Carbohydrate Research, 79 (1980) 165–192
© Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

A GENERAL AND SENSITIVE CHEMICAL METHOD FOR SEQUENCING THE GLYCOSYL RESIDUES OF COMPLEX CARBOHYDRATES*

BARBARA S. VALENT, ALAN G. DARVILL, MICHAEL MCNEIL, BØRRE K. ROBERTSEN, AND PETER ALBERSHEIM†
Department of Chemistry, University of Colorado, Boulder, Colorado 80309 (U.S.A.)
(Received March 30th, 1979; accepted for publication in revised form, June 10th, 1979)

ABSTRACT

This paper describes a new glycosyl-sequencing method. This method was made possible by the ability to fractionate complex mixtures of peralkylated oligosaccharides by reversed-phase, high-pressure liquid chromatography. The fractionation ability of the reversed-phase system allows the isolation and subsequent unambiguous identification by g.l.c.–m.s. of disaccharides, almost all trisaccharides, and, in some cases, tetrasaccharides generated by successive partial acid hydrolysis, reduction, and ethylation of a permethylated, complex carbohydrate. As these small oligosaccharides overlap within the unhydrolyzed, complex carbohydrate, the oligosaccharide sequences may be pieced together, and, with the glycosyl-linkage composition of the intact complex carbohydrate, can be used to determine the glycosyl sequence of the complex carbohydrate. The details of the sequencing method are illustrated by the elucidation of the glycosyl sequences of three complex carbohydrates. These examples demonstrate the wide variety of complex carbohydrates whose structures can be ascertained by the new sequencing technique. Two of the examples are the commercially available polysaccharides, lichenan and xanthan, whose structures have already been reported. The other example is a nonasaccharide derived from xyloglucan, a structural polymer of plant cell-walls. The glycosyl residues of the complex carbohydrates studied include hexosyl, deoxyhexosyl, pentosyl, glycosyluronic, and pyruvic acetal-substituted hexosyl residues. It will be demonstrated that the new glycosyl-sequencing technique is not compromised by the presence, in the carbohydrate to be analyzed, of glycosyl linkages possessing very different acid labilities. Two major advantages of this sequencing technique are that it is relatively rapid and that it requires only milligram quantities of carbohydrate.

---

*Supported by the U.S. Department of Energy, Contract No. EY-76-S-02-1426 and U.S. Department of Agriculture, Grant No. 616-53-73. One of us (B.K.R.) was supported by a Fellowship from The Norwegian Research Council for Science and the Humanities.

†To whom correspondence should be addressed.

INTRODUCTION

Several techniques are available for determining the glycosyl* compositions[1,2] and the glycosyl-linkage compositions of oligo- and polysaccharides[3,4]. Other techniques provide information about the glycosyl sequences of oligo- and polysaccharides[5]. These techniques include Smith degradation (periodate oxidation)[3,5,6], elimination of glycosyl linkages attached to C-4 of glycosyluronic residues[5,7,8], selective acetolysis[5,6,9,10], and keto degradation[11,12]. A general method that has been applied with some success to the structural analysis of oligo- and poly-saccharides involves partial, acid-catalyzed hydrolysis[5,6,13]. This method depends on the isolation and structural characterization of oligosaccharides produced by partial hydrolysis. This paper reports a refinement of the partial acid-hydrolysis method of sequencing oligo- and poly-saccharides that commends this technique for routine structural analysis. Experimental details of the sequencing technique are illustrated here with three examples.

RESULTS AND DISCUSSION

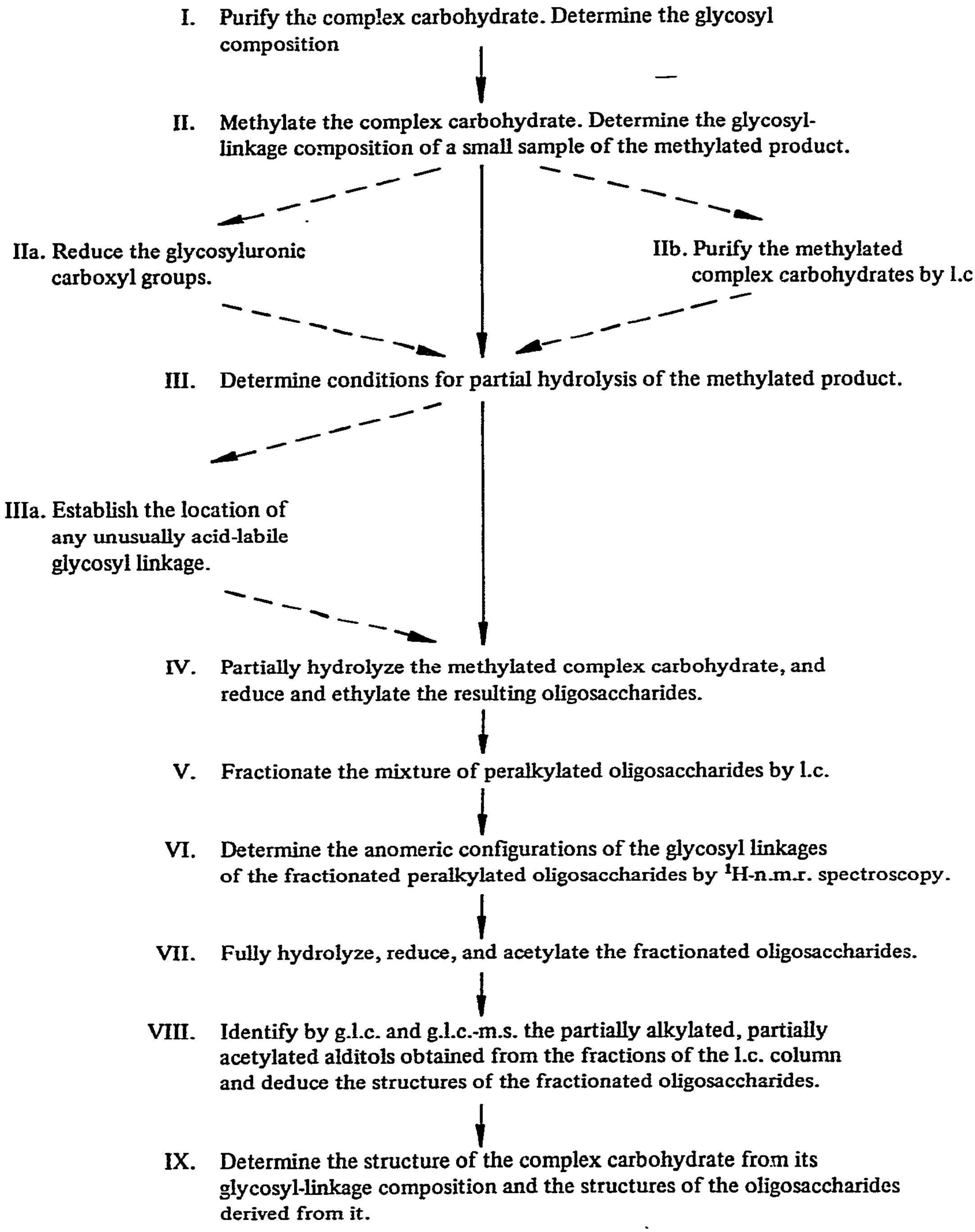
General introduction to the glycosyl-sequencing method. — The steps involved in the glycosyl-sequencing technique are listed in Fig. 1 (see facing page). Successful structural characterization of a complex carbohydrate requires that the carbohydrate be relatively pure (see Step I in Fig. 1). Large oligosaccharides may be purified after permethylation (Step IIb). This is achieved by chromatographing the permethylated oligosaccharide on the same reversed-phase l.c. (h.p.l.c.) column used for fractionation of the mixture of fully alkylated oligosaccharides produced after partial hydrolysis (Step V). Purification of a permethylated oligosaccharide is illustrated with the nonasaccharide of xyloglucan (Example 3).

The carboxyl groups of glycosyluronic residues, if present in the complex carbohydrate being structurally characterized, are reduced to the corresponding dideuterio-labelled primary alcohol (Step IIa). Methods are available for reducing carboxyl groups, either before methylation, by using a water-soluble carbodiimide and sodium borodeuteride[14], or after methylation, by using sodium borodeuteride[15] or lithium aluminum deuteride[16]. The reduction method chosen must be demonstrated to be effective with the particular complex carbohydrate being sequenced. The reduction of the glycosyluronic carboxyl groups of xanthan is presented in Example 2.

The glycose residue, at the reducing terminus of a complex oligosaccharide that is to be sequenced, is reduced by sodium borodeuteride before methylation (pre-reduced) in order to convert it into the corresponding alditol. This alditol then becomes uniquely labelled during the methylation procedure (Step II) as the hydroxyl groups at C-1 and C-5 (or C-4 for furanosyl residues) are methylated (see Example 3).

---

*The terminology and abbreviations used in this paper are described at the beginning of the Experimental section.



I. Purify the complex carbohydrate. Determine the glycosyl composition

II. Methylate the complex carbohydrate. Determine the glycosyl-linkage composition of a small sample of the methylated product.

IIa. Reduce the glycosyluronic carboxyl groups.

IIb. Purify the methylated complex carbohydrates by l.c

III. Determine conditions for partial hydrolysis of the methylated product.

IIIa. Establish the location of any unusually acid-labile glycosyl linkage.

IV. Partially hydrolyze the methylated complex carbohydrate, and reduce and ethylate the resulting oligosaccharides.

V. Fractionate the mixture of peralkylated oligosaccharides by l.c.

VI. Determine the anomeric configurations of the glycosyl linkages of the fractionated peralkylated oligosaccharides by ^{1}H-n.m.r. spectroscopy.

VII. Fully hydrolyze, reduce, and acetylate the fractionated oligosaccharides.

VIII. Identify by g.l.c. and g.l.c.-m.s. the partially alkylated, partially acetylated alditols obtained from the fractions of the l.c. column and deduce the structures of the fractionated oligosaccharides.

IX. Determine the structure of the complex carbohydrate from its glycosyl-linkage composition and the structures of the oligosaccharides derived from it.

Fig. 1. Glycosyl sequence analysis.

The reducing glycose residue of a polysaccharide to be characterized is usually pre-reduced in order to prevent degradation of the polysaccharide[5]. This reduced glycose residue will usually be present in too small a quantity to be detected during the sequencing procedures.

The purified, pre-reduced (when appropriate), and carboxyl-reduced (if glycosyluronic residues are present and if the reduction is conducted prior to methylation) carbohydrate is next permethylated (Step II). The glycosyl-linkage composition of a small portion of the permethylated carbohydrate is obtained by formation of the partially methylated alditol acetates and analysis by g.l.c. and by g.l.c.–m.s.[4]. The major portion of the permethylated carbohydrate is converted, by partial formolysis or hydrolysis, into a complex mixture of partially methylated oligosaccharides. Before the partial formolysis or hydrolysis can be performed, conditions that will afford high yields of di-, tri-, and tetra-saccharides must be ascertained. This is necessary because the acid labilities of the many types of glycosidic linkages usually present in complex carbohydrates vary greatly.

