A broad-host-range endosymbiont, Sinorhizobium sp. NGR234 is a component of several legume-symbiont model systems; however, there is little structural information on the cell surface glycoconjugates. NGR234 cells in free-living culture produce a major rough lipopolysaccharide (LPS, lacking O-chain) and a minor smooth LPS (containing O-chain), and the structure of the lipid A components was investigated by chemical analyses, mass spectrometry, and NMR spectroscopy of the underivatized lipids A. The lipid A from rough LPS is heterogeneous and consists of six major bisphosphorylated species that differ in acylation. Pentaacetyl species (52%) are acylated at positions 2, 3, 2’, and 3’; tetraacetyl species (46%) lack an acyl group at C-3 of the proximal glucosamine. In contrast to Rhizobium etli and Rhizobium leguminosarum, the NGR234 lipid A contains a bisphosphorylated β-(1’ − 6)-glucosamine disaccharide, typical of enterobacterial lipid A. However, NGR234 lipid A retains the unusual acylation pattern of R. etli lipid A, including the presence of a distal, amide-linked acyloxyacyl residue containing a long chain fatty acid (LCFA) (e.g. 29-hydroxytriacontanoate) attached as the secondary fatty acid. As in R. etli, a 4-carbon fatty acid, β-hydroxybutyrate, is esterified to (ω – 1) of the LCFA forming an acyloxyacyl residue at that location. The NGR234 lipid A lacks all other ester-linked acyloxyacyl residues and shows extensive heterogeneity of the amide-linked fatty acids. The N-acetyl heterogeneity, including unsaturation, is localized mainly to the proximal glucosamine. The lipid A from smooth LPS contains unique triacyl species (20%) that lack ester-linked fatty acids but retain bisphosphorylation and the LCFA-acyloxyacyl moiety. The unusual structural features shared with R. etli/R. leguminosarum lipid A may be essential for symbiosis.

Structural Characterization of the Lipid A Component of Sinorhizobium sp. NGR234 Rough and Smooth Form Lipopolysaccharide

DEMONSTRATION THAT THE DISTAL AMIDE-LINKED ACYLOXYACYL RESIDUE CONTAINING THE LONG CHAIN FATTY ACID IS CONSERVED IN RHIZOBIUM AND SINORHIZOBIUM SP.*[S]

Received for publication, October 13, 2002, and in revised form, November 15, 2002
Published, JBC Papers in Press, November 26, 2002, DOI 10.1074/jbc.M210491200

Seshu K. Gudlavalleti and Lennart S. Forsberg§
From the Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602

The family Rhizobiaceae includes the Rhizobium and Sinorhizobium, Gram-negative bacteria able to form nitrogen-fixing symbioses with legumes in a host-specific manner. Sinorhizobium sp. NGR234 is a fast growing, broad-host-range symbiont able to colonize a diverse range of commercially important legumes (1, 2), including both indeterminate and determinate nodule-forming hosts. Partly because of its agricultural role, the molecular genetics of NGR234 are of interest, and the symbiotic plasmid was recently sequenced (3). However, there is little complementary structural information on the cell surface macromolecules or the alterations that occur in these molecules during symbiotic infection and bacteroid differentiation.

Lipopolysaccharides (LPS)1 are the major structural and antigenic components of the rhizobial outer membrane (4–8) and are proposed to contribute to the biochemical processes that result in symbiotic infection (8–16). Rhizobial LPS structural mutants typically yield phenotypes with underdeveloped nodules (Nd− phenotype) in which nitrogen fixation is absent or diminished (Fix−) (5, 9, 15, 17–23).

The lipid A moieties of rhizobial LPS are of interest, because of their highly unusual structures and, by analogy to enterobacterial lipid A, because of their essential role in maintaining cell viability and membrane integrity. The structure of the Rhizobium etli-Rhizobium leguminosarum common lipid A-core region was recently elucidated from laboratory cultured cells (24–27) and was found to have an entirely different structure from the typical lipid A-core of enterobacterial LPS. The lipids A of R. etli-R. leguminosarum lack phosphate and instead have trisaccharide backbones containing a distal galacturonic acid residue in α1,4-linkage to glucosamine, linked β1,6 to a proximal 2-amino glucuronic acid residue (26). The latter is generated from normal glucosamine by a membrane-associated dehydrogenase (24). The acylation pattern of R. etli-R. leguminosarum lipid A also differs in a number of key respects from that of enteric bacteria (24, 26, 27). These features include extensive heterogeneity of the ester and amide-linked fatty acids, particularly on the proximal 2-amino glucuronate residue, the absence of ester-linked acyloxyacyl groups, and the presence of a single amide-linked acyloxyacyl residue, containing a unique long chain fatty acid (LCFA, e.g. 27-hydroxyoctaocoenoic acid).

1 The abbreviations used are: LPS, lipopolysaccharide; R-LPS, rough LPS; S-LPS, smooth LPS; Kdo, 3-deoxy-D-manno-2- octulosonic acid; LCFA, long chain fatty acid; 27-OH-28:0, 27-hydroxyoctacosanonic acid; GC-MS, gas-liquid chromatography-mass spectrometry; TMS, trimethylsilyl; MALDI-TOF, matrix assisted laser desorption ionization-time of flight; COSY, 1H-1H correlation spectroscopy; TOCSY, total correlation spectroscopy; HMBC, heteronuclear multiple bond coherence spectroscopy; NOE, nuclear Overhauser effect; EFS, extracellular polysaccharides; NOESY, NOE spectroscopy; HISOQC, heteronuclear single quantum coherence spectroscopy.

*[S] The on-line version of this article (available at http://www.jbc.org) contains additional figures.

† Present address: Emory University School of Medicine, Atlanta, GA 30303.
§ To whom correspondence should be addressed. Tel.: 706-542-4401; Fax: 706-542-4412; E-mail: SFORSB@ccrc.uga.edu.
Sanoic acid, carried as the secondary acyl group. These unique structures can be rationalized by considering their possible role in bacteroid survival and symbiotic infection. It has been proposed, for example, that replacement of phosphate with less acidic, negatively charged residues could help attenuate stimulation of the plant defense response to the invading rhizobia (27, 28). The occurrence of the C28 LCFA is particularly interesting, because its presence (but not specific location) has been demonstrated in a variety of bacteria that survive within intracellular host membrane-derived compartments, including all of the Rhizobiaceae studied to date, the plant pathogen Agrobacterium, and the phylogenetically related intracellular animal pathogen Brucella (29, 30). The C28 LCFA is long enough to span the entire lipid bilayer and has been suggested to promote membrane stability during the critical stages of symbiotic infection, when the bacterial and plant membranes appear to be in close proximity (i.e., during endocytosis-symbiosome formation and during the synchronized cell division of the bacterium-symbiosome (31)).

In contrast to *R. etli*-*R. leguminosarum*, there is no detailed information on the structures of any of the *Sinorhizobium* LPS. Compositional studies have indicated that sinorhizobial LPS core regions are substantially different from those of *R. etli* (8, 12, 32, 33) or from enteric bacteria and that the O-chains are often expressed in trace amounts or consist of simple homopolymers (8, 12), at least on vegetative cells. A single study on the lipid A moieties from the *Sinorhizobium meliloti* mutant 10406 indicated the presence of a bisphosphorylated glucosamine disaccharide backbone; however, detailed structures, component heterogeneity, and the locations of some fatty acids, including the LCFA, were not characterized (34). Interestingly, acyloxyacyl residues were not detected in the *S. meliloti* mutant lipid A, possibly because effective solvents for NMR analysis of under-derivatized lipids A were unavailable at the time.

In this report we describe structural analysis of the lipid A moiety from the LPS produced by *Sinorhizobium* sp. NGR234 grown in vegetative culture. Our results demonstrate several structural features that are shared with *Rhizobium etli*- *Rhizobium leguminosarum* lipid A, which may be important in establishing or maintaining symbiosis. These features include a remarkable similarity in the type and location of acyloxyacyl residues containing the LCFA, when compared with *R. etli*-*R. leguminosarum* (24). In marked contrast to *R. etli* lipid A, the NGR234 lipid A also shares certain features with the entero-bacterial lipids A, constituting a “hybrid” structure that may prove useful in endotoxin-elicitor studies. Structural information on the NGR234 lipid A will provide additional insight on the biosynthesis of rhizobial and sinorhizobial lipids A and facilitate comparisons with bacteroid-derived and other structurally altered lipids A.

