ELSEVIER

www.elsevier.nl/locate/carres

Carbohydrate Research 330 (2001) 103–114

CARBOHYDRATE RESEARCH

The occurrence of internal (1→5)-linked arabinofuranose and arabinopyranose residues in arabinogalactan side chains from soybean pectic substances

Miranda M.H. Huisman,[a,1] Lars P. Brüll[1,],[b,2] Jane E. Thomas-Oates,[c] Johan Haverkamp,[b] Henk A. Schols,[a] Alphons G.J. Voragen[a,*]

[a]*Laboratory of Food Chemistry, Department of Food Technology and Nutritional Sciences, Wageningen University, PO Box 8129, 6700 EV Wageningen, The Netherlands*

[b]*Department of Mass Spectrometry, Bijvoet Centre for Biomolecular Research, Faculty of Chemistry, Utrecht University, FAFC Wentgebouw, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands*

[c]*Departments of Chemistry and Biomolecular Sciences, Michael Barber Centre for Mass Spectrometry, UMIST, PO Box 88, Manchester M60 1QD, UK*

Received 4 January 2000; received in revised form 29 August 2000; accepted 10 October 2000

Abstract

CDTA-extractable soybean pectic substances were subjected to enzymatic digestion with arabinogalactan degrading enzymes yielding a resistant polymeric pectic backbone and arabino-, galacto-, and arabinogalacto-oligomers. The complex digest was fractionated using size-exclusion chromatography. Monosaccharide composition analysis, HPAEC fractionation and MALDI-TOF MS analysis of the resulting fractions showed that each contained a mixture of oligosaccharides of essentially the same degree of polymerisation, composed of only arabinose and galactose. MALDI-TOF MS analysis was used for molecular mass screening of oligosaccharides in underivatised HPAEC fractions. The monosaccharide sequence and the branching pattern of oligosaccharides (degree of polymerisation from 4 to 8) were determined using linkage analysis and ES-CID tandem MS analysis of the per-O-methylated oligosaccharides in each of the HPAEC fractions. These analyses indicated the presence of common linear (1→4)-linked galacto-oligosaccharides, and both linear and branched arabino-oligosaccharides. In addition, the results unambiguously showed the presence of oligosaccharides containing (1→4)-linked galactose residues bearing an arabino*pyranose* residue as the non-reducing terminal residue, and a mixture of linear oligosaccharides constructed of (1→4)-linked galactose residues interspersed with an internal (1→5)-linked arabino*furanose* residue. The consequences of these two new structural features of pectic arabinogalactan side chains are discussed.

Keywords: Soybean; Arabinogalacto-oligosaccharides; Internal arabinofuranose residues; Terminal arabinopyranose residues

* Corresponding author. Tel.: +31-317-482888; fax: +31-317-484893.

E-mail address: fons.voragen@chem.fdsci.wau.nl (A.G.J. Voragen).

[1] Contributed equally.

[2] Present address: Netherlands Organisation for Applied Scientific Research, TNO Food and Nutrition, Utrechtseweg 48, PO Box 360, 3700 AJ Zeist, The Netherlands.

0008-6215/01/$ - see front matter

PII: S0008-6215(00)00269-X

1. Introduction

Arabinose- or galactose- containing homoglycans are known to occur in nature, but heteropolysaccharides containing both types of monosaccharide residue are much more abundant. Arabinogalactans are often linked covalently to protein, or to pectic substances.[1] They can be subdivided into two main structural types: the arabino-(1→4)-galactans (type I), and the heavily branched arabino-(1→3),(1→6)-galactans (type II). Type I pectic L-arabino-D-galactans are arabinose-substituted derivatives of linear (1→4)-linked β-D-galactan. Ara*f* and Gal*p* groups form stubs linked via C-3 along the main chains. No association with protein has been reported for this group.[1–3] The second group of arabinogalactans, the type II L-arabinosyl-substituted branched (1→3),(1→6)-D-galactans, are widespread in plant tissues, tissue cultures, and exudate gums. They comprise a highly branched polysaccharide with ramified chains of β-D-Gal*p* residues joined by (1→3) and (1→6) linkages, the former predominantly in the interior and the latter in the exterior chains. β-D-Gal*p* residues terminate the bulk of the exterior chains containing L-Ara*f*, with, to a lesser extent, L-Ara*p* residues terminating some of the chains. In addition to Ara and Gal, type II arabino-(1→3),(1→6)-galactans contain a range of other monosaccharides, including D-Glc*p*A and its 4-*O*-methyl ether and D-Gal*p*A.[1–4] The type II arabinogalactans can also occur as pectic side chains, consisting of a (1→3)-linked galactan backbone to which (1→6)-linked galactosyl side chains are attached to C-6 of the backbone. These side chains possess Ara*f* side chains attached to C-3.[5]

In previous studies, cell wall material was isolated from soybean meal and sequentially extracted. The chelating agent soluble solids (ChSS fraction) contained the major fraction of the cell wall pectic substances.[6] The neutral side chains of these pectic substances could be degraded by the combined activity of endo-galactanase, exo-galactanase, endo-arabinanase and arabinofuranosidase B.[7] In the literature, the arabinogalactan is described as a β-(1→4)-linked Gal*p* chain with little branching. Ara*f* residues are present as (1→5)-linked side chains with an average length of two monosaccharide units attached to O-3 of the Gal residues.[8–11] However, research carried out by Labavitch and co-workers[12] indicated that the Ara residues are primarily present in large oligo- or polyarabinosides.