The relative rates of formolysis or hydrolysis of the individual glycosidic linkages in permethylated complex carbohydrates may be monitored by the method described in Example 1, Step III. This method utilizes g.l.c.–m.s. with a 10–50 μg sample of the methylated carbohydrate to ascertain the relative percent cleavage of the glycosidic linkages. By this procedure, the duration and temperature of formolysis or hydrolysis may be varied until optimal conditions are determined. Optimal formolytic or hydrolytic conditions for production of di-, tri-, and tetra-saccharides occur when $\sim$35% of the glycosyl linkages of each glycosyl residue have been cleaved. The method for determining the extent of cleavage of each glycosyl linkage also identifies any glycosyl linkages that may be extremely labile and, thereby, indicates the need to identify the location of these linkages in the oligo- or polysaccharide being characterized (Step IIIa). The procedure to identify particularly acid-labile linkages is illustrated in the xanthan example.

The mixture of partially methylated oligosaccharides produced by fragmentation of the permethylated, complex carbohydrate is reduced with sodium borodeuteride (Step IV). Each resulting reduced and partially methylated oligosaccharide contains free hydroxyl groups wherever other glycosyl residues were originally attached to that particular oligosaccharide when the oligosaccharide was part of the unfragmented carbohydrate. Free hydroxyl groups also occur at C-1 and C-5 (or C-4 for furanosyl residues) of each of the alditols present at what was the reducing end of each of the internal oligosaccharides. All of these are then labelled by ethylation (Step IV). The exact location of the ethoxyl groups is readily ascertained during subsequent g.l.c.–m.s. analysis[4,17]. The location of the ethoxyl groups identifies where, in the unfragmented complex carbohydrate, other glycosyl residues were linked to the small oligosaccharides being examined.

The mixture of totally alkylated small oligosaccharides produced from the complex carbohydrate is fractionated by reversed-phase l.c. (Step V). The effluent from the reversed-phase l.c. column is collected as a series of fractions. The oligo-

TABLE I

MOLE % OF GLUCOSYL RESIDUES IN METHYLATED LICHENAN, AND PERCENT CLEAVAGE OF THE GLUCOSYL LINKAGES OF THESE RESIDUES BY 90% FORMIC ACID ACTING AT 70° FOR VARIOUS TIME-PERIODS

Glucosyl residues	*Mole % of glucosyl residues*	*Duration of formolysis at 70° (min)*					
		0	*20*	*40*	*60*	*90*	*120*
		Formolysis of the glucosyl linkages (%)					
3-Glc	31	0	3	9	16	28	33
4-Glc	69	1	9	13	21	31	38

saccharide-containing fractions from the reversed-phase column may be subjected to ^{1}H-n.m.r. analysis for identification of anomeric protons[18-21] (Step VI). The oligosaccharides present in these fractions are derivatized (Step VII) and structurally identified by g.l.c.–m.s. analysis (Step VIII). The structures of the disaccharides, almost all of the trisaccharides, and, in some cases, the tetrasaccharides may be unambiguously determined by g.l.c.–m.s. analysis. As these small oligosaccharides overlap within the complex carbohydrate, the oligosaccharides can be pieced together and, in conjunction with the glycosyl-linkage composition of the complex carbohydrate, used to determine its structure (Step IX).

Example 1. Structural analysis of lichenan. — *Step I.* Lichenan, a commercially available β-D-glucan isolated from Icelandic moss[22-25], has been structurally investigated by several laboratories (see Fig. 4, p. 174). Glycosyl-composition analysis shows that lichenan is composed solely of glucosyl residues.

Step II. The purified lichenan was pre-reduced with sodium borodeuteride and then methylated as described in the Experimental section. Glycosyl-linkage composition analysis of lichenan shows that the polysaccharide is composed of 4-linked glucosyl residues and 3-linked glucosyl residues in the ratio of 2.2 to 1 (Table I).

Step III. Optimal conditions for producing di-, tri- and tetra-saccharides from permethylated lichenan were determined as follows. Samples of permethylated lichenan were treated with 90% formic acid at 70° for various time-periods. The glucose residues at the reducing ends of each oligosaccharide produced during the partial formolyses were converted into the corresponding partially methylated alditol (glucitol) by reduction with sodium boro*hydride*. The mixture of reduced, partially methylated oligosaccharides was desalted and the remaining glycosyl linkages were hydrolyzed to yield a mixture of partially methylated glucose and glucitol residues. The partially methylated glucose residues were converted into partially methylated glucitol residues by reduction with sodium boro*deuteride*. The resulting, partially methylated glucitol residues, some of which were labelled with a deuterium atom at C-1, were then acetylated.

The percent of formolysis, by the various conditions tested (Table I), of both the 3-linked and the 4-linked glucosyl residues of lichenan was estimated by capillary-column g.l.c.–m.s. analysis. Mass spectra were obtained of the 3- and the

4-linked glucosyl residues as these components eluted from the g.l.c. capillary column. The proportion of the glucosyl residues cleaved during the partial formolysis was estimated from the ratio of a C-1-containing fragment-ion reduced with hydrogen to the same C-1-containing fragment-ion reduced with deuterium. For example, a prominent C-1-containing fragment-ion of the partially methylated alditol acetate derived from a 3-linked glucosyl residue has *m/e* 117 when the 3-linked glucosyl residue was reduced with a hydrogen atom (following partial formolysis) and *m/e* 118 when the 3-linked glucosyl residue was reduced with a deuterium atom (following completion of the hydrolysis with trifluoroacetic acid). Thus, the ratio of the amounts of the ions at *m/e* 117 and 118 is an estimate of the cleavage of the 3-linked glucosyl residues.

The abundance of a particular fragment-ion, such as *m/e* 117, is obtained by integration of the curve produced by connecting the abundance values of *m/e* 117 in each mass spectrum obtained (~1 per 2 sec) during the elution of the 3-linked glucosyl derivative from the capillary column. The estimates of the percent of the glycosyl linkages cleaved are probably accurate to ±15%.

Samples of permethylated lichenan (20 μg) in 100 μL of 90% formic acid were heated at 70° for 20, 40, 60, 90, and 120 min. The percents of formolyses of the 3-linked and 4-linked glucosyl residues under these conditions are listed in Table I. The conditions selected for partial formolysis of permethylated lichenan (90% formic acid for 90 min at 70°) result in about 30% cleavage of both the 3-linked and 4-linked glucosyl residues.

Step IV. The permethylated lichenan (10 mg) was formolyzed in 1 mL of 90% formic acid under the conditions determined in Step III, that is, for 90 min at 70°. The formic acid was evaporated off in a stream of filtered air and the mixture of partially methylated, partially formylated oligosaccharides was reduced with sodium borodeuteride to yield the partially methylated, oligosaccharide alditols. This mixture of partially methylated and reduced oligosaccharides was desalted and then ethylated.

Step V. The mixture of partially methylated, partially ethylated lichenan oligosaccharides was fractionated by reversed-phase l.c. The elution pattern obtained with 3 mg of the partially methylated, partially ethylated lichenan oligosaccharides is shown in the bottom panel of Fig. 2. This pattern is reproducible; essentially the same elution-pattern has been obtained at least 20 times, over a period of 9 months, with 4 different preparations of peralkylated lichenan oligosaccharides.

Step VII. The peralkylated oligosaccharide-containing fractions from the reversed-phase column were successively hydrolyzed, reduced, and acetylated to yield the partially methylated, partially ethylated alditol acetates.

Step VIII. The structures of the partially methylated, partially ethylated alditol acetates obtained from the fractions emerging from the l.c. column were determined by g.l.c.–m.s. and quantitated by g.l.c. using a 25-m, open-tubular, glass capillary column containing SE-30. The structures of the lichenan oligosaccharides were deduced from the identity of the partially methylated, partially ethylated alditol acetates obtained from each oligosaccharide.

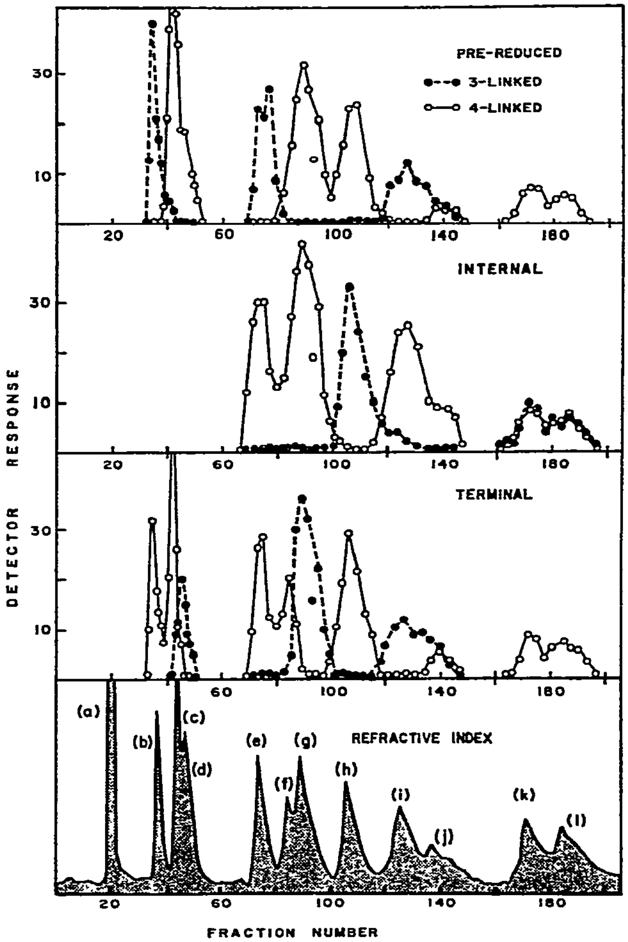


Fig. 2. Reversed-phase l.c. elution-profile, as detected by a refractive-index monitor (attenuation 2×), of 3 mg of the partially methylated, partially ethylated oligosaccharide fragments of lichenan (bottom panel). Chromatography was accomplished as described in the Experimental section. Oligosaccharides (*a*)–(*l*) were collected in a series of fractions and converted into their partially methylated, partially ethylated alditol acetates. G.l.c. and g.l.c.–m.s. analysis of each fraction identified the pre-reduced, the internal, and the terminal glucosyl residues (Example 1, Step VIII). The amounts of these alditol acetate derivatives were determined by integration of the areas of their g.l.c. peaks and comparison of these areas to the area of a peak containing a known amount of *myo*-inositol. This information for the derivatives of the pre-reduced glucosyl residues obtained from each l.c. fraction is plotted in the top panel. Similar plots (two middle panels) are shown for the derivatives of the internal and terminal glucosyl residues. The glucosyl residues contained in the lichenan oligosaccharides were either 3-linked (●) or 4-linked (○). The derivatization technique also labels the terminal glucosyl residues as having been, in the intact polysaccharide, either 3-linked or 4-linked (Example 1, Step VIII). By reading from the top to the bottom of the Figure, the oligosaccharides eluted from the l.c. column may be identified from the pre-reduced, internal, and terminal residues present in each fraction. Note that trisaccharides (*f*) and (*g*) are partially separated, even though they differ only by the point of attachment of the ethoxyl group on the terminal glucosyl residue.