**EXPERIMENTAL PROCEDURES**

**Growth of Bacteria—Agar slants of *Sinorhizobium* sp. strain NGR234 were provided by Dr. Peter van Berkum at the National Rhizobium Germplasm Collection Center in Beltsville, MD. The bacteria were transferred to liquid media (tryptone/yeast extract supplemented with CaCl₂) and grown in fermentor culture at 28 °C as described previously for related species (8, 35). The cells were harvested by centrifugation at late log phase (A₅₇₀ = 2.70) yielding 500 g (wet weight) of cells/100 liters of culture.**

**Isolation of LPS and Purification of Lipid A—LPSs were extracted by the hot phenol/water procedure (5, 36) and analyzed by deoxycholate-PAGE (37–39). Water layer extracts containing LPS were dialyzed and treated sequentially with ribonuclease, deoxyribonuclease, and proteinase K (5) and then redialyzed and subjected to size exclusion chromatography under dissociative conditions (0.25% sodium deoxycholate, 0.2 M NaCl, 1.0 mM EDTA, 10 mM Tris, pH 9.2) on a column of Sephadex G-150 (2.2 × 100 cm). This procedure separates the R-LPS from the S-LPS. 60 mg of water layer extract yielded ~23 mg of R-LPS and 1 mg of S-LPS per run. Portions (2–3 mg) of the G-150 purified R-LPS were subjected to further purification on Superose-12 HR 10/30 (Amersham Biosciences) under associative conditions (50 mM ammonium formate) at 0.3 ml/min.**

The R-LPS was hydrolyzed in 10 mM sodium acetate buffer, pH 4.5, containing 1% SDS, following a published procedure (40), to yield the free lipid A (R-lipid A₆₅₀,ω). For comparison, lipid A was also isolated by hydrolyzing R-LPS in 1.0% acetic acid (pH 2.8, 105 °C, 24 h), and the precipitated lipid A (R-lipid A₆₈₀,ω) was recovered by ultracentrifugation at 160,000 × g for 1 h (16). The resulting lipids A were analyzed by TLC, MALDI-TOF mass spectrometry, and GC-MS analysis of derivatives as described below. Selected lipid A species were isolated by preparative TLC (described below) and analyzed by two-dimensional NMR spectroscopy, MALDI-TOF MS, and GC-MS derivatives.

**Glycosyl Composition Analysis—**Glycosyl compositions of LPS and the intact and de-O-acylated lipid A samples were determined by GC-MS analysis (electron impact) of the TMS methyl glycosides (41) using a 30-m DB-5 fused silica capillary column (J & W Scientific). Neutral and amino sugars were also determined by GC-MS analysis of the alditol acetates using a 30-m SP2330 capillary column (Supelco) (42).

**Fatty Acid and N-Acylglucosamine Analysis—**Total fatty acids were released by hydrolysis of lipid A in 4 M HCl at 100 °C for 4 h (43). The liberated fatty acids were converted to methyl esters using methanolic 1 M HCl at 80 °C for 2 h, followed by trimethylsilylation and analysis by GC-MS using the DB-5 column. Esteryl-linked fatty acids were selectively released by de-O-acylation of intact lipid A with sodium methoxide (0.5 M) at 35 °C for 16 h (26). Labeled fatty acids were then methyl-esterified, trimethylsilylated, and analyzed by GC-MS. Amide-linked fatty acids were analyzed by hydrolyzing the de-O-acylated lipid A in 4 M HCl; the liberated fatty acids were then analyzed as the TMS methyl esters by GC-MS. Amide-linked fatty acids were also determined by mild methanolysis of the intact lipid A, followed by preparative TLC and GC-MS analysis of the N-acetylglucosamine TMS methylglycoside derivatives (44). Chemical analysis of amide-linked acylxyacyl groups was performed on lipid A₆₅₀,ω by methylation with silver oxide/methyl iodide (Krasca methylation), following the procedures of Wolgast and Rietsch (45).

**Thin Layer Chromatography—**TLC of lipid A was performed on Silica Gel 60 high performance TLC plates (Merck) using two solvent systems as follows: A, chloroform/methanol/water/triethylamine (30:12:2.01.1, v/v); and B, chloroform, pyridine, 88% formic acid, methanol, water (60:55:10:5.2, v/v) (27). Lipids A were visualized by spraying with 15% ethanolic sulfuric acid and charring, and by a molybdenum blue spray reagent (Sigma). Preparative TLC was performed in the same manner except bands were localized by misting the plate lightly with water and holding the plate under UV light, under which conditions the hydrophobic lipid A bands were readily visible. Individual bands containing lipid A were removed by scraping, and the lipid A was solubi-lized in chloroform/isopropanol alcohol/water (5:3:0.25 v/v). Silica gel and particulates were removed by passage through a small column (5 × 8 mm) of Sephadex LH-20 equilibrated in the same solvent.

**NMR Spectroscopy—**H and two-dimensional homonuclear and heteronuclear spectra of the intact lipid A₆₅₀,ω were recorded at 30 °C on a Varian Inova 500 NMR spectrometer using standard Varian software. The lipid A₆₅₀,ω was dissolved and analyzed in the ternary solvent CDCl₃/CD3OD/D₂O (2:3:1 v/v) (24, 46), yielding clear solutions at ~3 mg/ml; spectra were referenced to internal tetrathymisilane (0.00 ppm). COSY data were recorded in the absolute value mode with a 5.0-kHz spectral width, collecting 512 increments at 16 scans per increment. H-13C HSQC was performed with a mixing time of 80 ms and 2 sets of 512 time increments at 24 scans per increment. Phase-sensitive NOESY was recorded with a 500-mas mixing time using 2 arrays of 256 increments at 80 scans per increment. H-13C HMQC spectra was performed with an acquisition time of 0.199 s by collecting 2 arrays of 200 increments at 80 scans per increment. H-13C HMBC correlation spectra were recorded without 13C decoupling on the R-Lipid A₆₅₀,ω in the associated solvent system, demonstrating observed H-13C decoupling of intact lipid A₆₅₀,ω and lipid A₆₈₀,ω obtained by dissolving the lipid A in D₂O containing 2% sodium deoxycholate and 5 mM EDTA; pH adjustments were made with triethylamine. Spectra were recorded on a Varian 300 instrument at 121.51 MHz, and 13C chemical shifts were recorded relative to an external standard of 85% phosphoric acid at 0.00 ppm.

**Mass Spectrometry—**MALDI-TOF mass spectrometry was performed on a Kratos Analytical Compact SEQ instrument using delayed extraction, in both positive and negative ion modes. Samples were desorbed with a nitrogen laser (λ = 337 nm) and extraction voltage of 20 kV.
Lipid A samples were dissolved in CHCl3:isoamyl alcohol/water (5:3:0.25 v/v) and desalted by treating with OnGuard-H cation exchange resin (H⁺) (Dionex) for 20 min with shaking. The resin was removed by centrifugation (12,000 × g), and the supernatant containing lipid A was dried, adjusted to a concentration of 2 μg/μl in fresh solvent, and mixed in a 1:1 ratio with matrix solution (2,4,6-trihydroxyacetophenone, 0.5 M in methanol). 1 μl was applied to the MALDI stage. Mass calibration was performed with maltosiosecarbohydrates (positive ion mode) and with R. etli CE3 lipid A (negative ion mode). The predicted molecular weights of the various lipid A species were calculated from the average incremental masses of subunits, based on the atomic weights of the elements: 2-amino-2-deoxyhexose, 161.1577; phosphate, 79.9799; β-hydroxybutyrate, 86.0904; 3-OH-14:0, 228.3599; 3-OH-16:0, 254.4130; 3-OH-17:0, 282.4699; 3-OH-18:0, 294.4777; 3-OH-18:1, 280.4509; 3-OH-19:0, 296.4935; 3-OH-19:1, 276.4351; 3-OH-15:0, 218.3470; 3-OH-16:0, 246.3910; 3-OH-17:0, 274.4379; 3-OH-18:0, 282.4667; 3-OH-18:1, 268.4399; 3-OH-18:2, 278.4351; 3-OH-19:0, 296.4935; 3-OH-19:1, 294.4777; 3-OH-28:0, 422.7355; 3-OH-29:0, 450.7891; and free reducing end, 18.0153.

