. Components I, II, III, IV and VI are substituted by fucose only. Component V holds both fucose and arabinose. Methylation analysis of the relatively abundant component II indicated that fucose was linked to *O*-6.

The nature of the *N*-acyl part cannot be deduced directly from the mass spectra of intact Nod factors and fatty acids should be released from their amide linkage by hydrolysis. Hydrolysis by acid (2 M hydrochloric acid in methanol, overnight at 80 °C) afforded a mixture of fatty acids that was analysed by electron impact-GC-MS of their TMS derivatives. Fraction 1 liberated Δ-hydroxyoctadecanoic acid that was identified by peaks at *m*/*z* 233 and 313 resulting from α-cleavages at the Δ-*O*-TMS group. Fractions 2 and 5 liberated hexadecanoic and octadecanoic acids, respectively. Fatty acids from fraction 3 gave a wide GC peak with several shoulders corresponding to a mixture of poorly separated octadecenoic acids. Another octadecenoic acid with greater retention time and a small amount of the previous mixture were produced by hydrolysis of fraction 4.

Structures of the unsaturated fatty acids were identified by capillary GC separation of their pentafluorobenzyl derivatives, followed by dissociative electron capture. Each carboxylate anion was then dissociated by high-energy collision on helium and product ions were analysed by MIKE spectrometry [18]. Each carbon–carbon bond of the acyl chain cleaves and gives carboxylate-containing ions. Allylic cleavages were favoured over cleavages of vinylic and double bonds.

The wide GC peak produced by fatty acids of fraction 3 was due to a mixture of 9-, 10-, 11-, 12- and 13-octadecenoic acids (Fig. 5A), while the major component of fraction 4 was identified as 2-octadecenoic acid as follows (Fig. 6B). From high to low masses, the regularly spaced peaks indicated a polymethylenic chain starting from the methyl end. This series is terminated at *m*/*z* 84. This abundant fragment is the distonic ion $^{\bullet}CH2–CH{=}CH–COO^-$. Absence of fragments containing two and three carbon atoms confirms the presence of a double bond between carbons 2 and 3.

Strong acidic conditions used to cleave the acyl–amide bond may induce isomerisation of double bonds. To verify this assumption, alkaline hydrolysis was used to liberate fatty acids from their amide linkage. It was shown that fatty acids from fraction 3 were 9- and 11-octadecenoic acids, only (Figs. 5B and 6A). Thus other isomers are artefacts due to hydrolytic conditions. Fraction 4 still liberated 2-octadecenoic acid. These results are summarised in Table 2.

Assuming that *N*-methyl glucosamine is the terminal non-reducing residue, it can be now calculated from the masses of the B-1 ions that an additional substituent of mass 43 u (a carbamoyl group) is always present.

It was recently demonstrated that all substitutions on the non-reducing terminus can be identified and located by examining the high-energy collision or the metastable ion spectra of B-1 ions [19]. Presence of