The capillary-column chromatogram (Fig. 3) of the derivatives obtained from a glucotrisaccharide illustrates the separation of the three types (pre-reduced, internal, and terminal) of partially alkylated alditol acetates obtainable by the derivatization procedure. The partially methylated, partially ethylated alditol acetate eluting in peak (*a*) (Fig. 3) was identified as being derived from a pre-reduced 3-linked glucosyl residue[4,10]. This particular derivative is demonstrated to have been at the reducing end of the trisaccharide by the presence of ethoxyl groups at C-1 and C-5. The partially alkylated alditol acetate eluting in peak (*b*) was identified as being derived from the terminal glucosyl residue of the trisaccharide, as acetoxyl groups are found only at C-1 and C-5. This alditol derivative has an ethoxyl group at C-4, demonstrating that O-4 was involved in a glycosyl linkage in lichenan, as it was protected during methylation of intact lichenan. The partially methylated alditol acetate eluting in peak (*c*) was derived from a 4-linked glucosyl residue. This internal glucosyl residue, together with the 3-linked, pre-reduced glucosyl residue and the 4-ethoxyl, terminal glucosyl residue, define the glucotrisaccharide illustrated in Fig. 3.

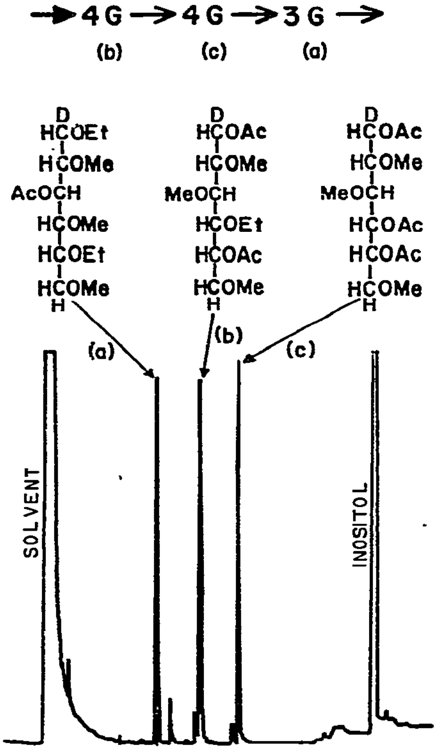


Fig. 3. Gas chromatogram of the partially alkylated alditol acetates obtained from one of the l.c. column fractions in peak (*e*) of Fig. 2. This chromatogram was achieved on an SE-30 capillary column. The gas chromatograph was temperature-programmed from 170–240° at 2° per min. *myo*-Inositol was added as an internal standard before acetylation. The sequence of the oligosaccharide (Example 1, Step VIII) contained in this l.c. fraction is illustrated (G = Glc).

The other partially methylated, partially ethylated, di-, tri-, and tetra-saccharides produced from lichenan were analyzed as was done for the trisaccharide in Fig. 3. Every partially methylated, partially ethylated alditol acetate obtained was derived from either a 3-linked or 4-linked glucosyl residue. The amount of each derivative in each fraction, relative to a known amount of *myo*-inositol introduced into each fraction before acetylation, was determined by integration of the g.l.c. peaks. The relative amounts of the derivatives detected in each l.c. column-fraction are plotted in the top three panels of Fig. 2. This method of plotting the data illustrates both the relative amounts of the oligoglucosides present in the fractions of the reversed-phase column and the structures of these oligoglucosides.

Peak (*a*) of the refractive-index profile (the bottom panel of Fig. 2) contains monosaccharide derivatives. Peaks (*b*), (*c*), and (*d*) contain disaccharides, which account for ~4% of the peralkylated lichenan fragments applied to the l.c. column. Peaks (*e*), (*f*), (*g*), and (*h*) contain trisaccharides, which account for ~8% of the lichenan fragments, and peaks (*i*), (*j*), (*k*), and (*l*) contain tetrasaccharides, accounting for ~6% of the lichenan fragments.

The lichenan di- and tri-saccharides are unambiguously defined. Several of the lichenan tetrasaccharides are also structurally defined. The first tetrasaccharide eluted, peak (*i*), has two identical internal residues, namely, two 4-linked glucosyl residues. An internal residue is attached to C-3 of the pre-reduced glucose residue. The terminal residue of the tetrasaccharide in peak (*i*) had another glucosyl residue attached to C-3 in the intact polysaccharide. Thus, tetrasaccharide (*i*) has the structure ➔3G→4G→4G→3G→.

At least two distinct tetrasaccharides are contained in peak (*j*). This is seen by the presence of both 3-linked and 4-linked pre-reduced glucosyl residues in the fractions containing this peak. Tetrasaccharides in peak (*j*) also contain the two possible types of terminal glucosyl residues. The two internal residues of the tetrasaccharides in peak (*j*) are the same, namely, 4-linked glucosyl residues. The tetrasaccharides in peak (*j*) cannot be the same as the tetrasaccharide in peak (*i*). Therefore, peak (*j*) must contain a mixture of tetrasaccharides, one having a 3-linked glucosyl residue at the reducing end and another having a 3-linked glucosyl residue at the non-reducing end. These tetrasaccharides are ➔4G→4G→4G→3G→ and ➔3G→4G→4G→4G→. Peak (*j*) may also contain small amounts of the tetrasaccharide ➔4G→4G→4G→4G→.

The tetrasaccharides in peaks (*k*) and (*l*) each contain two types of internal glucosyl residues. The tetrasaccharides in both peaks (*k*) and (*l*) contain terminal glucosyl residues that were originally 4-linked, and also pre-reduced residues that are 4-linked. Therefore, the tetrasaccharides in these peaks are ➔4G→4G→3G→4G→ and ➔4G→3G→4G→4G→. It is not known which of these tetrasaccharides is in peak (*k*) and which in peak (*l*).

Step IX. The structural features of lichenan may be deduced from the oligosaccharides obtained by partial formolysis. A summary of the arrangement of the 3-linked and 4-linked glucosyl residues of lichenan is presented in Fig. 4. Trisaccharides

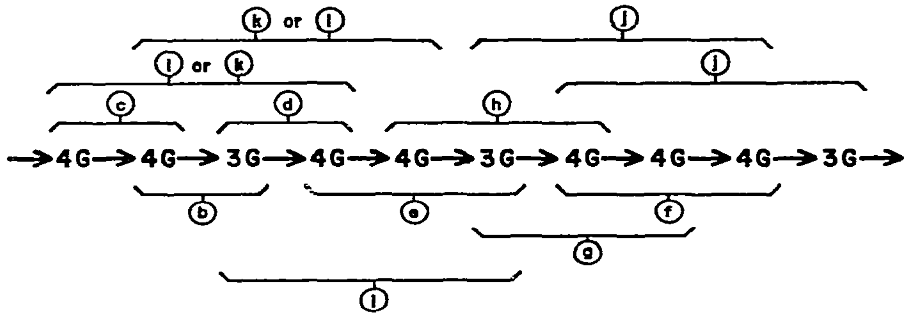


Fig. 4. Structural features of lichenan. For an explanation of the notation used, see the beginning of the experimental section. The identified oligosaccharide fragments of lichenan are also summarized (Example 1, Step VIII).

(*e*), (*g*), and (*h*) are consistent with a major portion of this molecule being a repeating unit of →4G→4G→3G→. Smaller amounts of the trisaccharide →4G→4G→4G→ (peak *f*) are also present. This structure for lichenan, including the regions of three contiguous 4-linked glucosyl residues, agrees well with the literature reports[22–24].

It is possible that a tetrasaccharide containing only 4-linked glucosyl residues (→4G→4G→4G→4G→) could be present in peak (*j*). However, the small size of peak (*j*) and the presence of at least two other tetrasaccharides in peak (*j*) make it unlikely that four consecutive 4-linked glucosyl residues are a significant structural feature of lichenan. Two consecutive 3-linked glucosyl residues were not found in any di-, tri-, or tetra-saccharide isolated from lichenan. Thus, it is likely that lichenan does not contain two consecutive, 3-linked glucosyl residues.

These results demonstrate that lichenan is not a random polymer for, on a random basis, two neighboring 3-linked glucosyl residues should occur in 10% of the disaccharides and four, consecutive, 4-linked glucosyl residues should occur in 21% of the tetrasaccharides. Neither of these oligosaccharides was detected; it is likely that these oligosaccharides would have been detected if they were present in lichenan in these amounts. It is also true that lichenan probably does not have a completely defined structure because the →4G→4G→4G→ units occur only about once for every four →4G→4G→3G→ units. This ratio is suggested by the relative areas of the trisaccharide peaks in the refractive-index profile (lower panel in Fig. 2). The relative abundance of the trisaccharides is quantitatively meaningful only because all of the glycosyl linkages of lichenan are cleaved at the same rate (see Table I). However, the possibility that lichenan has a regular, repeating-unit structure having ~15 glucosyl residues per repeating unit [(→4G→4G→3G→)$_{\sim 4}$(→4G→4G→4G→)$_1$] can neither be confirmed nor ruled out. The fact that lichenan is composed of only two different types of glycosyl residues makes the structural characterization of a large repeating-unit very difficult. This difficulty arises from the absence of unique structural features in the small oligosaccharides obtained by partial formolysis of permethylated lichenan. The absence of identifying structural features in the oligosaccharides makes it impossible to identify neighboring or overlapping oligosaccharides.

The ring forms of the glucosyl residues of lichenan are established by this glycosyl-sequencing method. This is important because glycosyl-linkage analysis does not distinguish between 4-linked pyranosyl residues and 5-linked furanosyl residues. If a 4-linked pyranosyl residue or a 5-linked furanosyl residue is at either the reducing or at the non-reducing end of an isolated oligosaccharide, its ring form can be unambiguously determined. For example, after hydrolysis and acetylation of disaccharide (*c*) in Fig. 2 (➡4G→4G→), the glucosyl residue at the reducing end of the disaccharide was converted into 4-(mono)-*O*-acetyl-1,5-di-*O*-ethyl-2,3,6-tri-*O*-methyl glucitol. The acetoxyl group at C-4 and the ethoxyl group at C-5 establish that, in lichenan, this residue was a 4-linked glucopyranosyl rather than a 5-linked glucofuranosyl residue. The partially alkylated alditol acetate from a 5-linked glucofuranosyl residue would have had an acetoxyl group at C-5 and an ethoxyl group at C-4. The glucosyl residue at the non-reducing end of the disaccharide was converted into 1,5-di-*O*-acetyl-4-(mono)-*O*-ethyl-2,3,6-tri-*O*-methyl glucitol. The ethoxyl group at C-4 and the acetoxyl group at C-5 show that this residue was also a 4-linked glucopyranosyl residue. All of the glucosyl residues in lichenan have been isolated at both the reducing and non-reducing ends of oligosaccharides and have been shown to be glucopyranosyl residues.

Example 2. Structural analysis of xanthan. — Step I. Xanthan is a polysaccharide (see Fig. 7 on p. 180) secreted by the bacterium *Xanthomonas campestris*. The structure of xanthan has been reported by several laboratories[10,13,26]. Glycosyl-composition analysis of xanthan shows that the polysaccharide is composed of glucosyl, mannosyl, and glucosyluronic acid residues in a ratio of 2:2:1. Xanthan has been shown to contain acetal-linked pyruvic acid and *O*-acetyl groups[26–30].

Step II. Xanthan was methylated as described in the Experimental section.

