Carbohydrate Research, 200 (1990) 9–31
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

Structural analysis of xyloglucan oligosaccharides by ^{1}H-n.m.r. spectroscopy and fast-atom-bombardment mass spectrometry*

William S. York, Herman van Halbeek, Alan G. Darvill, and Peter Albersheim†
Complex Carbohydrate Research Center, The University of Georgia, 220 Riverbend Road, Athens GA, 30602 (U.S.A.)

(Received April 5th, 1989; accepted for publication, August 4th, 1989)

ABSTRACT

A method to determine rapidly the identities and proportions of the oligosaccharide repeating-units in plant cell-wall xyloglucans by 1D ^{1}H-n.m.r. spectroscopy was developed. Six of the most commonly found xyloglucan oligosaccharide subunits (including three subunits that had not been fully characterized previously) were prepared by endo-(1→4)-β-D-glucanase digestion of xyloglucans from various plant species. The oligosaccharides were reduced to the corresponding oligoglycosyl-alditols, purified, and characterized by glycosyl composition and linkage analysis, ^{1}H-n.m.r. spectroscopy, and f.a.b.-mass spectrometry. Correlations between the ^{1}H-n.m.r. spectra and the structures of the oligoglycosyl-alditols can be used to identify oligoglycosyl-alditols derived from xyloglucans of unknown structure. The identities and relative amounts of the oligosaccharide subunits of xyloglucans isolated from tamarind seed and rapeseed hulls were determined on this basis.

INTRODUCTION

Xyloglucans (XGs) are an important class of plant polysaccharides, structurally related to cellulose and intimately associated with cellulose in the primary cell wall of plants[1]. They differ from cellulose in that they are branched, with α-D-Xyl*p*, β-D-Gal*p*-(1→2)-α-D-Xyl*p*, and α-L-Fuc*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p* side chains at C-6 of many of the β-D-Glc residues of the backbone[1–7]. XGs comprise a significant percentage of the dry weight of the cell walls of actively growing plant tissues. XGs and their oligosaccharide fragments are thought to play important roles in controlling cell-walls elongation during growth[8,9].

The relationship between the structure of XGs and their function in the plant cell wall has not been fully ascertained. XGs bind strongly to cellulose and appear to coat the surface of cellulose microfibrils in the walls of growing plant cells[10]. The metabolism (synthesis and degradation) of XGs is enhanced by auxins[11]. XGs may act as cell-wall stabilizing molecules that non-covalently cross-link cellulose microfibrils[10]. It has been suggested[12] that "hemicelluloses (including XGs) in the cell-wall matrix direct the

*This is number XXIX of the series entitled "The Structure of Plant Cell Walls".
†Author for correspondence.

0008-6215/90/$03.50

postion of insoluble cellulose molecules into a helicoidal arrangement". In addition, fucose-containing oligosaccharide fragments of XG inhibit[8,9] auxin-stimulated elongation of excised pea-stem segments *in vitro*.

The physical properties of XGs and XG oligosaccharides are likely to be determined by the conformations adopted by these molecules in solution or in the solid state. ^{1}H-N.m.r. spectroscopy is a powerful tool for determining both the sequence (primary structure)[13–15] and conformation (secondary and tertiary structure)[15–18] of complex carbohydrates. However, full assignment of the ^{1}H-n.m.r. spectra is necessary prior to their use in conformational analyses of complex carbohydrates[15–18].

We now report ^{1}H-n.m.r. data for an XG-derived diglycosyl-alditol (**1R**) and for the reduced forms (oligoglycosyl-alditols **2R**–**7R**) of the six most common oligosaccharides (**2**–**7**) produced by digestion of plant XGs with endo-(1→4)-β-D-glucanase. The oligoglycosyl-alditols afford simpler ^{1}H-n.m.r. spectra than the parent oligosaccharides because of the absence of reducing-end effects.

The characterization of these oligoglycosyl-alditols and assignment of their ^{1}H-n.m.r. spectra led to a method for the rapid identification and quantitation of XG oligosaccharides based on inspection of selected regions of the 1D ^{1}H-n.m.r. spectra. In order to generate a sufficiently comprehensive data base, the ^{1}H-n.m.r. spectra of oligoglycosyl-alditols prepared from the XGs that had been isolated from sycamore extracellular polysaccharides (SEPS)[2], tamarind seed[4], and rapeseed hulls[5] were analyzed. Oligosaccharides **2**, **6**, and **7** had previously been characterized[1,7] after isolation from SEPS XG and mung bean. However, the structures of oligosaccharide subunits **3**–**5**, released from the XGs of rapeseed hulls[5] and tamarind seeds[4], had not previously been fully ascertained. ^{1}H-N.m.r. spectroscopy, f.a.b.-mass spectrometry, and chemical analyses were used to determine the primary structures of the oligosaccharide fragments. The identities and relative amounts of the oligosaccharide subunits of the XGs from rapeseed hulls and tamarind seeds could then be readily deduced by inspection of the 1D ^{1}H-n.m.r. spectra of the oligoglycosyl-alditol fractions obtained by digestion with endo-(1→4)-β-D-glucanase followed by chromatography on Bio-Gel P-2 and reduction with borohydride.

RESULTS

Nomenclature for glycosyl residues of the oligoglycosyl-alditols. — Oligoglycosyl-alditols **2R**–**7R** were formed by borohydride reduction of oligosaccharides **2**–**7**, respectively. Their glycosyl residues are given a superscript (a–c) which specifies the proximity of the residue to the Glcol moiety (Fig. 1). Residues that are part of a side chain are given the superscript of the glucosyl residue to which the side chain is attached.

Preparation of the oligoglycosyl-alditols. — XGs isolated from the extracellular polysaccharides of suspension-cultured sycamore (*Acer pseudoplatanus*) cells[2], rapeseed hulls[5], and tamarind seed gum[4] were digested with *Trichoderma viride* endo-(1→4)-β-D-glucanase which selectively hydrolyzed the linkages of those 4-linked β-Glc residues in

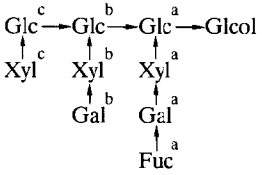


Fig. 1 Designation of the glycosyl residues in **1R**–**7R**; the letters a–c indicate the proximity of the residue to the Glcol residue.

β-D-Glc*p*-(1→4)-D-Glcol
6
↑
1
α-D-Xyl*p*

1R

β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-D-Glcol
6 6 6
↑ ↑ ↑
1 1 1
α-D-Xyl*p* α-D-Xyl*p* α-D-Xyl*p*

2R

β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-D-Glcol
6 6 6
↑ ↑ ↑
1 1 1
α-D-Xyl*p* α-D-Xyl*p* α-D-Xyl*p*
2
↑
1
β-D-Gal*p*

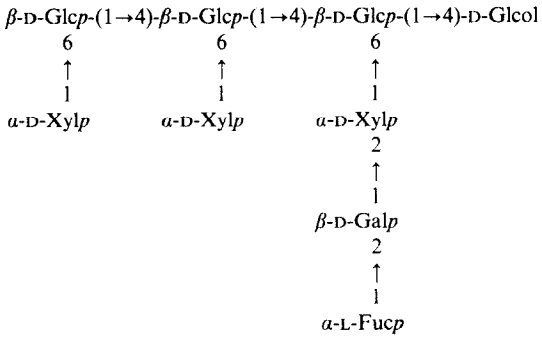
3R

β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-D-Glcol
6 6 6
↑ ↑ ↑
1 1 1
a-D-Xyl*p* α-D-Xyl*p* α-D-Xyl*p*
2
↑
1
β-D-Gal*p*

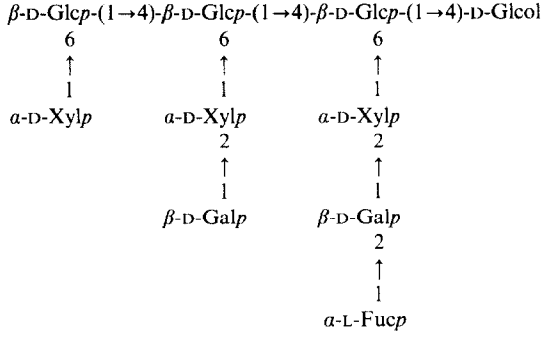
4R

β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p*-(1→4)-D-Glcol
6 6 6
↑ ↑ ↑
1 1 1
a-D-Xyl*p* α-D-Xyl*p* α-D-Xyl*p*
2 2
↑ ↑
1 1
β-D-Gal*p* β-D-Gal*p*

5R



6R



7R

the backbone that were not glycosylated at C-6. The released oligosaccharides were fractionated by gel-permeation chromatography on Bio-Gel P-2. The heterogeneous fractions were reduced with borohydride, and the products were further purified by reverse-phase h.p.l.c. The preparations of **2R** and **6R** (from SEPS), **5R** (from tamarind seed), and **7R** (from rapeseed hulls) were essentially homogeneous (as demonstrated by ^{1}H-n.m.r. spectroscopy, see Fig. 2), but **3R** and **4R** (from tamarind seed) could not be separated, and were analyzed as a mixture. The sequences of the oligoglycosyl-alditols were deduced from their glycosyl compositions and linkages (Table I), their negative ion-f.a.b.-mass spectra (Fig. 3 and Table II), and the positive ion-f.a.b.-mass spectra of their acetylated derivatives (Fig. 4 and Table III). The anomeric configurations, stereochemistry, and glycosylation sites of the constituent sugars of **1R**–**7R** were confirmed by 1D and 2D ^{1}H-n.m.r. spectroscopy (Figs. 2, 5, and 6; Tables IV–VI).

