Carbohydrate Research, 172 (1988) 97–112
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

THE COMPLETE STRUCTURE OF THE TRIFOLIIN A LECTIN-BINDING CAPSULAR POLYSACCHARIDE OF *Rhizobium trifolii* 843

RAWLE I. HOLLINGSWORTH, FRANK B. DAZZO*,
Department of Microbiology and Public Health, Michigan State University. East Lansing, MI 48824 (U.S.A.)

KLASS HALLENGA, AND BRIAN MUSSELMAN
Department of Chemistry, Michigan State University, East Lansing, Michigan 48824 (U.S.A.)

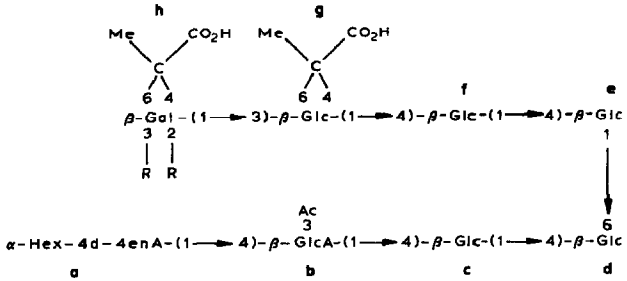
(Received January 12th, 1987; accepted for publication in revised form, May 20th, 1987)

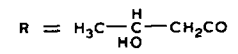
ABSTRACT

The complete structure of the acidic, extracellular, capsular polysaccharide of *Rhizobium trifolii* 843 has been elucidated by a combination of chemical, enzymic, and spectroscopic methods, confirming an earlier proposed sugar sequence and assigning the locations of the acyl substituents. The polysaccharide was depolymerized by a lyase into octasaccharide units which were uniform in carbohydrate composition and linkage. These units also contained a uniform distribution of acetyl and pyruvic acetal [*O*-(1-carboxyethylidene)] groups, and half of them were further acylated with D-3-hydroxybutanoyl groups. A much smaller proportion (<5%) of the oligomers was further acylated by a second D-3-hydroxybutanoyl group. The locations of the substituents were determined chemically and by *J*-correlated, ^{1}H-n.m.r. spectroscopy, proton nuclear Overhauser effect (n.O.e.) measurements, double-resonance ^{1}H-n.m.r. spectroscopy, and ^{13}C-n.m.r. spectroscopy. The composition and structure of the carbohydrate chain were determined by methylation analysis using g.l.c.–m.s. fast-atom-bombardment mass spectrometry, and n.m.r. studies on the reduced, deacylated oligomer. Structural studies were supplemented by n.m.r. analyses on the original polymer. The oligosaccharides were found to be branched octasaccharides with four sugar residues in each branch, and the carbohydrate sequence agreed well with that expected from earlier work. In the abbreviated sequence and structure (**1a**), the sugar residues are labelled "**a**" through "**h**". The main chain (**a**–**d**) is composed of a 4-deoxy-α-L-*threo*-hex-4-enopyranosyluronic acid group (**a**) that is linked to O-4 of a 3-*O*-acetyl-D-glucosyluronic acid residue (**b**) which is β-linked to O-4 of a D-glucosyl residue (**c**). Residue **c** is β-linked to O-4 of the branching D-glucose residue (**d**). The side chain consists of a substituted D-galactosyl group (**h**) which is β-linked to O-3 of residue **9** of a β-(1→4)-linked D-glucose trisaccharide (fragment **e**–**f**–**g**). The reducing end of the resulting tetrasaccharide (**e**–**f**–**g**–**h**) is β-linked to O-6 of the branching D-

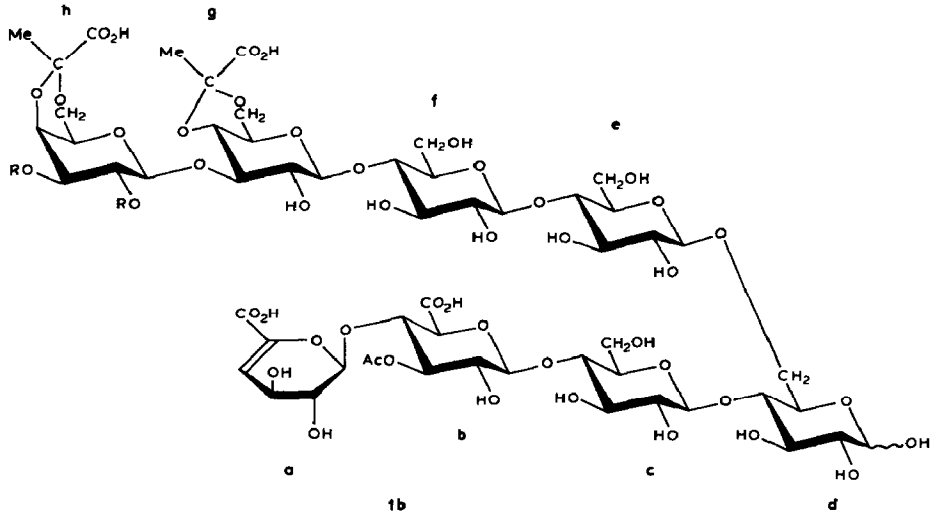
*To whom correspondence should be addressed.

0008-6215/88/$ 03.50





1a



1b

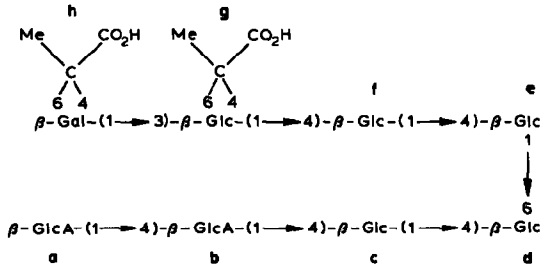
glucose residue (**d**). In the native polymer, this branching residue is α-linked to O-4 of the modified D-glucuronic acid residue (**a**) which is the unsaturated sugar in the oligomer. A small proportion of the O-2 atoms of the acetylated D-glucosyluronic acid residues is acetylated because of ester migration. The two terminal sugars (**g** and **h**) of the branch chain bear 4,6-*O*-(1-carboxyethylidene) groups. The D-galactosyl groups of half of the oligomers are acylated by D-3-hydroxybutanoyl groups at O-3. About 5% of the oligomers bear a second D-3-hydroxybutanoyl group at O-2 of the D-galactosyl group (**h**).