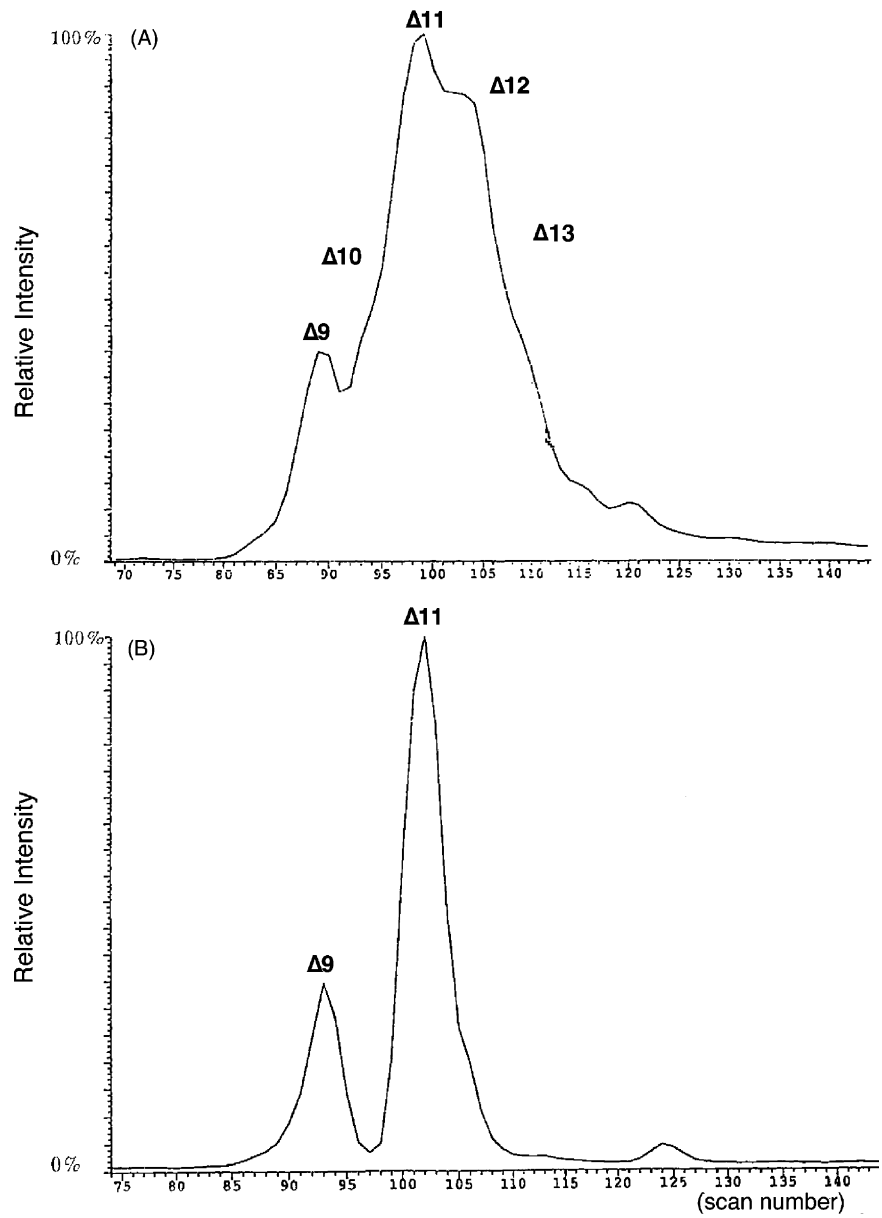


Fig. 5. Mass chromatogram of *m*/*z* 281 in the methane negative-ion GC-MS profile of fatty acids released by hydrolysis of Nod factor III from *Rhizobium* sp. *mus10* after derivation as pentafluorobenzyl esters. (A) Acid hydrolysis. (B) Alkaline hydrolysis. The symbol Δ means that the number which follows indicates position of double bonds from the carboxyl end.

an *N*-methyl substitution can be easily deduced from the intensity ratio of the diagnostic ions *m*/*z* 84 and 98. For all spectra the ratio *m*/*z* 84 and 98 was less than 1. Fig. 7A and B present CID spectra of B-1 ions at *m*/*z* 483 from Nod factors III and IV, respectively. In both spectra, the abundance of *m*/*z* 84 was negligible. This confirmed *N*-methylation of the terminal non-reducing glucosamine.