Structural details of the neutral side chains of soybean pectic substances have not yet been determined. Therefore, the present study of oligomers released during enzymatic degradation of soybean pectic arabinogalactan side chains has been performed. Arabino-, galacto-, and arabinogalacto-oligosaccharides were isolated by size-exclusion and anion-exchange chromatography and analysed by monosaccharide and linkage analyses, mass spectrometry, and enzymatic degradation studies.

2. Materials and methods

Enzymatic degradation of pectic arabinogalactan side chains from soybean.—A solution of soybean CDTA-extractable pectic substances (250 mg) in NaOAc buffer (0.05 M, 25 mL) was digested with a combination of endo-galactanase, exo-galactanase, endo-arabinanase and arabinofuranosidase B for 10 h at 30 °C, while continuously mixed 'head over tail'. The incubation was stopped by heating for 10 min at 100 °C.[7]

Size-exclusion chromatography.—The arabinogalacto-oligomers were separated from the polymeric pectic residue by fractionation on a Sephacryl S-100 HR column using a Hiload system (Amersham Pharmacia Biotech AB, Uppsala, Sweden). Elution was carried out using a NaOAc buffer (pH 5.0, 0.05 M) The elution rate was 0.4 mL/min.[7] The oligomers were pooled and concentrated, and further fractionated based on their size on a column (100 × 2.6 cm) of Bio-Gel P-2 (200–400 mesh, Bio-Rad, Richmond, CA, USA) using a Hiload system. Components were eluted with distilled water at 60 °C (flow rate was 0.5 mL/min) and monitored by refractive index detection using a Shodex RI-72 detector. Fractions (7.5 mL) were collected and fractions arising from individual peaks were pooled.

Off-line HPAEC-MALDI-TOF MS.—Bio-Gel pools 4–8 were further fractionated by HPAEC performed on a Dionex Bio-LC system.[13] The gradients were obtained by mixing solutions of 0.1 M NaOH and 1 M sodium acetate in 0.1 M NaOH. The gradient was optimised for each pool (Table 1). The CarboPac PA1 column (4 × 250 mm) was always equilibrated for 15 min before 20 μL of the sample was injected. After a run, the column was washed for 5 min with 1 M NaOAc in 0.1 M NaOH.

After passing through a Dionex PED detector operated in the pulsed amperometric detection (PAD) mode, the eluate of the CarboPac PA1 column was desalted on-line using a self-regenerating anion suppresser 4 mm unit (Dionex ASRS-ULTRA), and fractions (167 μL) were collected. The HPAEC fractions were then directly analysed using MALDI-TOF MS.

Enzymatic degradation.—Bio-Gel P-2 pool 6 was incubated with endo-galactanase. The incubations were performed in NaOAc buffer (pH 5.0, 50 mM) containing 0.01% NaN_3 at 30 °C for 7 h. Carbohydrate-degrading activities were determined by HPAEC analysis of the digest, using the optimised gradient for pool 6 (as described above).

Per-O-methylation of oligosaccharides.—Per-O-methylation[14] of the lyophilised oligosaccharides was performed by adding freshly-ground NaOH to the lyophilised oligosaccharide fractions dissolved in 200 μL Me_2SO. Aliquots of 250 μL MeI were added after 0, 10, and 30 min. The reaction was stopped 20 min after the final addition of methyl iodide by adding 1 mL sodium thiosulphate solution (100 mg/mL) and 1 mL chloroform. The chloroform layer was washed six times with water, after which the organic layer was evaporated to dryness under nitrogen.

Table 1
Gradients used for fractionation of pools 4–8 by HPAEC

Time (min)	Conc. NaAc (M)	Conc. NaOH (M)
dp 4 and 5		
0	0	0.1
40	0.4	0.1
45	1	0.1
dp 6		
0	0.07	0.1
5	0.07	0.1
40	0.14	0.1
45	1	0.1
dp 7 and 8		
0	0.07	0.1
5	0.07	0.1
40	0.11	0.1
50	1	0.1

Analytical methods

Neutral monosaccharide composition of the Bio-Gel P-2 pools. The neutral monosaccharide composition of the Bio-Gel P-2 pools was determined following release of the monosaccharides using methanolysis combined with TFA hydrolysis.[15] Samples were first dissolved in distilled water (1 mg/mL). An aliquot of 20 μL of this solution was dried under a stream of air. The dried sample was then submitted to methanolysis in 0.5 mL anhydrous 2 M HCl in absolute methanol for 16 h at 80 °C. After cooling, the liquid was evaporated under a stream of air and 0.5 mL of 2 M TFA solution was added and heated for 1 h at 121 °C. The samples were dried and 100 μl of distilled water was added. Analysis of the liberated monosaccharides was performed using HPAEC fractionation and PAD detection.[15]

Glycosidic linkage analysis. Glycosidic linkage analysis was performed following hydrolysis, reduction and O-acetylation of the per-O-methylated oligosaccharides.[16] GC–MS analyses were performed using a Fisons MD800 mass spectrometer fitted with a Carlo Erba GC8060 gas chromatograph and an on-column injector and using helium as the carrier gas. Monosaccharide derivatives were separated on a DB-5MS column (30 m × 0.32 mm i.d., J&W Scientific). Partially methylated alditol acetates (PMAAs) were injected in solution in CH_2Cl_2 (1 μL injected) and separated using the following temperature program: 50 °C for 2 min, 50–130 °C at 40 °C/min, 40 °C isothermal for 2 min, 130–230 °C at 4 °C/min and 230 °C isothermal for 15 min. Mass spectra were recorded under electron impact conditions in the positive ion mode with an electron energy of 70 eV and were recorded using linear scanning from m/z 55–400 over 0.9 s.