INTRODUCTION

The acidic, extracellular polysaccharides of *Rhizobium* are believed to be important determinants of the specificity between these bacteria and their respective plant hosts[1,2]. In view of this probability, the structures of these molecules constitute an area of rapidly expanding study (and much controversy).

The first detailed study of a *Rhizobium* acidic extracellular polysaccharide was on *Rhizobium trifolii* strain U226, by Jansson *et al.*[3]. They reported that this polysaccharide is composed of a branched heptasaccharide repeating-unit containing glucose, galactose, and glucuronic acid in the molar ratios of 5:1:1. They proposed that a lone glucuronic acid residue is in the main chain and that it is α-linked.

Subsequently, Robertson *et al.*[4] claimed that there is a β-(1→4)-linked glucuronic acid disaccharide instead of the single glucuronic acid residue reported by Jansson *et al.*[3]. Robertson *et al.*[4] also found that the sole α-linkage is at the branching sugar. At least one other study supported the conclusion of Jansson *et al.*[3] that only one uronic acid residue is present[5]. Both groups found that all of the sugars are in the pyranose form. The sequence of the repeat unit according to Robertson *et al.*[4] is shown in formula **2**. The one proposed by Jansson *et al.*[3] lacks residue **a**, and the linkage between **b** and **c** is α. In addition, the linkage at **d** is β. Neither of these groups identified the location and nature of the important acyl substituents in these polymers, and that was the major concern of the present investigation.



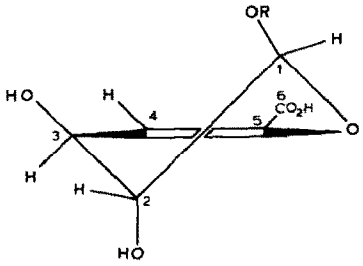
2

We are studying the molecular basis of the symbiotic specificity between *Rhizobium trifolii* and its host plant, and have been using bacteriophage-induced depolymerases to study the structures of rhizobial polysaccharides[6–8]. These enzymes produce oligosaccharide fragments from the parent polymer without disturbing the distribution of labile substituents. The fragments can interact specifically with a clover lectin, trifoliin A. When applied to clover seedlings, these fragments exert an infection-related biological activity that affects the number of clover-root hairs infected by *R. trifolii*[8]. During the course of our study, we identified 3-hydroxybutanoic acid as a component of these polymers. In a preliminary report[6], we speculated that this substituent might be ether linked because of its apparent stability to mild base. The ether bond of β-alkoxy acids can readily undergo elimination in extremes of acid or base; this is due to the facile formation of α,β-unsaturated systems. More recently, Kuo and Mort[9] found 3-hydroxybutanoic acid to be ester linked in one strain of *Rhizobium trifolii* that they examined, but absent from two others, one of which was strain 0403 (the one in

which it was discovered). In their study, Kuo and Mort[9] assumed the validity of the structure proposed by Robertson *et al.*[4] and assigned the acetyl group to unbranched glucosyl residue in the main chain. McNeil *et al.*[10] used phage depolymerases to compare the gross structural features of rhizobial polysaccharides; although they obtained fragments similar to ours, they did not elucidate the complete structures of these oligomers.

The variations in structures and compositions reported, even among different strains of the same species of *Rhizobium,* are large. Amemura *et al.*[11] claimed that the acidic, extracellular polysaccharide of *R. trifolii* 4S does not contain galactose and has only one *O*-(1-carboxyethylidene) group per repeating unit. Ghai *et al.*[12] reported that glucuronic acid is absent from the extracellular polysaccharide from[12] *R. trifolii* J60.

One major conclusion of the study by McNeil *et al.*[10] was that the acidic extracellular polysaccharides produced by several *Rhizobium* species are the same, and these polysaccharides must therefore not be determinants of host specificity (since each *Rhizobium* species specifically nodulates different plant hosts). This conclusion is certainly debatable in view of all the contradictions and discrepancies in the literature. The present report establishes a structure that concurs with the primary carbohydrate sequence (**2**) reported by Robertson *et al.*[4], and extends that study by unequivocally establishing the locations of the acyl substituents, as shown in **1b**.



3

MATERIALS AND METHODS

R. trifolii 843 (obtained from Dr. Barry Rolfe, Australian National University, Canberra, Australia) was grown on plates on a defined B3 medium for 5 days, and the acidic, extracellular polysaccharide was isolated, depolymerized by using a phage lyase, and the oligomers isolated and purified as described earlier[6–8]. Carboxyl reduction, methylation (Hakomori method using methyl iodide and methylsulfinyl anion), compositional analyses, and g.l.c.–m.s. analyses were conducted as described earlier[7]. N.m.r. spectra were recorded with a Bruker WM-250 spectrometer, at 250 MHz for ^{1}H, and at 62.9 MHz for ^{13}C. Deuterium oxide was used as the solvent for all analyses. The pulse width and relaxation delay were adjusted to suppress quaternary alkylidene-carbon resonances strongly. The ^{1}H–^{1}H

J-correlated spectrum was acquired by using a $90\text{–}t_1\text{–}D_1\text{–}45\text{–}D_1\text{–}t_2$ sequence, with appropriate phase-cycling to remove quadrature artifacts. The delay D_1 was chosen to emphasize small-to-medium couplings. All spectra are presented in the absolute value mode. Fast-atom-bombardment mass spectrometry was conducted in a JEOL HX-110 high-field instrument in the medium-resolution mode. 1-Thioglycerol was used as the matrix, and trifluoroacetic acid vapor was used to promote protonation of the analyte. The probe was unheated, and the accelerating voltage was 10 kV.