Step IIa. Two methods were tested for the ability to reduce the carboxyl groups of the glycosyluronic residues of xanthan. Treatment of permethylated xanthan with lithium aluminum deuteride in oxolane (tetrahydrofuran)[16] gave ~75% reduction of the carboxyl groups of the uronic acid, as determined by glycosyl-linkage composition analysis. Treatment of permethylated xanthan with sodium borodeuteride in 3:8 (v/v)[15] ethanol–oxolane gave >90% reduction of the uronic acid carboxyl groups, as determined by the same procedure. Therefore, the latter method was used in this work. The glycosyl-linkage composition of xanthan after methylation and subsequent reduction of the uronic acid carboxyl groups is presented in Table II.

Step III. The formolysis conditions for producing good yields of di-, tri-, and tetra-saccharides from methylated, carboxyl-reduced xanthan were determined by the method used for lichenan (Example 1, Step III). Samples of the derivatized xanthan (100 μg) in 100 μL of 90% formic acid were heated to 70° for either 60, 100, or 120 min. These conditions resulted in insufficient cleavage of the glycosyl linkages. However, terminal mannosyl residues, those not substituted with a pyruvic acid residue, were almost completely cleaved under these conditions (Table II). It was clear that the

TABLE II

MOLE PERCENT OF GLYCOSYL RESIDUES IN METHYLATED, CARBOXYL-REDUCED XANTHAN AND THE PERCENT CLEAVAGE OF THE GLYCOSYL LINKAGES OF THE RESIDUES UNDER VARIOUS FORMOLYSIS CONDITIONS

Glycosyl residues	*Mole % of glycosyl residues*	*Partial formolysis conditions*			
		70° 120 min	*90° 30 min*	*90° 45 min*	*90° 60 min*
		Formolysis of the glycosyl linkages (%)			
4-Glc	20	9	23	31	39
3,4-Glc	20	4	20	38	41
T-Man	10	81	91	87	85
2-Man	23	29	57	81	75
4,6-Man[a]	7	18	30	58	72
4-GlcA[b]	20	12	23	29	46

[a]This 4,6-mannosyl residue is derived[26] from a terminal mannosyl residue substituted with a pyruvic acid residue at O-4 and O-6. [b]The 4-linked glucosyluronic residue was converted into and detected as 1,4,5,6-tetra-*O*-acetyl-6,6-dideuterio-2,3-di-*O*-methylglucitol.

location of these residues would have to be established in a separate experiment (see Step IIIa).

Samples (100 μg) of the methylated, carboxyl-reduced xanthan, in 100 μL of 90% formic acid, were heated for 1 h at either 70, 80, 90, or 100°. The results of this experiment indicated that, under these conditions, 90° was closest to the optimal temperature. These conditions were further refined by subjecting 100-μg samples of the methylated, carboxyl-reduced xanthan to 90% formic acid at 90° for either 30, 45, or 60 min. The results are summarized in Table II. A 30-min reaction was chosen for large-scale fragmentation of xanthan.

The formolysis conditions chosen for fragmentation of xanthan result in complete cleavage of the glycosyl linkages of those terminal mannosyl residues not substituted by pyruvic acid. About half of the terminal mannosyl residues of xanthan possess[13,26] a pyruvic acid residue linked to O-4 and O-6. The pyruvic acid-substituted, terminal mannosyl residues are not formolyzed so quickly as the terminal mannosyl residues that are not substituted by pyruvic acid groups. Thus, it was expected that, in the mixture of oligosaccharides obtained by partial formolysis of xanthan, oligosaccharides would be identified which contained the pyruvic acid-substituted terminal mannosyl residues, but that no oligosaccharide would be identified which contained the unsubstituted, terminal, mannosyl residues. The results (see Step IX) are in agreement with these expectations.

Step IIIa. The location of the acid-labile, terminal, mannosyl residues, those not substituted by a pyruvic acid residue, was ascertained by selective cleavage of these glycosyl residues. A sample of methylated, carboxyl-reduced xanthan was subjected to mild formolysis by heating in 90% formic acid for 45 min at 70°. Assay of the percent cleavage of the various glycosidic linkages of xanthan showed that the terminal mannosyl residues were preferentially removed under these conditions. After removal of the formic acid by evaporation, the partially formolyzed xanthan sample was ethylated in order to label the glycosyl residue and, specifically, the carbon atom of the glycosyl residue to which the terminal mannosyl residue had been attached. The two most prevalent, partially alkylated alditol acetates containing ethoxyl groups were identified by g.l.c.–m.s. analysis as being derived from the carboxyl-reduced glucosyluronic residue. One derivative was a 6,6-dideuterio-labelled, 4-linked glucosyl residue having a single ethoxyl group attached to C-6; this is the derivative expected from a 4-linked carboxyl-reduced glucosyluronic residue, as the hydroxyl group attached to C-6 was formed by reduction of the carboxyl group after permethylation of xanthan. The other derivative containing ethoxyl groups was a 6,6-dideuterio-labelled, terminal glucosyl residue having ethoxyl groups attached to both C-4 and C-6; this is the derivative resulting from cleavage, during the mild formolysis, of the terminal mannosyl residue from C-4 of the 4-linked glucosyluronic residue.

Small amounts of other ethoxyl-substituted glycosyl residues were also found. These ethoxyl-substituted derivatives probably resulted from partial cleavage, during

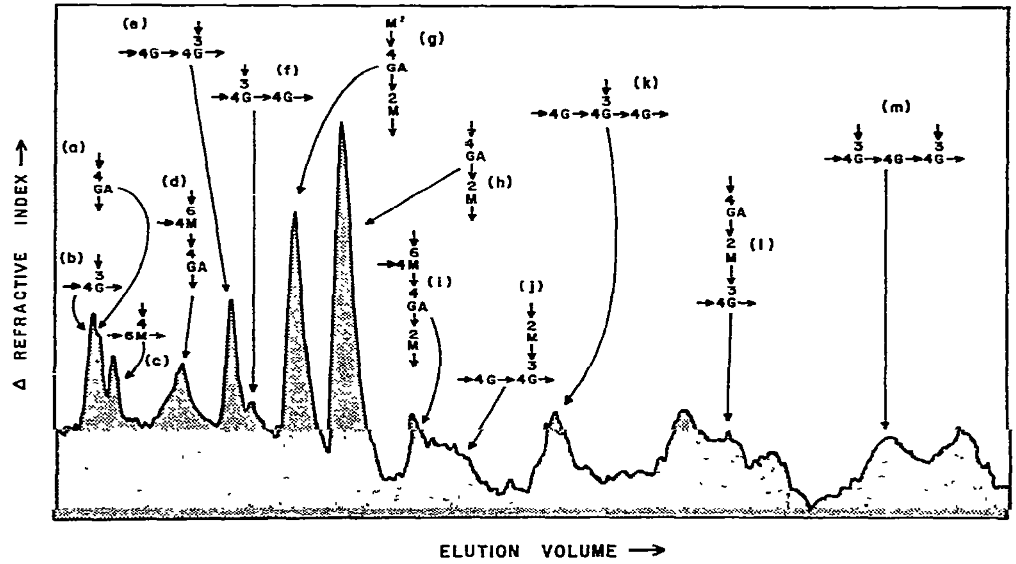


Fig. 5. The reversed-phase l.c. elution profile, as detected by a refractive-index monitor (attenuation 0.5×), of ~1 mg of partially methylated, partially ethylated oligosaccharide fragments of xanthan. Chromatography was accomplished as described in the Experimental section. The baseline is *unstable* at this high sensitivity-setting of the refractive-index monitor; this accounts for the roughness of the baseline throughout the chromatogram and for the apparent irregular shapes of the broad, late-eluting oligosaccharide peaks. The fractions containing the partially methylated, partially ethylated oligosaccharides were analyzed by g.l.c.–m.s. of their alditol acetate derivatives as illustrated in Fig. 6 for trisaccharide (*m*). The oligosaccharide notation used is defined in the experimental section. The terminal mannosyl residue of oligosaccharide (*g*) has been given the notation M^1 as it is substituted by a pyruvic acid derivative (Example 2, Step IX). Unlabelled peaks appearing in this l.c. elution-profile contained mixtures of oligosaccharides that were not identified. No oligosaccharides inconsistent with the determined structure of xanthan were detected.

the mild formolysis, of other glycosyl linkages, and from elimination, during the ethylation reaction, of small amounts of carboxyl-unreduced glucosyluronic residues^{5,7} (see Step IIa). The presence of small amounts of these other ethoxyl-substituted glycosyl residues did not interfere with assigning C-4 of the 4-linked glucosyluronic residues as the point of attachment of the terminal mannosyl residues not substituted by pyruvic acid.

Step IV. Large-scale ($\sim$2 mg), partial fragmentation of methylated, carboxyl-reduced xanthan was accomplished in 2 mL of 90% formic acid for 30 min at 90°. The partially methylated oligosaccharides obtained were reduced and ethylated.

Steps V and VII. The refractive-index profile of the peralkylated xanthan oligosaccharides separated by l.c. is shown in Fig. 5. The fractions from the reversed-phase l.c. column were sequentially hydrolyzed, reduced, and acetylated.

Step VIII. The structures of the partially methylated, partially ethylated alditol acetates obtained from the fractions of the l.c. column were determined by the same procedures as those described for lichenan (Example 1). An example of the structural elucidation of a trisaccharide obtained from xanthan is illustrated in Fig. 6.

M.s. analysis of the partially alkylated, partially acetylated alditol in peak (*a*)

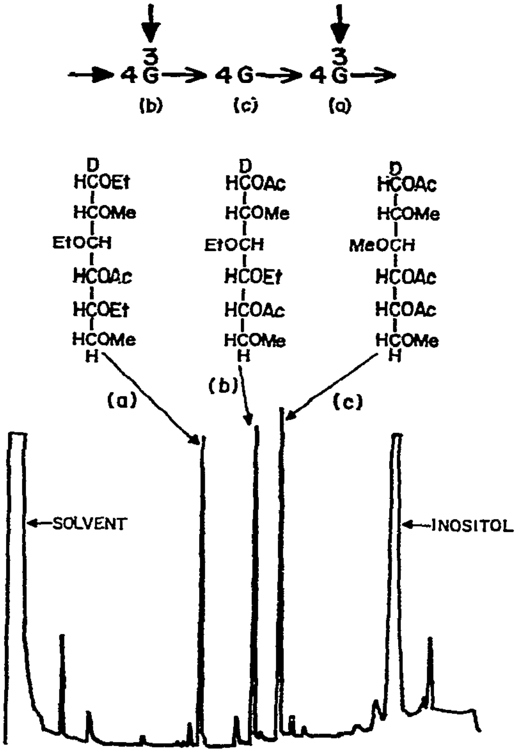


Fig. 6. The gas chromatogram of the partially alkylated alditol acetates obtained from peak (*m*) in Fig. 5. This chromatogram was obtained as described in Fig. 3. The analysis of the resulting oligosaccharides was accomplished as described in the text (Example 2, Step VIII).

of Fig. 6 unambiguously identified this component as the derivative of a pre-reduced 4-linked hexosyl residue having an ethoxyl group at C-3. The ethoxyl group at C-3 of this pre-reduced alditol acetate derivative demonstrates that, in xanthan, this hexosyl residue had another glycosyl residue attached to O-3. The 3-hydroxyl group was freed for ethylation during the partial formolysis of xanthan that led to formation of the trisaccharide. Thus, this residue was, in xanthan, a 3,4-linked hexosyl residue. As glycosyl-linkage analysis of xanthan established that the only 3,4-linked hexosyl residue is a 3,4-linked glucosyl residue, the residue at the reducing end of this trisaccharide is 3,4-linked glucose.