**Analytical Procedures**—Total phosphorus was determined by a colorimetric procedure (47). Column eluents were monitored by colorimetric assays for neutral carbohydrate (41) and Kdo (48) and by refractive index detection.

**RESULTS**

**Purification of LPS**—The major LPS produced by vegetative cultures of Sinorhizobium sp. NGR234 is a rough LPS (R-LPS), which lacks the polysaccharide O-chain (8, 49). This is typical of many Sinorhizobium strains examined to date (8, 9, 14, 32, 35) and contrasts with R. leguminosarum and R. etli, which synthesize a major smooth LPS (S-LPS, containing O-chain) when grown as free-living cultures (14, 16). In the case of strain NGR234, the majority of the total LPS (consisting of the major R-LPS and a minor S-LPS) was recovered in the water layer during hot phenol/water extraction. In Fig. 1, fractionation of the water layer by size exclusion chromatography under dissociative conditions afforded separation of the R-LPS and S-LPS, as well as other cell surface components, the capsular polysaccharides (K-antigens), extracellular polysaccharides (EPS), cyclic glucans, and derived fragments. The recovery of total LPS (R-LPS + S-LPS) was 340 mg per 40 g of dry cell weight (compare R. etli strain CE3, which yields ~600 mg of LPS from the same amount of cells (25)). Of the total LPS, over 96% consisted of R-LPS. Final purification of R-LPS was achieved by size exclusion chromatography under associative conditions on Superose-12 (Fig. 1). This step removed a low molecular weight population of K-antigen/EPS fragments which co-migrated with the R-LPS during dissociative chromatography.

**Isolation and Initial Characterization of Sinorhizobium sp. NGR234 Lipid A**—In previous studies on the lipid A from R. etli and the closely related R. leguminosarum, mild hydrolysis in 1% acetic acid was routinely used to cleave the core-lipid A ketosidic linkage, and the free lipid A was isolated without extensive degradation (26, 30). This procedure was applied to NGR234 R-LPS, and the resulting R-lipid AHOAc was analyzed by TLC (see figures in the Supplemental Material). The mobility of the major components suggested the presence of monophosphorylated species, which typically migrate near the solvent front in this solvent. Composition analysis of lipid AHOAc (Table I), indicated a total phosphate/glucosamine ratio of ~1:2, also consistent with monophosphorylated species.

The presence of trace components having low Rf values suggested that some bisphosphorylated or other more highly phosphorylated lipid A species might be present. This suggested that the 1% HOAc treatment may have caused substantial cleavage of labile glycosidic phosphate or other polar head groups and that more highly phosphorylated lipid A species may be the true biosynthetic products. Additional lipid A was therefore isolated using the mild hydrolytic procedure of Caroff et al. (40) (pH 4.5, 1% SDS), and the products (R-lipid ASDS) were analyzed and compared (Table I and see figures in the Supplemental Material). This procedure yielded a markedly different lipid A profile in which the most abundant components had lower Rf values, typical of those reported for various bisphosphorylated lipids A (45, 50). Analysis for phosphate (Table I) indicated a molar ratio of ~1:1 relative to glucosamine, also suggesting that R-lipid AHOAc contained bisphosphorylated species. In addition to elevated levels of phosphate, GC-MS analysis indicated that R-lipid ASDS had elevated levels of several fatty acids (3-OH-14:0, 27-OH-28:0, and 29-OH-30:0) compared with R-lipid AHOAc (Table I).

Analysis of lipid AHOAc and lipid ASDS for total fatty acids indicated a high degree of acyl group heterogeneity. As with R. etli, the Sinorhizobium sp. NGR234 R-lipid A contained the unusual LCFA 27-OH-28:0; in addition to the common series of 3-hydroxy-14- to 18-carbon fatty acids. The NGR334 lipid A also contained substantial amounts of 29-OH-30:0, and several unsaturated hydroxy fatty acids, not found in R. etli/R. leguminosarum lipids A. In contrast to the lipid A of R. etli/R. leguminosarum, glucosamine was the only carbohydrate detected by GC-MS analysis of the alditol acetates and TMS-methyl glycosides; 2-amino-glucosate and galactouronic acid, both components of the R. etli lipid A (26), were not detected. Mild hydrolysis of R-lipid ASDS in 0.2 M HCl, conditions which
The chemical shifts of both signals are characteristic of these two signals, in addition to a new signal at the TMS methyl glycosides with response factor correction from the response factor of authentic 3-hydroxymyristic acid. The location of double bonds in unsaturated fatty acids was not investigated.

Phosphate was determined by a colorimetric procedure (47).

### TABLE I

| Component          | Amount in mol/2 mol GlcN |
|--------------------|--------------------------|
| **Lipid AHOAc**    |                          |
| 3-OH-12.0°         | Tr                       |
| 3-OH-12.0°         | 0.03                     |
| 3-OH-14.0°         | 1.0                      |
| 3-OH-15.0°         | 0.05                     |
| 3-OH-16.0°         | 0.42                     |
| 3-OH-17.0°         | 0.13                     |
| 3-OH-18.0°         | 0.92                     |
| 3-OH-19.0°         | Tr                       |
| 3-OH-18.1          | 0.26                     |
| 3-OH-18.2          | 0.02                     |
| 3-OH-19.1          | 0.13                     |
| 3-OH-20.1          | 0.02                     |
| 27-OH-28:0         | 0.71                     |
| 29-OH-30:0         | 0.28                     |
| GlcNH₂             | 2                        |
| Phosphate⁻         | 0.78                     |
| **Lipid ASDS**     |                          |
| 3-OH-12.0°         | Tr                       |
| 3-OH-12.0°         | 0.06                     |
| 3-OH-14.0°         | 1.57                     |
| 3-OH-15.0°         | 0.07                     |
| 3-OH-16.0°         | 0.45                     |
| 3-OH-17.0°         | 0.16                     |
| 3-OH-18.0°         | 0.90                     |
| 3-OH-19.0°         | Tr                       |
| 3-OH-18.1          | 0.25                     |
| 3-OH-18.2          | 0.02                     |
| 3-OH-19.1          | 0.11                     |
| 3-OH-20.1          | Tr                       |
| 27-OH-28:0         | 0.82                     |
| 29-OH-30:0         | 0.31                     |
| GlcNH₂             | 2                        |
| Phosphate⁻         | 1.72                     |

3-Hydroxy fatty acids were analyzed as the TMS methyl esters by GC-MS. Nanomoles were calculated from the GC-MS total ion current peak areas, using docosanoic acid as internal standard with correction to the response factor of authentic 3-hydroxymyristic acid. The location of double bonds in unsaturated fatty acids was not investigated.

Phosphate was determined by a colorimetric procedure (47).

**Analysis of Phosphate Residues**—The number and type of phosphate residues in lipid AHOAc, and lipid ASDS from R-LPS was investigated by 31P NMR of the intact lipids. A. In Fig. 2A, R-lipid AHOAc yielded two sharp signals when analyzed at pH 10.6. The chemical shifts of both signals are characteristic of ester-linked, unsubstituted monophosphates. The signal at δ 4.804 was attributed to ester-linked phosphate on molecules that are lacking some of their O-linked fatty acids; loss of ester-linked fatty acids has been shown to produce 31P signal heterogeneity of this type (34). The majority of partially de-acylated species was probably generated during 1% HOAc hydrolysis and is a major component of R-lipid AHOAc, as shown below by MALDI-TOF and chemical analyses. The signal at δ 4.090 arises from fully (or more highly) O-acylated species. The chemical shifts of both signals are in close agreement with the two signals observed for the 4-monophosphate ester of R. mellioti 10406 lipid A (34). In Fig. 2B, R-lipid ASDS retained these two signals, in addition to a new signal at δ 2.352, characteristic of a glycosidically linked phosphomonooester (34, 51). As shown below, lipid ASDS contains a higher proportion of fully O-acylated species than does lipid AHOAc, accounting for the decreased intensity of the δ 4.804 signal relative to the δ 4.121 signal. In Fig. 2C, all signals showed a pH-sensitive chemical shift, typical of unsubstituted monophosphate residues, both ester- and glycosidically linked.