Cyanoborohydride reduction of the oligomer. — The oligomer (100 mg) was dissolved in dry methanol (50 mL) containing acetic acid (2 mL). Sodium cyanoborohydride (0.5 g) was added followed by dropwise addition of 1% hydrogen chloride in methanol until effervescence started. The mixture was stirred for 24 h at room temperature, and then evaporated to dryness under vacuum. The residue was chromatographed on a column (3 × 60 cm) of Biogel P2 in water and the voided peak (detected by u.v. monitoring at 215·nm) was collected and lyophilized.

Reductive amination of the oligomer. — The oligosaccharide (100 mg) was dissolved in dry *N,N*-dimethylformamide (20 mL) and acetic acid (0.1 mL) was added. 2-Aminopyridine (0.5 g) and sodium cyanoborohydride (0.5 g) were then added, and the mixture was stirred for 60 h at room temperature, evaporated to dryness under vacuum, and the product purified on a P2 column as before.

Deacetylation of the oligomer. — The oligosaccharide was dissolved in 0.05M sodium hydroxide and kept for 4 h at room temperature, the base neutralized with acetic acid, and the product isolated by chromatography on Biogel P2 using water as the eluant.

RESULTS AND CONCLUSIONS

The compositional analyses of the oligosaccharide products, using alditol acetate derivatives, gave glucose, galactose, and glucuronic acid in[7] the molar ratios of 5:1:1.

After methylation of the reducing and the carboxylic acid groups with methanolic hydrogen chloride, and deuterium labelling of uronic acid residues by carboxyl reduction with sodium borodeuteride, methylation analysis of the purified oligomer indicated the presence of one 4,6-substituted galactosyl residue, one 3,4,6-substituted glucosyl residue, one (4→6)-linked glucosyl residue, three 4-substituted glucosyl residues, and one 4-substituted glucosyluronic residue. On deacetalation, the (4→6)-linked galactosyl residue was converted into a terminal, unsubstituted galactosyl groups and the 3,4,6-substituted glucosyl residue was converted into a 3-substituted glucosyl residue. In addition, the (4→6)-linked glucosyl residue was converted into a (4→6)-linked alditol on pre-reduction of the reducing terminus. This confirmed that the *O*-(1-carboxyethylidene) groups spans O-4 and O-6 of glucose and of galactose, and that the reducing sugar is substituted at O-4 and O-6. The branching sugar is therefore the site of enzymic cleavage of the polysaccharide.

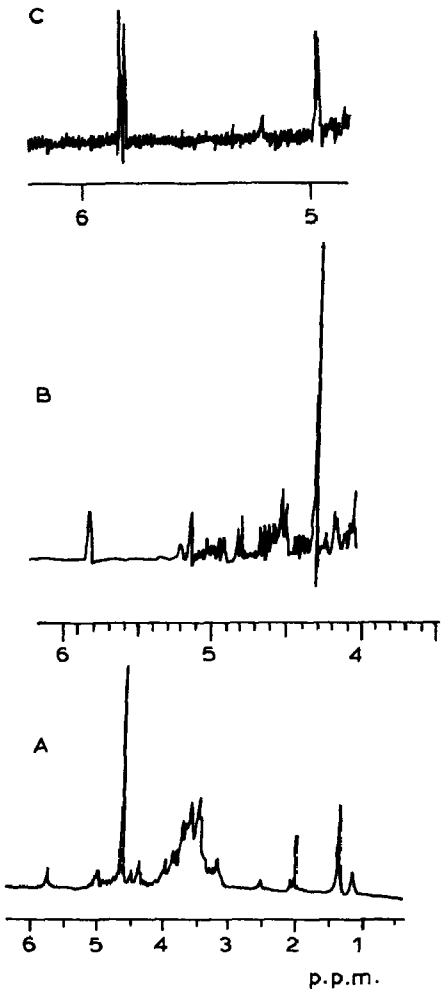


Fig. 1. (A) 250-MHz ^{1}H-n.m.r. spectrum of octasaccharides enzymically produced at 28°, showing acetyl (δ 2.0–2.2), *O*-(1-carboxyethylidene) (δ 1.3–1.5), and 3-hydroxybutanoyl (δ 1.21 and 2.61) groups. The signal at δ 5.84 is the resonance of H-4 of the α-L-*threo*-hex-4-enopyranosyluronic acid group, and the multiplets at δ 5.13 and 5.19 are components of the signal of the anomeric proton of the same group. (B) Partial spectrum of the same sample at 80°, showing downfield resonances. The H-1 and H-4 signals of the unsaturated sugar are now clear doublets. The smaller H-1 component is the result of an anisotropic shift from a neighboring acetyl group. (C) Partial spectrum after deacylation, showing the long-range coupling of H-4 and H-1 of the unsaturated residue. The small signal between the two major resonances is from the α anomer of the reducing sugar unit.

The ^{1}H-n.m.r. spectrum of the depolymerized product (see Fig. 1A) contained readily identifiable signals due to *O*-(1-carboxyethylidene) (δ 1.42–1.49), acetyl (δ 2.10 and 2.18) and 3-hydroxybutanoyl groups. The methyl group of this latter substituent appeared as a doublet at δ 1.21 (J 5.8) and the methylene group as a multiplet at δ 2.61. The downfield region of the spectrum was too complex for spectral assignments to be made, partly because of incomplete substitution and

anisotropic and coupling differences induced by the mixed anomeric forms of the reducing unit. In addition, the residual-water line obscured many of the signals in this region. The overlap in the downfield region of the ^{1}H-n.m.r. spectrum was significantly lessened by recording the spectrum at 80°. The water line was shifted upfield and no longer obscured the signals for the anomeric protons. The splittings of the signals in this region appeared to be ~7 Hz (see Fig. 1B).