The partially alkylated alditol acetate that eluted in peak (*b*) (Fig. 6) is a terminal hexosyl residue having ethoxyl groups at C-3 and C-4. Thus, this residue of xanthan is a 3,4-linked hexosyl residue, or more specifically, a 3,4-linked glucosyl residue.

The partially methylated alditol acetate eluting in peak (*c*) (Fig. 6) is derived from a 4-linked hexosyl residue; as the only 4-linked hexosyl residue of xanthan is a glucosyl residue, this partially alkylated alditol acetate was derived from a 4-linked glucosyl residue. Thus, the trisaccharide of Fig. 6 is unambiguously defined.

Step IX. Many of the peralkylated oligosaccharides derived from xanthan and separated by reversed-phase l.c. have been identified by g.l.c.–m.s. analysis of the derived partially alkylated, partially acetylated alditols (Fig. 5). Trisaccharide (*m*) in Fig. 5 is the trisaccharide illustrated in Fig. 6. This xanthan trisaccharide has branched, 3,4-linked glucosyl residues at each end and a 4-linked glucosyl residue in the middle. Trisaccharide (*k*) in Fig. 5 has 4-linked glucosyl residues at each end and a 3,4-linked glucosyl residue in the middle. Therefore, trisaccharides (*k*) and (*m*) establish that xanthan contains a 4-linked glucosyl backbone with sidechains attached to O-3 of every second glucosyl residue. Oligosaccharides (*e*), (*f*), and (*j*) (Fig. 5) support this structural feature of xanthan.

Trisaccharides (*j*) and (*l*) of Fig. 5 identify the 2-linked mannosyl residue as the glycosyl residue that is attached to C-3 of the 3,4-linked glucosyl residue of the backbone. Trisaccharides (*g*), (*i*), and (*l*), and disaccharide (*h*) establish that the 4-linked glucosyluronic residue is attached to O-2 of the 2-linked mannosyl residue.

Those terminal mannosyl residues not substituted by a pyruvic acid residue are linked to O-4 of the 4-linked glucosyluronic residues (Step IIIa). Those terminal mannosyl residues substituted by a pyruvic acid residue are also attached to O-4 of 4-linked glucosyluronic residues. This conclusion was established by oligosaccharides (*d*), (*g*), and (*i*) (Fig. 5).

The pyruvic acid residues were partially, but not fully, removed from the terminal mannosyl residues during formolysis. The reason that all of the pyruvic acid residues were not removed is unknown, but it may have been incomplete reduction of the pyruvic acid carboxyl groups in Step IIa. The ketosidically linked pyruvic acid residues, whose carboxyl groups were not reduced, would be more resistant to formolysis. The partially alkylated alditol acetates derived from the terminal mannosyl residues of oligosaccharides (*d*) and (*i*) (Fig. 5) were substituted with ethoxyl groups at C-4 and C-6. In these cases, the pyruvic acid residues were cleaved from the mannosyl residues during the partial formolysis of methylated, glycosyluronic-reduced xanthan. The partially alkylated alditol acetate derived from the terminal mannosyl residue of oligosaccharide (*g*) is substituted by acetoxyl groups at C-4 and C-6. In this instance, the pyruvic acid residue was present during ethylation and was only removed during the total acid hydrolysis leading to formation of the partially alkylated, partially acetylated alditols. This unusual, terminal mannosyl residue of oligosaccharide (*g*) is labelled M^1 in Fig. 5.

Glycosyl-linkage analysis of xanthan established that the terminal mannosyl and the 2-linked mannosyl residues are in the pyranosyl rather than the furanosyl ring-form. However, glycosyl-linkage analysis does not establish the ring forms of the remaining glycosyl residues of xanthan. The ring forms of these glycosyl residues was established by isolating one or more oligosaccharides in which each of these residues was at either the reducing or the non-reducing end of an oligosaccharide. Oligosaccharides (*e*), (*f*), and (*j*) (Fig. 5) demonstrate that both the 4-linked and 3,4-linked glucosyl residues are in the pyranoid ring-form. Oligosaccharides (*d*), (*h*), and (*l*) demonstrate that the 4-linked glucosyluronic residue is in the pyranoid ring-form. Oligosaccharides (*d*) and (*i*) establish that the pyruvic acid-substituted terminal mannosyl residue is also pyranoid.

The glycosyl sequences of xanthan and the oligosaccharide fragments used in elucidating the structure of xanthan are summarized in Fig. 7; this structure is identical to that established by other methods[10,13,26].

Example 3. Structural analysis of the nonasaccharide of xyloglucan. — Steps I and II. Xyloglucan is a hemicellulosic polysaccharide of plant cell-walls. It is reported to be composed of a backbone of glucosyl residues with xylosyl residues attached to approximately 3 out of every 4 of the glucosyl residues[31,32]. Some of the xylosyl residues are substituted by fucosyl-galactosyl side-chains. A nonasaccharide, released from purified xyloglucan by enzymic degradation, consists[31] of glucosyl, xylosyl, galactosyl, and fucosyl residues in the ratio of 4 : 3 : 1 : 1. The xyloglucan nonasaccharide (10 mg) was purified as reported[31] and prepared for methylation by pre-reduction

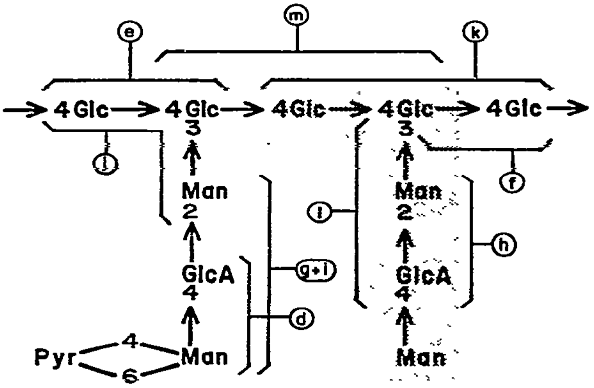


Fig. 7. A summary of the repeating-unit structure of xanthan (see shaded area) and the oligosaccharide fragments (see Fig. 5) that defined the sequence of xanthan.

with sodium borodeuteride. The nonasaccharide was methylated as described in the experimental section.

Step IIb. The permethylated nonasaccharide was further purified by reversed-phase l.c. The elution profile of 3 mg of the permethylated nonasaccharide, in 60% (v/v) acetonitrile–water, is illustrated in Fig. 8. Several, quantitatively minor, oligosaccharides were separated from the nonasaccharide. These oligosaccharide contaminants are similar to, but distinct from, the nonasaccharide in their glycosyl-linkage compositions. Aliquots taken from various locations in the nonasaccharide peak (shaded area in Fig. 8) had identical glycosyl-linkage compositions. The nonasaccharide-containing fractions (shaded area) were pooled for further structural characterization. The glycosyl-linkage composition of the l.c.-purified xyloglucan nonasaccharide is given in Table III.

Step III. Optimal conditions for partial hydrolysis of the xyloglucan nonasaccharide were determined as for the polysaccharides in Examples 1 and 2. The

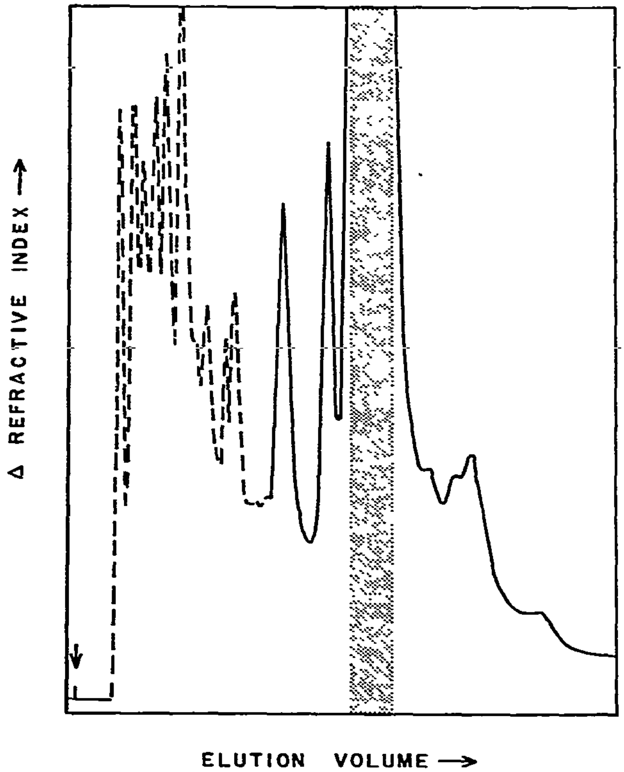


Fig. 8. Purification of 3 mg of permethylated xyloglucan nonasaccharide by reversed-phase l.c. (Example 3, Step IIb). Elution of the oligosaccharides from the column was monitored with the refractive-index detector set at an attenuation of 0.5×. Peaks that did not contain carbohydrate are illustrated with a dashed line; such peaks are routinely detected during l.c. of alkylated carbohydrates. The smaller, solid-line peaks contain oligosaccharides having glycosyl-linkage compositions similar to, but not identical with, those found in the xyloglucan nonasaccharide. The portion of the nonasaccharide peak that was pooled for glycosyl sequence analysis is indicated by the shading.

TABLE III

MOLE PERCENT OF GLYCOSYL RESIDUES IN METHYLATED XYLOGLUCAN NONASACCHARIDE AND THE PERCENT HYDROLYSIS OF THE GLYCOSYL LINKAGES OF THOSE RESIDUES IN 2M TRIFLUOROACETIC ACID FOR 40 MIN AT 95°

Glycosyl residues	*Mole % of glycosyl residues*	*Hhydrolysis of the glycosyl linkages (%)*
6-Glc	11	26
4,6-Glc	22	17
4-pr-Glc[a]	10	—
2-Gal	11	71
T-Xyl	23	45
2-Xyl	10	49
T-Fuc	13	90

[a] This 4-linked, pre-reduced glucose residue [4-(mono)-*O*-acetyl-1,2,3,5,6-penta-*O*-methylglucitol] is derived from the reducing terminus of the intact nonasaccharide.

conditions chosen for partial acid hydrolysis of the nonasaccharide, 2M trifluoroacetic acid for 40 min at 90°, result in the percent hydrolyses listed in Table III. The nonasaccharide of xyloglucan contains glycosyl linkages having different susceptibilities to acid hydrolysis (Table III). Thus, the degree of hydrolysis chosen was a compromise between over-hydrolysis of the 2-linked galactosyl residues (~71% hydrolyzed) and under-hydrolysis of the 4,6-linked glucosyl residues (~17% hydrolyzed). The terminal fucosyl residues (~90% hydrolyzed) actually appear to have been completely hydrolyzed by the partial hydrolysis conditions chosen.

Step IIIa. The xyloglucan nonasaccharide contains a highly acid-labile, terminal fucosyl residue. This residue has previously been shown, by a method similar to that used for determining the location of the terminal mannosyl residue of xanthan (Example 2, Step IIIa), to be glycosidically linked to O-2 of the 2-linked galactosyl residue[31,32].