Determination of the glycosyl sequence of the oligoglycosyl-alditols by f.a.b.-m.s. and chemical analysis. — One of the motivations for developing spectroscopic techniques for the structural analysis of XG oligosaccharides is that quantitative recovery of derivatives is not generally achieved during linkage analysis. Undermethylation, the differential resistance of the various glycosidic linkages to hydrolysis, and the differ-

TABLE I

Compositions[a] of **1R–7R**

Residue	**1R**	**2R**	**3R/4R**[b]	**5R**	**6R**	**7R**
Glc[c]	2.1	4.1	3.8	4.1	3.8	3.2
Xyl	1.0	3.0	3.0	3.0	3.0	3.0
Gal	0.0	0.0	1.0	2.3	1.0	1.7
Fuc	0.0	0.0	0.0	0.0	1.0	1.0

Residue	*Positions methylated*[d]	*Deduced linkage*	**1R**	**2R**	**3R/4R**[b]	**5R**	**6R**	**7R**
Glc	1,2,3,5,6	4 Glcol	0.5	0.6	0.5	0.6	0.6	0.6
Glc	2,3,4	6 Glc	1.0	1.0	1.0	1.0	1.0	1.0
Glc	2,3	4,6 Glc	0.0	2.0	2.3	2.2	2.3	2.2
Xyl	2,3,4	T Xyl	0.8	2.7	1.5	1.1	1.7	1.2
Xyl	3,4[e]	2 Xyl	0.1	0.2	0.9	1.6	1.2	2.2
Gal	2,3,4,6	T Gal	0.0	0.0	0.6	1.4	0.1	0.8
Gal	3,4,6	2 Gal	0.0	0.0	0.0	0.0	0.7	0.7
Fuc	2,3,4	T Fuc	0.0	0.0	0.0	0.0	0.8	0.8

[a] Molar ratios obtained using hydrolysis conditions that were a compromise between complete release of the β-Glc residues and degradation of Xyl residues. Therefore, the ratios are not always integers. See also footnote *e*. [b] Analyzed as a mixture. [c] Includes glucitol residue. [d] The glycosyl residues were converted into alditols and then acetylated. [e] Some 2,3-di-*O*-methylxylitol triacetate, which co-eluted with 3,4-di-*O*-methylxylitol triacetate, was also detected. This product was attributed to undermethylation, and contributed to the nonstoichiometric recoveries of the various derivatives.

ential susceptibility of the various glycosyl residues to degradation by acid or base contribute to this problem. Although the recovery of derivatives can be optimized, the yields are never quantitative. The semi-quantitative nature of residue and linkage analyses is evident in the data in Table I, where no special effort was made to optimize the conditions of methylation or hydrolysis. Nevertheless, the chemical analyses (Table I), when combined with the positive ion- and negative ion-f.a.b.-m.s. data (see below), provided enough information to deduce unambiguously the sequences of **1**–**7**, given that a backbone consisting of (1→4)-linked β-Glc residues was assumed. All previously reported XG structures have such a backbone[1–7], and this structural feature was substantiated by subsequent ^{1}H-n.m.r. analysis. The sequences **1R–7R** were also consistent with data[19] obtained using an α-D-xylosidase that was isolated from pea seedlings treated with 2,4-dichlorophenoxyacetic acid.

The negative ion-f.a.b. mass spectra of the oligoglycosyl-alditols **1R–7R** were recorded (*e.g.*, Fig. 3) using 1-amino-2,3-dihydroxypropane (aminoglycerol, AG) as the ionization matrix[20]. These spectra included abundant $[M-H]^-$ ions and "Type B" fragment[21] ions (Table II), and were superior to those obtained with glycerol, thioglycerol, or diethanolamine. The presence of "double-cleavage" ions prevented the glycosyl sequences of XG oligoglycosyl-alditols **2R–7R** from being deduced solely from the negative ion-f.a.b.-mass spectra[21,22], but the glycosyl sequence of the linear oligoglycosyl-alditol **1R** could be determined unambiguously from its linkage composition and negative ion-f.a.b.-mass spectrum.

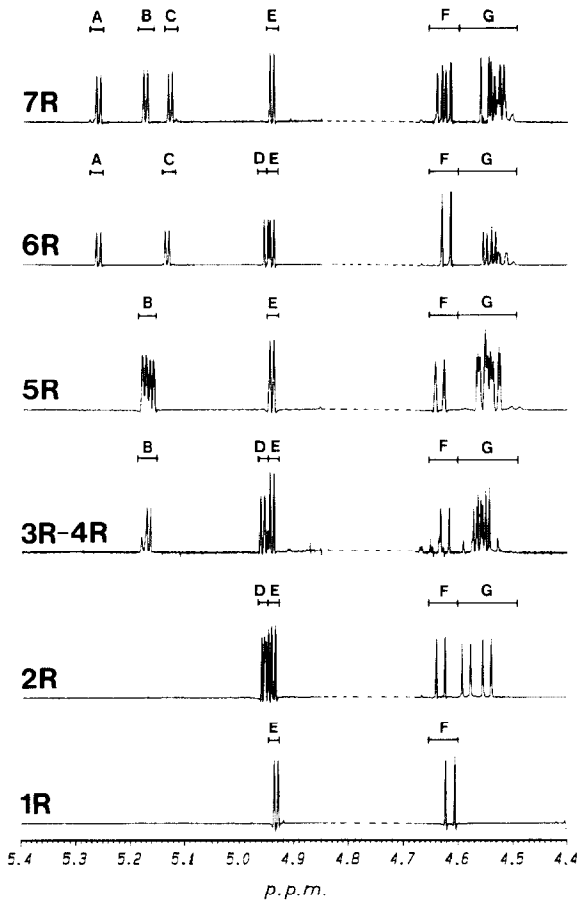


Fig. 2 H-1 resonances of oligoglycosyl-alditols **1R**–**7R**. The HDO signal at δ 4.75 has been deleted for clarity (dashed lines). The H-1 resonances are assigned as follows. A, α-Fuc; B, α-Xyl with a β-Gal substituent at C-2; C, α-Xyl with an α-Fuc-(1→2)-β-Gal substituent at C-2; D, non-reducing terminal α-Xyl^a or α-Xyl^b (see text and Fig. 1 for nomenclature); E, non-reducing terminal α-Xyl^c (or the lone α-Xyl in **1R**); F, β-Glc^a and 2-linked β-Gal^a; G, β-Glc^b, β-Glc^c, non-reducing terminal β-Gal^a and β-Gal^b. Region G also includes H-5 of the α-Fuc residue.

The positive ion-f.a.b.-mass spectra of acetylated[21,22] **1R**–**7R** were recorded using thioglycerol as the matrix (Fig. 4). Intense signals corresponding to $[M + NH_4]^+$ ions[23] and A^+ type[24] (non-reducing end) fragment ions were observed, along with low-intensity signals corresponding to $[M + H]^+$ and $[J2]^+$ ions[24]. The observation of abundant $[M + Na]^+$ ions when NaOAc was added to the thioglycerol matrix confirmed the assignment of the $[M + NH_4]^+$ ions[21,22].

The positive ion-f.a.b.-mass spectra of acetylated **2R**–**7R** also included low-abundance "double-cleavage" ions[21]. For example, the A^+ ion at m/z 835 corresponds to a $[Hex_2Pent]^+$ structure that cannot be formed by a single fragmentation. It was

TABLE II

Negative ion-f.a.b.-m.s. of **2R**–**7R**

Nominal mass	*Type*[a]	*Ion composition*	*Parent oligoglycosyl-alditol*[b]					
475	B^-	XylGlcGlcol	**2R**		**4R**[c]			
637	B^-	GalXylGlcGlcol		**3R**[c]		**5R**		
769	B^-	Xyl_2Glc_2Glcol	**2R**					
783	B^-	FucGalXylGlcGlcol					**6R**	**7R**
931[d]	B^-	Xyl_2Glc_3Glcol	**2R**	**3R**[c]	**4R**[c]		**6R**	
931[d]	B^-	$GalXyl_2Glc_2Glcol$		**3R**[c]	**4R**[c]			
1063	B^-	Xyl_3Glc_3Glcol		**3R**[c]	**4R**[c]		**6R**	
1077	B^-	$FucGalXyl_2Glc_2Glcol$					**6R**	
1093[d]	B^-	$GalXyl_2Glc_3Glcol$		**3R**[c]	**4R**[c]			
1093[d]	B^-	$Gal_2Xyl_2Glc_2Glcol$				**5R**		
1093[d]	B^-	$GalXyl_2Glc_3Glcol$				**5R**		**7R**
1225	B^-	$GalXyl_3Glc_3Glcol$				**5R**	**6R**	**7R**
1239[d]	B^-	$FucGalXyl_2Glc_3Glcol$					**6R**	**7R**
1239[d]	B^-	$FucGal_2Xyl_2Glc_2Glcol$						**7R**
1255	B^-	$Gal_2Xyl_2Glc_3Glcol$				**5R**		
1371	B^-	$FucGalXyl_3Glc_3Glcol$						**7R**
1387	B^-	$Gal_2Xyl_3Glc_3Glcol$						**7R**
1401	B^-	$FucGal_2Xyl_2Glc_3Glcol$						**7R**
1063	$[M-H]^-$	Xyl_3Glc_3Glcol	**2R**					
1225	$[M-H]^-$	$GalXyl_3Glc_3Glcol$		**3R**[c]	**4R**[c]			
1371	$[M-H]$	$FucGalXyl_3Glc_3Glcol$					**6R**	
1387	$[M-H]^-$	$Gal_2Xyl_3Glc_3Glcol$				**5R**		
1533	$[M-H]^-$	$FucGal_2Xyl_3Glc_3Glcol$						**7R**

[a] Ions labelled "B^-" arise from the alditol end by a single cleavage, and correspond in mass to "Pathway B" ions described by Dell[21]. $[M-H]^-$ ions were distinguished from "B^-" ions by their greater intensity. [b] Only single-cleavage ions are included (see text). [c] **3R** and **4R** were analyzed as a mixture. [d] Single-cleavage "B^-" ions, such as those at *m/z* 931, 1093, and 1239, could be formed from different parts of the parent molecule, and thus have been assigned more than one glycosyl composition.