High-performance anion-exchange chromatography. High-performance anion-exchange chromatography (HPAEC) was performed on a Dionex Bio-LC system (as described above). Different gradients were used for determination of the monosaccharide composition after methanolysis combined with TFA hydrolysis,[15] and for the elution of arabinogalacto-oligomers.

For the determination of the arabinogalacto-oligomers, the CarboPac PA1 column was equilibrated with 0.1 M NaOH. The sample (20 μL) was injected, and a linear gradient from 0 to 0.4 M NaOAc in 40 min was applied. The column was washed for 5 min with 1 M NaOAc, and then equilibrated again for 15 min with 0.1 M NaOH. Calibration was performed with a series of Gal- and Ara-oligomers, obtained on enzymatic degradation of linear galactan by endo-galactanase and enzymically-debranched arabinan[17] by endo-arabinanase.

Matrix-assisted laser-desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS). The matrix solution was prepared by dissolving 9 mg 2,5-dihydroxybenzoic acid and 3 mg 1-hydroxy-isoquinoline in 700 μL distilled water and 300 μL acetonitrile. A 1 μL volume of this solution was placed on the sample plate together with 1 μL of the sample solution and allowed to dry at room temperature. The sample plate was then placed in the instrument.

MALDI-TOF MS were recorded on a Voyager-DE RP Biospectrometry Workstation (PerSeptive Biosystems, Inc., Framingham, MA, USA) (Department of Food Technology and Nutritional Sciences, Wageningen University) equipped with a nitrogen laser operating at 337 nm (3-ns pulse duration), a single stage reflector, and delayed extraction. The accelerating voltage used was 12 kV and the delay time setting was 200 ns. Each spectrum was produced by accumulating data from 100 to 256 laser shots. Mass spectra were calibrated with an external standard containing GalA-oligomers (degree of polymerisation 2–9).

Tandem mass spectrometry. Collision induced dissociation (CID) tandem mass (MS–MS) spectra were obtained using a Micromass Q-TOF hybrid tandem mass spectrometer (Department of Mass Spectrometry, Utrecht University) equipped with a Z-Spray sample introduction system and gold coated glass capillaries in a nanoelectrospray ionisation source. Argon was used as collision gas and a collision energy of 60 eV was employed. Cone voltage and skimmer off-set were set at approximately 75 and 5 V, respectively with a capillary voltage of 2100 V. Ten percent of the native sample was used for methylation. The native and per-O-methylated products were dissolved in 100 μL 1:1 methanol–water and 1 μL of the sample was introduced into the glass capillary. Spectra were acquired with the TOF analyser over a mass range that is dependent on the molecular mass of the analyte, data were integrated every 2.3 s, and processed using the MASSLYNX software, version 3.0.

3. Results and discussion

Enzymatic degradation of pectic arabinogalactan side chains from soybean.—Soybean pectic substances obtained by CDTA extraction of water-unextractable cell wall material (ChSS fraction), were digested with endo-galactanase, exo-galactanase, endo-arabinanase and arabinofuranosidase B, as described by Huisman et al.[7] However, in the current experiments, the objective of the incubation was to obtain oligosaccharides large enough to provide information about the structure of the arabinogalactan side chains. This was achieved by incubating for only 10 h. It should be pointed out that the oligosaccharides in the digest are not necessarily limit-digest products. The first step in isolating the oligosaccharides was the removal of the remaining polymeric material by fractionation on a Sephacryl S-100 HR column. The oligosaccharide-containing pool, on fractionation using HPAEC with PAD detection yielded a very complex chromatogram, indicating that this pool contains a wide variety of different oligomers (Fig. 1).

The oligomers in the oligosaccharide pool fraction were fractionated by Bio-Gel P-2 chromatography (not shown). The numbers of the pools (1–11) correspond to the degree of

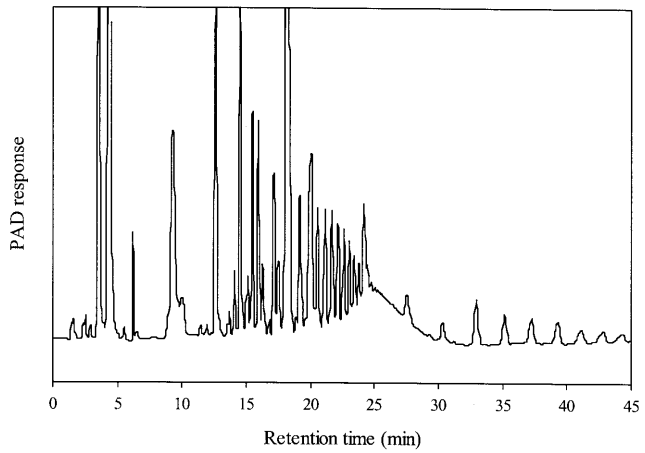

Fig. 1. HPAE chromatogram for the oligosaccharide-containing pool from the ChSS digest obtained after incubation of the soybean ChSS fraction with endo-galactanase, exo-galactanase, endo-arabinanase, and arabinofuranosidase B.