Steps IV and V. The pre-reduced, permethylated xyloglucan nonasaccharide (3 mg) was heated in 250 μL of 2M trifluoroacetic acid for 40 min at 90°. The resulting partially methylated, fragment-oligosaccharides were reduced with sodium borodeuteride and ethylated. These partially methylated, partially ethylated oligosaccharides were purified by chromatography on LH-20 and fractionated by reversed-phase l.c. The elution profile, as detected by the refractive-index monitor, is illustrated in Fig. 9.

Step VII. The peralkylated oligosaccharides in the fractions obtained from the l.c. column were sequentially fully hydrolyzed, reduced, and acetylated.

Step VIII. The partially alkylated, partially acetylated alditols obtained from the fragment oligosaccharides of the xyloglucan nonasaccharide were identified by g.l.c.–m.s., as in Examples 1 and 2. Quantitative g.l.c.-analysis of the partially methylated, partially ethylated alditol acetate derivatives, obtained from the l.c.

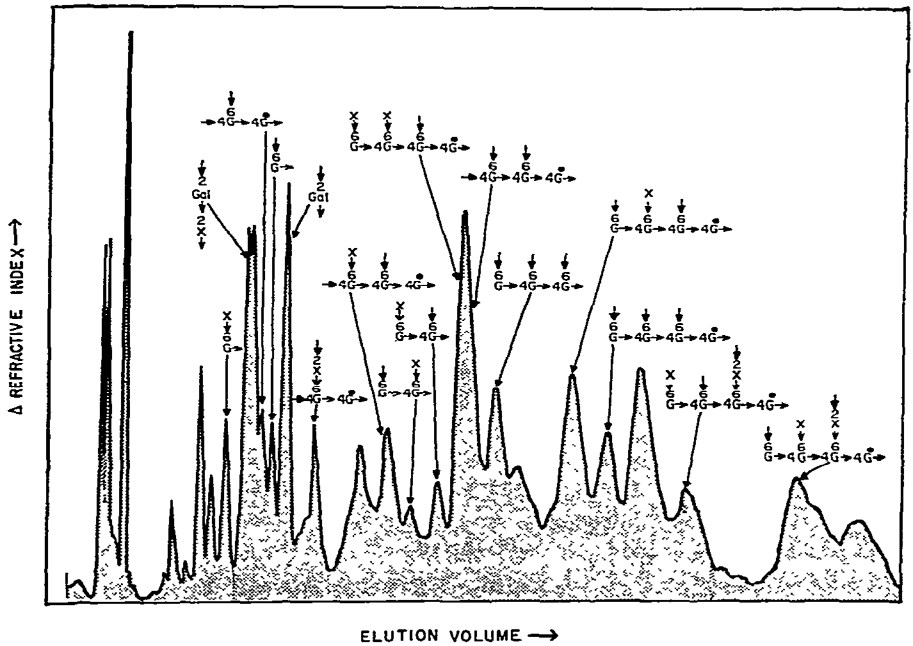


Fig. 9. The reversed-phase l.c. elution profile, as detected by a refractive-index monitor (attenuation 0.5×), of 2.5 mg of the partially methylated, partially ethylated oligosaccharide fragments of the l.c.-purified xyloglucan nonasaccharide. The oligosaccharides identified are summarized in Fig. 10. The unlabelled peaks in this chromatogram contained mixtures of oligosaccharides that were not identified.

column-fractions, was accomplished on a g.l.c. column[1] containing 0.2% ethylene glycol succinate, 0.2% ethylene glycol adipate, and 0.4% XF-1150 on Gas-chrom P. The glass-capillary column used in Examples 1 and 2 does not separate the derivatives of the terminal xylosyl and the pre-reduced 4-linked glucosyl residues present in the xyloglucan nonasaccharide.

One of the partially alkylated, partially acetylated alditols is particularly important for characterization of the nonasaccharide. This derivative is that obtained from the pre-reduced glycose residue, the residue at the reducing end of the unhydrolyzed nonasaccharide. This residue is derived from 4-linked glucose. This pre-reduced, 4-linked glucose residue is uniquely labelled by having methoxyl rather than ethoxyl groups at C-1 and C-5. Nine of the fourteen nonasaccharide fragments that were structurally characterized contain this pre-reduced, 4-linked glucose residue.

Step IX. The glycosyl sequence of the xyloglucan nonasaccharide is completely defined by the sequences of those fragment-oligosaccharides that have been characterized (Fig. 9). The nonasaccharide-derived, fragment-oligosaccharides and the nonasaccharide defined by these fragment-oligosaccharides are illustrated in Fig. 10. The nonasaccharide sequence may be elucidated, in conjunction with its glycosyl-linkage composition, with any of several different combinations of the fragment-

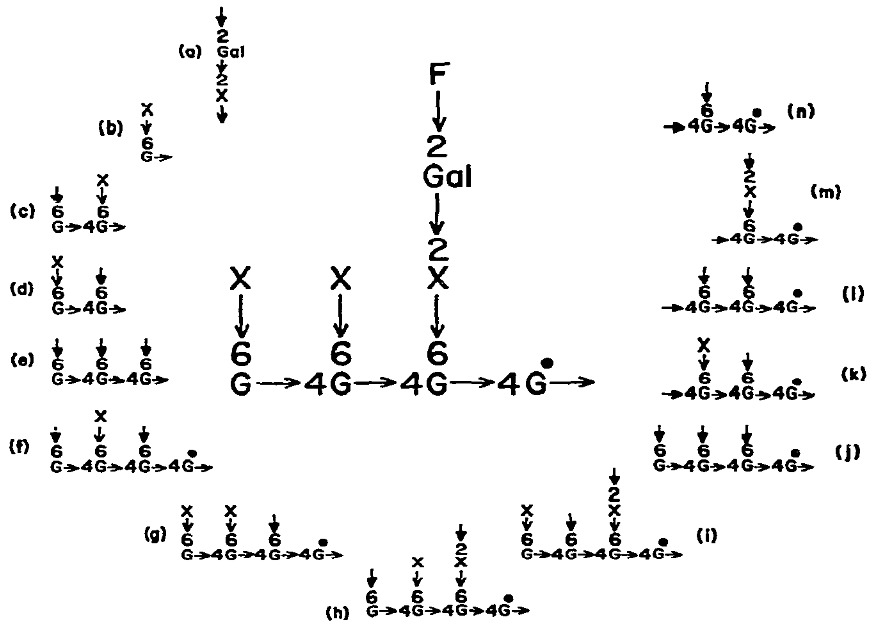


Fig. 10. The glycosyl sequence of the xyloglucan nonasaccharide and the fragment-oligosaccharides obtained from it (Example 3, Steps VIII and IX). The terminology is described at the beginning of the Experimental section.

oligosaccharides. All of the oligosaccharides obtained from the nonasaccharide preparation are consistent with the proposed nonasaccharide sequence.

One combination of fragment-oligosaccharides that can be used to define the sequence of the nonasaccharide is oligosaccharides (*a*), (*j*), and (*n*). Oligosaccharide (*j*) is an example of a tetrasaccharide whose structure is totally defined by its partially methylated, partially ethylated alditol acetates. This is true for this particular tetrasaccharide because its two internal glycosyl residues are identical. Tetrasaccharide (*j*) demonstrates that the four glucosyl residues of which it is composed are linked as a linear, 4-linked glucosyl backbone.

Three of the glucosyl residues of tetrasaccharide (*j*) are substituted by other glycosyl residues at O-6. This is established by the presence of ethoxyl groups at C-6 of the two internal and the non-reducing terminal glucosyl residues. Trisaccharide (*m*) establishes that the 2-linked xylosyl residue is attached to O-6 of the glucosyl residue adjacent to the reducing glucose of the nonasaccharide. Disaccharide (*a*) establishes that the 2-linked xylosyl residue has a 2-linked galactosyl residue attached to it at O-2. By using the knowledge that the terminal fucosyl residue is attached to O-2 of the 2-linked galactosyl residue (Step IIIa), the nonasaccharide is completely defined, as the remaining two terminal xylosyl residues must be attached to O-6 of the remaining two glucosyl residues. The locations of these xylosyl residues is directly established by trisaccharides (*c*) and (*d*) (Fig. 10).

The glycosyl sequences of oligosaccharides (*f*), (*g*), (*h*), (*i*), and (*k*) in Fig. 10

can be completely defined only with knowledge of the structures of other oligosaccharides that were, themselves, completely defined. Nevertheless, these more-complex oligosaccharides add support to the proposed structure of the nonasaccharide.

All of the glycosyl residues of the xyloglucan nonasaccharide, except for the terminal fucosyl residue, may be placed at defined positions in the nonasaccharide by using the sequences of the fragment-oligosaccharides. By deduction, the terminal fucosyl residue must be attached to O-2 of the 2-linked galactosyl residue. Thus, the position of this residue has been confirmed indirectly as well as more directly by the method outlined[31] in Step IIIa.

Glycosyl-linkage analysis has established that at least 7 of the 9 glycosyl residues of the xyloglucan nonasaccharide are in the pyranosyl ring-form. Glycosyl-linkage analysis could not establish the ring forms of the two 4,6-linked glucosyl residues. Oligosaccharides (*c*), (*d*), (*l*), and (*k*) each establish that the 4,6-linked glucosyl residue adjacent to the 6-linked glucosyl residue is in the pyranosyl ring-form. Oligosaccharides (*n*), (*m*), and (*e*) each establish that the 4,6-linked glucosyl residue that is adjacent to the 4-linked glucose residue is also in the pyranosyl ring-form. Thus, all of the glycosyl residues of the xyloglucan nonasaccharide are pyranoid rather than furanoid.

GENERAL DISCUSSION

A general and sensitive chemical method for deciphering the glycosyl sequence of complex carbohydrates has been described. An important feature is that it permits discrimination between 4-linked pyranosyl and 5-linked furanosyl residues. All three complex carbohydrates studied in this paper illustrate this point. This problem has proved, at times, difficult to resolve by methods previously available.

A structural feature of complex carbohydrates not elucidated by the methods described in this paper is the absolute configuration (D or L) of the component sugar residues. Two g.l.c. assays recently developed for determining the absolute configuration of sugars require only microgram quantities for analysis[33,34]. Thus, these methods complement the glycosyl-sequencing method described here.

One structural feature of complex carbohydrates that is difficult to ascertain is the anomeric configurations, α or β, of the glycosyl linkages. A chemical method has been used for this purpose[35]; ^{1}H-n.m.r. spectroscopy is also useful. The peralkylated oligosaccharides obtained by the procedure described here (Steps V and VI) are ideally suited for ^{1}H-n.m.r. analysis. Most of them contain only one or two glycosyl linkages and are soluble in deuterated organic solvents; ^{1}H-n.m.r. spectroscopy can establish the anomeric configuration of 5-μg quantities of such peralkylated oligosaccharides (unpublished results).

Rapid and efficient methods for obtaining the entire primary structure of complex carbohydrates are thus now available. Such analyses allow determination of the glycosyl sequence of the carbohydrate, the ring-forms of the glycosyl residues, the absolute configuration of the monosaccharide constituents, the anomeric configuration of the glycosyl linkages, and even the location of pyruvyl[12,36] and acetyl[27] groups.