We had demonstrated[7] that the mode of action of the enzyme is by elimination in the sugar residue substituted at O-4 of a glucuronic acid unit, to yield an α-L-*threo*-hex-4-enopyranosyluronic acid residue (**3**). The doublet in the ^{1}H-n.m.r. spectrum at δ 5.84 (*J* 3.1 Hz) (see Fig. 1B) was shown to be due to H-4 of the unsaturated sugar, and the other doublet, at δ 5.13 (*J* 3.9 Hz), to H-1 of the same residue[7]. The small multiplet at δ 5.19 is a component of the latter signal, shifted downfield by an anisotropic effect of a non-stoichiometrically substituted neighboring group. Only one component for this anomeric proton was observed when the oligosaccharide ws completely deacetylated with base (see Fig. 1C). In addition, the doublets at δ 5.84 and 5.13 both exhibited a slight upfield shift and, under optimal resolution conditions, could be resolved into doublets of doublets because of long-range couplings (see Fig. 1C). The H-4 atom of the unsaturated uronic acid residue (**3**) must therefore be in a "W" relationship with H-2, and H-1 and H-3 must be in the same relationship to each other. This confirmed the α-*threo* configuration. The glucuronic acid residue at the site of cleavage was, therefore, originally in the β configuration in the parent polymer. The comparatively large chemical shift of the anomeric proton of the α-L-*threo*-hex-4-enopyranosyluronic acid residue is due to its quasi-equatorial orientation in the $^1H_2$ conformation (**3**). The ^{1}H-n.m.r. spectrum of the deacetylated polysaccharide is shown in Fig. 2. The most downfield resonance occurred at δ 5.38 and, because of its large chemical shift and small coupling (*W*/2 <4 Hz), was attributed to an α-anomeric proton. This signal is noticeably absent from the spectrum of the oligomer. The α-linked residue is therefore at the site of cleavage by the enzyme and it is also the branching sugar.

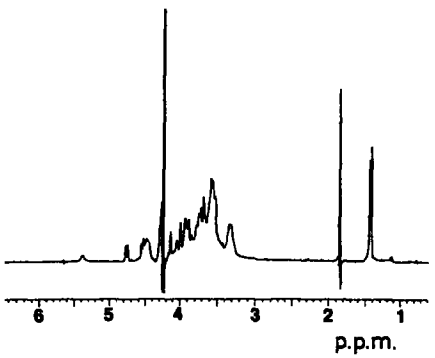


Fig. 2. 250-MHz ^{1}H-n.m.r. spectrum at 65° of the deacylated polysaccharide. [The singlet at δ 1.83 is from residual free acetate. The broad singlet at δ 5.38 is from the sole α-proton of the branching residue.]

In the ^{1}H-n.m.r. spectrum of the oligosaccharide (see Fig. 1A), there are several extra splittings and contributions because of the mixed anomeric nature of the reducing unit; the small doublet at δ 5.24 is due to the α-anomeric proton thereof. The signal of the β anomer is not resolved from the other anomeric signals. The intensity and chemical shift of this signal was found to be temperature- and concentration-dependent, as expected. The complications in the ^{1}H-n.m.r. spectrum from the different anomeric contributions of the reducing unit were removed by reducing its hemiacetal function with sodium cyanoborohydride in methanol containing acetic acid and a trace of hydrogen chloride. This procedure did not disturb any of the non-carbohydrate substituents of the oligosaccharides, and significantly simplified the spectrum (see Fig. 3A). The triplet at δ 5.0 (J 9.2 Hz) was assigned to H-3 of a residue having the *gluco* configuration. This is the proton on the carbon atom bearing the acetoxyl group. Although H-2 of a β-gluco- or -galacto-pyranose residue would also give a triplet with a similar splitting, this was ruled out from the result of irradiating the triplet; this caused no perturbation of signals in the anomeric region of the spectrum. Acetylation at O-3 of galactose should give a doublet of doublets, with the smaller coupling being ~2–3 Hz. The triplet at δ 5.0 is also quite noticeable in the spectrum of the oligosaccharides before reduction (see Figs. 1A and 1B).

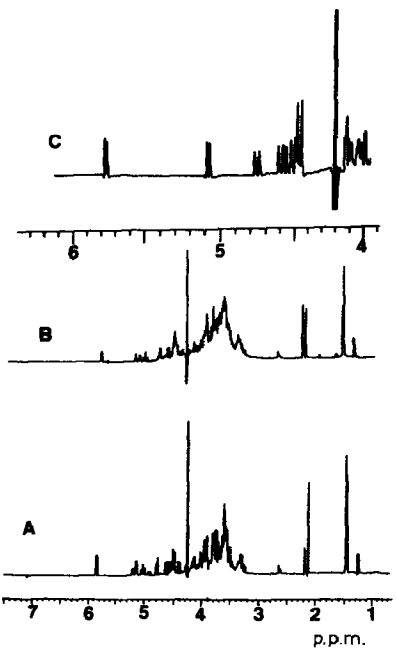


Fig. 3. (A) 250-MHz ^{1}H-n.m.r. spectrum at 80° of cyanoborohydride-reduced oligosaccharides. The triplet at δ 5.00 is from the proton at the site of acetylation. The smaller acetyl singlet, at δ 2.18 is due to migration. (B) Spectrum after equilibrating the sample in 0.01M sodium hydrogencarbonate for 24 h; the signals of both acetyl groups and both H-1 components of the unsaturated sugar now have equal intensities. (C) Partial spectrum at 80° after complete deacylation of the cyanoborohydride-reduced product. A total of six doublets with splittings of ~7 Hz (two coinciding at δ 4.45, and resolved from one another at δ 4.48) are observed. The signal H-1 of the unsaturated residue now appears as one doublet; the triplet at δ 5.00 has disappeared.

