Rhizobium fredii synthesizes an array of lipooligosaccharides, including a novel compound with glucose inserted into the backbone of the molecule

Marie-Pierre Bec-Ferté[a], Hari B. Krishnan[b], Arlette Savagnac[a], Steven G. Pueppke[b,*], J.-C. Promé[a]

[a]*Institut de Pharmacologie et de Biologie Structurale du Centre National de la Recherche Scientifique, 205 Route Norbonne, 31077 Toulouse Cedex, France*
[b]*Department of Plant Pathology, University of Missouri, Columbia, MO 65211, USA*

Received 29 July 1996

Abstract **Flavonoid cues from the plant host cause symbiotic, nitrogen-fixing rhizobia to synthesize lipochitooligosaccharides (LCOs), which act as return signals to initiate the nodulation process. *Rhizobium fredii* strain USDA257 is known to produce four LCOs, all substituted with vaccenic acid ($C_{18:1}$). We show here that a mutant strain can replace vaccenic acid with a $C_{16:0}$ side chain, and that strain USDA191 synthesizes additional LCOs that differ from one another in the length and unsaturation of their fatty acyl substituents. USDA191 and 257DH4 both produce a novel LCO with glucose substituted for *N*-acetyl-D-glucosamine in the backbone of the molecule.**

Key words: Lipochitooligosaccharides; Nitrogen fixation; Nod factors; Nodulation; *Rhizobium fredii*

1. Introduction

Compatible interactions between the roots of legume plants and symbiotic bacteria lead to dedifferentiation of root cortex cells and the programmed appearance of a new structure, the root nodule. *Rhizobium* cells live within these symbiotic organs, where they fix atmospheric nitrogen and deliver it to the plant as ammonia. The ability of rhizobia and legumes to recognize one another is selective to varying degrees [1]. Some strains, such as the broad host range *Rhizobium* sp. NGR234, are compatible with a wide variety of diverse legume species. The host ranges of other rhizobia, including the biovars of *R. leguminosarum*, are restricted to just a few, often closely related legumes. In extreme cases, individual bacterial strains may enter into cultivar-specific associations with their host legumes [2–5].

The flow of molecular signals between the plant and bacteria is a key factor in defining the specificity of the nodulation process [6]. Rhizobia synthesize arrays of LCOs, or Nod factors, in response to flavonoid cues from the plant. At subnanomolar concentrations, these molecules can substitute for intact *Rhizobium* cells in inducing a number of host nodulation responses, in a manner that matches the specificity of the producing *Rhizobium* strain. We previously characterized the LCOs of *R. fredii* USDA257, a strain that originally was isolated from nodules of wild soybean, *Glycine soja* [4]. This organism has a wide host range for legume species, but it is specific for certain cultivars of soybean [4,5]. The LCOs of strain USDA257 comprise a series of β-1,4-linked oligomers of *N*-acetyl-D-glucosamine, with degrees of polymerization ranging from three to five [7]. These molecules are *N*-substituted with vaccenic acid ($C_{18:1}$) on the nonreducing glucosamine residue and contain 2-*O*-methylfucose (or occasionally fucose), α-linked to carbon 6 of the reducing *N*-acetyl-D-glucosamine residue.

We now have examined the LCOs of 257DH4, a transposon mutant of strain USDA257, as well as those of *R. fredii* strain USDA191 [8,9]. These organisms share the wide host range of USDA257, but they nodulate soybean cultivars nonspecifically. Strain USDA191 elaborates all of the LCOs produced by strain USDA257, but it also synthesizes an array of additional compounds, including a novel molecule with glucose incorporated into its backbone. Although the LCOs of mutant 257DH4 resemble those of its parental strain, it also produces the glucose-containing molecule.

2. Materials and methods

2.1. Purification of LCOs

LCOs were prepared from wild-type *R. fredii* strain USDA191 [4] and the mutant 257DH4 [8,9], essentially as described [7]. Briefly, cells were grown overnight at 30°C in glutamate-mannitol medium (10 g mannitol, 1.1 g sodium glutamate, 0.22 g K_2HPO_4, 0.1 g NaCl, 0.1 g $MgSO_4$, 20 mg $FeCl_3$, and 0.5 mg biotin per liter of water) supplemented with 10 μM genistein (Sigma Chemical Co.) as flavonoid inducer. Bacteria were removed by centrifugation, and the culture supernatants batch-extracted with Amberlite XAD-4 resin. LCOs were eluted from the resin with 70% aqueous methanol and then with methanol. The eluants were pooled and dried, redissolved in 20% aqueous acetonitrile, and then purified on an open C_{18} reversed-phase column (140×8 mm, Lichroprep RP-18, particle size 25–40 μm, Merck). LCOs were eluted from the column with 50% aqueous acetonitrile. Fractions of interest were purified twice by C_{18} reversed-phase HPLC (7.5×250 mm, Spherisorb, ODS2, particle size 5 μm), first for 50 min with a gradient of 20% aqueous acetonitrile to 100% acetonitrile and then for 40 min with 36–68% aqueous acetonitrile [7].