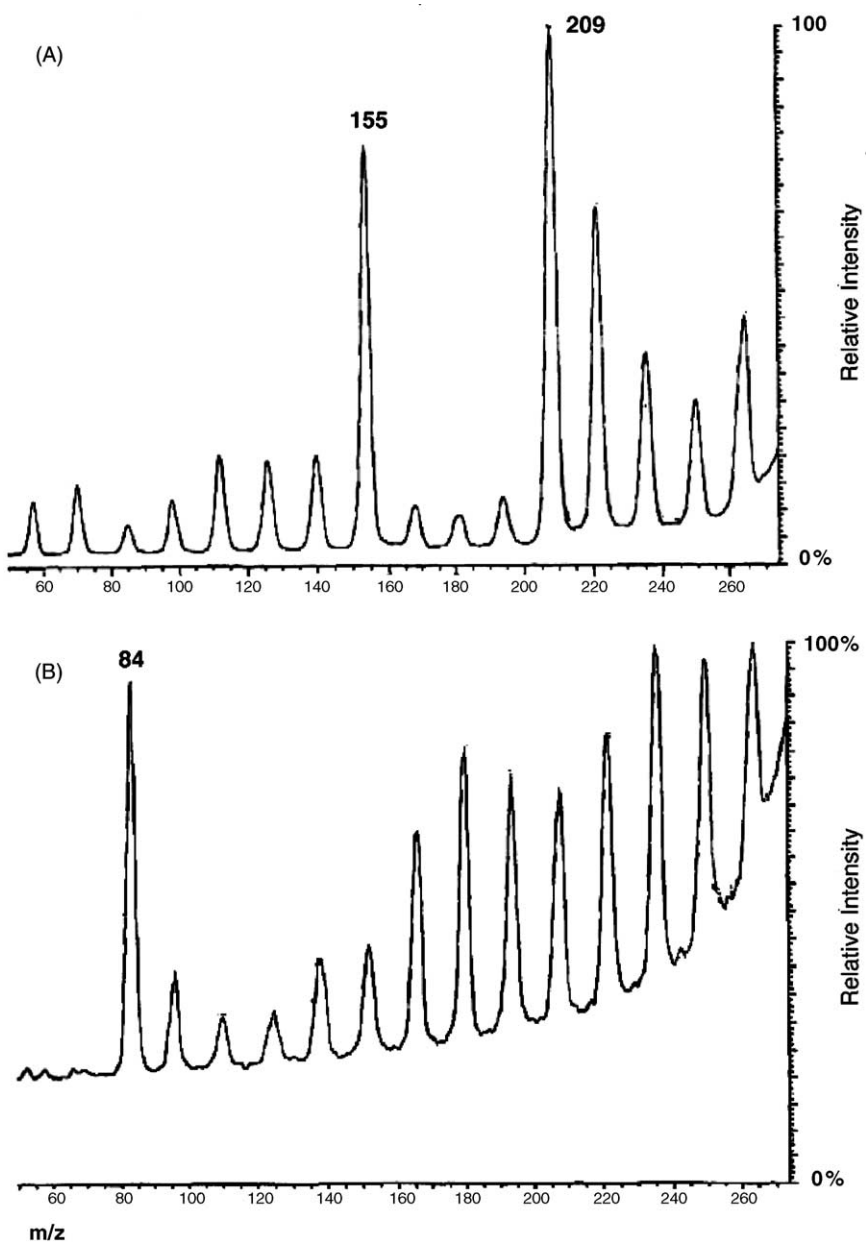


Fig. 6. CID–MIKE spectra of carboxylate anions at *m*/*z* 281 recorded during a negative-ions GC-MS–MS run of fatty acids pentafluorobenzyl derivatives. (A) Spectrum of the C18:1 (Δ11) carboxylate produced by fatty acids released by alkaline hydrolysis of Nod factor III. (B) Spectrum of the C18 (Δ2) carboxylate formed from fatty acids of Nod factor IV.

Position of the carbamoyl group on the sugar ring can be deduced from the relative abundances of ions issued from losses of water or/and of carbamic acid, either in the metastable or CID spectra of the B-1 ions. An *O*-6-carbamoyl group is characterised by loss of water much more abundant than loss of carbamic acid. This characteristic was clearly observed for Nod factors I, II, III (Fig. 7A) and VI. However, this criterion

Table 2
Structural characteristics of Nod factors from *Rhizobium* sp. *mus10*

Nod factor	Molecular mass (Da)	Reducing end	Non-reducing end
I	1476	Fucose	Cb, *N*-Me C18:0 (OH)
II	1432	Fucose	Cb, *N*-Me C16:0
III	1458	Fucose	Cb, *N*-Me C18:1 (Δ9, Δ11)
IV	1458	Fucose	Cb, *N*-Me C18:1 (Δ2)
V	1590	Fucose Arabinose	Cb, *N*-Me C18:1
VI	1460	Fucose	Cb, *N*-Me C18:0