The location of phosphate in lipid ASDS was unequivocally determined by 1H-31P HMBC (Fig. 2D), using a ternary solvent system (46). The lipid ASDS, determined by chemical analysis and mass spectrometry to consist mainly of bisphosphorylated species (see below), revealed 3-bond couplings indicating two sites of phosphate attachment, at C-1 of the proximal GlcN residue (glycosidic attachment) and at C-4' of the distal GlcN residue. In addition, a 4-bond coupling between the glycosidic phosphorous and H-2 of the proximal GlcN residue was observed. The identities of protons H-1 (δ 5.44), H-2 (δ 4.12), and H-4' (δ 4.17 and δ 4.15) were determined from 1H-1H COSY and TOCSY experiments (described below). The C-1 phosphorus signal was located at δ −2.21, and the C-4' phosphorus was found at δ −0.65 (major) and δ −0.22 (minor). Heterogeneity of the H-4'-P coupling was again attributed to partial de-O-acylation of lipid ASDS, which appeared to occur to a minor extent in the solvent during the analysis. The downfield location of the C-4' phosphorus signal relative to C-1 phosphorus is in agreement with previous studies (46) on enterobacterial lipid A using this ternary solvent; however, the absolute values reported here for NGR234 lipid ASDS are somewhat different, probably due to differences in temperature, pH, solvent ratios, and acylation.

**Determination of Ester- and Amide-linked Fatty Acids**—Ester-linked fatty acids were released from R-lipid AHOAc and analyzed by GC-MS (Table II). The 3-OH-14:0 fatty acid was exclusively ester-linked, accounting for 54% of the base-labile fatty acids. All other fatty acids were acylated species. The δ 4.090 signal was attributed to ester-linked phosphate on molecules that are lacking some of their O-linked fatty acids; loss of ester-linked fatty acids has been shown to produce 31P signal heterogeneity of this type (34). The majority of partially de-acylated species was probably generated during 1% HOAc hydrolysis and is a major component of R-lipid AHOAc, as shown below by MALDI-TOF and chemical analyses. The signal at δ 4.090 arises from fully (or more highly) O-acylated species. The chemical shifts of both signals are in close agreement with the two signals observed for the 4-monophosphate ester of R. mellioti 10406 lipid A (34). In Fig. 2B, R-lipid ASDS retained these two signals, in addition to a new signal at δ 2.352, characteristic of a glycosidically linked phosphomonooester (34, 51). As shown below, lipid ASDS contains a higher proportion of fully O-acylated species than does lipid AHOAc, accounting for the decreased intensity of the δ 4.804 signal relative to the δ 4.121 signal. In Fig. 2C, all signals showed a pH-sensitive chemical shift, typical of unsubstituted monophosphate residues, both ester- and glycosidically linked.

The location of phosphate in lipid ASDS was unequivocally determined by 1H-31P HMBC (Fig. 2D), using a ternary solvent system (46). The lipid ASDS, determined by chemical analysis and mass spectrometry to consist mainly of bisphosphorylated species (see below), revealed 3-bond couplings indicating two sites of phosphate attachment, at C-1 of the proximal GlcN residue (glycosidic attachment) and at C-4' of the distal GlcN residue. In addition, a 4-bond coupling between the glycosidic phosphorous and H-2 of the proximal GlcN residue was observed. The identities of protons H-1 (δ 5.44), H-2 (δ 4.12), and H-4' (δ 4.17 and δ 4.15) were determined from 1H-1H COSY and TOCSY experiments (described below). The C-1 phosphorus signal was located at δ −2.21, and the C-4' phosphorus was found at δ −0.65 (major) and δ −0.22 (minor). Heterogeneity of the H-4'-P coupling was again attributed to partial de-O-acylation of lipid ASDS, which appeared to occur to a minor extent in the solvent during the analysis. The downfield location of the C-4' phosphorus signal relative to C-1 phosphorus is in agreement with previous studies (46) on enterobacterial lipid A using this ternary solvent; however, the absolute values reported here for NGR234 lipid ASDS are somewhat different, probably due to differences in temperature, pH, solvent ratios, and acylation.
Sinorhizobium sp. Strain NGR234 Lipid A Structure

Table II

Relative recovery of ester- and amide-linked fatty acids during de-O-acetylation of lipid \( A_{\text{HOAc}} \) from Sinorhizobium sp. NGR234 rough-LPS

| Fatty acid | Ester-linked\(^a\) | Amide-linked\(^b\) | Total |
|-----------|-------------------|--------------------|-------|
|           | Amount in nmol/mg lipid A (mol/2 mol GlcN) and the mol % | Amount in nmol/mg lipid A (mol/2 mol GlcN) and the mol % |          |
| 3-OH-12:0 | 4.1 (Tr)          | 0.2                | 4.1 (Tr) |
| 3-OH-13:0 | 40.1 (0.06)       | 2.2                | 40.1 (0.06) |
| 3-OH-14:0 | 958 (1.41)        | 53.6               | 970.7 (1.43) |
| 3-OH-15:0 | 52.1 (0.08)       | 2.9                | 52.1 (0.08) |
| 3-OH-16:0 | 11.3 (0.02)       | 0.6                | 274.3 (0.41) |
| 3-OH-17:0 | 1.2 (Tr)          | 0.1                | 95.1 (0.14) |
| 3-OH-18:0 | 9.0 (Tr)          | 0.5                | 635.6 (0.93) |
| 3-OH-19:0 | 2.7 (Tr)          | 0.2                | 13.7 (0.02) |
| 3-OH-18:1 | 0.7 (Tr)          | 0.04               | 113.2 (0.16) |
| 3-OH-18:2 | —                 | —                  | 56.1 (0.08) |
| 3-OH-19:1 | —                 | —                  | 175.0 (0.26) |
| 3-OH-20:1 | —                 | —                  | 12.3 (0.02) |
| 27-OH-28:0 | 503 (0.74)       | 28.2               | 503 (0.74) |
| 29-OH-30:0 | 205 (0.30)       | 11.5               | 205 (0.30) |

\(^a\) 3-Hydroxy fatty acids were analyzed as the TMS methyl esters by GC-MS analysis. Nanomoles were calculated from the GC-MS total ion current peak areas, using docosanoic acid as internal standard with correction to the response factor of authentic 3-hydroxymyristic acid.

\(^b\) Amount in nmol/mg lipid A (mol/2 mol GlcN), and the mol %.

Fatty acids, and the two LCFAAs accounted for 40%. Excluding LCFAA, 3-OH-14:0 composed 88% of the ester-linked fatty acids. Analysis of the de-O-acetylated lipid A precipitate (DOA-lipid \( A_{\text{HOAc}} \)) indicated that 3-OH-18:0 composed 46% of the amide-linked fatty acids and that unsaturated 3-hydroxy fatty acids composed 26%.

The N-acyl heterogeneity was also assessed by GC-MS analysis of the N-acetylglucosamine derivatives, prepared by mild methanolation (44) of R-lipid \( A_{\text{HOAc}} \). Four different N-fatty acyl derivatives of glucosamine were detected (ratios from GC-MS ion current peak areas): GN-3-OH-16:0 (56); GN-3-OH-17:0 (32); GN-3-OH-18:0 (72), and GN-3-OH-18:1 (0.08). Other derivatives were not detected; however, these ratios are in reasonably good agreement with the amide-linked fatty acid ratios determined by total acid hydrolysis of the DOA-lipid \( A_{\text{HOAc}} \) (Table II).