2.2. Mass spectrometry

Spectra were measured on an Autospec Instrument (Fisons, VG Analytical, Loughborough, UK). The acceleration voltage was 8 kV in all ionization modes. For LSIMS, a 25 keV cesium ion beam was focused on the target that had been loaded with 1 μl of the matrix (1:1, v/v, *m*-nitrobenzyl alcohol/glycerol) and spiked with either 1% trichloroacetic acid in water or 1% NaI in water. An electron impact ionization source was utilized for GC/MS experiments. The electron

*Corresponding author. Present address: Fachbereich Biologie/Botanik, Philipps Universität, D-35032 Marburg/Lahn, Germany. Fax: (49) (6421) 288997.

Abbreviations: LCOs, lipochitooligosaccharides; LSIMS, liquid secondary ionization mass spectrometry; GC/MS, gas chromatography/mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; CID, collision-induced association

PII S0014-5793(96)00903-9

energy was 70 eV, and the trap current was 100 μA. The instrument was coupled with a Hewlett Packard model 5890 gas chromatograph fitted with an OV1 capillary column (30 m×0.32 mm). Matrix-assisted laser desorption/ionization (MALDI) spectra were measured on a Fisons TofSpec Instrument fitted with a 337 nm nitrogen laser. The matrix was 2,5-dihydroxybenzoic acid.

2.3. Chemical analysis

LCOs were hydrolyzed in aqueous 3 M HCl for 3 h at 80°C for carbohydrate determination. The resulting monosaccharides were identified as alditol acetates or by methanolysis as methyl glycoside peracetates. Methanolysis was with 1 M methanol/HCl for 18 h at 80°C. Fatty methyl esters were extracted with dichloromethane and analyzed by capillary GC (Girdal Series 30 instrument equipped with a 30 m×0.32 mm OV1-bound capillary column) with a temperature gradient of 3°C per min from 100°C to 280°C. Peak assignments were confirmed by GC/MS as described above. Sugars formed by acid hydrolysis were reduced with $NaBD_4$ and permethylated by the method of Ciucanu and Kerek [10]. The methylated LCOs were extracted with chloroform and hydrolysed with aqueous 2 M trifluoroacetic acid for 2 h at 110°C. The resulting partially methylated sugars then were reduced with $NaBH_4$, acetylated for 20 min at 110°C with 1:1 (v/v) acetic anhydride/pyridine, and analyzed by GC/MS in both the electron impact and chemical ionization modes.

3. Results

3.1. R. fredii *strains synthesize differing arrays of LCOs*

LCOs were extracted from supernatants of large-scale, genistein-induced cultures onto Amberlite XAD-4 resin and then passed over an open C_{18} reversed-phase column, with eliminates hydrophilic and highly hydrophobic components. The LCO-containing fraction was further purified by two successive HPLC runs. Six HPLC fractions were collected and analyzed by LSIMS and MS/MS methods, which revealed structures similar to those of *R. fredii* strain USDA257 [7]. There was one major difference: LCO preparations from USDA191 and 257DH4 contained arrays of fatty acids that were not present in preparations from USDA257. This was confirmed by fatty acid analysis after methanolysis, a procedure that identified a mixture of $C_{16:0}$, $C_{16:1}$, $C_{18:0}$, and $C_{18:1}$ moieties. Fucose and 2-*O*-methylfucose were characterized from hydrolysis products, and the corresponding substituents were assigned to the reducing glucosamine residues on the basis of MS/MS and normal mass spectra [7]. Fig. 1a illustrates the tri-, tetra-, and pentameric LCO structures that were identified, and Table 1 summarizes all of the known LCOs of *R. fredii*. It can be seen that USDA191, USDA257, and 257DH4 all produce three basic LCO molecules. Mutant 257DH4 replaces one of the tetramers produced by the parental strain with a new molecule, NodRf-IV($C_{16:0}$, MeFuc). Strain USDA191 synthesizes a total of 11 LCOs, seven of which could not be detected in genistein-induced cultures of USDA257 or 257DH4. These molecules include tri-, tetra-, and pentamers, and they differ from one another primarily in the length and degree of unsaturation of their fatty acyl side chains.

3.2. A novel glucose-containing LCO from strain USDA191 and mutant 257DH4

Careful examination of the LSIMS spectra of one of the HPLC fractions from either USDA191 or 257DH4 revealed a weak ion at $m/z = 1375$ (Fig. 2). This species was absent in preparations from USDA257, and its presence could not be explained by variations of known LCO structures, by fragmentation, or by adduct formation. To determine if this ion corresponds to a protonated molecule possessing a structure related to those of previously characterized LCOs, we obtained a CID spectrum at constant B/E ratio and analyzed