Abbreviations: Cb $=(O)CO–NH_2$ group located on *O*-6; the nature of the *N*-linked fatty acids is shown on the second line, with the carbon chain length and the number of double bonds. Their position from the carboxyl end is indicated between brackets.

cannot be used for component I as the huge peak due to loss of water was mainly due to the presence of a hydroxylated fatty acid. For component IV substituted by 2-octadecenoic acid, the ion corresponding to water loss at *m*/*z* 465 had a relatively low intensity (Fig. 7B). This observation did not rule out *O*-6 substitution for component IV, since elimination of carbamic acid was negligible. However, it was seen on models that eliminations of *O*-ester substituents are poorly favoured when the *N*-acyl group holds an αβ-unsaturation and it cannot be excluded that the ion at *m*/*z* 465 is provided by the isomeric B-1 ions formed from low amounts of component V in the same fraction. To precise location of carbamoyl groups on components I and IV, the relative abundances of the **a**, **b** and **c** fragments at *m*/*z* 140, 152 and 158, respectively, were measured. Indeed, it was shown on models that the intensity ratios between **a** and **b** and **a** and **c** differentiate all the positional isomers [19]. In spectra presented in Fig. 7A and B, the **a** fragment was much more abundant than **b** and **c** (**c** cannot be detected in the background). This feature was characteristic of an *O*-6 substitution. All other nod factors exhibited the same behaviour. The relative intensity of fragments **a**, **b**, and **c** was lower for component I whose spectrum was dominated by a water loss, but their abundance was high enough to provide a reliable location of the carbamoyl group by that way.

Taking together these results, it was demonstrated that Nod factors II, III, V and VI from *Rhizobium* sp. *mus10* are almost the same as those produced by *Sesbania*-nodulating strains isolated in Africa (Fig. 8). This result confirms the hypothesis of a mutual adaptation of bacteria and plants, independent of the geographical distance of growth that precludes gene transfer between West African and Indian strains during Evolution. However, the abundance of Nod factor V with fucose and arabinose substitutions is particularly low in the *Rhizobium* sp. *mus10* strain.

The major structural differences between Nod factors from African and Indian strains are the nature of *N*-linked fatty acids that possessed a Δ-hydroxy group (I) or one carbonyl-conjugated double bond (IV).

Up to now, Nod factors with carbonyl-conjugated double bonds were exclusively found in strains that nodulate plants of the *galegae* tribe [13]. This tribe does not include *Sesbania* species. This variety of fatty acids is produced by the *nodFE* genes, whose products slow down or stop the enoyl-reductase activity in the last cycles of the fatty acid biosynthesis. It seemed likely that Δ-hydroxy fatty acids can be formed in the same way, by slowing down the speed of the alcohol dehydrogenase which acts in the same catalytic cycle. Indeed, the *nodF* gene encodes a new acyl-carrier protein (NodF) that replaces the normal one. The NodF product is a substrate poorly recognised by several enzymes of the fatty acid synthesis. Thus, it seems likely that a NodF-like protein should be synthesised by the *Rhizobium* sp. *mus 10* strain.

This particularity is unique. Neither *N*-methylation of the terminal non-reducing glucosamine (*nodS*) nor fucosylation (*nodZ*) or arabinosylation (*nolK* and *noeC*) were found in strains possessing the *nodFE* genes. Fatty acids resulting from the action of *nodFE* genes were never found in Nod factors from strains with the *nodS* genes. It can be postulated that the peculiar structural features of *mus10* Nod factors induce some new and not yet known host-range extension of the Indian strain.