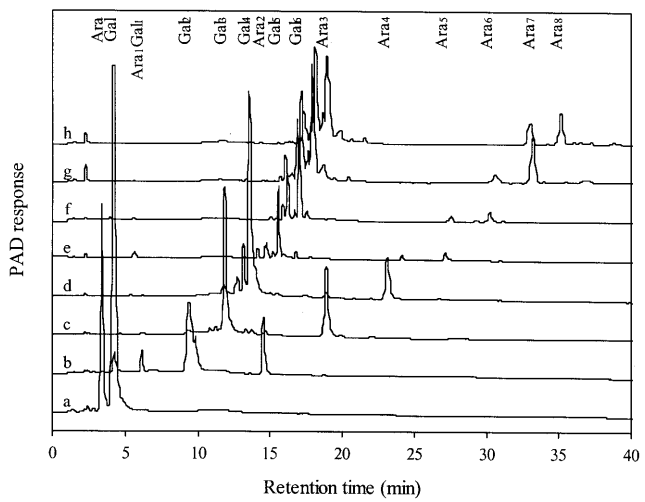

Fig. 2. HPAE chromatogram for arabinogalacto-oligomers in Bio-Gel P-2 pools 1–8 (traces a–h, respectively).

polymerisation of the oligosaccharides, as determined by MALDI-TOF MS analysis of the oligosaccharides (see below). Pool 12 eluted in the void volume and contained oligomers with a degree of polymerisation of 12 and higher.

Characterisation of the Bio-Gel P-2 pools.— The Bio-Gel P-2 pools were subjected to HPAEC and MALDI-TOF MS analyses. Both techniques showed that the pools contain mixtures of oligosaccharides. The MALDI-TOF MS indicated that the oligosaccharides have essentially the same degree of polymerisation (n). In addition, as n increases the presence of homologues with a degree of polymerisation of $n+1$ and $n-1$ is also detected. This is a direct result of the decreasing resolution of the Bio-Gel P-2 column with increasing n, and of the differences in hydrodynamic volume of Ara and Gal residues.

The HPAE chromatograms obtained from the Bio-Gel P-2 pools are shown in Fig. 2. Pool 1 contains Ara and Gal monomers only (trace a). The presence of four compounds is suggested by the HPAE chromatogram obtained from pool 2 (trace b). From the retention times of the eluting compounds it was concluded that they represent the Gal monomer, the dimer(s) Ara_1Gal_1, Gal_2, and Ara_2, respectively.

As the degree of polymerisation of the pool increases, the number of oligomers in the pool increases and the chromatogram becomes more complex (Fig. 2). Traces c–h in Fig. 2 show first a cluster of components eluting within 20 min, co-eluting with Gal_n-oligomers, followed by one or two components co-eluting with Ara_n-oligosaccharides. Monosaccharide composition analyses of pools 3–8 revealed that Ara and Gal are the only neutral monosaccharide residues present in these pools.

The HPAE chromatograms suggest a large diversity of oligomers within the pools, but do not give information about their composition. The molecular masses of the compounds in the pools analysed by MALDI-TOF MS are indicative of the compositions of the different oligomers with respect to the number of Ara and Gal residues present. As an example, the MALDI-TOF MS of the arabinogalacto hexamers is shown in Fig. 3. The sodium-cationised $[M+Na]^+$ ions are the dominant species observed in the spectra of the Bio-Gel P-2 pools, although some ions are accompanied by the potassium-cationised $[M+K]^+$ species. MALDI-MS analysis of Bio-Gel P-2 pool 6 yielded a spectrum containing ions for all possible Gal_xAra_y hexasaccharide compositions. The most abundant ion in the MALDI-TOF MS obtained from Bio-Gel P-2 pool 6 is observed at m/z 1013 and corresponds to sodium-cationised Gal_6. In addition to compounds with $n=6$, the spectrum shows the presence of

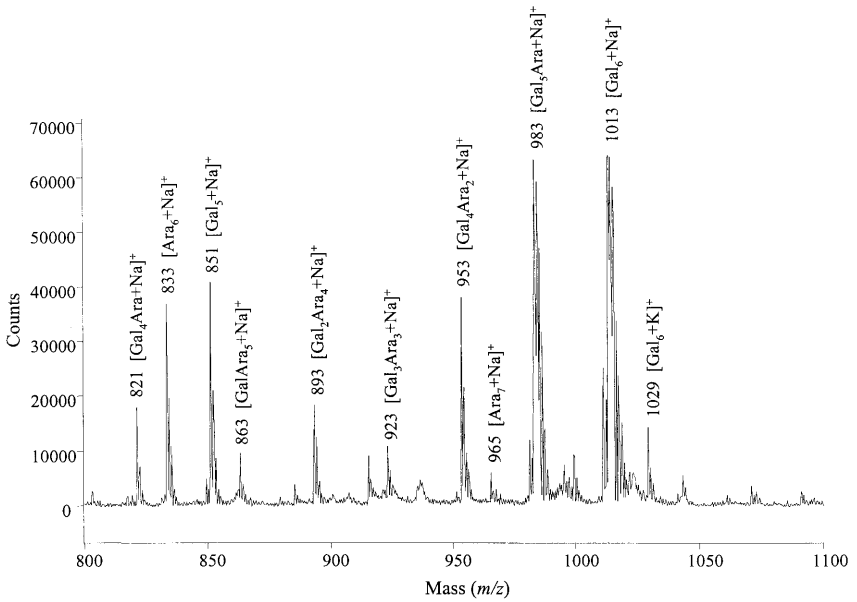

Fig. 3. MALDI-TOF MS of Bio-Gel P-2 pool 6 containing oligomers released from soybean pectic substances by arabinogalactan-degrading enzymes.

two different pentamers (Gal_5 and Gal_4Ara) and one heptamer (Ara_7).