important to determine whether or not the ion at *m/z* 835 was a "double-cleavage" ion, because the presence of a "single-cleavage" ion at *m/z* 835 would indicate that Xyl^c, rather than Xyl^b, was substituted with a Gal residue in **5R** and **7R**. The absence of a single-cleavage ion at *m/z* 835 is the *only* feature of the f.a.b.-mass spectra that distinguishes **5R** and **7R** from the alternative structures. The possibility that the ion at *m/z* 835 had the structure $[Gal^cXyl^cGlc^c]^+$, resulting from the presence of an extended side-chain on Glc^c, was rejected for the following reasons. (*a*) The ion at *m/z* 835 was observed in the f.a.b.-mass spectra of acetylated **2R**–**7R**, inclusive. However, the structure of **2R**, in which all the α-Xyl residues are terminal, is well established[1–7], and could not lead to the formation of an ion at *m/z* 835 by a single-cleavage event. (*b*) The abundance of the ion at *m/z* 835 was much lower than would be expected for an ion produced by a single cleavage[21] (Fig. 4). (*c*) Another signal (*m/z* 1123) of low intensity in the f.a.b.-mass spectra of acetylated **5R** and **7R** corresponds to the structure $[Hex_3\text{-}Pent]^+$, which cannot be formed by a single fragmentation, regardless of where the terminal Gal residues are located. Both of the ions (at *m/z* 835 and 1123) can be

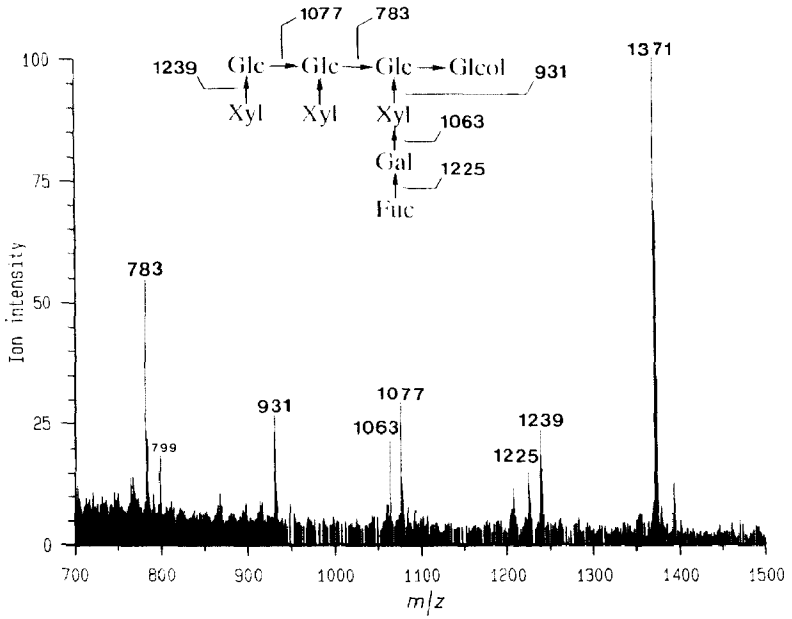


Fig. 3 Negative ion-f.a.b.-mass spectrum of oligoglycosyl-alditol **6R**. Fragment ions are assigned as having formed by cleavage of the C-1–O-1 bond, releasing the alditol-containing fragment as an oxanion[21]. The low intensity ion at *m/z* 799 ([XylGlc$_3$Glcol]$^-$) is assigned as a "double-cleavage" ion[21] formed by the loss of two side chains from the [M – H]$^-$ ion. Ions are labelled with the nominal mass of the isotopomer containing only ^{12}C, ^{1}H, and ^{16}O atoms.

explained by the loss of a single side-chain by a second fragmentation accompanied by acetyl migration, forming, for example, [Glc^c–(Xyl^b)–Glc^b]$^-$ (*m/z* 835) from acetylated **2R**, [Xyl^c–Glc^c–Glc^b]$^-$ (*m/z* 835) from acetylated **2R**, **5R**, and **7R**, and [Glc^c–(Gal^b–Xyl^b)–Glc^b]$^-$ (*m/z* 1123) from acetylated **5R** and **7R**. (*d*) The ions at *m/z* 793 and 1081 (*i.e.*, "double-cleavage" ions having a single unsubstituted hydroxyl group) were more abundant than the corresponding fully acetylated ions at *m/z* 835 and 1123. This finding is consistent with the observation[21] that "double-cleavage" ions in the f.a.b.-mass spectra of acetylated oligosaccharides are more likely to bear an unsubstituted hydroxyl group than to be fully acetylated (as a result of acetyl migration), and further substantiated that the ions at *m/z* 835 and 1123 form as a result of two separate fragmentation events.

One of the fractions obtained by Bio-Gel P-2 chromatography of the endo-(1→4)-*β*-D-glucanase-released fragments of tamarind seed XG was composed of octasaccharides **3** and **4** in the ratio ~3:1. The negative ion-f.a.b.-mass spectrum of the reduced mixture included an intense [M – H]$^-$ ion at *m/z* 1225, indicating the presence of one or more Hex$_4$Pent$_3$Hexol structures. The spectrum also contained fragment ions at *m/z* 475 and 637, suggesting the presence of two oligoglycosyl-alditols having a terminal *β*-Gal residue on side chains "a" and "b", respectively. The positive ion-f.a.b.-mass spectrum of the reduced, acetylated mixture contained a [M + NH$_4$]$^+$ ion at *m/z*

TABLE III

Positive ion-f.a.b.-m.s. of acetylated oligoglycosyl-alditols

Nominal mass	*Type*[a]	*Ion composition*	*Parent oligoglycosyl-alditol*[b]					
259	$[Al]^+$	Xyl	**2R**	**3R**[c]	**4R**[c]	**5R**	**6R**	**7R**
273	$[Al]^+$	Fuc					**6R**	**7R**
331	$[Al]^+$	Gal		**3R**[c]	**4R**[c]			**7R**
547	$[Al]^+$	XylGlc	**2R**	**3R**[c]	**4R**[c]	**5R**	**6R**	**7R**
561	$[Al]^+$	FucGal					**6R**	**7R**
777	$[Al]^+$	FucGalXyl					**6R**	**7R**
1051	$[Al]^+$	Xyl_2Glc_2	**2R**	**3R**[c]			**6R**	
1339	$[Al]^+$	$GalXyl_2Glc_2$			**4R**[c]	**5R**		**7R**
1555	$[Al]^+$	Xyl_3Glc_3	**2R**					
1843	$[Al]^+$	$GalXyl_3Glc_3$		**3R**[c]	**4R**[c]			
2073	$[Al]^+$	$FucGalXyl_3Glc_3$					**6R**	
2131	$[Al]^+$	$Gal_2Xyl_3Glc_3$				**5R**		
2361	$[Al]^+$	$FucGal_2Xyl_3Glc_3$						**7R**
375	$[J2]^+$	Glcol	**2R**	**3R**[c]	**4R**[c]	**5R**	**6R**	**7R**
1964	$[M + NH_4]^+$	Xyl_3Glc_3Glcol	**2R**					
2252	$[M + NH_4]^+$	$GalXyl_3Glc_3Glcol$		**3R**[c]	**4R**[c]			
2482	$[M + NH_4]^+$	$FucGalXyl_3Glc_3Glcol$					**6R**	
2540	$[M + NH_4]^+$	$Gal_2Xyl_3Glc_3Glcol$				**5R**		
2770	$[M + NH_4]^+$	$FucGal_2Xyl_3Glc_3Glcol$						**7R**
1969[d]	$[M + Na_3]^+$	Xyl_3Glc_3Glcol	**2R**					
2257[d]	$[M + Na_3]^+$	$GalXyl_3Glc_3Glcol$		**3R**[c]	**4R**[c]			
2487[d]	$[M + Na_3]^+$	$FucGalXyl_3Glc_3Glcol$					**6R**	
2545[d]	$[M + Na_3]^+$	$Gal_2Xyl_3Glc_3Glcol$				**5R**		
2775[d]	$[M + Na_3]^+$	$FucGal_2Xyl_3Glc_3Glcol$						**7R**

[a] Ions labelled $[Al]^+$ are derived from the non-reducing end, and that labelled $[J2]^+$ is derived from the reducing end of the saccharide[24]. [b] Only single-cleavage ions are included (see text). [c] **3R** and **4R** were analyzed as a mixture. [d] Strong $[M + Na]^+$ ions were observed after adding NaOAc.

2252 and a $[M + Na]^+$ ion at *m/z* 2257 when NaOAc was added to the matrix, confirming the Hex_4Pent_3Hexol composition. In addition, the A^+ ions $[Xyl^c–Glc^c–(Xyl^b)–Glc^b]^+$ (*m/z* 1051) and $[Xyl^c–Glc^c–(Gal^b–Xyl^b)–Glc^b]^+$ (*m/z* 1339) were observed in the ratio of ~3:1. ^{1}H-N.m.r. spectroscopy of the tamarind heptaglycosyl-alditol fraction confirmed that **3R** and **4R** were present in the ratio of ~3:1 (see below).

^{1}H-N.m.r. analysis of the oligoglycosyl-alditols.— The 500-MHz ^{1}H-n.m.r. analysis (Tables IV and V) confirmed the stereochemistry and the anomeric configurations of the glycosyl residues of **1R–7R**. Well-resolved resonances were assigned by evaluating 1D ^{1}H-n.m.r. spectra in light of the deduced glycosyl sequences (see above). 2D-N.m.r. techniques, including double-quantum filtered COSY[25] (Fig. 5) and homonuclear Hartmann–Hahn (HOHAHA)[26] spectroscopy (Fig. 6) allowed nearly all of the ^{1}H signals of **1R–7R** to be assigned. Oligoglycosyl-alditols **1R** and **2R** were also analyzed by ^{1}H-detected (^{1}H,^{13}C) correlation spectroscopy (HMQC)[28], which allowed ^{13}C signals to be assigned on the basis of one-bond scalar coupling to assigned protons (Table VI). The ^{13}C assignments were consistent with published data[29,30] for related di- and tri-saccharides. Finally, the linkage patterns and sequences of the glycosyl residues of **1R–7R** were correlated with the chemical shifts (Table IV) of various well-resolved ^{1}H resonances.