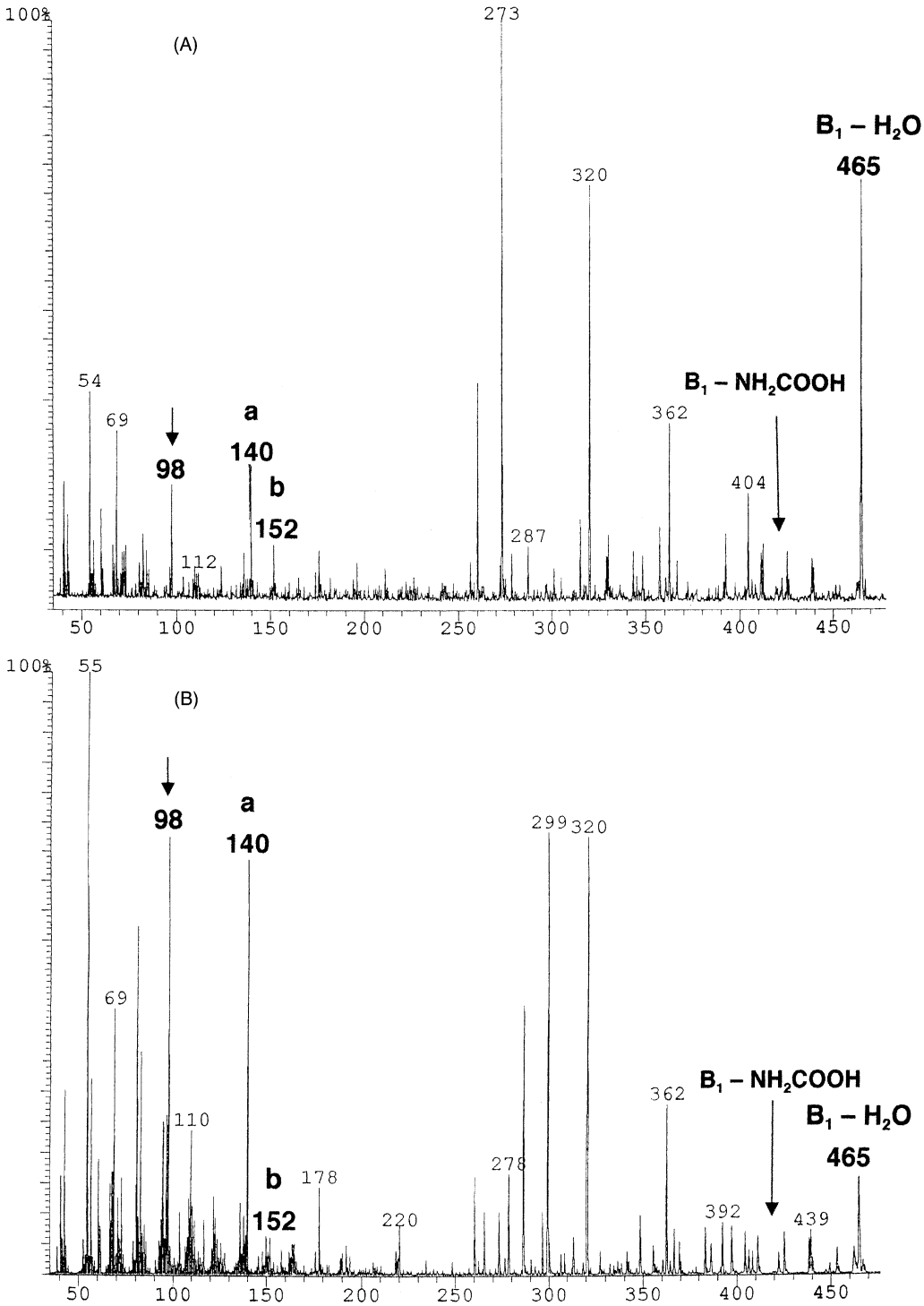
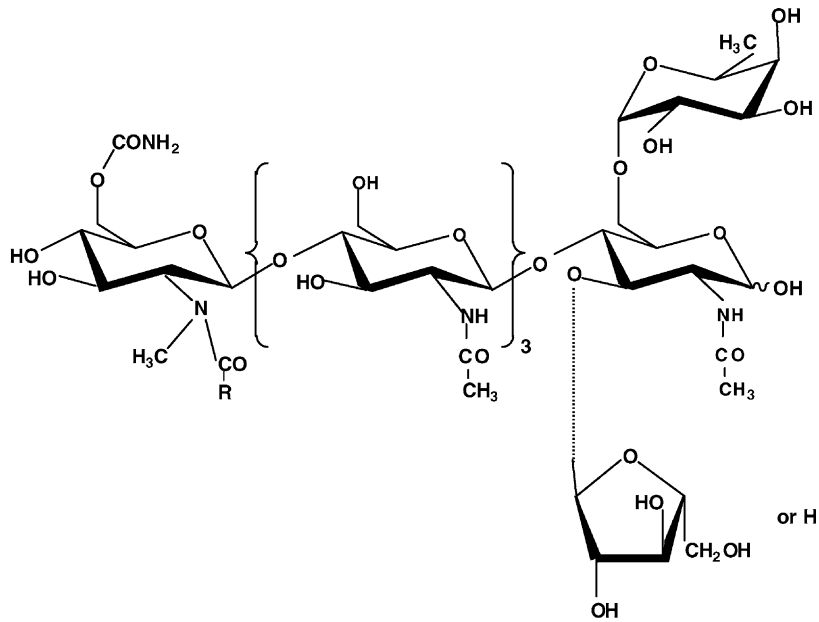