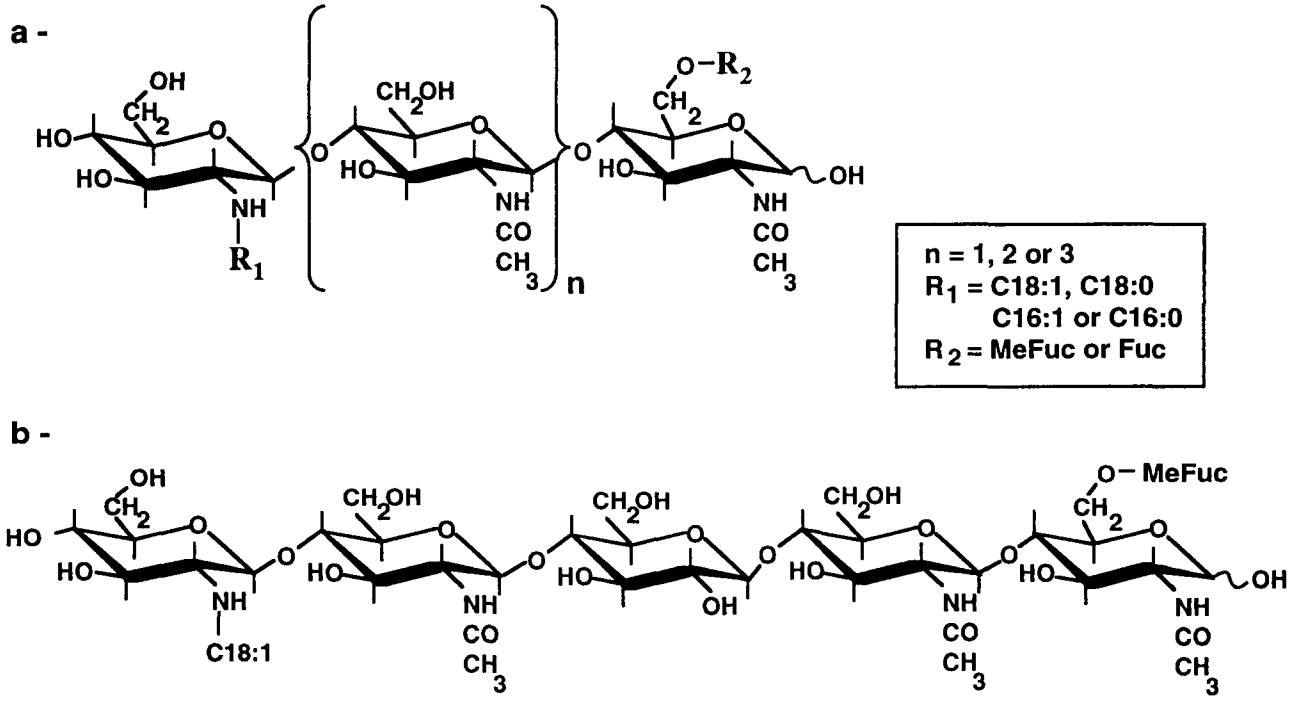

Fig. 1. The lipochitooligosaccharides of *Rhizobium fredii*. a. Schematic structure of the basic LCO molecule (see Table 1). b. Complete structure of the novel LCO, NodRf-V-Glc($C_{18:1}$, MeFuc).