The other Bio-Gel P-2 pools were also analysed by MALDI-TOF MS. Pools 3–5 also contain the whole range of possible Gal_xAra_y compositions for that particular degree of polymerisation. The pools containing oligomers with higher values for *n* contain only those arabinogalacto-oligosaccharides in which the majority of monosaccharide residues are Gal.

Characterisation of the arabinogalacto-oligomers by off-line HPAEC MALDI-TOF MS analysis.—The elution order of the oligosaccharides from HPAEC can not be predicted, and little structural information can therefore be deduced from these experiments. Pools 4–8 were fractionated using an analytical Carbopac PA1 column to allow structural studies of the oligosaccharides to be carried out. The high sodium concentration in the mobile phase eluting from the column was reduced on-line using a self-regenerating anion suppresser 4 mm unit (Dionex ASRS-ULTRA). The 'desalted' HPAEC fractions were then directly analysed using MALDI-TOF MS.

As an example, the HPAE chromatogram obtained from Bio-Gel P-2 pool 6 using an optimised gradient for this mixture is shown in Fig. 4. The most abundant component in the HPAE chromatogram (fraction 6.4, retention time 12.5 min) yielded an ion at *m/z* 1013 on MALDI-TOF MS analysis, corresponding to sodium-cationised Gal_6. This is consistent with its co-elution with the linear β-(1→4)-linked galacto-hexasaccharide in the standard. Fraction 6.1 contains Gal_5 (*m/z* 851); fractions 6.2 and 6.3 contain hexasaccharides composed of one Ara and five Gal residues (*m/z* 983). The MALDI-TOF MS of fraction 6.5 is identical to the spectrum of fraction 6.4, indicating Gal_6.

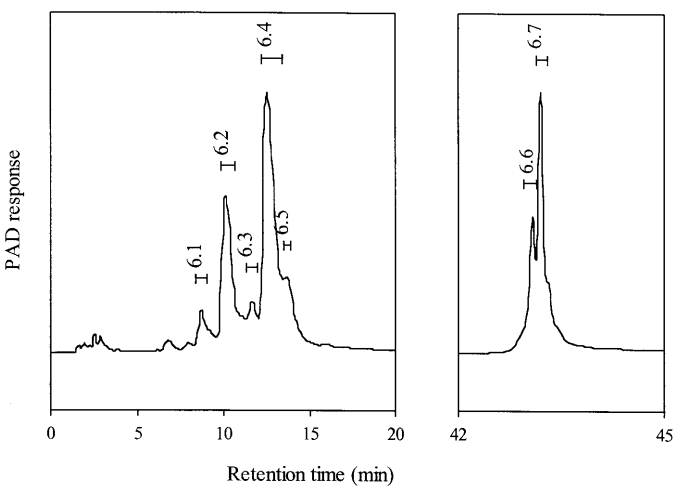

Fig. 4. HPAE chromatogram for Bio-Gel P-2 pool 6 from the soybean ChSS digest.

Fractions 6.6 and 6.7 both show ions at m/z 833, corresponding to sodium-cationised Ara_6. Their retention times are consistent with the elution behaviour of a linear α-(1→5)-linked Ara_6 standard.

The HPAE chromatograms of the Bio-Gel P-2 pools 4–8 show a general elution order: Gal_{n-1}, $Gal_{n-1}Ara$, Gal_n, followed by the arabino-oligosaccharides. With increasing degree of polymerisation of the pool, the number of peaks in the HPAE chromatogram increases and the chromatogram becomes more complex. The HPAE chromatogram of Bio-Gel P-2 pool 4 thus has only four peaks corresponding to compounds with $n=4$: one Gal_3Ara, two Gal_4 isomers and one Ara_4.[18] The HPAE chromatogram of Bio-Gel P-2 pool 8, in contrast, contains ten peaks corresponding to compounds with $n=8$ (not shown): three Gal_7Ara, two Gal_8, and five Ara_8 isomers.

In the MALDI-TOF MS of the Bio-Gel P-2 pools ($dp=n$) every possible oligosaccharide composition from Ara_0Gal_n up to Gal_0Ara_n was present (Fig. 3). On HPAEC fractionation, however, only Gal_n, $Gal_{n-1}Ara$, Gal_{n-1}, and Ara_n were detected. This can be explained by the fact that the peaks from the higher Ara-containing species in the MALDI-TOF MS are less intense than those from Gal_n, $Gal_{n-1}Ara$, Gal_{n-1}, and Ara_n, and MALDI-TOF MS is much more sensitive than the pulsed amperometric detector. The possibility that the ions assigned as corresponding to the higher Ara-containing species could actually correspond to fragment ions derived from the higher mass species (i.e. those with more Gal in them) cross-ring-cleavage can be ruled out, since ions arising by cross-ring-cleavage would then also be expected in the MALDI-TOF MS of the isolated components (HPAEC fractions) and they are not observed.