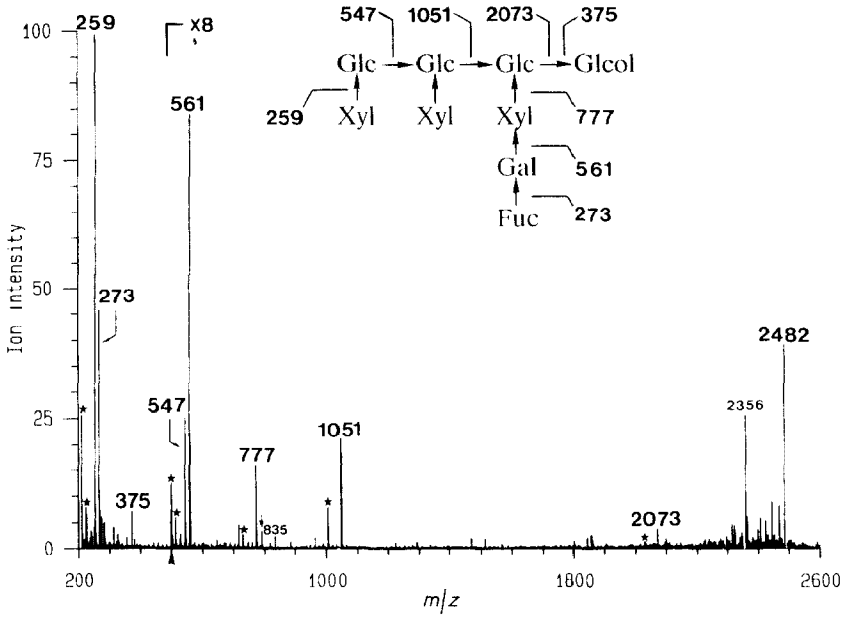


Fig. 4 Positive ion-f.a.b.-mass spectrum of acetylated oligoglycosyl-alditol **6R**. Ions that differ in mass from primary fragment ions by 42 a.m.u. (*i.e.*, ketene) are marked with a *. High-mass ions include $[M + NH_4]^+$ at *m/z* 2482, $[M + H]^+$ at *m/z* 2465, and ions derived from $[M + NH_4]^+$ (*m/z* 2440, 2398, 2356) and $[M + H]^+$ (*m/z* 2423, 2381) by consecutive losses of 42 a.m.u. (ketene). These psuedomolecular ions were suppressed and the $[M + Na]^+$ ion (*m/z* 2487) was predominant when NaOAc was added to the matrix[21,22].

Assignment of the resonances of the geminal protons in the glucitol residues. — The resonances of the geminal protons (*i.e.*, H-1/1′ and H-6/6′) in the glucitol residues of **1R**–**7R** exhibited similar coupling patterns (Table V), and were not easily distinguished. Therefore, the ^{1}H-n.m.r. spectrum of non-stereospecifically C-1 deuterio-labelled **1R** (see Experimental) was subtracted from that of **1R**. The difference spectrum included two dominant positive multiplets (dd, δ 3.743, $J_{1,2}$ 4.2, $J_{1,1'}$ −11.7 Hz; dd, δ 3.674, $J_{1',2}$ 6.2, $J_{1,1'}$ −11.7 Hz), corresponding to H-1 and H-1′ of the glucitol moiety, and two dominant negative multiplets (d, δ 3.722, $J_{1,2}$ 4.1 Hz; d, δ 3.655, $J_{1',2}$ 6.2 Hz), corresponding to H-1 and H-1′ of the diastereomeric deuterio-labelled glucitol moieties. The observed isotope effect (*i.e.*, an upfield shift of ~0.02 p.p.m. for the resonances of H-1 and H-1′) for the deuterio-labelled **1R** is consistent with these assignments[31]. These results were used to distinguish and assign the H-1/1′ and H-6/6′ resonances for the glucitol residues of **2R** **7R** (Tables IV and V).

Assignment of anomeric resonances. — Many of the anomeric resonances of **1R**–**7R** could be assigned (Fig. 2, Tables IV and V; see also ref. 3 for the ^{1}H-n.m.r. assignments for some corresponding oligosaccharides) by analyzing their 1D ^{1}H-n.m.r. spectra in the light of their deduced sequences and linkage patterns. Correlations between the structures of **1R**–**7R** and the chemical shifts of the *α*-anomeric proton resonances were observed. For example, the doublet at δ 4.94 ($J_{1,2}$ 3.6 Hz, see Fig. 2) in

TABLE IV

Chemical shift assignments of ^1H resonances

	1R								
Residue[a]	*H-1'*	*H-1*	*H-2*	*H-3*	*H-4*	*H-5*	*H-5e*	*H-6*	*H-6'*
Glcol	3.676	3.742	3.922	3.886	3.898	3.945	-----	3.744	3.865
β-Glc[a]	-----	4.614	3.349	3.508	3.558	3.644	-----	3.744	3.962
α-Xyl[a]	-----	4.934	3.541	3.677	3.623	3.544	3.711	-----	-----
	2R								
Glcol	3.676	3.744	3.923	3.89	3.90	3.944	-----	3.747	3.868
β-Glc[a]	-----	4.631	3.413	3.666	3.79	3.76	-----	3.832	4.059
β-Glc[b]	-----	4.586	3.403	3.679	3.732	3.835	-----	3.906	4.006
β-Glc[c]	-----	4.548	3.338	3.517	3.523	3.703	-----	3.781	3.935
α-Xyl[a]	-----	4.951	3.547	3.671	3.63	3.57	3.73	-----	-----
α-Xyl[b]	-----	4.956	3.547	3.732	3.62	3.58	3.73	-----	-----
α-Xyl[c]	-----	4.939	3.542	3.737	3.60	3.54	3.72	-----	-----
	3R[b]								
Glcol	n.a.[c]	n.a.	n.a.	n.a.	n.a.	3.964	-----	3.747	3.869
β-Glc[a]	-----	4.624	3.434	3.674	n.a.	n.a.	-----	n.a.	n.a.
β-Glc[b]	-----	4.562	3.401	3.674	3.736	3.827	-----	3.900	4.015
β-Glc[c]	-----	4.550	3.344	3.52	3.525	3.70	-----	3.777	3.936
α-Xyl[a]	-----	5.167	3.678	3.836	n.a.	n.a.	n.a.	-----	-----
α-Xyl[b]	-----	4.960	3.548	n.a.	n.a.	n.a.	n.a.	-----	-----
α-Xyl[c]	-----	4.943	3.544	3.733	n.a.	n.a.	n.a.	-----	-----
β-Gal[a]	-----	4.559	3.616	3.666	3.927	n.a.	-----	n.a.	n.a.
	4R[b]								
Glcol	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-----	n.a.	n.a.
β-Glc[a]	-----	4.643	3.428	n.a.	n.a.	3.76	-----	3.833	4.059
β-Glc[b]	-----	4.582	3.425	n.a.	n.a.	n.a.	-----	n.a.	n.a.
β-Glc[c]	-----	4.535	3.342	n.a.	n.a.	n.a.	-----	n.a.	n.a.
α-Xyl[a]	-----	4.953	3.545	n.a.	n.a.	n.a.	n.a.	-----	-----
α-Xyl[b]	-----	5.175	3.678	n.a.	n.a.	n.a.	n.a.	-----	-----
α-Xyl[c]	-----	4.943	3.544	n.a.	n.a.	n.a.	n.a.	-----	-----
β-Gal[b]	-----	4.565	3.624	n.a.	n.a.	n.a.	n.a.	-----	-----
	5R								
Glcol	3.716	3.783	3.939	n.a.	n.a.	3.968	-----	3.749	3.869
β-Glc[a]	-----	4.635	3.440	3.688	3.656	3.87	-----	3.973	3.92
β-Glc[b]	-----	4.545	3.419	3.67	3.66	3.89	-----	3.97	3.91
β-Glc[c]	-----	4.535	3.341	3.51	3.52	3.697	-----	3.782	3.943
α-Xyl[a]	-----	5.161	3.679	3.840	3.73	3.59	3.67	-----	-----
α-Xyl[b]	-----	5.175	3.672	3.923	3.73	3.57	3.66	-----	-----
α-Xyl[c]	-----	4.943	3.545	3.740	3.622	3.547	3.716	-----	-----
β-Gal[a]	-----	4.559	3.618	3.665	3.927	3.683	-----	3.766	3.817
β-Gal[b]	-----	[d]	[d]	[d]	[d]	[d]	-----	[d]	[d]

TABLE IV continued

	6R								
Glcol	3.680	3.761	3.917	n.a.	3.89	3.962	-----	3.754	3.862
β-Glc[a]	-----	4.621	3.418	3.680	3.597	3.89	-----	3.94[e]	3.90[e]
β-Glc[b]	-----	4.539	3.409	3.674	3.734	3.839	-----	3.893	4.010
β-Glc[c]	-----	4.546	3.341	3.52	3.52	3.698	-----	3.783	3.938
α-Xyl[a]	-----	5.132	3.674	3.782	3.648	3.58	3.69	-----	-----
α-Xyl[b]	-----	4.952	3.546	3.737	3.62	3.57[e]	3.72[e]	-----	-----
α-Xyl[c]	-----	4.941	3.546	3.735	3.61	3.55[e]	3.72[e]	-----	-----
β-Gal[a]	-----	4.621	3.715	3.854	3.917	3.666	-----	3.759	3.812
α-Fuc[a]	-----	5.257	3.796	3.885	3.807	4.519	-----	1.256	-----
	7R								
Glcol	3.678	3.758	3.922	n.a.	n.a.	3.965	-----	3.765	3.860
β-Glc[a]	-----	4.632	3.425	3.699	3.607	3.91	-----	3.9[e]	3.9[e]
β-Glc[b]	-----	4.526	3.429	3.66	3.66	3.90	-----	3.9[e]	3.9[e]
β-Glc[c]	-----	4.535	3.341	3.52	3.52	3.690	-----	3.784	3.946
α-Xyl[a]	-----	5.129	3.674	3.781	3.646	3.587	3.707	-----	-----
α-Xyl[b]	-----	5.174	3.669	3.916	3.66[e]	3.568	3.73	-----	-----
α-Xyl[c]	-----	4.944	3.543	3.736	3.615	3.546	3.717	-----	-----
β-Gal[a]	-----	4.623	3.713	3.853	3.916	3.666	-----	3.76	3.810
β-Gal[b]	-----	4.554	3.617	3.663	3.924	3.683	-----	3.76	3.805
α-Fuc[a]	-----	5.260	3.796	3.885	3.807	4.523	-----	1.259	-----