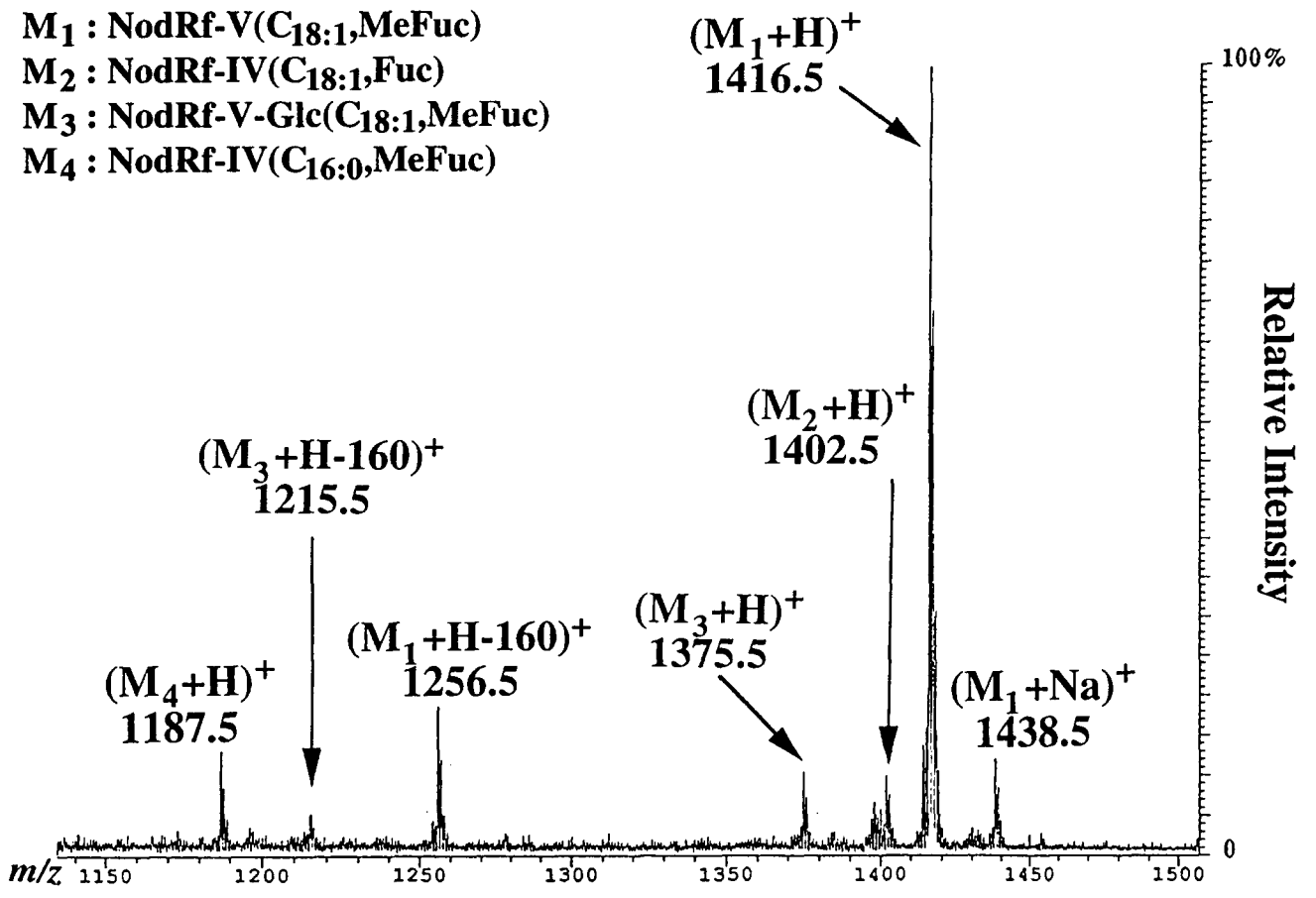

Fig. 2. LSIMS spectrum of the HPLC fraction containing the novel LCO from strain USDA191. The matrix was a 1:1 mixture of glycerol and *m*-nitrobenzyl alcohol spiked with 1% trichloroacetic acid.

fragmentations at different glycosidic linkages (Fig. 3). The loss of a 2-*O*-methylfucose residue (160 mass units) gave an ion at *m/z* 1215. This residue is linked to a reducing *N*-acetyl-D-glucosamine, as shown by the presence of an abundant fragment ion at *m/z* 994, which corresponds to a loss of 381 mass units. Then, as is typically evident in the CID spectra of LCOs, another fragment separated by 203 mass units provided evidence for an additional linked *N*-acetyl-D-glucosamine residue (*m/z* = 791). The following interval, leading to *m/z* 629, is unusual because it is 162 mass units instead of

Table 1
Lipochitooligosaccharides of *Rhizobium fredii*

Molecule	Produced by			Fatty acid (R_1)[a]	Sugar residue (R_2)[a]	Chain length (N)[a]	Pseudomolecular ion (MH^+)
	USDA257	257DH4	USDA191				
NodRf-V($C_{18:1}$, MeFuc)	+	+	+	$C_{18:1}$	Methylfucose	3	1416
NodRf-IV($C_{18:1}$, MeFuc)	+	+	+	$C_{18:1}$	Methylfucose	2	1213
NodRf-III($C_{18:1}$, MeFuc)	+	+	+	$C_{18:1}$	Methylfucose	1	1010
NodRf-IV($C_{18:1}$, Fuc)	+	−	+	$C_{18:1}$	Fucose	2	1199
NodRf-IV($C_{16:0}$, MeFuc)	−	+	+	$C_{16:0}$	Methylfucose	2	1187
NodRf-V($C_{18:1}$, Fuc)	−	−	+	$C_{18:1}$	Fucose	3	1402
NodRf-V($C_{18:0}$, MeFuc)	−	−	+	$C_{18:0}$	Methylfucose	3	1418
NodRf-V($C_{16:1}$, MeFuc)	−	−	+	$C_{16:1}$	Methylfucose	3	1388
NodRf-V($C_{16:0}$, MeFuc)	−	−	+	$C_{16:0}$	Methylfucose	3	1390
NodRf-IV($C_{18:0}$, MeFuc)	−	−	+	$C_{18:0}$	Methylfucose	2	1215
NodRf-IV($C_{16:1}$, MeFuc)	−	−	+	$C_{16:1}$	Methylfucose	2	1185
NodRf-III($C_{16:1}$, MeFuc)	−	−	+	$C_{16:1}$	Methylfucose	1	982

[a]Numbers in these columns refer to features of the structures given in Fig. 1a.