Fig. 7. Product ions in the CID spectra of the B-1 ion at *m/z* 483 formed by LSIMS ionisation of Nod factors III and IV from *Rhizobium* sp. *mus10*. The full 6-sector instrument was used with the collision cell (4th FFR) floating at half the accelerating voltage. (A) Nod factor III. (B) Nod factor IV.



	RCO
Azorhizobium caulinodans	C18:0 C18:1Δ11
Sinorhizobium saheli	C16:0 C18:1Δ11
Sinorhizobium bv. sesbania	C16:0 C18:1Δ11
Rhizobium sp. mus 10	C16:0 C18:1Δ9 C18:1Δ11 C18:1Δ2 C18:0 (Δ-OH)

Fig. 8. Structures of Nod factors produced by African and Indian Rhizobium strains nodulating *Sesbania* species.

4. Conclusion

Several authors used FAB and electrospray ionisation mass spectrometry with an extensive use of MS–MS to characterise most of the structural particularities of Nod factors that specify their activity on plants in the rhizobium–legume symbiosis (see, e.g., [1,4,13–16,19–23]. An important difficulty when

studying Nod factors from wild-type, low-producing strains is the elimination of inorganic salts from the samples that induces formation of "cationised" species not suitable for MS–MS studies. This is the one reason that limits the use of MALDI ionisation, despite its high sensitivity, that produced quite exclusively alkali-ion attached molecules. The high stability of these species precludes prompt fragmentations and post source decays. Another difficulty is the suppression ionisation effects when studying complex LCOs mixtures. For example, the doubly sulphated species D, present in the Nod factor mixture purified from *M. Huakii*, was only detected in the electrospray negative-ion mode and our attempts to cleave its negatively double-charged ion did not provide useful structural information.

The anomeric configurations of the linkage residues, needed to establish complete structures, are not provided by mass spectrometry. However, the stereospecificity of the NodC enzymes that construct the *N*-acetyl glucosaminyl oligomeric cores is well established: they have biosynthetic properties close to that of chitin synthases and build β1,4-linked GlcNac oligomers. Hydrolytic enzymes of known stereospecificity, such as α-fucosidase, can be used in association with mass spectrometry to provide the linkage stereochemistry of fucose [15]. Determination of the linkage modes by NMR required a relatively larger amount of pure material.

Recent improvements in LC-electrospray MS–MS, such as capillary LC-coupling and TOF or ion-trap ion analysers will be very helpful in the future for routine determinations of LCOs from very low-producing strains grown on small-scale cultures.