The three complex carbohydrates identified in this paper feature hexosyl, deoxyhexosyl, pentosyl, and glycosyluronic residues among their components. In developing this glycosyl-sequencing method, we realized that the previously observed, differential susceptibility to acid-catalyzed hydrolysis exhibited by such diverse glycosyl linkages might prove to be a problem. Although it is necessary to take this characteristic of glycosyl linkages into account, the method is not compromised by the presence of glycosyl linkages possessing very different acid labilities. This was established by isolating, from the xyloglucan nonasaccharide (Fig. 9), the oligosaccharides required for glycosyl sequencing. The necessary oligosaccharides were isolated, even though 71% of the glycosyl bonds of the 2-linked galactosyl residues had been hydrolyzed and only 17% of the glycosyl bonds of the 4,6-linked glucosyl residues had been hydrolyzed. Even complete hydrolysis of the terminal fucosyl residue did not prevent sequencing of the xyloglucan nonasaccharide, nor did complete hydrolysis of the terminal mannosyl residue prevent sequencing of xanthan.

Some peralkylated oligosaccharides are co-eluted or nearly co-eluted during l.c. In many such instances, it is still possible to structurally define the co-eluting oligosaccharides by analyzing successive fractions across the peak. The partially alkylated alditol acetates obtained from each fraction of the peak are quantitated. It is then possible to determine which partially alkylated alditol acetates are derived from the same oligosaccharide, because the amounts of these derivatives in the various fractions remain constant relative to each other but vary in relation to the derivatives obtained from other oligosaccharides. Of course, oligosaccharides present in roughly equal amounts and which are co-eluted exactly cannot be identified in this manner.

The degree of retention of a partially methylated, partially ethylated oligosaccharide on the reversed-phase l.c. column is dependent on the number of ethoxyl groups present in the oligosaccharide. In fact, substituting an ethoxyl group for a permethylated, terminal xylosyl residue results in increased retention of the oligosaccharide [compare oligosaccharides (*g*), (*f*), and (*j*) in Figs. 9 and 10]. This chromatographic characteristic of the peralkylated oligosaccharides may be used to separate oligosaccharides that otherwise would be co-eluted to such an extent that their structures could not be deciphered. For example, co-eluting oligosaccharides might be separated if the number of ethoxyl groups in the oligosaccharides were changed. This might be accomplished by first ethylating the complex carbohydrate (Step II in Fig. 1) and then methylating the oligosaccharide fragments produced by partial hydrolysis of the ethylated, carbohydrate (Step IV in Fig. 1). Another method that might separate co-eluting oligosaccharides is to label, with trideuteriomethoxyl rather than ethoxyl groups, the free hydroxyl groups of the partially methylated oligosaccharides produced by partial hydrolysis of the methylated carbohydrate (Step IV).

The xyloglucan nonasaccharide described in Example 3 has been sequenced with 10 mg of starting material. The amount of complex carbohydrate required for structural analysis will be lowered as technological advances in instrumentation occur.

For example, a mass spectrometer coupled directly to a liquid chromatograph would constitute a much more sensitive and descriptive detector (than the refractive-index detector) for oligosaccharides eluted from an l.c. column. Refractive-index detectors are not compatible with gradient elution of l.c. columns, but m.s. detectors permit gradient elution, which would sharpen the oligosaccharide peaks and, thereby, make detection more efficient; such l.c.–m.s. systems are being perfected. Injection systems for g.l.c. capillary columns, which do not require sample splitting, are now available. The use of these injection systems will immediately lower the amount of complex carbohydrates required for analysis by about a factor of ten. As demand rises, l.c. columns will be specifically developed to give more efficient and different separations of peralkylated oligosaccharides. It seems likely that the amount of a complex carbohydrate required for complete chemical identification will be decreased to 0.1 mg or less as instrumentation and technology now being developed becomes available.

EXPERIMENTAL

Terminology used in this paper. — Glycosyl and glycose residues. A sugar residue glycosidically linked to another residue is called a glycosyl residue. Sugars not glycosidically attached, whether or not they have other glycosyl residues attached to them, are called glycose residues. The linkage notation indicates which oxygen atoms of glycosyl and glycose residues have other glycosyl residues linked to them. Thus, a 2-linked mannosyl residue (2-Man) has some other glycosyl residue linked to it at O-2. A terminal glycosyl group (T-Glc) has no other glycosyl residues linked to it. The linkage notation is determined from the arrangement of methoxyl and acetoxyl groups on the alditol acetate formed after sequential methylation, hydrolysis, reduction, and acetylation of the complex carbohydrate. Thus, 2-Man represents 1,2,5-tri-*O*-acetyl-3,4,6-tri-*O*-methylmannitol.

Glycosyl sequence. The glycosyl sequence of an oligo- or poly-saccharide is the exact arrangement of all of its glycosyl residues. Thus, in an oligosaccharide, the location of each individual glycosyl residue with respect to all other glycosyl residues (for instance, the other eight glycosyl residues in a nonasaccharide) is determined by its glycosyl sequence. Polysaccharides often are composed of a repeating sequence of glycosyl residues[18–20,26]. The sequence of such polysaccharides is known when the arrangement of the glycosyl residues in the repeating unit has been determined. Some polysaccharides are not composed of a regular repeating-unit (for instance, amylopectin). The glycosyl sequence of such polysaccharides would be determined only when the arrangement of all glycosyl residues of the polysaccharide is known.

Abbreviations. The standard three-letter sugar abbreviations are mainly used. However, in some of the Figures and in the shorthand notation for oligosaccharide structures in the text, the following abbreviations are used: G = Glc = glucosyl; Gal = galactosyl; M = Man = mannosyl; X = Xyl = xylosyl; F = Fuc = fucosyl; GA = GlcA = glucosyluronic acid; and Pyr = pyruvic acid.

Shorthand notation for oligosaccharide structures. The symbol ➧ indicates an ethoxyl-substituted carbon atom of a glycosyl residue. The symbol → indicates a glycosidic linkage. An →, not pointing to another glycosyl residue, identifies a glycose residue at the reducing terminus of an oligosaccharide produced by fragmentation of a complex carbohydrate. These glycose residues become labelled with ethoxyl groups at C-1 and C-5. The symbol $\dot{\rightarrow}$ identifies the unique glycose residue at the reducing terminus of an intact complex oligosaccharide being sequenced. These glycose residues become labelled with methoxyl groups at C-1 and C-5.

Liquid chromatography. — A Waters Model 6000A solvent-delivery system and a Waters Model R401 differential refractometer were used for l.c. The differential refractometer cells were maintained at 30°. Sample injection was accomplished by using a Valco Model CV-6-UHPa-N60, 7000 lb.in^{-2} injection-valve, with a 100-μL injection loop. Fractions were collected by using a Gilson Model FC-80K microfractionator.

A Dupont, prepacked Zorbax$^{\text{TM}}$ ODS column (4.6 mm i.d. × 25 cm) was used for reversed-phase l.c. of peralkylated oligosaccharides. Guard columns (precolumns) (2 mm i.d. × 7 cm) were dry-packed with a reversed-phase, pellicular packing-material supplied by Whatman (Co:Pell ODS). The solvents used for l.c. were mixtures of acetonitrile and water. Distilled water was filtered through 0.22-μm Millipore filters. Certified A.C.S. grade acetonitrile (Fisher Scientific) was filtered through 0.5-μm Fluoropore filters (Millipore). All solvents were degassed with a water aspirator.

L.c. was performed with 1–3 mg of the peralkylated oligosaccharides. The oligosaccharides were injected, in 50 μL of solvent, onto the Dupont Zorbax$^{\text{TM}}$ ODS column. Unless otherwise stated, the solvent was 50% (v/v) acetonitrile–water. The chromatography was completed in less than 100 min with a flow rate of 0.5 mL per min. The column pressure was about 400 lb.in^{-2}. Fractions of 0.5 mL were collected.

Purification of the complex carbohydrates. — Lichenan (ex *Cetraria islandica*, Sigma Chemical Company, 200 mg) was dissolved in 40 mL of hot water and then precipitated by freezing and melting the solution. The dissolution, freezing, and thawing of lichenan was repeated a total of three times to yield 180 mg of lichenan. This procedure has been reported to remove isolichenan, a glucan containing 3-linked and 4-linked α-D-glucosyl residues[23].

Gum xanthan (Grade II, Sigma Chemical Company, 500 mg) was dissolved in distilled water (1 L). The solution was filtered through two layers of glass-fiber paper (Whatman) and then through a 1.2-μm Millipore filter. The polysaccharide was purified by precipitation with a quaternary ammonium salt[37] as follows. Sodium sulfate was added to the polysaccharide-containing solution to a final concentration of 10mM. The polysaccharide was precipitated with 3% cetyltrimethylammonium bromide. The precipitate was collected by centrifugation, washed with water, and then dissolved in 10% sodium chloride. The xanthan was precipitated from this solution with 2 vols of acetone, redissolved in 10% sodium chloride, dialyzed extensively against 1% sodium chloride and then against water, and, finally, lyophilized.

The xanthan was deionized on a cation-exchange column (Rexyn 101) prior to methylation[38].

The Bio-Gel P-2-purified, xyloglucan nonasaccharide was prepared, according to the methods of Bauer *et al.*[31], from the xyloglucan secreted into the culture medium of suspension-cultured sycamore cells.

Reduction with sodium borodeuteride. — The aldehyde groups at the reducing termini of lichenan and xyloglucan nonasaccharide were reduced with sodium borodeuteride (10 mg/mL) in M ammonia prior to methylation. The carbohydrate was dissolved in the sodium borodeuteride solution at a concentration of approximately 10 mg per mL. The solution was kept for 1–4 h at room temperature and then for 1 h at 50°. Reduction was terminated by dropwise addition of acetic acid until effervescence ceased. The ammonia and water were evaporated from the reduced carbohydrate and the borate was removed as its trimethyl ester by sequential addition and evaporation of 0.5 mL of methanol acidified with 10% acetic acid. Four evaporations of acidified methanol from the carbohydrate were followed by four evaporations of pure methanol.

Lichenan was desalted by extensive dialysis against water, and then lyophilized. The xyloglucan nonasaccharide was desalted by dissolution in 0.5 mL of distilled water and passed through a column (0.5 mL) of Rexyn 101 cation exchanger (H^+ form) to convert sodium acetate into acetic acid. The acetic acid and water were evaporated under a stream of filtered air. Less than 10% of the nonasaccharide was lost by this desalting procedure.

Alkylated or partially alkylated oligosaccharides were reduced as just described except that the reduction was conducted in M ammonia in ethanol instead of in water. Alkylated or partially alkylated oligosaccharides were desalted on columns of Rexyn 101 equilibrated with 50% ethanol (v/v).

Methylation. — Carbohydrates were methylated by a modification of the method of Hakomori[39] as adapted by Sandford and Conrad[16]. Potassium dimethylsulfinyl anion was used instead of sodium dimethylsulfinyl anion because, in our hands, the former gives more-complete methylation. The potassium anion is made as reported for the sodium anion[39] except that the hexane was evaporated with a stream of nitrogen rather than with a water aspirator, and heating was omitted during preparation of the anion. Extra care had to be taken to keep the potassium hydride dry and to work only with small amounts of the potassium hydride (3 g of a 22% suspension) in order to avoid fires. Oligo- or poly-saccharides to be methylated were dried *in vacuo* over phosphorus pentaoxide for a minimum of 10 h at 45°.