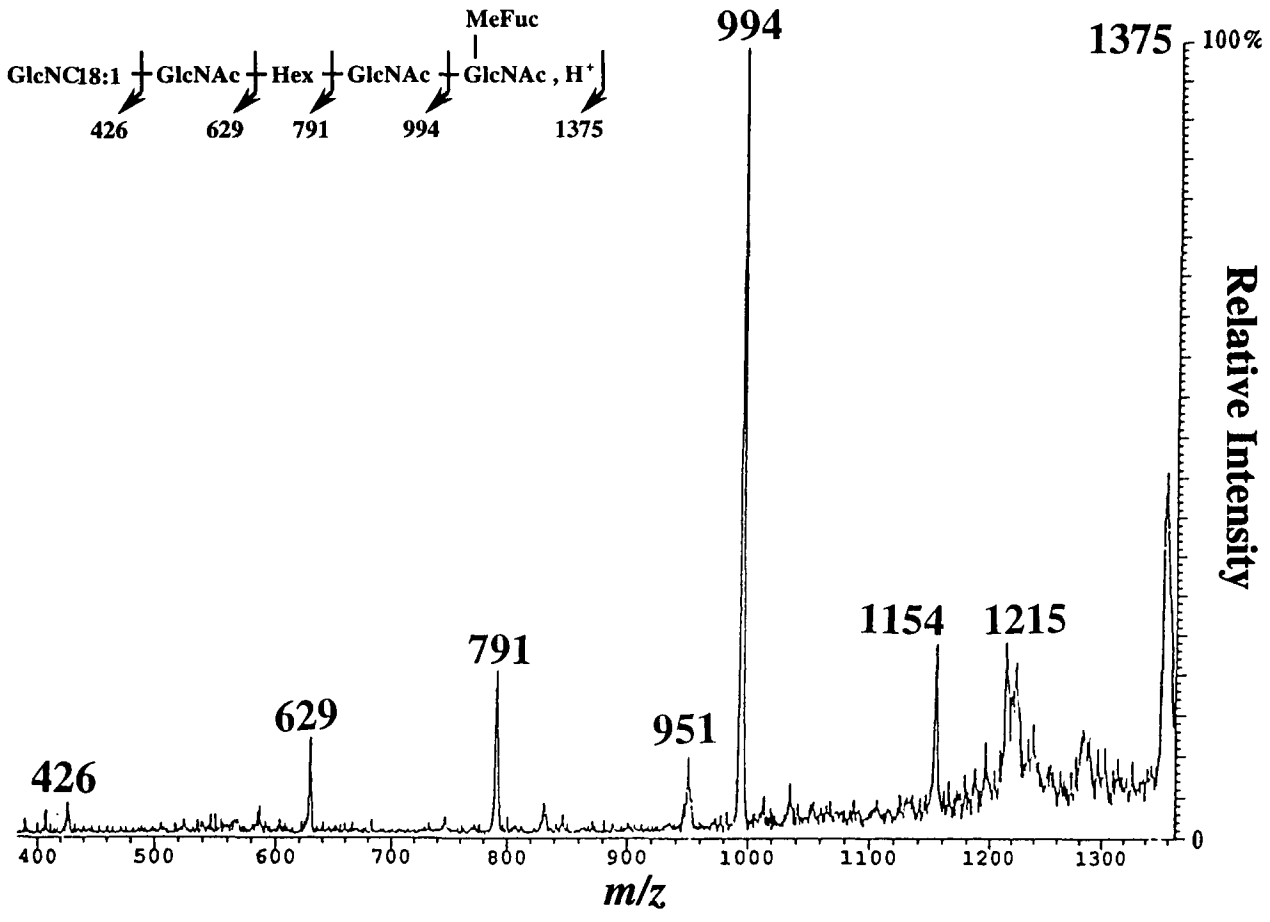

Fig. 3. CID spectrum at constant B/E scans of the ion $m/z = 1375$, which corresponds to the MH^+ ion of NodRf-V-Glc($C_{18:1}$, MeFuc).

203. A further interval of 203 mass units gave $m/z = 426$, which corresponds to the oxenium ion B1 with a vaccenoyl moiety attached to the nonreducing glucosaminyl terminus.

The 162 mass unit interval suggested that a hexose residue is inserted within the chitooligomeric chain as illustrated in Fig. 1b, and indeed, acid hydrolysis revealed the presence of small amounts of glucose. We NaBD4-reduced, permethylated, and hydrolysed the fraction and then used GC/MS to identify the resulting sugars as partially methylated alditol acetates. Although we found the expected derivatives of glucosamine and fucose (Fig. 4), e.g. the alditol acetates of 2,3,4-tri-*O*-methylfucose (branched fucose), *N*-methyl-3,6-dimethylglucosamine (internal glucosamine residues), *N*-methyl-3,4,6-trimethylglucosamine (nonreducing terminal glucosamine), and *N*-methyl-1,3,5-trimethylglucosamine (terminal reducing glucosamine), we also detected a small amount of the alditol acetate of 2,3,6-tri-*O*-methylglucose (Fig. 4). This is characteristic of either a 1,4-linked glucopyranosyl or a 1,5-linked glucofuranosyl residue.

Although the relative proportion of the glucose-containing compound varied from one culture to another, it never exceeded a small percentage of the total, and all attempts to separate it from the bulk of the conventional LCOs failed. We examined the novel compound by digestion with 0.5 units of an aqueous chitinase solution from *Streptococcus griseus* (Sigma Chemical Co.), which was added directly to the HPLC fraction. After 18 h at 28°C, a 1 μl aliquot of the reaction mixture was combined with 1 μl of 50 mM 2,4-dihydroxybenzoic acid (in 70% acetonitrile) on the MALDI target, dried, and then analyzed. We detected a mixture of acylated and nonacylated oligosaccharides, all of which were smaller than the intact LCOs. Ions characteristic of the vaccenoyl chitobiose portion of the molecule were present (MNa^+ and MK^+ at m/z 669 and 685, respectively), as were weaker peaks due to vaccenoyl chitotriose (m/z 872 and 888). The main nonacylated oligosaccharide was identified as 2-*O*-methylfucosyl chitobiose (MNa^+ and MK^+ at m/z 607 and 623, respectively). 2-*O*-Methylfucosyl chitotriose gave weaker ions at m/z 809 and 825. Characteristic ions attributable to the cleavage by chitinase of the glucose-containing molecules were just detectable at m/z 768 and 784, positions corresponding to chitobiose bearing both 2-*O*-methylfucosyl and glucosyl substituents (Fig. 5).