Amide-linked acylxoyacyl residues were investigated by the Kraska methylation procedure as described (43). Two acylxoyacyl derivitatives were detected, consistent with those of 27-O-(\( \beta \)-hydroxybutyroxy)-C28:0 and 29-O-(\( \beta \)-hydroxybutyroxy)-C30:0. The EI spectrum of the former product was identical to that identified previously in \( R. \) etli lipid A (26); the spectrum of the latter showed an increase of 28 mass units for ions containing the LCFA moiety, i.e. \( m/z \) 437 \( \rightarrow \) 465, and \( m/z \) 453 \( \rightarrow \) 481. The detection of these products indicated that the \( \omega -1 \) hydroxyl groups of the LCFA are esterified with a \( \beta \)-OH-butyrate residue; as shown below, this substitution is not stoichiometric. No other derivatives were detected, indicating that the base-labile fatty acids (with the exception of LCFA) were not secondary components of amide-linked acylxoyacyl residues. As a positive control, \( E. \) coli lipid A yielded the expected methyl ester derivatives of [14:0:3-0(14:0)] and [14:0:3-0(12:0)], during the Kraska procedure. The failure of the LCFA to yield an acylxoyacyl derivative during Kraska methylation was also observed during the original analysis of the \( R. \) etli lipid A (26).

Characterization of NGR234 Lipid A Species by MALDI-TOF Mass Spectrometry—Analysis of R-lipid \( A_{\text{HOAc}} \) by negative ion MALDI-TOF MS revealed a complex pattern, which was found to consist of 6 major species, all having molecular ions corresponding to monophosphorylated lipids \( A_{\text{MONO}} \) (Fig. 3A). Analysis of R-lipid \( A_{\text{SDS}} \) also revealed 6 major components (Fig. 3B), each 80 mass units higher than the corresponding lipid \( A_{\text{HOAc}} \) species, confirming their identification as bisphosphorylated species (\( A_{\text{BIS}} - F_{\text{BIS}} \)). Table III summarizes the observed molecular ions, proposed composition, and calculated mass of the different lipid \( A_{\text{SDS}} \) species. Comparison of lipid \( A_{\text{HOAc}} \) and lipid \( A_{\text{SDS}} \) (Fig. 3, A and B) also showed that lipid \( A_{\text{HOAc}} \) is enriched in partially de-O-acetylated species (species A–D) in comparison to lipid \( A_{\text{SDS}} \), which has a higher proportion of fully acylated species (E and F). Species A and B each lack a single copy of a LCFA (either 27-OH-28:0 or 29-OH-30:0), resulting in a decrement of 423 (or 451) mass units compared with species C and E, respectively. Species A, C, and D each lack a single 3-OH-14:0 fatty acid (\( \Delta \) mass 226) in comparison to species B, E, and F, respectively. This mass decrement represents the absence of an ester-linked 3-OH-14:0 residue (as shown below, in lipid \( A_{\text{SDS}} \) the loss of this residue occurs specifically at O-3 of the proximal residue, rather than at O-3'). Another incremental difference of 86 mass units exists between species E and F, and between species C and D. This mass difference is attributed to partial substitution by \( \beta \)-hydroxybutyrate, esterified to some of the LCFA acyl chains at the \( \omega -1 \) hydroxyl group. Based on the relative intensities of the MALDI-TOF molecular ions, it was estimated that 96% of the R-lipid \( A_{\text{SDS}} \) components were bisphosphorylated, and \( ~4 \% \) were monophosphorylated. The lipid \( A_{\text{SDS}} \) from R-LPS consisted of pentaacylated species (52%), tetraacylated species (46%), and triacylated species (2%), and 93% of the R-lipid \( A_{\text{SDS}} \) components contained a single LCFA substituent.

The major mono- and bisphosphorylated components were isolated from lipid \( A_{\text{HOAc}} \) and lipid \( A_{\text{SDS}} \) by preparative TLC, and the components were analyzed by negative ion MALDI-TOF MS (Fig. 4). The major upper band from lipid \( A_{\text{HOAc}} \) was enriched in species \( E_{\text{MONO}} \) and \( F_{\text{MONO}} \), and the major lower band from lipid \( A_{\text{SDS}} \) was enriched in species \( E_{\text{BIS}} \) and \( F_{\text{BIS}} \). These results confirm the degree of phosphorylation for both the high and low mobility TLC species.

Localization of the Long Chain Fatty Acyl Substituent—In previous studies (27) of \( R. \) etli lipid A, positive ion MALDI-TOF MS was an effective method for localizing fatty acyl substituents to the distal or proximal glucosamine residues. At high laser power settings, cleavage of the GlcN-\( \beta \) (1-6)-GlcN glycosidic linkage occurred readily and specifically, yielding prompt fragment ions of the \( B^{-} \) type (52) carrying the respective (distal) acyl substituents. When applied to NGR234 R-lipid \( A_{\text{SDS}} \), two distinct families of \( B^{-} \) ions were observed, along with an
attenuation of the parent molecular ions, as shown in Fig. 5. All $B_1^+$ ions had masses corresponding to LCFA attachment, indicating that the LCFA must be located on the distal GlcN residue. The two $B_1^+$ ion families differed by 86 mass units, consistent with the presence or absence of a $\beta$-hydroxybutyrate moiety, which occurs in acyloxyacyl linkage to either of the two LCFAAs (27-OH-28:0 and 29-OH-30:0) as discussed above. Molecular ions ($M + Na$)$^+$ for the intact lipid $A_{\text{SDS}}$ species were also observed (Fig. 5), each 24 mass units higher than the corresponding ($M - H$)$^-1$ ions shown in Fig. 3B, thus confirming the mass assignments for these species. The two $B_1^+$ ion families originate from glycosidic cleavage of the four parent species, C–F. The $B_1^+$ ion family at $m/z$ 1173 originates from parent species C and E, and the $B_1^+$ family at $m/z$ 1273 originates from parent species D and F. Additional $\pm 14$ mass unit heterogeneity arises from partial endogenous O-methylation of some of the $\beta$-hydroxybutyrate residues (sometimes found as 3-methoxybutyrate by GC-MS). In Fig. 5, it is also noted that there are no $B_1^+$ ions having a decrement of 226 mass units compared with the observed $B_1^+$ ions, i.e. ions at $m/z$ 1047 and 947 were not observed. This indicates that there are no $B_1^+$ fragments (distal fragments) lacking a 3-OH-14:0 residue, even though parent molecular species are clearly present which lack this fatty acid (species C and D in Figs. 3B or 5). This in turn indicates that the absence of the 3-OH-14:0 fatty acid in species C and D occurs at a specific location, on the proximal residue (presumably C-3), in the case of lipid $A_{\text{ADP}}$. At lower power settings, the $B_1^+$ cleavage products were not observed, and the ($M + Na$)$^+$ ions displayed increased intensity (not shown).

Additional information on the nature of the LCFA linkage was obtained by further examination of the $B_1^+$ ions. In order to accommodate the mass of a LCFA component in any of the $B_1^+$ ions, the LCFA would need to be esterified as the secondary component of an acyloxyacyl residue, the most likely location being the (distal) amide-linked fatty acid. This is suggested by the following rationale. 1) Both 3-OH-14:0 and LCFA are totally released from lipid A by sodium methoxide or mild hydrazine treatments (Table II), indicating they are both ester-linked. 2) 3-Methoxy fatty acid derivatives were not detected during sodium methoxide-catalyzed de-O-acylation of lipid A, indicating that ester-linked acyloxyacyl residues were not present (43); thus, the LCFA could not be attached as a secondary acyl group to an ester-linked 3-OH-14:0 residue. 3) The methyl ester derivative of 3-(3-hydroxytetradecanoyloxy)-octadecanoic acid, which if present would be released from its amide-linkage during Kraska methylation of lipid $A_{\text{ADP}}$, was not observed. This indicates that 3-OH-14:0 was not esterified as a secondary fatty acid to an amide-linked 3-OH-18:0 residue. 4) The majority of the LCFA is esterified at the $\omega - 1$ hydroxyl with a $\beta$-hydroxy-butyryl residue, precluding attachment of a 14 carbon fatty acid (or any other fatty acid) at this location. These points also rule out the possibility that LCFA could be esterified directly to the 3'-position. All of these points indicate that, on the distal glucosamine residue, the LCFA is in acyloxyacyl linkage to amide-linked 3-OH-18:0 and that 3-OH-14:0 is esterified to the 3'-position. NMR analysis of the intact R-lipid $A_{\text{ADP}}$ (described below) also indicated the existence of a single 3-acyloxyacyl residue; however, its location (distal or proximal) could not be unambiguously assigned.