[a] [a,b,c] Indicate the position of the residue in the oligosaccharide, see Fig. 1. [b] **3R** and **4R** were analyzed as a mixture. [c] Not assigned. [d] The ^{1}H signals of the β-galactosyl residues of **5R** could not be distinguished. [e] A more accurate chemical shift measurement for certain protons was not possible, due to presence of overlapping signals or to higher order effects.

the ^{1}H-n.m.r. spectra of **1R**–**7R**, inclusive, was assigned to a structural element common to **1R**–**7R**, namely, H-1 of the terminal α-Xyl residue linked to C-6 of the unbranched β-Glc residue. The anomeric protons of terminal α-Xyl residues linked to C-6 of branched Glc residues (present in **2R**–**4R**, and **6R**, but not in **1R**, **5R**, or **7R**) yielded signals at δ 4.95. When a β-Gal substituent was present at C-2 of an α-Xyl residue (as in **3R**–**7R**), the resonance of H-1 of the α-Xyl residue was shifted downfield by 0.18–0.23 p.p.m. relative to the corresponding resonance for the terminal α-Xyl residue, yielding a signal at δ 5.12–5.18. The magnitude of this chemical shift effect depended on whether or not the β-Gal residue was substituted with an α-Fuc residue (as in **6R** and **7R**), and to a lesser extent on the position of the α-Xyl residue along the β-D-glucan backbone (Fig. 2 and Table IV). The signal at δ 5.26 (d, $J_{1,2}$ 4 Hz) in the ^{1}H-n.m.r. spectra of **6R** and **7R** was assigned to H-1 of the α-Fuc residue[13]. These assignments were substantiated by using 2D techniques which allowed the ^{1}H chemical shifts and $J_{H,H}$ values of the resonances of the ring protons coupled directly or indirectly to each of the α-anomeric protons described above to be assigned (Tables IV and V).

It was not possible to assign the signals for H-1 of all of the β-glycosyl residues of **1R**–**7R** by comparing their 1D ^{1}H-n.m.r. spectra and, therefore, 2D techniques were

TABLE V

$J_{H,H}$ values for the oligoglycosyl-alditols

Residue[a]	$J_{1,1'}$	$J_{1,2}$	$J_{1',2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5a}$	$J_{4,5e}$	$J_{5e,5a}$	$J_{5,6}$	$J_{5,6'}$	$J_{6,6'}$
	1R										
Glcol	−11.8	4.3	6.3	5.0	2.4	8.5	----	----	5.5	3.1	−11.5
β-Glc[a]	----	7.9	----	9.4	9.4	9.4	----	----	2.3	4.4	−11.3
α-Xyl[a]	----	3.7	----	9.8	9.1	10.2	5.1	−11.1	----	----	----
	2R										
Glcol	−11	4	6	n.a.[b]	n.a.	n.a.	----	----	6	4	−12
β-Glc[a]	----	7.9	----	9.5	10	11	----	----	<2	4.4	−12.1
β-Glc[b]	----	8.0	----	9.4	10	10	----	----	<2	6	−11.9
β-Glc[c]	----	8.1	----	9.5	s.c.[c]	10	----	----	<2	6	−11
α-Xyl[a]	----	3.6	----	10	10	n.a.	n.a.	n.a.	----	----	----
α-Xyl[b]	----	3.7	----	10	10	n.a.	n.a.	n.a.	----	----	----
α-Xyl[c]	----	3.4	----	10	10	n.a.	n.a.	n.a.	----	----	----
	3R[d]										
Glcol	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	----	----	6	4	−12
β-Glc[a]	----	8.6	----	9.8	9	n.a.	----	----	n.a.	n.a.	n.a.
β-Glc[b]	----	7.4	----	9.8	10	10	----	----	<2	5.8	−12.0
β-Glc[c]	----	8.6	----	9.8	s.c.	10	----	----	<2	6	−12
α-Xyl[a]	----	3.6	----	9.5	9	n.a.	n.a.	n.a.	----	----	----
α-Xyl[b]	----	3.7	----	9.8	n.a.	n.a.	n.a.	n.a.	----	----	----
α-Xyl[c]	----	3.7	----	9.8	9	n.a.	n.a.	n.a.	----	----	----
β-Gal[a]	----	8.5	----	9.8	3.6	<2	----	----	n.a.	n.a.	n.a.
	4R[d]										
Glcol	n.a.	n.a.	n.a.	n.a.	n.a	n.a.	----	----	n.a.	n.a.	n.a.
β-Glc[a]	----	7.3	----	9.8	9	n.a.	----	----	<2	4.2	−12.0
β-Glc[b]	----	8.5	----	9.8	n.a.	n.a.	----	----	n.a.	n.a.	n.a.
β-Glc[c]	----	8.5	----	9.7	n.a.	n.a.	----	----	n.a.	n.a.	n.a.
α-Xyl[a]	----	3.6	----	9.8	9	n.a.	n.a.	n.a.	----	----	----
α-Xyl[b]	----	3.6	----	9.8	n.a.	n.a.	n.a.	n.a.	----	----	----
α-Xyl[c]	----	3.7	----	9.8	n.a.	n.a.	n.a.	n.a.	----	----	----
β-Gal[b]	----	8.5	----	9.8	n.a.	n.a.	----	----	n.a.	n.a.	n.a.
	5R										
Glcol	−12	4	6	~4	n.d.	n.d.	----	----	6	3	−12
β-Glc[a]	----	7.5	----	9	9	9	----	----	<2	s.c.	−11
β-Glc[b]	----	7.5	----	9	s.c.	9	----	----	<2	s.c.	−12
β-Glc[c]	----	7.5	----	10	s.c.	9	----	----	~2	5	−11
α-Xyl[a]	----	3.6	----	9	10	10	5	−11	----	----	----
α-Xyl[b]	----	3.6	----	10	10	10	5	−11	----	----	----
α-Xyl[c]	----	3.6	----	10	9	10	5	−11	----	----	----
β-Gal[a]	----	7.5[e]	----	10[e]	3.6[e]	<2[e]	----	----	5[e]	7[e]	−11[e]
β-Gal[b]	----	7.5[e]	----	10[e]	3.6[e]	<2[e]	----	----	5[e]	7[e]	−11[e]

TABLE V continued

	6R										
Glcol	−12	4	6	n.a.	n.a.	n.a.	----	----	6	4	−12
β-Glc[a]	----	7.8	----	9.4	9	10	----	----	<2	s.c.	−11
β-Glc[b]	----	8.0	----	9.5	9	10	----	----	<2	5.8	−11.6
β-Glc[c]	----	7.9	----	9.6	s.c.	10	----	----	<2	5.1	−11.3
α-Xyl[a]	----	3.5	----	10	9	10	5	-12	----	----	----
α-Xyl[b]	----	3.6	----	10	9	10	5	n.a.	----	----	----
α-Xyl[c]	----	3.6	----	10	9	10	5	n.a.	----	----	----
β-Gal[a]	----	7.8	----	10	3.5	<2	----	----	5	7	−12
α-Fuc	----	4.0	----	10	3.5	1.0	----	----	6.5	----	----
	7R										
Glcol	−12	4	6	n.a.	n.a.	n.a.	----	----	6	4	−12
β-Glc[a]	----	8	----	10	9	10	----	----	s.c.	s.c.	s.c.
β-Glc[b]	----	8	----	10	s.c.	s.c.	----	----	s.c.	s.c.	s.c.
β-Glc[c]	----	8	----	10	s.c.	10	----	----	<2	6	−11
α-Xyl[a]	----	4	----	10	9	10	5	−11	----	----	----
α-Xyl[b]	----	4	----	10	9	10	5	−11	----	----	----
α-Xyl[c]	----	4	----	10	9	10	5	−11	----	----	----
β-Gal[a]	----	7.5	----	10	3	<2	----	----	5	7	−12
β-Gal[b]	----	8	----	10	3	<2	----	----	5	7	−12
α-Fuc	----	4	----	10	3	<2	----	----	6	----	----

[a a,b,c] Indicate the position of the residue in the oligosaccharide, see Fig. 1. [b] Not assigned. [c] Strongly coupled. [d] **3R** and **4R** were analyzed as a mixture. [e] The ^{1}H-n.m.r. signals of the β-galactosyl residues of **5R** were indistinguishable.

TABLE VI

Assignment of ^{13}C resonances

Residue[a]		*C-2*	*C-3*	*C-4*	*C-5*	*C-6*
	1R					
Alditol	63.5	73.4	70.5	81.0	72.4	63.2
β-Glc	103.9	74.4	76.9	70.5	75.5	67.0
α-Xyl	99.6	72.7	74.3	70.5	62.4	----
	2R					
Alditol	63.6	73.5	70.5	80.9	72.3	63.2
β-Glc[a]	103.5	74.2	75.2	80.0	74.5	67.3
β-Glc[b]	103.6	73.8	75.2	80.6	74.5	67.3
β-Glc[c]	104.1	74.2	76.7	70.6	75.5	67.2
α-Xyl[a]	100.4	72.6	74.3	70.5	62.7	----
α-Xyl[b]	100.0	72.6	74.2	70.6	62.8	----
α-Xyl[c]	99.5	72.6	74.2	70.7	62.4	----

[a a,b,c] Indicate the position of the residue in the oligoglycosyl-alditol (see Fig.1).