4. Discussion

The LCOs from a number of *Rhizobium* spp. have been characterized, and they are known to be chitin oligomers with an *N*-linked fatty acyl side chain on the nonreducing glucosamine residue and various other substituents located elsewhere in the molecule [6,11]. Although there is a substantial body of evidence that these molecules trigger many of the

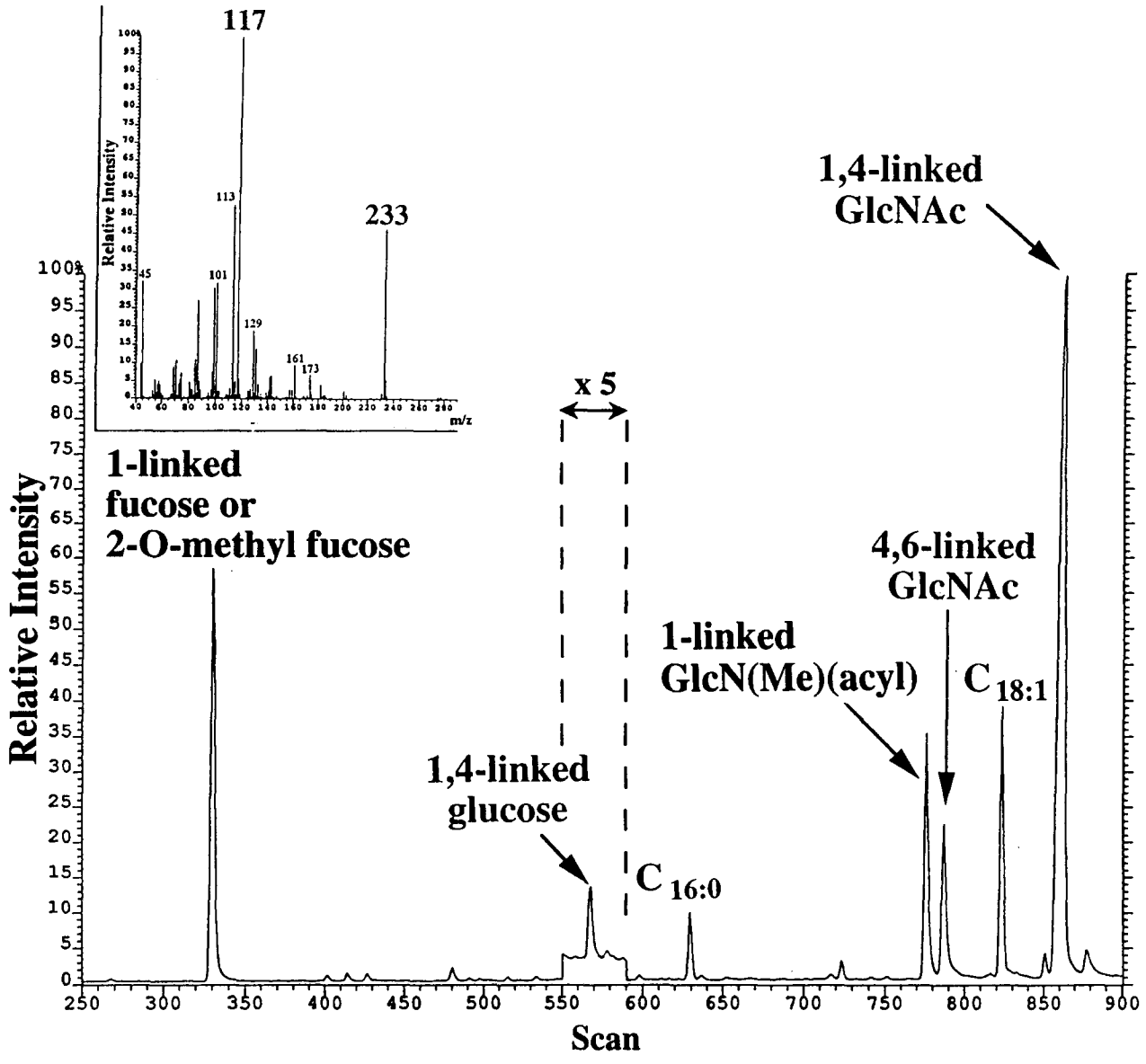

Fig. 4. Total ion current-gas chromatogram (electron impact mode) of partially methylated alditol acetates arising from the fraction containing NodRf-V-Glc($C_{18:1}$, MeFuc). The insert shows the mass spectrum of 1,4,5-tri-*O*-acetyl-2,3,6-tri-*O*-methylglucitol. Linkages of the different monosaccharides within the LCO are labeled. The peaks labeled $C_{16:0}$ and $C_{18:0}$ represent the methyl esters of palmitic and vaccenic acids, respectively.

cellular responses that lead to nodulation, structure-activity relationships are just beginning to be sorted out [12]. In at least one case, the interaction of *R. leguminosarum* bv. viciae strain TOM with Afghanistan pea, cultivar specificity has been attributed to these signal molecules. The appearance of nodules in this interaction depends on a positively acting gene, *nodX*, which is present in TOM but absent in strain of *R. leguminosarum* bv. viciae that cannot nodulate this cultivar [13]. This gene, which probably encodes a specific *O*-acetyl transferase, causes a pentameric LCO of strain TOM to be *O*-acetylated on carbon 6 of the reducing glucosamine residue.