The different behaviour of oligosaccharides of identical composition on HPAEC indicates that their structures are not identical; they might differ in their reducing terminal residue, in the type of glycosidic linkages, or in branching pattern.[19] To determine the structural details of these oligosaccharides, additional types of analysis are required, such as linkage analysis, tandem mass spectrometry, NMR spectroscopy, and digestion of the oligosaccharides by specific enzymes. Fractions 6.2–6.7 were purified, but the amounts available were insufficient for analysis by NMR spectroscopy. The results of linkage analysis and tandem mass spectrometry are described below.

Linkage analysis.—Analysis of the partially methylated alditol acetates (PMAAs) from fractions 6.4 and 6.5 allowed identification of derivatives indicative of terminal and (1→4)-substituted galactose residues, indicating the presence of linear (1→4)-linked Gal_6 in both fractions. To explain their different elution behaviours on HPAEC, additional structural information about these galacto-hexasaccharides is required.

Linkage analysis of permethylated fraction 6.2 ($AraGal_5$) indicated the presence of (1→4)-substituted and terminal Gal, and (1→5)-substituted Ara*f*. A small peak that was observed corresponding to a PMAA derived from (1→4, 6)-disubstituted galactose could be indicative of the presence of branched galactose residues, but is more likely to be the result of undermethylation, since tandem mass spectrometry of fraction 6.2 shows the presence of linear oligosaccharides only (see below).

Fraction 6.3, having the same monosaccharide composition as fraction 6.2, yielded different linkage analysis data. It gave PMAA derivatives indicative of (1→4)-substituted Gal, terminal Gal, terminal arabino*pyranose*, and (1→5)-substituted arabinofuranose residues. Since only Gal_5Ara was demonstrated in this fraction, it must contain a mixture of at least two compounds.

Both ring forms of arabinose occur in arabinogalactan type II, in which L-arabinofuranosyl, and to a lesser extent L-arabinopyranosyl residues, terminate some of the side chains.[4] However, in soybean arabinogalactan we are dealing with linear (1→4)-linked β-D-galactan isolated from the pectic substances, which is type I arabinogalactan. The existence of arabinopyranose residues in pectic arabinogalactan is uncommon, since the presence of only arabinofuranose residues is generally reported. Only pectic substances iso-

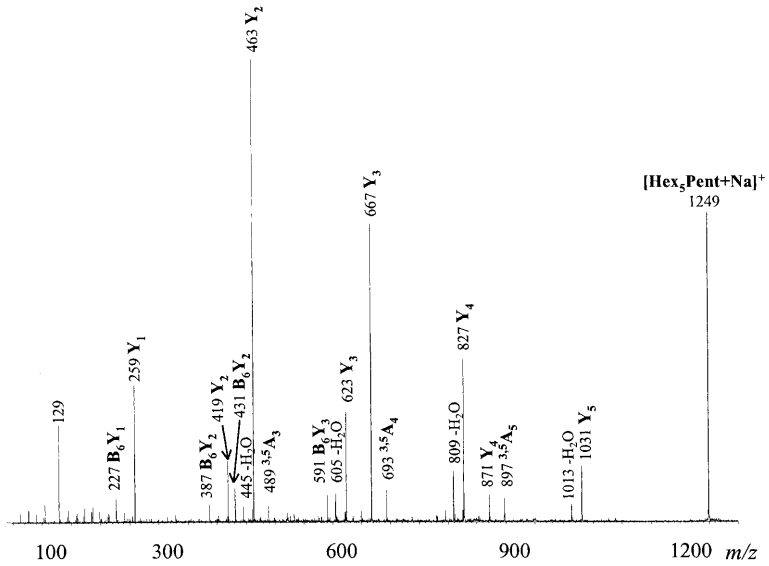

Fig. 5. Positive ES CID tandem mass spectrum of per-O-methylated $[Hex_5Pent + Na]^+$ from fraction HPAEC 6.2, released from soybean pectic substances by arabinogalactan-degrading enzymes.

lated from the roots of *Angelica acutiloba* Kitagawa were shown to contain a small proportion of arabinopyranose.[20]

Tandem mass spectrometry.—Per-O-methylated fraction 6.4 was analysed using tandem mass spectrometry. In the tandem mass spectrum obtained from the sodium-cationised pseudomolecular ion of Hex_6, at m/z 1293 $[Hex_6 + Na]^+$ (not shown), a complete series of Y_n ions and the absence of $Y_n - 14$ ions are indicative of a linear hexasaccharide.

Ions at m/z 227, 431 and 635 are the result of double cleavage events. The ions at m/z 445, 649, 853 and 1057 are the result of the loss of water from the Y_n ions, which according to the nomenclature of Domon and Costello[21] are Z_n ions. A series of cross-ring cleavage ions, observed at m/z 315, 329, 519, 533, 723, 737, and 941, is indicative of $(1\rightarrow4)$-substitution of the monosaccharide residues. Since the only hexose present in pool 6 is galactose, the tandem mass spectrum of the per-O-methylated $[Hex_6 + Na]^+$ indicates a linear $(1\rightarrow4)$-linked galacto-hexasaccharide, which is consistent with the results from linkage analysis of fraction 6.4.