Pre-reduced lichenan (10 mg) was suspended in 1 mL of dry dimethyl sulfoxide and dissolved by stirring for 1 h at room temperature, with subsequent sonication for 1 h. Sonication was performed in a Heat Systems–Ultrasonics, Inc., small cleaning sonicator. Potassium dimethylsulfinyl anion (5.0M, 130 μL) was added and the solution was stirred for 1 h at room temperature. Methyl iodide (40 μL, 1.05 equiv. based on the amount of dimethylsulfinyl anion) was added and the solution was stirred for 1 h at room temperature. The treatments were repeated except that the second addition of methyl iodide was 240 μL instead of 40 μL. The excess of methyl iodide was then evaporated off under a stream of nitrogen. The permethylated lichenan was dialyzed extensively against distilled water and lyophilized.

Xanthan was prepared for methylation by dissolving 20 mg in 40 mL of dry dimethyl sulfoxide. Dimethylsulfinyl anion (2.0M, 5 mL) was added and the solution was stirred overnight. An excess of methyl iodide (~3 mL) was added and the solution was stirred for 18 h. The excess of methyl iodide was then evaporated under a stream of nitrogen. The permethylated xanthan was dialyzed extensively against distilled water and lyophilized.

The xyloglucan nonasaccharide (10 mg) was dissolved in 0.5 mL of dry dimethyl sulfoxide by sonication for 4 h at ~60° followed by stirring for 1 h at room temperature. Dimethylsulfinyl anion (3.0M, 100 μL) was added and this solution was sonicated for 1 h. Methyl iodide (18.5 μL) was added and the solution was stirred for 1 h at room temperature. Another 100 μL of 3.0M dimethylsulfinyl anion was added, and the solution was sonicated again for 1 h. Methyl iodide (18.5 μL) was added and the solution was stirred overnight at room temperature. The addition of dimethylsulfinyl anion and methyl iodide was repeated one more time, an excess of methyl iodide (150 μL) being added in this step. The permethylated oligosaccharide was chromatographed on a column (1 × 50 cm) of LH-20 with 1:1 (v/v) chloroform–methanol as solvent. The permethylated carbohydrate was identified in the column effluent by the anthrone assay[40].

Ethylation. — The carbohydrate sample was dried *in vacuo* over phosphorus pentaoxide for at least 10 h at 45°. The sample was then dissolved in dry dimethyl sulfoxide at a concentration of ~10 mg/mL. Potassium dimethylsulfinyl anion (3.0M, 140 μL) was added to each sample and the solutions were sonicated for 1 h at ~60°. Ethyl iodide (35 μL) was added to the sample and the solution was stirred for 1 h at room temperature. The ethylation procedure was repeated again, except that 500 μL of ethyl iodide was added, and the solution was stirred overnight. The partially methylated, partially ethylated oligosaccharides were purified by chromatography on LH-20 in 1:1 (v/v) chloroform–methanol. The location of these carbohydrate derivatives in the LH-20 effluent was determined by the anthrone assay.

Reduction of carboxyl groups of uronic acids. — Methylated xanthan (~20 mg) was dissolved in 7 mL of 95% ethanol, and oxolane (17 mL) and 60 mg of sodium borodeuteride were added[15]. The solution was incubated for 18 h at room temperature, and then for 1 h at 70°. Reduction was terminated by the dropwise addition of acetic acid until effervescence ceased. Acetic acid in methanol (10%, 10 mL) was added and the mixture was evaporated at room temperature by rotary evaporation. The addition of acidified methanol and evaporation to dryness was repeated 3 times. The methylated, reduced xanthan was dissolved in 5 mL of 1:1 (v/v) chloroform–methanol and purified by chromatography on LH-20.

Glycosyl-composition analysis. — Glycosyl compositions were determined by the alditol acetate method[1].

Analysis of glycosyl-linkage composition. — The glycosyl-linkage composition of methylated, complex carbohydrates was performed as described[41], except that the reduction following hydrolysis was performed with sodium borodeuteride in ethanol (rather than in water) that had been made M with respect to ammonia.

Preparation of partially methylated, partially ethylated alditol acetates from peralkylated fragment-oligosaccharides. — The partially methylated, partially ethylated oligosaccharides obtained by fragmentation of methylated, complex carbohydrates were converted into partially alkylated alditol acetates as described in the preceding paragraph. Care was taken during this procedure not to evaporate the particularly volatile, partially alkylated alditols derived from the pre-reduced and terminal residues of the oligosaccharides. All evaporations were performed at room temperature and for the minimum time required to bring the samples to dryness.

G.l.c. and g.l.c.–m.s. — G.l.c. of the alditol acetates and of the partially alkylated alditol acetates was performed on either a packed glass column (2 mm i.d. × 180 cm) that contained[1] 0.2% ethylene glycol succinate, 0.2% ethylene glycol adipate, and 0.4% XF-1150 on Gas-chrom P, or on a 25-m, open-tubular, glass capillary column containing SE-30 (LKB, Broma, Sweden). All g.l.c.–m.s. analyses were carried out on the capillary column with a Hewlett–Packard Model 5710 gas chromatograph, 5980A mass spectrometer, and a Model 5933A data system.

The areas of the g.l.c. flame-ionization detector peaks were quantitated with a Spectra-Physics Minigrator. The g.l.c. flame-ionization responses were corrected to molar responses as described by Sweet *et al.*[42].

ACKNOWLEDGMENTS

The authors thank David Gollin for preparing the Bio-Gel P-2-purified xyloglucan nonasaccharide, and Pat Schuller for extraordinarily competent and generous assistance in preparation of this manuscript. We also thank Steve McGaughey of Waters Associates for his expert assistance during the introduction of h.p.l.c. to this laboratory.

REFERENCES

1 P. ALBERSHEIM, D. J. NEVINS, P. D. ENGLISH, AND A. KARR, *Carbohydr. Res.*, 5 (1967) 340–345.
2 C. C. SWEELEY, R. BENTLEY, M. MAKITA, AND W. W. WELLS, *J. Am. Chem. Soc.*, 85 (1963) 2497–2507.
3 N. SHARON, *Complex Carbohydrates, Their Chemistry, Biosynthesis and Functions*, Addison–Wesley, Boston, Mass., 1975.
4 H. BJÖRNDAL, C. G. HELLERQVIST, B. LINDBERG, AND S. SVENSSON, *Angew. Chem. Int. Ed. Engl.*, 9 (1970) 610–619.
5 B. LINDBERG, J. LÖNNGREN, AND S. SVENSSON, *Adv. Carbohydr. Chem. Biochem.*, 31 (1975) 185–240.
6 I. DANISHEFSKY, R. L. WHISTLER, AND F. A. BETTELHEIM, in W. PIGMAN AND D. HORTON (Eds.), *The Carbohydrates*, Vol. IIA, Academic Press, New York, 1970, pp. 375–412.
7 B. LINDBERG, J. LÖNNGREN, AND J. L. THOMPSON, *Carbohydr. Res.*, 28 (1973) 351–357.
8 G. O. ASPINALL AND K.-G. ROSELL, *Carbohydr. Res.*, 57 (1977) C23–C26.
9 T. S. STEWART, P. B. MENDERSHAUSEN, AND C. E. BALLOU, *Biochemistry*, 7 (1968) 1843–1854.
10 C. J. LAWSON AND K. C. SYMES, *Carbohydr. Res.*, 58 (1977) 433–438.
11 S. SVENSSON, *Methods Enzymol.*, 100 (1978) 33–38.

12 P.-E. Jansson, L. Kenne, B. Lindberg, H. Ljunggren, J. Lönngren, U. Rudén, and S. Svensson, *J. Am. Chem. Soc.*, 99 (1977) 3812–3815.
13 L. D. Melton, L. Mindt, D. A. Rees, and G. R. Sanderson, *Carbohydr. Res.*, 46 (1976) 245–257.
14 R. L. Taylor and H. E. Conrad, *Biochemistry*, 11 (1972) 1383–1388.
15 G. G. S. Dutton, K. L. Mackie, A. V. Savage, and M. D. Stephenson, *Carbohydr. Res.*, 66 (1978) 125–131.
16 P. A. Sandford and H. E. Conrad, *Biochemistry*, 5 (1966) 1508–1517.
17 D. P. Sweet, R. H. Shapiro, and P. Albersheim, *Biomed. Mass Spectrom.*, 1 (1974) 263–268.
18 G. G. S. Dutton and K. L. Mackie, *Carbohydr. Res.*, 62 (1978) 321–335.
19 B. Lindberg, B. Lindqvist, J. Lönngren, and W. Nimmich, *Carbohydr. Res.*, 58 (1977) 443–451.
20 G. G. S. Dutton and M.-T. Yang, *Carbohydr. Res.*, 59 (1977) 179–192.
21 J.-P. Joseleau, G. Chambat, M. Vignon, and F. Barnoud, *Carbohydr. Res.*, 58 (1977) 165–175.
22 S. Peat, W. J. Whelan, and J. G. Roberts, *J. Chem. Soc.*, (1957) 3916–3924.
23 N. B. Chanda, E. L. Hirst, and D. J. Manners, *J. Chem. Soc.*, (1957) 1951–1958.
24 A. S. Perlin and S. Suzuki, *Can. J. Chem.*, 40 (1962) 50–56.
25 H. Suzuki and T. Kaneko, *Agric. Biol. Chem.*, 40 (1976) 577–586.
26 P.-E. Jansson, L. Kenne, and B. Lindberg, *Carbohydr. Res.*, 45 (1975) 275–282.
27 A. N. de Belder and B. Norrman, *Carbohydr. Res.*, 8 (1968) 1–6.
28 I. R. Siddiqui, *Carbohydr. Res.*, 4 (1967) 284–291.
29 P. A. J. Gorin, T. Ishikawa, J. F. T. Spencer, and J. H. Sloneker, *Can. J. Chem.*, 45 (1967) 2005–2008.
30 J. H. Sloneker and A. Jeanes, *Can. J. Chem.*, 40 (1962) 2066–2071.
31 W. D. Bauer, K. W. Talmadge, K. Keegstra, and P. Albersheim, *Plant Physiol.*, 51 (1973) 174–187.
32 M. McNeil, A. G. Darvill, and P. Albersheim, *Progr. Chem. Org. Nat. Prod.*, 37 (1979) 191–249.
33 G. J. Gerwig, J. P. Kamerling, and J. F. G. Vliegenthart, *Carbohydr. Res.*, 62 (1978) 349–357.
34 K. Leontein, B. Lindberg, and J. Lönngren, *Carbohydr. Res.*, 62 (1978) 359–362.
35 J. Hoffman, B. Lindberg, and S. Svensson, *Acta Chem. Scand.*, 26 (1972) 661–666.
36 Y. M. Choy and G. G. S. Dutton, *Can. J. Chem.*, 51 (1973) 198–207.
37 J. E. Scott, *Methods Carbohydr. Chem.*, 5 (1965) 38–44.
38 A. S. Chaudhari, C. T. Bishop, and W. F. Dudman, *Carbohydr. Res.*, 28 (1973) 221–231.
39 S. Hakomori, *J. Biochem. (Tokyo)*, 55 (1964) 205–208.
40 Z. Dische, *Methods Carbohydr. Chem.*, 1 (1962) 477–514.
41 K. W. Talmadge, K. Keegstra, W. D. Bauer, and P. Albersheim, *Plant Physiol.*, 51 (1973) 158–173.
42 D. P. Sweet, R. H. Shapiro, and P. Albersheim, *Carbohydr. Res.*, 40 (1975) 217–225.