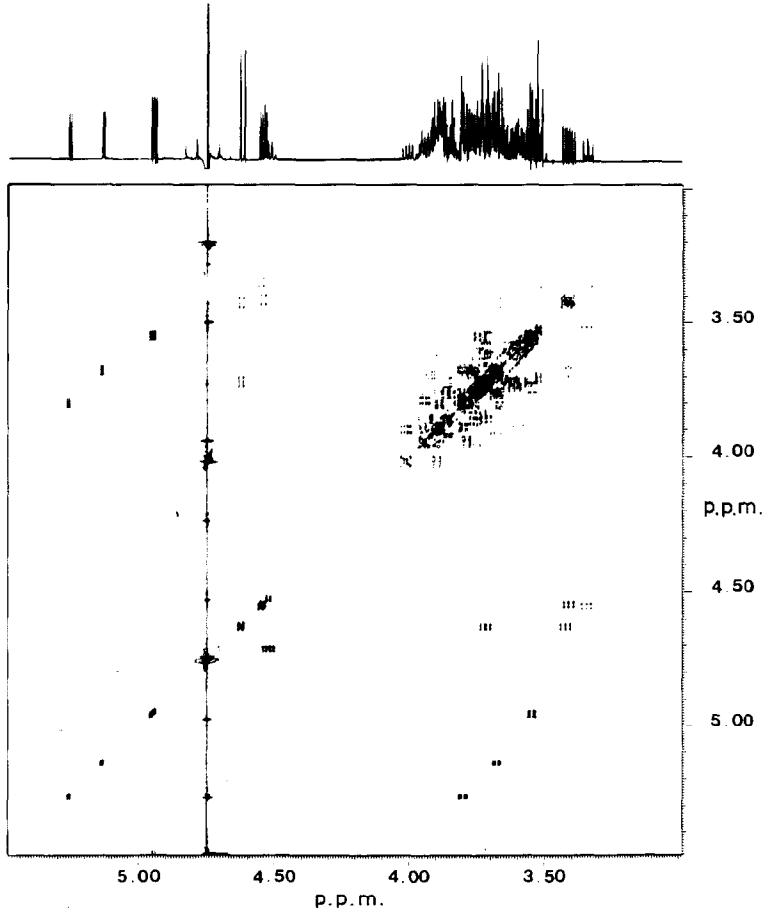


Fig. 5 Double-quantum filtered COSY spectrum of **6R** recorded in the phase-sensitive mode. The H-6 signal of the α-Fuc residue, outside of the narrow spectral width used to record this spectrum, was inserted into the spectrum at δ 4.71.

applied. Initially, the H-1 resonance of the β-Glc^a residue was assigned by comparing the spectra of **2R–7R** and **2–7** (see Fig. 2, Table IV, and ref. 3). The conversion of each oligosaccharide into an oligoglycosyl-alditol resulted in the significant deshielding of a single anomeric proton, shifting its signal away from the bulk of the β-anomeric proton signals into the region δ 4.61–4.64. The H-1 resonances of the β-Glc residues directly connected to Glcol are likely to be affected significantly by the reduction, and the deshielded β-anomeric signals were therefore assigned to Glc^a.

Additional evidence for the assignment of H-1 of the Glc^a residues was obtained by 2D-n.m.r. techniques. The signals of the non-anomeric protons coupled (directly or indirectly) to the (deshielded) protons assigned as H-1 of Glc^a were examined by

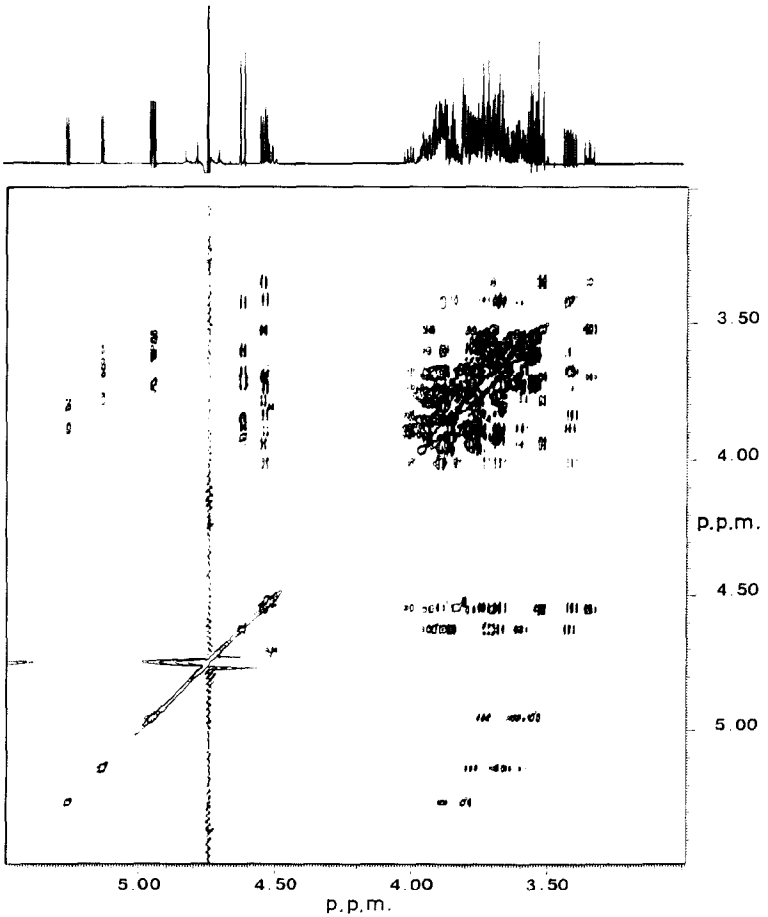


Fig. 6 2D-HOHAHA (TOCSY) spectrum of **6R** recorded in the phase-sensitive mode. Cross-sections through H-1 of glycosyl residues with the *galacto* configuration included only signals of H-1 through H-4, due to the small magnitude of $J_{4,5}$ in these residues. The H-6 signal of the α-Fuc residue, outside of the narrow spectral width used to record this spectrum, was inserted into the spectrum at δ 4.71.

double-quantum filtered COSY and HOHAHA spectroscopy. The chemical shifts and J values of the H-2 through H-6/6′ signals of the residues with deshielded β-anomeric protons indicated that they arose from a 6-linked β-Glc residue in the spectrum of **1R** and from 4,6-linked β-Glc residues in the spectra of **2R**–**7R** (Tables IV and V). The former could be distinguished from the latter because glycosylation at C-4 causes significant deshielding of H-4. This distinction confirmed the assignment of the resonances of H-1 of Glc^a residues, and allowed Glc^b to be distinguished from Glc^c in the ^{1}H-n.m.r. spectra of **2R**–**7R**.

The same 2D techniques were used to assign the remaining β-anomeric proton signals in the ^{1}H-n.m.r. spectra of **2R**–**7R**. The 2-linked and terminal β-Gal residues

could be distinguished from each other (by the glycosylation shift of the resonance of H-2) and from β-Glc residues (by differences in chemical shifts and *J* values of non-anomeric resonances) This approach allowed H-1 of the Glc^a residue (δ 4.63) to be distinguished from H-1 of the Gal^a residue (δ 4.62) of **7R**, and to establish, for example, that the H-1 resonances of Glc^a and Gal^a are coincident (δ 4.62) in the spectrum of **6R**.

Assignment of signals in the mixture of **3R** *and* **4R**. — Chemical and mass-spectrometric data (see above) strongly indicated that **3R** and **4R** were present in the ratio 3:1 in the reduced octasaccharide-containing fraction from tamarind seed XG. This conclusion was confirmed by ^{1}H-n.m.r. spectroscopy. The intensity of the doublet at δ 4.943, assigned to H-1 of the terminal Xyl^c for *both* **3R** and **4R** (see above), was equal to the sum of the intensities of the two partially resolved doublets at δ 5.167 and 5.175, assigned to H-1 of the 2-linked Xyl^a residue in **3R** and H-1 of the 2-linked Xyl^b in **4R** , respectively. The ratio of the intensities of these partially resolved doublets was 3:1. Two other doublets, assigned to H-1 of the terminal Xyl^b (δ 4.960) of **3R**, and H-1 of the terminal Xyl^a (δ 4.953) of **4R**, also had a ratio of intensities of 3:1. These assignments illustrate how the chemical shifts of the H-1 resonances of the α-Xyl residues can be correlated with their substitution pattern (*i.e.*, terminal *vs.* 2-linked α-Xyl) and location (*i.e.*, Xyl^a *vs.* Xyl^b).

Procedure for the analysis of polymeric XGs. — The XG oligoglycosyl-alditol structures **2R**–**7R** correspond to the most common products released when XGs isolated from a wide range of plant species are digested with an endo-(1→4)-β-D-glucanase that specifically hydrolyzes unbranched (1→4)-β-D-glucosyl residues[1-7]. Analysis of **2R**–**7R** made it possible to assign the structures of XG oligoglycosyl-alditols by examination of their 1D ^{1}H-n.m.r. spectra, and a general method to determine the structures and relative proportions of the oligosaccharide subunits of XGs was developed on this basis. The method relies on the well-established techniques of digestion of the XG with endo-(1→4)-β-D-glucanase[1-4], isolation of the oligosaccharide products by gel-permeation chromatography[1-3,8,9], reduction with borohydride, and ^{1}H-n.m.r. spectroscopy of the resulting oligoglycosyl-alditols. Because **2**–**7** are the most common oligosaccharide subunits of XGs, the spectra of unknown reduced oligosaccharide subunits can usually be analyzed by comparison with the ^{1}H-n.m.r. spectra of **2R**–**7R** presented here. Comparison of the behavior of the oligosaccharides to that of authentic standards during high-resolution chromatography on Bio-Gel P-2 provides another criterion for their identification. The anthrone assay[32] is used before reduction to quantitate the molar ratio of the XG oligosaccharide subunits eluted from Bio-Gel P-2 (Fig. 7). Relative molar anthrone-response factors[32] of Glc 1.0, Gal 0.6, Fuc 0.6, and Xyl 0.06 result in relative molar oligosaccharide-anthrone-response factors of: **2**, 4.18; **3**, 4.78; **4**, 4.78; **5**, 5.38; **6**, 5.38; and **7**, 5.98. The relative amounts of oligoglycosyl-alditols in any fraction that contains more than one component is estimated by ^{1}H-n.m.r. spectroscopy. For example, the intensities of the α-anomeric proton signals of the 2-linked Xyl residues (δ 5.167 in **3R** and δ 5.175 in **4R**) indicated that the octasaccharide