Our data show that *R. fredii* strains with differing abilities to nodulate soybean cultivars can produce highly divergent, overlapping arrays of LCOs. Although most of these compounds represent variations on a common theme that also appears in two other soybean symbionts, *Bradyrhizobium japonicum* [14] and *B. elkanii* [15], strain USDA191 and mutant 257DH4 appear to be unique in their ability to synthesize a pentamer with a glucose inserted into the center of the oligochitin backbone. Although a pentameric LCO with mannose at its reducing terminus was recently identified in *R. tropici* [16], *R. fredii* synthesizes the first example of an LCO with a neutral sugar sandwiched between *N*-acetyl-D-glucosamine residues. In accordance with established nomenclature [16], we have designated the novel glucose-containing LCO as NodRf-V-Glc($C_{18:1}$, MeFuc). All attempts to chromatographically purify it failed, so we subjected it to chitinase digestion. Subsequent analysis by MALDI confirmed cleavage at both penultimate glycosidic bonds, a reaction which yielded vaccenoyl chitobiose and fucosylated chitobiose, as well as small amounts of the corresponding chitotrioses. Since the glucose-containing LCO produced both chitobiose breakdown

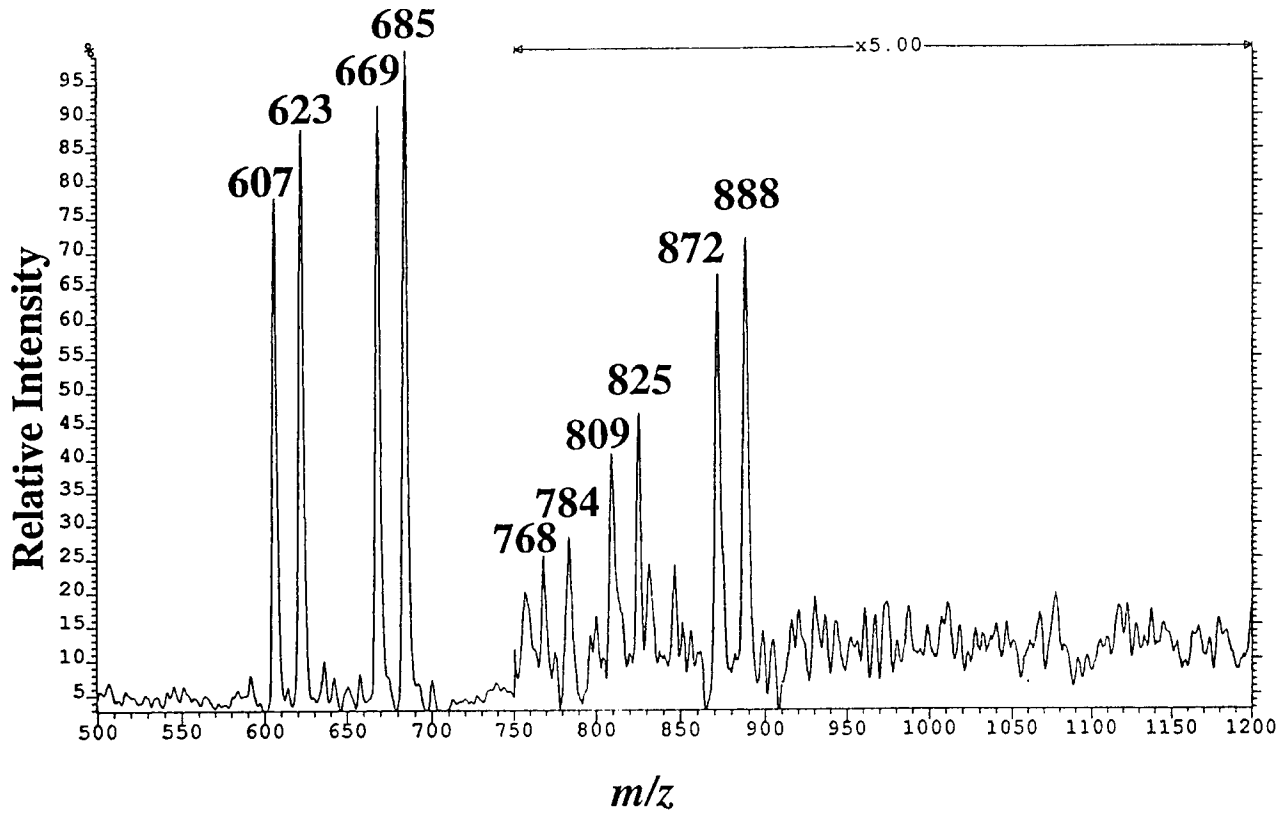

Fig. 5. MALDI mass spectrum of the chitinase hydrolytic products from the HPLC fraction containing NodRf-V-Glc($C_{18:1}$, MeFuc). See text for analysis of peak assignments.

products, the bond between *N*-acetyl-D-glucosamine and glucose also is sensitive to chitinase and is therefore likely to be a β-linkage as illustrated in Fig. 1b.