Localization of Amide-linked Fatty Acids by MALDI-TOF Analysis of the De-O-acylated Lipid A—To assess further N-acyl heterogeneity, the de-O-acylated lipid $A_{\text{HOAc}}$ was analyzed by MALDI-TOF mass spectrometry (Fig. 6). The DOA-lipid $A_{\text{HOAc}}$ yielded a family of molecular ions ($M + Na$)$^+$ having $m/z$ values and peak ratios in close agreement with the types of species predicted on the basis of amide-linked fatty acid composition (Table II). The most abundant ions and proposed species (Table IV) were at $m/z$ 1021.0, P-GN$_{[3-18:0]}$ + 3-19:0; $m/z$ 1019.7, P-GN$_{[3-18:0]}$ + 3-19:1; and $m/z$ 1007.6, P-GN$_{[3-18:0]}$ + 3-18:0, where P-GN$_n$ designates the n-monophosphoryl glucosamine disaccharide. At high laser power, $B_1^+$ fragment ions were again observed, representing the distal 4'-phosphogluco- disaccharide and its attached N-fatty acyl chain. Major $B_1^+$ ions (Table IV) were $m/z$ 565, P-GN$_{[3-18:0]}$; and $m/z$ 537, P-GN$_{[3-16:0]}$.

Two-dimensional NMR Analysis of the Intact NGR234 Lipid $A_{\text{ADP}}$—The intact R-lipid $A_{\text{ADP}}$ was enriched for species $E_{\text{HIS}}$ and $F_{\text{HIS}}$ by preparative TLC (Fig. 4B) and analyzed by homo- and heteronuclear NMR spectroscopy. The lipid A was analyzed without derivatization, using a ternary solvent system reported to give well resolved spectra of intact and underivatized lipids A (24, 53). $^1$H-$^1$H COSY (not shown) and TOCSY (Fig. 7) revealed two glycosyl ring systems, as well as five major spin systems arising from the hydroxylated fatty acyl chains.
The H-1 proton of the proximal GlcN residue was assigned to a doublet-doublet resonance downfield at δ 5.45. This proton also showed a strong connectivity with the (glycosidic) phosphate during 1H-31P HMBC (Fig. 2). The chemical shift of this proton is consistent with glycosidic phosphate substitution and the α-anomeric configuration for the proximal GlcN residue. H-1 showed a strong COSY cross-peak with a proton at δ 4.12 (H-2), which showed coupling to H-3 (δ 5.20). The remaining glycosyl protons of the proximal residue were readily assigned from TOCSY connectivities, and the complete proton assignments are listed in Table V.

Analysis of the distal GlcN residue was also initiated at the anomeric proton. A COSY doublet at δ 4.60 was assigned to H-1, indicating the α-anomeric configuration of the distal GlcN residue.

### Table III

**Summary of the major lipid A<sub>NSDS</sub> species derived from Sinorhizobium sp. NGR234 rough-LPS and smooth-LPS**

The lipids A were isolated by hydrolyzing the R-LPS or S-LPS in NaOAc buffer (pH 4.5) containing 1% SDS.

| Component<sup>a</sup> | Observed ion<sup>b</sup> | Calculated mass and proposed formula<sup>c</sup> | Predicted ion<sup>d</sup> | Acylation | % Abundance<sup>e</sup> |
|----------------------|-------------------------|---------------------------------|-------------------------|-----------|------------------|
| E<sub>BIS</sub>      | 1936.8                  | C<sub>10</sub>H<sub>20</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1939.6                  | Penta     | 32               |
| F<sub>BIS</sub>      | 2024.2                  | C<sub>10</sub>H<sub>20</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 2025.7                  | Penta     | 20               |
| C<sub>BIS</sub>      | 1710.8                  | C<sub>10</sub>H<sub>17</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1713.3                  | Tetra     | 22               |
| D<sub>BIS</sub>      | 1797.4                  | C<sub>10</sub>H<sub>18</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1799.4                  | Tetra     | 16               |
| B<sub>BIS</sub>      | 1515.4                  | C<sub>10</sub>H<sub>15</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1516.9                  | Tetra     | 4                |
| C<sub>MON</sub>      | 1629.8                  | C<sub>10</sub>H<sub>17</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1633.3                  | Tetra     | 4                |
| A<sub>BIS</sub>      | 1290.5                  | C<sub>6</sub>H<sub>12</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1290.6                  | Tri       | 2                |
| Tri-C<sub>BIS</sub>  | 1485.9                  | C<sub>6</sub>H<sub>17</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1486.9                  | Tri       | 14               |
| Tri-D<sub>BIS</sub>  | 1572.3                  | C<sub>6</sub>H<sub>18</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1573.0                  | Tri       | 4                |
| Tri-C<sub>MON</sub>  | 1406.9                  | C<sub>6</sub>H<sub>17</sub>N<sub>2</sub>O<sub>5</sub>P<sub>2</sub> | 1407.0                  | Tri       | 2                |

<sup>a</sup> The letter designations are based on the ion families observed during MALDI-TOF MS analysis, as in Fig. 3. Species designated Tri-C<sub>BIS</sub>, Tri-D<sub>BIS</sub>, and Tri-C<sub>MON</sub> are triacylated versions of the parent species.

<sup>b</sup> All ions are of the formula (M – H)<sup>+</sup>.

<sup>c</sup> Calculated from the average incremental masses of the component glycosyl and fatty acyl residues as described under “Experimental Procedures.”

<sup>d</sup> The predicted (M – H)<sup>+</sup> ion as determined from the proposed formula.

<sup>e</sup> Estimated by summing the ion peak intensities in the negative ion MALDI-TOF spectra, as shown in Fig. 3B for the R-LPS.
for the distal GlcN. H-1’ showed strong COSY cross-peaks with H-2’ (δ 3.93) and a proton subsequently identified as H-5’ (δ 3.48). H-4’ (δ 4.17) was identified from TOCSY cross-peaks with H-2’ (δ 3.93), H-3’ (δ 5.20), and H-5’ (δ 3.48). The downfield shift of H-4’ is consistent with coupling to an ester-linked phosphate, observed during 1H-31P HMQC (Fig. 2). H-5’ showed cross-peaks with H-4’ (δ 4.17), H-6a’ (δ 3.94), and H-6b’ (δ 3.79); strong cross-peaks with the remaining (axial) ring protons were also observed, typical of the glucos-configuration.

A distinct feature of these spectra is that both H-3 and H-3’ are located downfield (coincident at δ 5.20) indicating that both the proximal (C-3) and distal (C-3’) positions are acylated. The proton chemical shifts for both glycosyl ring systems were found to be quite similar to those reported for the intact, bisphosphorylated lipid A from R. etli (24). Although their carbohydrate backbones are totally different, both lipids A contain a single acyloxyacyl residue. The cross-peak at (δ 5.35, δ 2.03) was assigned to the vinylic and allylic proton couplings of the unsaturated 3-hydroxy fatty acids; further TOCSY connectivities were traced to upfield secondary and terminal alkyl protons (not shown).

TABLE IV
Selected ions from MALDI TOF mass spectrometry of the de-O-acylated lipid A<sub>HOMA</sub> from the rough-LPS of Sinorhizobium sp. NGR234

| Observed (M+Na)<sup>+</sup> ion<sup>a</sup> | Observed β<sup>b</sup> ion<sup>c</sup> | Calculated mass and proposed fatty acyl<sup>d</sup> | relative abundance<sup>d</sup> |
|----------------------------------------|---------------------------------|---------------------------------|-------------------|
| 1021.4                                 | 565.9                           | 1022.2                          | 14 %              |
| 1019.7                                 | 565.9                           | 1020.2                          | 16 %              |
| 1007.6                                 | 565.9                           | 1008.2                          | 19 %              |
| 1005.7                                 | 565.9                           | 1006.2                          | 10 %              |
| 565.9                                  | 1004.2                          |                                | 3 %               |
| 563.9                                  | 1003.6                          |                                | 3 %               |
| 561.9                                  | 1002.2                          |                                | 3 %               |

<sup>a</sup> Quasimolecular ion, arising from the de-O-acylated lipid A.