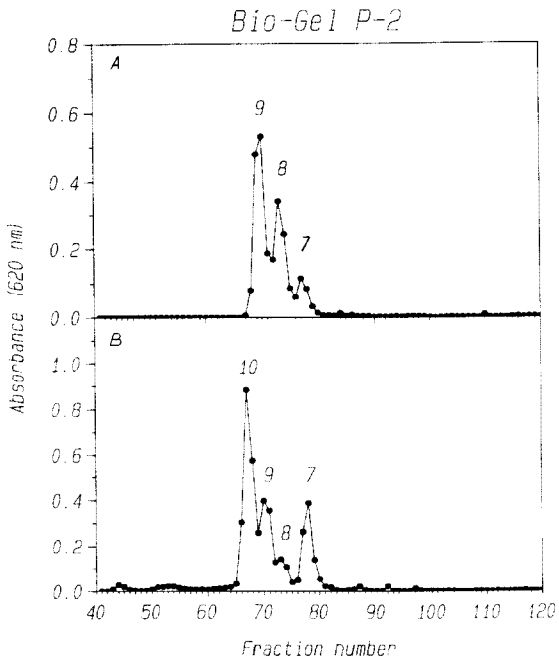


Fig. 7 High-resolution chromatography on Bio-Gel P-2 of the endoglucanase digests of **A**, tamarind seed XG; and **B**, rapeseed hull XG. Hexose content of the fractions was monitored by the anthrone–sulfuric acid assay[32]. Numbers above each peak indicate the d.p. of the major component(s). Oligosaccharide-containing fractions were combined for each peak and analyzed by 1D ^{1}H-n.m.r. spectrosopy after reduction to the corresponding oligoglycosyl-alditols.

fraction from tamarind XG contained **3** and **4** in the ratio 3:1, whereas the corresponding fraction from rapeseed XG contained the same two components in the ratio 1:3.

Application of this procedure (Fig. 7) to the XG from rapeseed indicated that oligosaccharide subunits **2**, **3**, **4**, **6**, and **7** are present in the ratios 11.5:1:3:9.4:16.9. The XG from tamarind seed is composed of subunits **2**–**5** in the ratios 1.4:3:1:5.4. No other oligosaccharide subunits were detected in these XGs.

The ^{1}H-n.m.r. assignments reported here should facilitate analysis of the conformations of XGs in solution, and preliminary information in this regard has been obtained. The C-5–C-6 rotamer populations of β-glycosyl residues can be correlated to chemical shift and $^3J_{H,H}$ data[17,18]. The *gauche–gauche* and *gauche–trans* rotamers (Fig. 8) are the most favorable for 4,6-linked β-Glc residues[17,18], with little contribution of the *trans–gauche* rotamer (see, for example, data for compound III of ref. 17, an analogue of the branched β-Glc residues of XGs). The $^3J_{H,H}$ values for H-6 and H-6′ of the XG β-Glc residues substituted with a *terminal* α-Xyl side-chain are consistent with average rotamer populations that include both the *gauche–gauche* and *gauche–trans* rotamers[17,18].

Significant deshielding of H-5 and shielding of H-4 of the β-Glc residues having a disaccharide or trisaccharide side-chain at C-6 is observed relative to analogous residues having terminal α-Xyl residues at C-6 (Table IV). This chemical shift effect is probably due (at least in part) to changes in the C-5–C-6 rotamer populations for the β-Glc residue that occur when the side chain at C-6 is extended. Accurate measurement of $^3J_{H,H}$ of H-6 and H-6′ is necessary in order to estimate the C-5–C-6 rotamer populations[17,18]. Unfortunately, the $^3J_{H,H}$ values for H-6 and H-6′ of the XG β-Glc residues bearing an extended side-chain could not be measured accurately due to higher order interactions. We are currently attempting to obtain accurate estimates of these $^3J_{H,H}$ values by 2D-n.m.r. techniques in conjunction with spectral simulation and are studying the ^{1}H-n.m.r. spectra of larger XG oligomers in order to examine conformational effects in polymeric XGs.

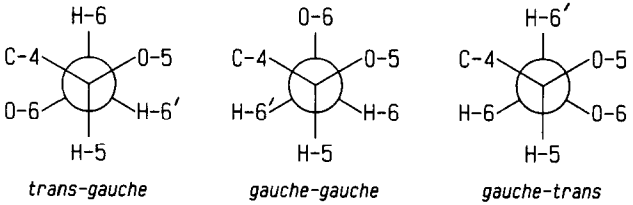


Fig. 8 Newman projections for three C-5–C-6 rotamers. The *trans–gauche* rotamer is energetically unfavorable for 4,6-linked β-Glc residues.

EXPERIMENTAL

Enzymes. — An endo-(1→4)-β-D-glucanase (EC 3.2.1.4) that specifically hydrolyzes unbranched (1→4)-β-D-Glc residues in XGs was isolated[2] from cultures of *Trichoderma viride*. A preparation containing both isoprimeverose-producing oligoxyloglucan hydrolase (IPP-OXG hydrolase)[6] and β-D-galactosidase was prepared from Sanzyme 1000 (Sankyo) as follows. A solution of Sanzyme (4 g) in 20mM NaOAc (60 mL, pH 5.2) was dialyzed against the same buffer (2 L), then loaded on a column of QAE Sephadex Q25-120 (50 mL) equilibrated in the same buffer. The column was washed with buffer (100 mL), and eluted with an NaCl gradient (200 mL; 0.0→0.4M NaCl in the same buffer). Fractions were assayed for β-D-galactosidase and for IPP-OXG hydrolase[6], using *p*-nitrophenyl β-D-galactopyranoside and **2R**, respectively, as substrates. The IPP-OXG hydrolase was eluted at ~0.1M NaCl. The fractions containing the IPP-OXG hydrolase (100 mL) also contained β-D-galactosidase.

Purification of XGs. — SEPS XG was prepared[2] from the culture filtrate of suspension-cultured *Acer pseudoplatanus* cells. Tamarind seed XG[4] was prepared as follows. A suspension of tamarind seed powder (15 g, Dycol Chemicals, Inc.) in water (3 L) was stirred (3 h), then centrifuged (20 000*g*, 30 min), and the supernatant solution was made 70% with ethanol and chilled (4°) for 16 h. The precipitated XG was collected by centrifugation (20 000*g*, 30 min), dissolved in water (500 mL), and lyophilized.

Rapeseed hull XG[5] was obtained by the following procedure. A suspension of pin-milled, hexane-extracted rapeseed (*Brassica napus*) hulls (150 g, a generous gift of John D. Jones of the Food Research Institute, Ottawa, Ontario, Canada) in M imidazole (1.5 L) containing 0.1M EDTA (final pH 7.8 was stirred overnight, then centrifuged (16000*g* for 15 min). The pellet was washed with H_2O (0.5 L) and centrifuged again. A suspension of the pellet in M NaOH (1.5 L) was stirred overnight and then filtered through sintered glass. The solid residue was washed with H_2O (1 L), then suspended overnight in aq. 24% KOH (1.5 L), and filtered through sintered glass, and the pH of the filtrate was brought to 5.2 with glacial acetic acid. The precipitate that formed upon acidification of the filtrate was removed by centrifugation and the supernatant solution was dialyzed against water. The pH of the retentate (2 L) was adjusted to 7.0 by dropwise addition of M imidazole. A portion (1 L) of this solution was passed through a column of DEAE Sephadex (1.5 L, Pharmacia) equilibrated with 20mM imidazole/HCl (pH 7). The pH of the eluate was adjusted to 7.0 with NaOH, and the solution was passed through a second column of DEAE Sephadex (100 mL), equilibrated as before. The eluant was dialyzed and lyophilized, yielding crude XG (800 mg). A portion (400 mg) of this material was dissolved in 50mM NaOAc (25 mL, pH 5.2) and chromatographed on a column (2.3 × 95 cm) of BioRad agarose A5.0m. The partially included hexose-rich fractions were combined, dialyzed, and lyophilized to yield purified rapeseed XG (100 mg).

Preparation of oligoglycosyl-alditols. — α-D-Xyl*p*-(1→6)-β-D-Glc*p*-(1→4)-D-Glcol (**1R**) and its glucitol-*1-d* analogue were prepared as follows. Purified tamarind seed XG (65 mg) was digested (24 h, 30°, in 0.1M NaOAc, pH 5.2, 50 mL) with *T. viride* endo-(1→4)-β-D-glucanase (1.33 units; 1 unit produces 1 μmol of reducing D-glucose per min under these conditions, using carboxymethylcellulose as a substrate.) The products were passed through a column of Amberlite MB3 (50 mL) and then concentrated. The residue was dissolved in M NH_4OH (10 mL), and the solution was divided into two aliquots (8 and 2 mL) to which solid $NaBH_4$ and $NaBD_4$, respectively, were added to a final concentration of 5 mg.mL^{-1}. After 1 h, acetone (50 μL) was added to the solution which was then passed through a column of Dowex 50 (H^+) resin (4 or 1 mL). The solvents were evaporated, and borate was removed from the residue as trimethyl borate by co-distillation with methanol[33]. Each resulting oligoglycosyl-alditol preparation was then treated (24 h, 30°) with excess (8 or 2 mL) of IPP-OXG hydrolase–β-D-galactosidase mixture. Each sample was passed through a column of Amberlite MB3 resin (10 or 2 mL) in order to remove enzymes and salts, and then chromatographed on Bio-Gel P-2 in H_2O to give three fractions, identified as **1R** (H- or D-form), isoprimeverose, and a galactose-rich mixture of monosaccharides.

Larger oligosaccharides were prepared by digestion of purified XGs with the *T. viride* endoglucanase[2] and chromatography of the products on Bio-Gel P-2[2,3,8]. After reduction with $NaBH_4$, the oligoglycosyl-alditols were purified by reverse-phase h.p.l.c., as described below.