We have not yet examined the biological activities of the new LCOs, so we do not know if the differing arrays of compounds contribute to the observed patterns of cultivar specificity [4,5] of the producing strains. Our observations raise another interesting issue as well: how does *R. fredii* insert glucose into the center of the oligochitin backbone of one of its LCOs? Chitooligomers of rhizobia are synthesized by the product of the *nodC* gene, which is known to be a *N*-acetylglucosaminyl transferase. This enzyme employs UDP-*N*-acetylglucosamine as substrate to lengthen the oligochitin chain [17–20]. It seems likely that the *R. fredii* enzyme is able to replace this intermediate with UDP-glucose, the normal substrate for the synthesis of β-glucans [21]. Alternatively, another transferase may be involved in the assembly of the novel LCO. Coordinated molecular analysis of strains USDA191 and USDA257 should allow these possibilities to be resolved in the future.

Acknowledgements: This research was supported by competitive research grants from the United States Department of Agriculture (S.G.P. and H.B.K.), by the European Communities BIOTECH Program, as part of the 'Project of Technical Priority', and by the Centre National de la Recherche Scientifique (J.-C.P.). We thank Region Midi Pyrénées for assistance in acquiring the mass spectrometers and Prof. Dietrich Werner of Philipps University, Marburg, Germany, for facilitating preparation of the manuscript. This is Journal Series No. 12527 of the Missouri Agricultural Experiment Station.

References

[1] Pueppke, S.G. (1996) Crit. Rev. Biotechnol. 16, 1–51.
[2] Lie, T.A. (1978) Ann. Appl. Biol. 88, 462–465.
[3] Gibson, A. (1968) Aust. J. Agric. Res. 19, 907–918.
[4] Keyser, H.H., Bohlool, B.B., Hu, T.S. and Weber, D.F. (1982) Science 215, 1631–1632.
[5] Heron, D.S. and Pueppke, S.G. (1984) J. Bacteriol. 160, 1061–1066.
[6] Dénarié, J., Debellé, F. and Rosenberg, C. (1992) Annu. Rev. Microbiol. 46, 497–531.
[7] Bec-Ferté, M.-P., Krishnan, H.B., Promé, D., Savagnac, A., Pueppke, S.G. and Promé, J.-C. (1994) Biochemistry 33, 11782–11788.
[8] Heron, D.S., Érsek, T., Krishnan, H.B. and Pueppke, S.G. (1989) Mol. Plant-Microbe Interact. 2, 4–10.
[9] Meinhardt, L.W., Krishnan, H.B., Balatti, P.A. and Pueppke, S.G. (1993) Mol. Microbiol. 9, 17–29.
[10] Ciucanu, I. and Kerek, F. (1984) Carbohydr. Res. 131, 209–217.
[11] Carlson, R.W., Price, N.P.J. and Stacey, G. (1995) Mol. Plant-Microbe Interact. 7, 684–695.
[12] Heidstra, R. and Bisseling, T. (1996) New Phytol. 133, 25–43.
[13] Firmin, J.L., Wilson, K.E., Carlson, R.W., Davies, A.E. and Downie, J.A. (1993) Mol. Microbiol. 10, 351–360.
[14] Sanjuan, J., Carlson, R.W., Spaink, H.P., Bhat, R., Barbour, M.W., Glushka, J. and Stacey, G. (1992) Proc. Natl. Acad. Sci. USA 89, 8789–8793.
[15] Carlson, R.W., Sanjuan, J., Bhat, R., Glushka, J., Spaink, H.P., Wijfjes, A.H.M., Van Brussel, A.A.N., Stokkermans, T.J.W., Peters, K. and Stacey, G. (1993) J. Biol. Chem. 268, 18372–18381.
[16] Folch-Mallol, J.L., Marroquí, S., Sousa, C., Manyani, H., López-Lara, I.M., van der Drift, K.M.G.M., Haverkamp, J., Quinto, C., Gil-Serrano, A., Thomas-Oates, J., Spaink, H.P. and Megías, M. (1996) Mol. Plant-Microbe Interact. 9, 151–163.

[17] Geremia, R.A., Mergaert, P., Geelen, D., van Montagu, M. and Holsters, M. (1994) Proc. Natl. Acad. Sci. USA 91, 2669–2673.
[18] 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) Mol. Microbiol. 13, 821–831.
[19] Inon de Iannino, N., Pueppke, S.G. and Ugalde, R.A. (1995) Mol. Plant-Microbe Interact. 8, 292–301.
[20] Mergaert, P., D'Haeze, W., Geelen, D., Promé, D., van Montagu, M., Geremia, R.A., Promé, J.-C. and Holsters, M. (1995) J. Biol. Chem. 270, 29217–29233.
[21] Breedveld, M.W. and Miller, K.J. (1994) Microbiol. Rev. 58, 145–161.