In the ^{1}H-n.m.r. spectra of the oligosaccharides (see Fig. 1A, 1B, and 3A), the signal for the anomeric proton of the unsaturated sugar is perturbed, to give a smaller component at δ 5.19. This was attributed to an effect of non-stoichiometric substitution at a neighboring site. The spectra also indicated a small amount of acetylation at one site by the presence of a much smaller acetyl signal at δ 2.18. An observation that led us to propose[6] an ether linkage for the 3-hydroxybutanoyl groups was that the acetyl groups could be removed in preference to the 3-hydroxybutanoyl groups on short exposure of the oligomer to mild base. However, Kuo and Mort[9] subsequently demonstrated that the 3-hydroxybutanoyl groups in the polysaccharide from another strain of *R. trifolii* could be removed under stronger conditions without elimination, and were therefore ester linked. We have confirmed this observation with the polysaccharide used in this study (see Fig. 2). It thus seemed likely that the acetyl groups could be induced to migrate under mildly basic but non-saponifying conditions without disturbing the 3-hydroxybutanoyl groups, which are conceivably more sterically hindered and form with more difficulty the obligatory, hemi-orthoester intermediate required for ester migration. When a solution of the cyanoborohydride-reduced oligomer was evaporated for 24 h at room temperature in a 0.01M solution of sodium hydrogencarbonate, the two acetyl signals in the ^{1}H-n.m.r. spectrum became almost equal in intensity (see Fig. 3B). This was accompanied by a parallel equilibration of the two components of the anomeric proton of the unsaturated sugar to equal intensity as well as by a decrease in the intensity of the triplet at δ 5.0. This confirmed that the small acetyl signal arose from acetyl migration to O-2 of the acetylated residue. It also confirmed that the residue which was acetylated was close to the unsaturated sugar. A proton n.O.e.-difference n.m.r. experiment in which the signal for the major acetyl group in the cyanoborohydride-reduced oligomer was irradiated led to diminution in the intensity of the resonance for the anomeric proton of the unsaturated sugar residue; this confirmed the relative spatial orientations of these groups. A negative n.O.e. might be expected as a molecule of this size should exhibit a relatively long correlation time. In the ^{1}H-n.m.r. spectrum of the cyanoborohydride-reduced oligomer, the downfield component of the anomeric-proton signal of the unsaturated sugar appeared at the same position as the entire proton if the oligomer was completely deacetylated (see Fig. 3C). The triplet at δ 5.0 also disappeared. The neighboring acetyl group must therefore exert a shielding effect *via* its carbonyl group or influence the orientation of the unsaturated sugar about its glycosidic linkage. In any event, the n.O.e. difference-spectrum, along with the chemical-shift change of the anomeric proton on induced acetate migration or deacylation, confirm that the acetyl group is attached to O-3 of the sugar unit adjacent to the unsaturated sugar.

The positions of esterification of the oligosaccharide by 3-hydroxybutanoic acid and acetic acid were also determined by exhaustively oxidizing the oligosaccharide fragments with periodate. Were the esterification at O-2 or O-3 of a residue, this should lead to an enrichment of that residue, as esterification would

make two vicinal hydroxyl groups unavailable (methylation analysis showed that the O-4 atoms of all of the sugar residues are involved in a linkage). The resulting oxidation product contained glucose and galactose in the ratio of 2:1. This meant that 3-hydroxybutanoic acid was esterified to the galactose residue as the stoichiometry of substitution by this group was 0.5 per repeating unit. It also meant that the acetyl group was attached to a uronic acid residue, because the glucose residue bearing the *O*-(1-carboxyethylidene) group is linked at O-3 and should therefore resist periodate oxidation. This result also indirectly confirmed that the uronic acid residue is linked to the unsaturated sugar, as it had already been established that the site of acetylation is the residue adjacent to the unsaturated sugar. The ^{1}H-n.m.r. spectrum of cyanoborohydride-reduced oligomer (see Fig. 3A) contained a doublet of doublets superimposed on a small doublet (J 7.1 Hz) at δ 4.90. This small doublet was a component of the larger doublet immediately upfield of it. The downfield shift of the smaller component was attributed to a deshielding effect of a neighboring substituent which was present in less than a

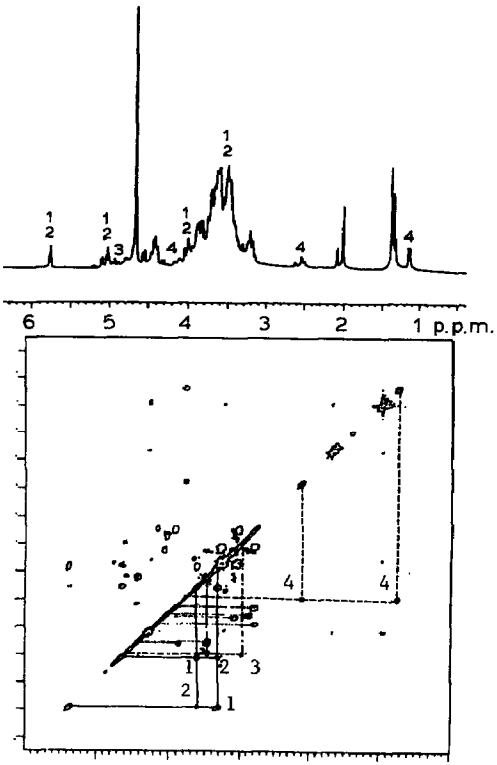


Fig. 4. 2-D J-correlated, ^{1}H-n.m.r. spectrum of the oligosaccharides, with the 1-D proton spectrum shown above for clarity. Resonances of connected spin systems are joined with the same line types and the contours are labelled with the same numbers. In the 1-D spectrum the positions of signals in a system are indicated by using the contour numbers in the 2-D plot. Anomeric resonances are joined to their contours by dotted lines but connectivities to the corresponding H-2 resonances are omitted for clarity.

stoichiometric proportion. The larger splitting of the doublet of doublets was ~7 Hz, and the smaller splitting was <3 Hz. Both signals disappeared when the sample was deacylated (see Fig. 3C). The doublet of doublets was assigned to H-3 of the esterified galactose residue. The signal of H-2 of the galactose residue should be a triplet (*J* 7–9Hz).