Chromatography of the endoglucanase-generated fragments of SEPS XG on Bio-Gel P-2 gave an elution pattern similar to that obtained by Bauer *et al.*[2]. Fucosylated oligosaccharides **6** and **7** were eluted as a single peak, which was resolved from the peak containing **2**. The oligosaccharide fractions (3–5 mg) were each reduced with $NaBH_4$ in M NH_4OH (1 mL), passed through a column of Dowex-50 (H^+) resin (1 mL), and concentrated in a stream of filtered air. Residual borate was removed as above. Reverse-phase h.p.l.c. of the fucosylated oligoglycosyl-alditols on a semipreparative Hibar Lichrosorb RP-18 column (10 × 250 mm) in aq. 11% methanol (5 mL.min^{-1}) yielded **6R** (major product, *T* 17 min) and **7R** (minor product, *T* 14 min). Chromatography of the oligoglycosyl-alditols that lacked Fuc residues in aq 9% methanol yielded **2R** (*T* 11 min) as the major product. Oligoglycosyl-alditol **7R** (and a small amount of **6R**) were also prepared by the same techniques from XG isolated from rapeseed hull.

High-resolution chromatography on Bio-Gel P-2 (see Fig. 7A) of the endoglucanase-generated fragments of XG from tamarind seed yielded three fractions, reduction of which gave **5R**, a mixture of **3R** and **4R**, and **2R**, respectively. Reverse-phase h.p.l.c. failed to resolve **3R** and **4R**.

High-resolution chromatography on Bio-Gel P-2. — Two columns (1.6 × 95 cm, Amicon) packed with Bio-Gel P-2 (−400 mesh) in H_2O were connected in series. Samples (20 mg) were loaded with a Rheodyne low-pressure LC injector and eluted at ambient temperature in H_2O (10 mL.h^{-1}, delivered with a peristaltic pump).

Chemical analysis of oligoglycosyl-alditols. — The glycosyl components of **1R–7R** were analyzed as alditol acetates[32,33], and the glycosyl linkages were determined by analysis of partially methylated alditol acetate derivatives[32,34] as described, except that methylation was catalyzed[35] by solid NaOH in Me_2SO.

F.a.b.-mass spectrometry. — A VG ZAB-SE mass spectrometer operating at low resolution (1:1000) with an accelerating voltage of 8 kV was used. Each oligoglycosyl-alditol was dissolved in H_2O (10 μg.μL^{-1}), and a 1-μL aliquot of this solution was mixed on the probe tip with 1-amino-2,3-dihydroxypropane (2 μL) for negative ion-f.a.b.-m.s. Samples were acetylated[36] using trifluoroacetic anhydride and acetic acid prior to positive ion-f.a.b.-m.s. Each acetylated oligoglycosyl-alditol was dissolved in methanol (10 μg.μL^{-1}), and 1 μL of this solution was mixed on the probe tip with thioglycerol (2 μL). The average masses of non-resolved high-mass ion clusters were converted into nominal masses using the CARBOMASS[37] software.

N.m.r. spectroscopy. — The oligoglycosyl-alditols were passed through Amberlite MB3 resin prior to recording their n.m.r. spectra in order to remove inorganic contaminants that appeared to give rise to spurious signals by interaction with the oligoglycosyl-alditols. Exchangeable protons were replaced with deuterons, and the samples were dissolved (1–10mM) in D_2O (99.96% D). The ^{1}H- and ^{13}C-n.m.r. spectra

were recorded with a Bruker AM 500 spectrometer. Double-quantum filtered (^{1}H, ^{1}H) COSY spectra[25] were recorded in the phase-sensitive mode, using a spectral width of 1250 Hz. The evolution time (t_1) was increased in steps of 400 μs to obtain 512 free induction decays (f.i.d.s), each consisting of the sum of 64 transients acquired in 2k data points. Quadrature detection in t_1 was accomplished by the time-proportional phase-increment (TPPI) method[27]. 2D (^{1}H, ^{1}H) Hartmann–Hahn (HOHAHA) spectra were recorded using the MLEV-17[26] sequence during the mixing period (200 ms), increasing t_1 in steps of 400 μs to obtain 256 f.i.d.s, each consisting of the sum of 48 transients acquired in 2k data points. ^{1}H-Detected (^{1}H, ^{13}C) one-bond shift correlation (HMQC) spectra were recorded as described by Bax *et al.*[28], using a spectral width of 1250 Hz, increasing t_1 in steps of 85 μs to obtain 256 f.i.d.s, each consisting of the sum of up to 320 transients acquired in 1k data points. This resulted in an acquisition time of ~410 ms per transient. The ^{1}H-magnetization of protons attached to ^{12}C was suppressed by the BIRD pulse[28]. The delay between the BIRD pulse and the generation of multiple quantum coherence was 350 ms. ^{13}C-Decoupling during acquisition (t_2) was accomplished using the GARP pulse sequence[38].

ACKNOWLEDGEMENTS

The authors thank Dennis Warrenfeltz (Complex Carbohydrate Research Center) for invaluable technical support. This research was supported by the United States Department of Energy Grant DE-FG09-87ER13819 as part of the USDA/DOE/NSF Plant Science Centers program and also by DOE grant DE-FG09-85ER13424.

REFERENCES

1 M. McNeil, A. G. Darvill, S. C. Fry, and P. Albersheim, *Annu. Rev. Biochem.*, 53 (1984) 625–663.
2 W. D. Bauer, K. W. Talmadge, K. Keegstra, and P. Albersheim, *Plant. Physiol.*, 51 (1973) 174–187.
3 W. S. York, J. E. Oates, H. van Halbeek, A. G. Darvill, P. Albersheim, P. R. Tiller, and A. Dell, *Carbohydr. Res.*, 173 (1988) 113–132.
4 P. Kooiman, *Recl. Trav. Chim. Pays-Bas*, 80 (1961) 849–865.
5 G. O. Aspinall, T. N. Krishnamurthy, and K.-G. Rosell, *Carbohydr. Res.*, 55 (1977) 11–19.
6 Y. Kato, J. Matsushita, T. Kubodera, and K. Matsuda, *J. Biochem. (Tokyo)*, 97 (1985) 801–810.
7 J. Matsushita, Y. Kayo, and K. Matsuda, *Agric. Biol. Chem.*, 49 (1985) 1533–1534.
8 W. S. York, A. G. Darvill, and P. Albersheim, *Plant Physiol.*, 75 (1984) 295–297.
9 G. J. McDougall and S. C. Fry, *Planta*, 175 (1988) 412–416.
10 T. Hayashi and G. Maclachlan, *Plant Physiol.*, 75 (1984) 596–604.
11 T. Hayashi and G. Maclachlan, *Plant Physiol.*, 76 (1984) 739–742.
12 A. C. Neville, *J. Theor. Biol.*, 131 (1988) 243–254.
13 H. van Halbeek, *Biochem. Soc. Trans.*, 12 (1984) 601–605.
14 J. Dabrowski, in W. R. Croasmun and R. M. K. Carlson (Eds.), *Methods in Stereochemical Analysis (Two Dimensional NMR Spectroscopy)*, VCH Publishers, New York, 1987, pp. 349–386.
15 C. A. Bush, *Bull. Magn. Reson.*, 10 (1988) 73–95.
16 J. P. Carver and R.-J. Brisson, in V. Ginsburg and P. W. Robbins (Eds.), *Biology of Carbohydrates*, Vol. 2, Wiley, New York, 1984, pp. 289–331.
17 Y. Nishida, H. Hori, H. Ohrui, H. Meguro, S. Zushi, J. Uzawa, and T. Ogawa, *Agric. Biol. Chem.*, 52 (1988) 1003–1011.
18 S. W. Homans, R. A. Dwek, J. Boyd, M. Mahmoudian, W. G. Richards, and T. Rademacher, *Biochemistry*, 25 (1986) 6342–6350.

19 R. A. O'Neill, P. Albersheim, and A. G. Darvill, *J. Biol. Chem.*, in press.
20 J. L. Gower, *Biomed. Mass Spectrom.*, 12 (1985) 191–196.
21 A. Dell, *Adv. Carbohydr. Chem. Biochem.*, 45 (1987) 19–72.
22 H. Egge, A. Dell, and H. von Nicolai, *Arch. Biochem. Biophys.*, 224 (1983) 235–253.
23 M. A. Baldwin and K. J. Welham, *Org. Mass Spectrom.*, 21 (1986) 235–237.
24 N. K. Kochetkov and O. S. Chizhov, *Adv. Carbohydr. Chem.*, 21 (1966) 39–92.
25 M. Rance, O. W. Sorensen, G. Bodenhausen, G. Wagner, R. R. Ernst, and K. Wuthrich, *Biochem. Biophys. Res. Commun.*, 117 (1983) 479–485.
26 A. Bax and D. G. Davis, *J. Magn. Reson.*, 65 (1985) 355–360.
27 D. Marion and K. Wuthrich, *Biochem. Biophys. Res. Commun.*, 113 (1983) 967–974.
28 A. Bax and S. Subramanian, *J. Magn. Reson.*, 67 (1986) 565–569.
29 K. Bock, C. Pedersen, and H. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 42 (1984) 193–225.
30 K. Bock and H. Pedersen, *J. Carbohydr. Chem.*, 3 (1984) 581–592.
31 R. A. Bernheim and H. Batiz-Hernandez, *J. Chem. Phys.*, 45 (1966) 2261–2264.
32 W. S. York, A. G. Darvill, M. McNeil, T. T. Stevenson, and P. Albersheim, *Methods Enzymol.*, 118 (1986) 3–40.
33 P. Albersheim, D. J. Nevins, P. D. English, and A. Karr, *Carbohydr. Res.*, 5 (1967) 340–345.
34 T. J. Waeghe, A. G. Darvill, M. McNeil, and P. Albersheim, *Carbohydr. Res.*, 123 (1983) 281–304.
35 I. Ciucanu, and F. Kerek, *Carbohydr. Res.*, 131 (1984) 209–217.
36 A. Dell and P. R. Tiller, *Biochem. Biophys. Res, Commun.*, 135 (1985) 1126–1134.
37 W. S. York, R. S. Doubet, A. G. Darvill, and P. Albersheim, *Int. Carbohydr. Symp., XIVth, Stockholm, Sweden*, 1988, Abstr. A9.
38 A. J. Shaka, P. B. Barker and R. Freeman, *J. Magn. Reson.*, 64 (1986) 547–552.