<sup>b</sup> β<sup>+</sup> fragment ions arising from cleavage of the glycosidic bond.

<sup>c</sup> The calculated average mass (including sodium) for the molecular ion (M + Na)<sup>+</sup>. The diagrams depict the possible N-linked fatty acid combinations that are compatible with both the molecular ion and the specified B<sup>+</sup> ion. All fatty acids are 3-hydroxylated (this feature is not represented), and all species are sodiated (not shown). Abbreviations: 18, 18:1, etc., = fatty acid chain length and number of double bonds.

<sup>d</sup> The relative abundance of the parent de-O-acylated lipid A species, estimated from the relative intensities of the (M + Na)<sup>+</sup> ions.
**Fig. 7.** 500-MHz $^1$H–$^1$H TOCSY spectrum of the intact lipid A$_{SGS}$ from NGR234 R-LPS. The two glycosyl-ring systems are identified, with H-1 of the proximal residue resonating downfield at $\delta$ 5.45 due to substitution by $\alpha$-linked PO$_4$. H-1′ ($\delta$ 4.60) of the distal residue indicates the $\beta$-anomeric configuration for that residue. Both H-3 and H-3′ are found downfield (coincident at $\delta$ 5.20) indicating that both the proximal (C-3) and distal (C-3′) positions are acylated. A single, prominent acyloxyacyl moiety was indicated by a downfield $\beta$-oxymethine proton ($\delta$ 5.13) having cross-peaks with the $\alpha$-methylene protons ($\delta$ 2.57, $\delta$ 2.45) and the $\gamma$-methylene protons ($\delta$ 1.58, 1.31) of a $3\text{-}O$-acylated fatty acyl residue.

**Table V**

$^1$H and $^13$C chemical shifts of NGR234 lipid A$_{SGS}$ components

The lipid A$_{SGS}$ sample was prepared as described and analyzed in CDCl$_3$/CD$_3$OD/D$_2$O (2:3:1 v/v) (24) at 30 °C relative to internal TMS. Proton chemical shifts ($\delta$ ppm) were assigned from COSY and TOCSY experiments; carbon chemical shifts ($\delta$ ppm) were assigned from HSQC experiments; "BHB" = $\beta$-OH-butyrate.

| Spin system/residue | Proton $\delta$H | Carbon $\delta$C |
|---------------------|-----------------|-----------------|
| GlcN (proximal)     |                 |                 |
| H-1 $\alpha$        | 5.45            | C-1 96.2        |
| H-2                 | 4.12            | C-2 54.5        |
| H-3                 | 5.20            | C-3 76.3        |
| H-4                 | 3.64            | C-4 69.9        |
| H-5                 | 4.03            | C-5 74.0        |
| H-6a                | 4.13            | C-6 67.9        |
| H-6b                | 3.84            |                 |
| GlcN (distal)       |                 |                 |
| H-1′                | 4.80            | C-1′ 104.4      |
| H-2′                | 3.93            | C-2′ 56.5       |
| H-3′                | 5.20            | C-3′ 76.3       |
| H-4′                | 4.17            | C-4′ 73.3       |
| H-5′                | 3.48            | C-5′ 78.2       |
| H-6a′               | 3.94            | C-6′ 62.9       |
| H-6b′               | 3.79            |                 |
| $\beta$-OH fatty acid |              |                 |
| H-\(3\text{-}OH\)   | 2.30            | C-$\beta$ 69.1 |
| H-\(2\text{-}OH\)   | 3.92            | C-$\beta$ 70.7 |
| H-\(3\text{-}OH\)   | 2.41            | C-$\beta$ 70.7 |
| H-\(4\text{-}OH\)   | 3.98            | C-$\beta$ 70.7 |
| H-\(2\text{-}OH\)   | 2.51            |                 |
| H-\(2\text{-}OH\)   | 5.13            | C-$\beta$ 73.7 |
| H-3′                | 2.46            |                 |
| H-3′                | 3.99            | C-$\beta$ 70.9 |
| H-\(\text{acetylated}\) | (1.42–1.51) |                 |
| Vinylic-H            | 5.35            |                 |
| Allylic-H            | 2.03            |                 |
| CH$_3$              | 0.89            |                 |
| 27-OH:28:0 (−BHB)   |                 |                 |
| H-\(\text{acetylated}\) | 1.34          |                 |
| H-\(\text{acetylated}\) | 3.72          |                 |
| H-\(\text{acetylated}\) | 1.13          |                 |
| 27-OH:28:0 (+BHB)   |                 |                 |
| H-\(\text{acetylated}\) | 1.54          |                 |
| H-\(\text{acetylated}\) | 4.92          |                 |
| H-\(\text{acetylated}\) | 1.24          |                 |
| $\beta$-OH-butyrate |                 |                 |
| H-\(\alpha\)        | 2.50            |                 |
| H-\(\beta\)         | 4.06            |                 |
| H-\(\gamma\)        | 1.46            |                 |


**DISCUSSION**

In free-living culture, the major LPS synthesized by *Sinorhizobium* sp. NGR234 is a rough-LPS, which yields a heterogeneous lipid A consisting of six major components that differ in acylation (Fig. 8). As estimated from the relative intensities of molecular ions observed during mass spectrometry, 52% of the lipid A components are pentaacylated (components E and F) and 46% are tetraacylated (components B–D). Over 90% of these R-lipid A molecules contain a single copy of an LCFA, carried as the secondary fatty acid of an acyloxyacyl substituent on the distal glucosamine residue. The absence of the labile $\beta$-hydroxybutyrate moiety (components C and E) could arise from degradation or from endogenous heterogeneity in acyla-
tion. As discussed below, the specific absence of a 3-O-acyl substituent (components C and D) may reflect a biosynthetic origin, as reported for *R. leguminosarum* (54). At the present time, detailed structures are known for three rhizobial lipids A, the *Rhizobium etli-R. leguminosarum* common lipid A (24–27) and that of *Sinorhizobium* sp. NGR234.

It is generally recognized that all *Rhizobiaceae* examined to date contain LCFA in their lipid A (14, 29, 30), at least when the LPS are extracted from free-living, vegetative cells. However, because of the lack of detailed structural studies on the lipids A from these organisms, and because of difficulties associated with analyzing LCFA, the location, stoichiometry, and type of attachment of these substituents have not been reported. It is also not clear if LCFA is restricted to a specific location among all rhizobial lipids A or is variable in location, perhaps characteristic of different strains or genera. Although this latter point is still not answered, the present study shows that there is a striking similarity in the overall acylation pattern, including the specifics of LCFA attachment, for the lipids A of *Sinorhizobium* sp. NGR234 and *R. etli-R. leguminosarum* (24, 26, 27). Features of lipid A acylation that are common to both but that are not characteristics of *E. coli* or other enteric

**FIG. 8.** Structural diagrams of the R-lipid A<sub>DS</sub> and S-lipid A<sub>DS</sub> components isolated from *Sinorhizobium* sp. NGR234 R-LPS and S-LPS, by mild hydrolysis in pH 4.5 buffer containing SDS. The triacylated C and D species were detected only in the S-lipid A. The proposed formulas and relative abundance of individual species are summarized in Table III.
lipids A include the following: 1) a high degree of acyl chain length heterogeneity of the amide-linked fatty acids, primarily localized to the proximal carbohydrate residue; 2) variability in the occurrence of an ester-linked fatty acid at C-3 of the proximal residue; 3) the presence of a single amide-linked acyloxyacyl substituent, on the distal glucosamine, in which the LCFA is attached as the secondary fatty acid; 4) the absence of any ester-linked acylxyacyl residues, with the exception of 5) the occurrence, in at least 50% of the molecules, of a β-hydroxybutyryl moiety, attached in ester linkage to the (ω = 1) hydroxyl of the LCFA, thus forming an acyloxyacyl moiety at that location (Fig. 8, species D and F).