The ^{1}H-n.m.r. spectrum of the cyanoborohydride-reduced, deacylated oligomer (see Fig. 3C) contained clearly resolved resonances for all of the anomeric protons. The large coupling constants of the signals (generally ~7 Hz) confirmed that all sugar residues (except for the unsaturated sugar) were in the β-pyranose configuration. The other anomeric protons appeared at δ 4.75 (*J* 6.50 Hz), 4.59 (*J* 7.3 Hz), and 4.48 (*J* 6.8 Hz), and there were two overlapping resonances at δ 4.45 (*J* 6.9 Hz). This gave a total of seven anomeric resonances, and thus, there are octasaccharide fragments, because one anomeric resonance is lost on converting the reducing unit into an alditol residue.

The oligosaccharide fragments were also studied by 2-D *J*-correlated ^{1}H-n.m.r. spectroscopy (see Fig. 4). Contours that are part of an interconnected spin system are joined with the same type of line and have the same number. The contours that define connectivities between protons on the L-*threo*-hex-4-enopyranosyluronic acid group are labelled "1" or "2" and are connected with solid lines. The positions of coupled nuclei are also denoted on the 1-D spectrum by the same numbers. For the unsaturated uronic acid system, the contours arising from normal 3-bond couplings are labelled "1" and those from long-range couplings are labelled "2". The H-4 of the unsaturated sugar (the most downfield signal) is connected to H-3 by a contour labelled "1", but is also connected to H-2 by a contour labelled "2". The position of H-2 is defined by the connectivity through a contour labelled "1" from H-3. There is a connectivity between H-2 and H-1 (labelled "1") from a normal 3-bond coupling and a long-range connectivity between H-1 and H-3 through a contour labelled "2". The contour arising from the triplet due to H-3 of the acetylated glucuronic acid residue is labelled "3". The two nuclei to which this proton are coupled are both upfield of the anomeric protons confirming that the triplet is not due to a proton on C-2. The contours from the connectivities of the 3-hydroxybutanoyl group are labelled "4". The proton on the carbon atom to which this group is attached did not show a connectivity, because it occurred too close to the water line (which had to be suppressed by irradiation during the experiment). The absence of more than one connectivity for the other nuclei downfield of δ 4.2 confirmed that they are all due to anomeric protons.

The ^{13}C-n.m.r. spectrum of the deacylated polysaccharide contained seven signals in the region where primary hydroxylated carbon atoms resonate (see Fig. 5A). The most downfield signal in this region appeared at δ 68.0, and the most upfield signal, at δ 62.1. One of the resonances is due to the hydroxylated carbon atom of 3-hydroxybutanoate. The free acetate and free 3-hydroxybutanoate were not removed from the mixture after deacylation. Six of the resonances were assigned to the primary carbon atoms of hexose residues. This confirmed the

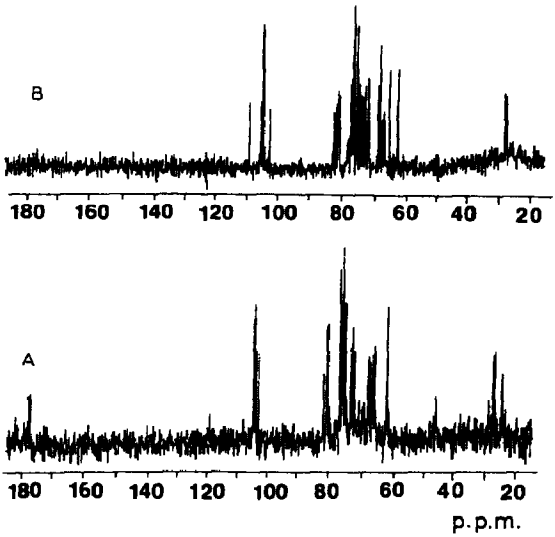


Fig. 5. 62.5-MHz broad-band-decoupled ^{13}C-n.m.r. spectrum of base hydrolyzed, neutralized polysaccharide. A total of six resonances due to free and substituted hydroxymethyl (C-6) resonances are observed. (B) Spectrum of the cyanoborohydride-reduced deacylated oligosaccharide. There are seven free and substituted primary carbon resonances (δ 60–72), one of which is derived from the reduced hemiacetal group. Seven anomeric carbon resonances are also observed. Conditions were adjusted to null the quaternary alkylidene carbon resonances.

presence of six hexoses (with their hydroxymethyl groups) in the polymer. The non-glycosidic ring-carbon atom having the unusually large chemical shift (δ 81.5) was assigned to C-3 of the 3-substituted glucose residue. This carbon resonance normally occurs even farther downfield[13,14], at 86–87 p.p.m., but is shifted upfield by the presence of the (1-carboxyethylidene) group at O-4. The resonances in the anomeric region (δ 100–105) were not sufficiently resolved to permit determination of how many anomeric carbon atoms were present. However, the signal at δ 103.4 was assigned to the carbon atom of the branching residue because it is not present in the ^{13}C-n.m.r. spectrum of the cyanoborohydride-reduced, deacetylated oligomer (see Fig. 7B). The ^{13}C-n.m.r. spectrum of the reduced deacylated oligomer (see Fig. 5B) contained seven resonances in the region of the signals from the anomeric carbon atoms. One anomeric resonance is lost on reduction so hence there were originally eight resonances (and eight sugar residues) in the oligomer. One anomeric-carbon signal had an unusually large chemical shift (δ 109.4) and was assigned to the unsaturated glucuronic acid unit. The glucuronic acid residue undergoes a major conformational change to a half-chair ($^1H_2$) form (**3**) on elimination. The other anomeric residues occurred between δ 102.6 and 105.5. The signals due to the alkylidene carbon atoms of the *O*-(1-carboxyethylidene) groups also appeared in the region of the anomeric carbon atoms, but were effectively nulled by choosing an appropriate relaxation delay. However, in fully relaxed spectra, these signals still did not interfere with the other resonances, and appeared at δ 103.3 and 104.2. The resonance at δ 105.5 was assigned to the anomeric carbon