References

[1] P. Lerouge, P. Roche, F. Maillet, G. Truchet, J.C. Promé, J. Denarié, Nature 344 (1990) 781.
[2] N. Demont, M. Ardourel, F. Maillet, D. Promé, M. Ferro, J.C. Promé, J. Denarié, EMBO J. 13 (1994) 2139.
[3] M. Schulze, B. Duiclet-Sire, E. Kondorosi, H. Virelizier, J.N. Glushka, G. Endre, S.D. Gero, A. Kondorosi, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 192.
[4] P. Roche, P. Lerouge, C. Ponthus, J.C. Promé, J. Biol. Chem. 266 (1991) 10933.
[5] J. Denarié, F. Debellé, J.C. Promé, Annu. Rev. Biochem. 65 (1996) 503.
[6] W.E. Newton, in: E. Martinez, G. Hernandez (Eds.), Highlights of Nitrogen Fixation Research, Kluwer, New York, 1999.
[7] M. Hanin, S. Jabbouri, W.J. Broughton, R. Fellay, D. Quesada-Vincens, Plant Microbe Interact. 4 (1999) 1.
[8] R.F. Fisher, S.R. Long, Nature 357 (1992) 655.
[9] J.A. Downie, Trends Microbiol. 2 (1994) 318.
[10] J.C. Promé, Curr. Opin. Struct. Biol. 6 (1996) 671.
[11] M. Schulze, E. Kondorosi, P. Ratet, M. Buiré, A. Kondorosi, Int. Rev. Cytol. 156 (1994) 1.
[12] B. Relic, X. Perret, M.T.M. Estrada-Garcia, J. Kopcinska, W. Golinowski, H.B. Krishnan, S.G. Pueppke, W.J. Broughton, Mol. Microbiol. 13 (1994) 171.
[13] G.-P. Yang, F. Debellé, A. Savagnac, M. Ferro, O. Schiltz, F. Maillet, D. Promé, M. Treilhou, C. Vialas, K. Lindstrom, J. Denarié, J.C. Promé, Mol. Microbiol. 34 (1999) 227.
[14] P. Mergaert, M. Van Montagu, J.C. Promé, M. Holsters, Proc. Natl. Acad Sci. U.S.A. 90 (1993) 1551.
[15] P. Mergaert, W. D'Haeze, M. Fernandez-Lopez, D. Geelen, K. Goethals, J.C. Promé, M. Van Montagu, M. Holsters, Mol. Microbiol. 21 (1996) 409.
[16] J. Lorquin, G. Lortet, M. Ferro, N. Mear, B. Dreyfus, J.C. Promé, C. Boivin, Mol. Plant Microbe Interact. 10 (1997) 879.
[17] I. Ciacanu, F. Derek, Carbon Res. 131 (1984) 209.
[18] J.C. Promé, H. Aurelle, F. Couderc, A. Savagnac, Rapid Commun. Mass Spectrom. 1 (1987) 50.
[19] M. Treihou, M. Ferro, C. Monteiro, V. Poinsot, S. Jabbouri, C. Kanony, D. Promé, J.C. Promé, J. Am. Soc. Mass Spectrom. 11 (2000) 301.
[20] H.P. Spaink, D.M. Sheeley, A.A.N. van Brussel, J. Gluska, W.S. York, T. Tak, O. Geiger, E.P. Kennedy, V.N. Reinhold, B.J.J. Lugtemberg, Nature 354 (1991) 125.
[21] R.W. Carlson, J. Sanjuan, U.R. Bhat, J. Gluska, H.P. Spaink, A.H.M. Wijfjes, A.A.N. van Brussel, T.J.W. Stokkermans, N.K. Peters, G. Stacey, J. Biol. Chem. 268 (1993) 18372.
[22] S. Philip-Hollingsworth, G.G. Orgambide, J.J. Bradford, D.K. Smith, R.I. Hollingsworth, F.B. Dazzo, J. Biol. Chem. 270 (1995) 20968.
[23] M.M.A. Olsthoorn, I.M. Lopez-Lara, B.O. Petersen, K. Bock, J. Haverkamp, H.P. Spaink, J.E. Thomas-Oates, Biochemistry 37 (1998) 9024.