These common structural features have implications for both the biological function and biosynthetic pathway of rhizobial and sinorhizobial lipids A. Considering the structure of the carbohydrate backbone of Sinorhizobium sp. NGR234 lipid A, the early stage enzymes required for the initial synthesis of the Õ-lipid IVₐ bisphosphorylated precursor are most likely present (in homologous forms), as they are in Rhizobium sp. and E. coli (55). The structures described here further suggest that several of the enzymes involved in lipid A acylation, unique to R. etli-R. leguminosarum (i.e. not found in enteric bacteria), may also exist in Sinorhizobium sp. NGR234. These would include a 3-O-deacylase, which in R. etli selectively removes ester-linked 3-OH-14:0 from the proximal glucosamine (54), and a homologous system for delivering and attaching the LCFA to the 2′ amide-linked fatty acid, forming the distal acyloxyacyl residue of the completed lipid A. The sinorhizobial LCFA transfer system would include a special long chain acyltransferase (LpxXL) and a dedicated acyl carrier protein (AcpXL), homologous to those identified in R. etli (56). Compositional analyses indicate that in NGR234, the putative LpxXL and AcpXL proteins would have slightly altered chain length specificities compared with those of R. etli, because −30% of the LCFA in NGR234 lipid A is C30, whereas C26 is the sole component in R. etli, under the described growth conditions (26, 30, 54). The structures reported here also show that prior removal of the 4′- and 1-phosphates from the Õ-lipid IVₐ precursor is not required for the transfer of LCFA to substrate, in the case of NGR234 lipid A. It was previously reported that the 3'-deacylated species (Fig. 8) was not detectable in S. meliloti Rm1021 (54). The possible existence of a 3-O-deacylase in NGR234 could represent one of the variables in lipid A structure, among the many strains of S. meliloti and the related Sinorhizobium fredii. The fine structural differences in lipid A acylation among these bacteria have not been characterized; however, it is noteworthy that the profile of lipid A components derived from S. fredii USDA257 is considerably different from the profile reported here for Sinorhizobium sp. NGR234. Alternatively, the 3-O-deacylated species (C and D) could arise from degradation of lipid A during isolation. However, all of these components quantitatively retained the very labile C-1 glycosidic phosphate, and based on the observed B₁⁺ ions (Fig. 5), 3'-O-deacylated species were not observed in the rough lipid A, suggesting that loss of 3-O-acyl groups occurred specifically. Finally, in NGR234, distinct acyltransferase/acyl carrier proteins may be required for attachment of the unsaturated N-linked fatty acids, which are not components of the R. etli lipid A (26, 27). These unsaturated fatty acids are primarily localized to the proximal residue, as evidenced from MALDI-TOF analysis of the de-O-acylated lipid A (Table IV and Fig. 6).

The fact that key features (i.e. the occurrence, type, and location of acyloxyacyl residues), which distinguish the R. etli acylation pattern from that of E. coli, are retained almost identically in NGR234 further strengthens the supposition that these features are in some way essential to successful symbiotic infection. During R. leguminosarum colonization of Pismus sativum, Kappenberg and Carlson (51) observed an increase in the mol % of LCFA in bacteroid-derived LPS, compared with LPS from free-living cells. The LCFA was proposed to stabilize the bacteroid membrane, which is closely surrounded by the plant-derived symbiosome membrane during endocytosis, symbiosome formation, and symbiosome/bacterial cell division. Recently, an acpXL mutant of R. leguminosarum did not incorporate C28 LCFA into its lipid A and showed delayed nodulation in pea plants. This lipid A mutant also demonstrated less competitive growth in vegetative culture.

The NGR234 lipid A isolated from smooth LPS appeared to have a different acylation profile than the R-lipid A components. In addition to molecular species C–F, several unusual triacylated lipid A components were observed, which appeared to lack all ester-linked fatty acids (Fig. 8). The absence of ester substituents may have occurred through degradation; however, it is interesting that all of these species retained the other labile substituents, i.e. phosphorylation at C-1, and the LCFA acyloxyacyl residue. Recent studies with Sinorhizobium sp. NGR234 (49, 57) have indicated that the proportion of S-LPS is increased relative to R-LPS, when vegetative cells are cultured in the presence of apigenin or other plant flavonoids (inducers of bacterial nod and rol genes). These “flavonoid-induced LPS” appear to have an altered lipid A and core region structure compared with normal (uninduced) LPS (49). Smooth LPS may also be the dominant form of LPS on the bacteroid surface in the case of Sinorhizobium sp. NGR234 and in R. leguminosarum 3841 (31). The biological significance of these observations requires further structural analysis of both the vegetative and bacteroid-derived forms.

Acknowledgments—We thank Dr. John Glushka for advice and assistance with NMR and Dr. Brad Reuhs for discussions. The Complex Carbohydrate Research Center was supported in part by Department of Energy Grant DE-FG09-93ER20097.

REFERENCES
1. Pueppke, S. G., and Broughton, W. J. (1999) Mol. Plant-Microbe Interact. 12, 293–318
2. Broughton, W. J., Jabbouri, S., and Perret, X. (2000) J. Bacteriol. 182, 5641–5652
3. Freiberg, C., Fellay, B., Bairoch, A., Broughton, W. J., Rosenthal, A., and Perret, X. (1997) Nature 387, 404–407
4. Carrion, M., Bhat, U. R., Reuhs, B., and Carlson, R. W. (1999) J. Bacteriol. 172, 1725–1731
5. Carlson, R. W., Kaledzhba, S., Turwski, D., Pachori, P., and Noel, K. D. (1987) J. Bacteriol. 169, 4923–4928
6. Petrovics, G., Putnoky, P., Reuhs, B., Kim, J., Thorup, T. A., Noel, D., Carlson, R. W., and Kondorosi, A. (1993) Mol. Microbiol. 9, 1063–1094
7. Reuhs, B. L., Williams, M. N. V., Kim, J. S., Carlson, R. W., and Défée, F. (1995) J. Bacteriol. 177, 4289–4296
8. Reuhs, B. L., Geller, D. P., Kim, J. S., Fox, J. E., Koli, V. S. K., and Pueppke, S. G. (1998) Appl. Environ. Microbiol. 64, 4920–4928
9. Niehaus, K., Lagaes, A., and Puhler, A. (1998) Mol. Plant-Microbe Interact. 11, 906–914
10. Reuhs, B. L., Stephens, S. B., Geller, D. P., Kim, J. S., Glenn, J., Pretyzky, J., and Ojaneen-Reuhs, T. (1999) Appl. Environ. Microbiol. 65, 5186–5191
11. Forsberg, L. S., and Reuhs, B. L. (1997) J. Bacteriol. 179, 5366–5371
12. Reuhs, B. L., Kim, J. S., Bredtigg, A., and Carlson, R. W. (1994) Mol. Plant-Microbe Interact. 7, 240–247
13. Williams, M. N. V., Hollingsworth, R. I., Klein, S., and Signer, E. R. (1990) J. Bacteriol. 172, 2623–2626
14. Kappenberg, E. L., Reuhs, B. L., Forsberg, L. S., and Carlson, R. W. (1998) in The Rhizobiaceae (Spaink, H. P., Kondorosi, A., and Hooykaas, P. J. J., eds) pp. 119–154, Kluwer Academic Publishers Group, Dordrecht, Netherlands
15. Noel, K. D., and Ducilli, D. M. (2000) in Prokaryotic Nitrogen Fixation: A Model

2 L. S. Forsberg and S. Gudlavalleti, unpublished MALDI-TOF and TLC results.

3 V. Vedam, E. L. Kannenberg, J. G. Haynes, D. J. Sherrier, A. Datta, and R. W. Carlson, submitted for publication.

4 B. L. Reuhs, B. Relic, C. Marie, S. Jabbouri, T. Ojaneen-Reuhs, S. Stevens, L. S. Forsberg, C.-H. Wong, and W. J. Broughton, submitted for publication.