The tandem mass spectrum obtained from the $[Hex_6 + Na]^+$ ion in fraction 6.5 (not shown) does not significantly differ from the spectrum obtained from fraction 6.4. This indicates that fraction 6.5 also contains a linear $(1\rightarrow4)$-linked galacto-hexasaccharide, as was already suggested by its linkage composition. Separation on HPAEC may well point to the presence of differences in the anomeric configuration of (one or more) glycosidic linkages.[19]

In the CID tandem mass spectrum of per-O-methylated $[Hex_5Pent + Na]^+$ in fraction 6.2, an intense series of Y_n ions is observed at m/z 259, 463, 667, 827, and 1031 (Fig. 5). These ions are indicative of a linear oligosaccharide corresponding to structure II (Fig. 6). The ion observed at m/z 827 is a Y_4 ion indicative of a structure containing three

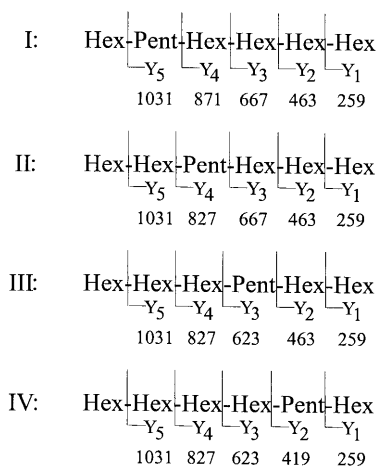

Fig. 6. Possible structures for Hex_5Pent isomers with an internal pentose present in fraction 6.2, based on results from tandem mass spectrometry.

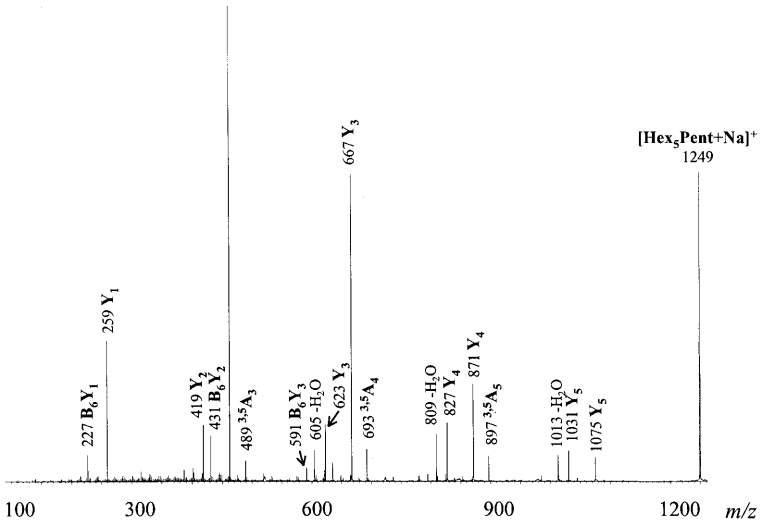

Fig. 7. Positive ES CID tandem mass spectrum of per-O-methylated $[Hex_5Pent + Na]^+$ from fraction HPAEC 6.3, released from soybean pectic substances by arabinogalactan-degrading enzymes.

hexose and one pentose residues. The absence of either a Y_1 ion at m/z 215 indicative of a reducing pentose or a Y_5 ion at m/z 1075 indicative of a non-reducing terminal pentose indicates that the compound that yields this Y_4 ion must have an internal pentose. In the tandem mass spectrum of $[Hex_5Pent + Na]^+$ further Y_n ions for structures that bear a pentose are also present: at m/z 419 (PentHex) and 623 ($PentHex_2$) and 1031 ($PentHex_4$). The absence of the Y_n ions at m/z 215 and 1075, in combination with the pentose-bearing Y_n ions m/z 1031, 827, 623 and 419, means that none of the oligosaccharide structures I–IV (Fig. 6) can be ruled out as being present in fraction 6.2.

Other ion series as indicated in the spectrum (Fig. 5) result from double cleavage or cross-ring cleavage events and corroborate the interpretation given above. They indicate the presence of either $(1\rightarrow4)$ or $(1\rightarrow6)$ linkages. Since the ion observed at m/z 489 is composed of a Hex, a Pent, and the remainder of the next Hex residue the internal Pent is substituted to the C-6 or C-4 of the cleaved Hex.

From these data, taken together with the monosaccharide composition of pool 6 and the linkage analysis results for fraction 6.2, it can be deduced that fraction 6.2 contains a mixture of $(1\rightarrow4)$-linked galacto-hexasaccharide isomers with an internal $(1\rightarrow5)$-linked arabinofuranose residue. Internal arabinose has been described once before in arabinogalactan type II from larch, where arabinose was shown to be present in the side chains as well as in the backbone of the molecule.[2,22,23] The presence of an internal arabinofuranose residue in a pectic arabinogalactan chain in cell wall polysaccharides has not been reported previously, either in soybean, or in other fruit or vegetable cell walls.

The tandem mass spectrum of per-O-methylated $[Hex_5Pent + Na]^+$ in fraction 6.3 (Fig. 7) is significantly different from that obtained from the ion at the same m/z in fraction 6.2 (Fig. 5). The most noticeable difference between the spectra is found in the presence of an ion at m/z 1075 in the spectrum obtained from fraction 6.3. This ion corresponds to a Y_5 ion composed of five hexose residues. Since this Y_5 ion can only result from the loss of a pentose from the parent ion and since 'internal residue loss' has not been observed from sodium-cationised oligosaccharides[24] the Pent residue in this oligosaccharide is assigned as being in a terminal position.