atom of the galactosyl group because this unit normally has the largest chemical shift of all present[14]. It was, however, split by 0.4 p.p.m. in the ^{13}C-n.m.r. spectrum of the oligomer before deacylation, indicating that there was partial substitution, by an acyl group of O-3 of that residue. An acyl substituent on O-2 would have led to a much larger splitting. This also supported the conclusion that the 3-hydroxybutanoyl group is attached to the galactosyl group. In the ^{13}C-n.m.r. spectrum of the cyanoborohydride-reduced, deacylated oligomer (see Fig. 5B), there were seven signals in the region where hydroxymethyl and substituted hydroxymethyl resonances appear. One of these signals (δ 64.9) appeared only after reduction of the reducing unit, again confirming that there were six primary carbon atoms before reduction of the hemiacetal group.

The results from the chemical, n.m.r., and g.l.c.–m.s. analyses on the oligosaccharide fragments were confirmed by fast-atom-bombardment mass spectrometry (f.a.b.-m.s.) of the product formed by reductive amination of the reducing unit of the oligomer with cyanoborohydride and 2-aminopyridine. ^{1}H-N.m.r. spectral analysis of the purified reductive-amination product showed that none of the non-carbohydrate substituents and glycosyl residues were disturbed by the process. The f.a.b.-m.s. analysis was conducted by using a 1-thioglycerol matrix and trifluoroacetic acid vapor to promote protonation of the pyridine group. The mass spectrum (see Fig. 6) contained an ion cluster, at *m/z* 1585, corresponding to $(M + H)^+$ for octasaccharide fragments lacking 3-hydroxybutanoyl groups. Another ion, at *m/z* 1671, corresponding to the oligomers having 3-hydroxybutanoyl groups, was also observed. In addition, another ion, at *m/z* 1757, was observed, and this ion suggested the presence of a second 3-hydroxybutanoyl group on some of the oligomers. The ^{1}H-n.m.r. spectrum of the oligosaccharide in D_2O contained no data to support the suggestion of a second acyl substituent on any of the octasaccharide units. However, when spectra were measured in deuterium

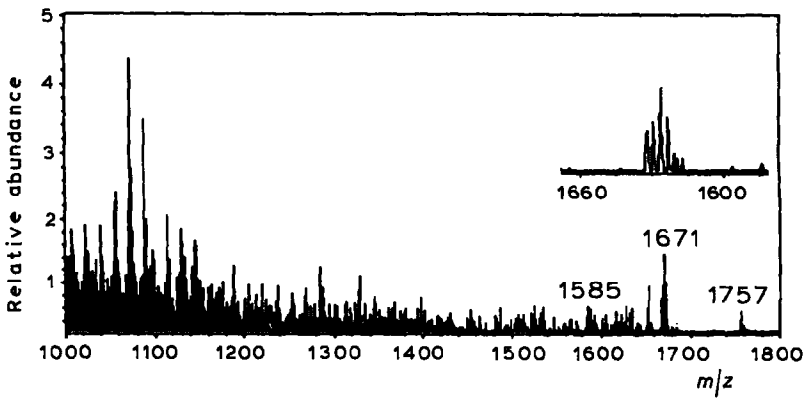


Fig. 6. F.a.b.-mass spectrum of reductive-amination product of the oligosaccharides with 2-aminopyridine. The ions at *m/z* 1585, 1671, and 1757 are quasi-molecular ions for oligomers respectively having zero, one, and two 3-hydroxybutanoyl substituents.

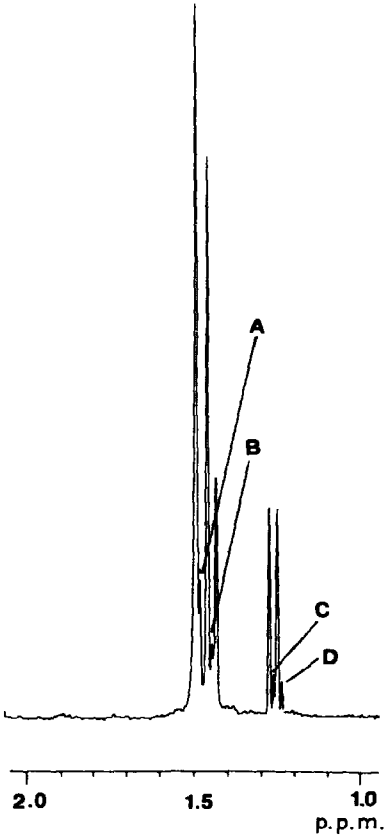


Fig. 7. Partial ^{1}H-n.m.r. spectrum of the octasaccharide in 1:9 pyridine–deuterium oxide. The resonances labelled C and D are from the methyl group of the second 3-hydroxybutanoyl group. Note the effect on the methyl resonances of the acetal groups giving rise to A and B. The signal immediately upfield of B is due to the methyl group of the *O*-(1-carboxyethylidene) group on the galactosyl group of octasaccharides bearing 3-hydroxybutanoyl groups. The most downfield resonance is due to the acetal group on a glucosyl residue.

oxide containing 10% of pyridine, the signals for the methyl group of 3-hydroxybutanoate showed two small, additional signals, indicating the presence of a second such substituent. However, an estimate of the integrals of the signals indicated that <5% of the units bore two such substituents. The sample was lyophilized and recording the n.m.r. spectrum in deuterium oxide was repeated, in order to ensure that no deacylation had occurred. A partial n.m.r. spectrum of the oligosaccharide in 10% pyridine is shown in Fig. 7. There is a pronounced anisotropic effect of the second substituent on the methyl signals of the acetal groups. This indicated that this second substituent is also localized on the terminal galactosyl group in close proximity to both acetal groups (see Fig. 7).

There were not many sequence ions in the f.a.b.-mass spectrum. A fragment at *m/z* 627 was observed for glycosyl cleavage (with charge retention at the glycosyl carbon atom) of the three terminal residues of the side chain. Another fragment, at *m/z* 713, was observed for the same fragment having a 3-hydroxybutanoyl substituent attached. Fragments at *m/z* 1133 and 883, expected from elimination of the branching derivatized alditol from the nearest glycosyl residue on the main chain and side chain (**d–e–f–g–h** and **a–b–c–d** in formula **1a**) respectively (with charge retention on the alditol fragment[15]) were observed, but were of low intensity. This did, however, confirm that there were four sugar residues in the side chain and four in the main chain.

One major question that arose in the rationalization of the mode of action of the enzyme was the basis for the selectivity of cleavage between the branching sugar and the uronic acid residue in preference to cleavage between the two uronic acid residues (**a–b**). This selectivity can be explained by the unique configuration of the glucose residue attached to the uronic acid residue that is involved in the cleavage. Another explanation is the functionalization of the second uronic acid residue (**b**) by the acetyl group. The enzyme apparently does not recognize this modified uronic acid group (**a**) as a potential site of cleavage.

The efficiency of the phage depolymerase method in generating oligosaccharides with an unchanged distribution of non-carbohydrate substituents (compared to the parent polymer) cannot be approached by any known chemical methods. Kuo and Mort[9] used liquid hydrogen fluoride to cleave a similar polysaccharide into oligomer units, but they encountered loss of *O*-(1-carboxyethylidene) groups, galactosyl residues, and 3-hydroxybutanoyl substituents. It also led them to the erroneous conclusion that 3-hydroxybutanoate is not present in the acidic extracellular polysaccharide from *R. trifolii* 0403. The capsular polysaccharide of other rhizobial strains contains[9] 3-hydroxybutanoyl groups in the D configuration.

In an earlier study[8], we identified quantitative growth-phase-dependent changes in the relative molar proportions of non-carbohydrate components of oligosaccharides similar to those used in this study. These changes can now be assigned to progressive substitution of the galactosyl group in the side chain by 3-hydroxybutanoyl and *O*-(1-carboxyethylidene) groups as the culture advances in stationary phase.

Now that the complete structure of this polysaccharide is known, and definitive, unambiguous assignments of the signals in the n.m.r. spectra of the enzymically produced oligosaccharide can be made, it will be possible to determine whether this structure is different in *Rhizobium* species incapable of nodulating clover; this can be achieved by using heterologous species of *Rhizobium,* and specific mutations in *Rhizobium trifolii* genes governing host specificity. This structure also contains crucial information needed to permit understanding the molecular recognition of *R. trifolii* by trifoliin A, and cell–cell signal transduction from this bacterial symbiont to infectable, clover-root hairs.

ACKNOWLEDGMENTS

This work was supported by N.S.F. grant 80-21906 and by N.I.H. grants 34331-01, 02, and 03. We thank L. Lee, N. R. Nirmala, and I. Jardine for technical assistance, and L. Anderson, C. C. Sweeley, W. Reusch, R. Reusch, and S. Philip-Hollingsworth for reading the manuscript. This is article No. 12450 of the Michigan Agricultural Experiment Station.

REFERENCES

1 H. LJUNGGREN AND G. FAHRAEUS, *Nature (London)*, 184 (1959) 1578–1579.
2 F. DAZZO AND D. HUBBELL, *Appl. Microbiol.*, 30 (1975) 1017–1033.
3 P.-E. JANSSON, B. LINDBERG, AND H. LJUNGGREN, *Carbohydr. Res.*, 75 (1979) 207–220.
4 B. K. ROBERTSON, P. ÅMAN, A. G. DARVILL, M. MCNEIL, AND P. ALBERSHEIM, *Plant Physiol.*, 67 (1981) 389–400.
5 A. CHAUDHARI, C. T. BISHOP, AND W. F. DUDMAN, *Carbohydr. Res.*, 28 (1973) 221–231.
6 R. I. HOLLINGSWORTH, M. ABE, F. B. DAZZO, AND K. HALLENGA, *Carbohydr. Res.*, 133 (1984) c1–c4.
7 R. I. HOLLINGSWORTH, M. ABE, J. E. SHERWOOD, AND F. B. DAZZO, *J. Bacteriol.*, 160 (1984) 510–516.
8 M. ABE, J. E. SHERWOOD, R. I. HOLLINGSWORTH, AND F. B. DAZZO, *J. Bacteriol.*, 160 (1984) 517–523.
9 M. KUO AND A. J. MORT, *Carbohydr. Res.*, 145 (1986) 247–265.
10 M. MCNEIL, J. DARVILL, A. G. DARVILL, P. ALBERSHEIM, R. VAN VEEN, P. HOOYKAAS, R. SCHILPEROORT, AND A. DELL, *Carbohydr. Res.*, 146 (1986) 307–326.
11 A. AMEMURA, T. HARADA, M. ABE, AND S. HIGASHI, *Carbohydr. Res.*, 113 (1983) 165–174.
12 S. K. GHAI, M. HISAMATSU, A. AMEMURA, AND T. HARADA, *J. Gen. Microbiol.*, 122 (1981) 33–40.
13 D. E. DORMAN AND J. D. ROBERTS, *J. Am. Chem. Soc.*, 92 (1970) 1355–1361.
14 K. BOCK AND C. PEDERSEN, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 27–66.
15 J. P. KAMERLING, W. HEERMA, J. F. G. VLIEGENTHART, B. N. GREEN, L. A. S. LEWIS, G. STRECKER, AND G. SPIK, *Biomed. Mass Spectrom.*, 10 (1983) 420–425.