The Y_n ion series in the spectrum obtained from $[Hex_5Pent + Na]^+$ from HPAEC fraction 6.3 present at m/z 259, 463, 667, 871, and 1075 is not accompanied by corresponding $Y_n - 14$ fragment ions, indicating that these are linear oligosaccharides. The most plausible structural explanation for the ion at m/z 1249 in fraction 6.3 is a linear hexasaccharide bearing the pentose on the non-reducing terminus. The remainder of the spectrum is very similar

to that from fraction 6.2, consistent with fraction 6.3 also containing a mixture of isomeric structures with an internal pentose. This is consistent with the results of linkage analysis, showing both a terminal pentopyranose and (1→5)-substituted pentofuranose. The hexoses are present as either (1→4)-substituted or as terminal hexoses, as in fraction 6.2.

The presence of isomers with an internal arabinose residue is most probably caused by the fact that fractions 6.2 and 6.3 are not fully resolved, and the relative amount of fraction 6.2 in the mixture is much larger than of fraction 6.3.

In the linear arabinogalacto-oligosaccharide bearing the arabinopyranose residue on the non-reducing terminus, this arabinose most probably does not originate as the remains of a side chain, because side chains are usually linked to C-3 along the main chain and not to C-4. This terminal arabinopyranose was probably present as a terminal arabinose in the polymer, since none of the enzymes used in the degradation of soybean ChSS is able to hydrolyse the glycosidic linkage between a galactose or arabinose residue and an arabinopyranose residue, and oligomers with an internal arabinopyranose residue would also have been present in the digest.

The mass spectra obtained from fractions 6.6 and 6.7 (not shown), contain ions at m/z 1029, which correspond to per-O-methylated sodium-cationised $Pent_6$. In the tandem mass spectra obtained from m/z 1029 from both fractions (not shown), a complete series of Y_n ions is present (m/z 215, 375, 535, 695, and 885), indicative of a linear $Pent_6$ oligosaccharide. Ions observed at m/z 361, 521, and 681 can be described as $Y_n - 14$ ions, which are the result of two glycosidic bond cleavages both accompanied by proton transfer and which are indicative of cleavage at a branched residue. The ion observed at m/z 507 is indicative of a triply substituted residue. So these fractions contain mixtures of linear and various branched arabinohexasaccharides. The characterisation of two co-eluting branched arabinotetrasaccharides was described previously by Brüll et al.[18] Labavitch et al.[12] have described the presence of large arabinan chains in soybean cell walls. The present research extends the knowledge of the structures of the arabinan side chains as being heavily branched.

Similar fractionation and tandem mass spectrometric analyses of the fractions from pools 4 to 8 were performed. All these pools were shown to contain linear (1→4)-linked galactooligosaccharides, (1→4)-linked galacto-oligosaccharides bearing an arabino*pyranose* residue at the non-reducing terminus, a mixture of linear oligosaccharides constructed of (1→4)-linked galactose residues interspersed with one internal (1→5)-linked arabino*furanose* residue, and both linear and branched arabino-oligosaccharides. The presence of linear β-(1→4)-linked galacto-oligosaccharides is consistent with published structures of soybean (arabino)galactan.[11] Arabinose-containing side chains might have been present, but could have been removed during enzymatic degradation by arabinofuranosidase B. Prior to this analysis it had been anticipated that arabinogalacto-oligosaccharides with a galactan main chain and (residual) arabinofuranose residues as side chains would be isolated. The presence of arabino-oligosaccharides (which can be degraded further by arabinofuranosidase B) demonstrates that arabinofuranosidase B had not digested the mixture to completion. Therefore, it is remarkable that all arabinose-containing side chains appear to have been removed from the galactan main chain. A possible explanation is that arabinose-containing side chains were not present in the original arabinogalactan pectic side chains in the ChSS extract. Another possible explanation, that arabinose substituents attached to a galactan main chain might somehow be more accessible to the enzyme than arabinose substituents attached to an arabinan main chain, is not very likely.

Further enzymatic degradation of the arabino-galacto-oligosaccharides.—Confirmation of the structures of the oligosaccharides identified by ES tandem MS, using enzymatic digestion is difficult. The main reason is that suitable pure enzymes are not available. Incubation of pool 6 with arabinofuranosidase B, indeed resulted in further degradation of the compounds in 6.6 and 6.7. The enzyme that is

required for the degradation of compound 6.3 would be an arabinopyranosidase. Such an enzyme, β-L-arabinopyranosidase, is poorly described in the literature.[25,26] The enzyme described by Dey is a true β-L-arabinopyranosidase isolated from *Cajanus indicus* seeds, and is unable to hydrolyse *p*-nitrophenyl α-D-galactoside, *p*-nitrophenyl α-D-fucoside or *p*-nitrophenyl β-D-galactoside.[25,26] This enzyme would probably be able to release the terminal arabinopyranose from the Ara*p*-(1→4)-Galp_n oligomers.

Further characterisation of the arabino-galacto-oligosaccharides using endo-galactanase, exo-galactanase or β-galactosidase is complicated. A major complication is the presence of mixtures of compounds in the Bio-Gel P-2 pools and of different isomers in the HPAEC fractions (particularly fraction 6.2). Due to the incomplete resolution of fractions 6.2 and 6.3 on HPAEC it is impossible to obtain the arabinopyranose-containing oligosaccharides in a pure form. A further difficulty is that the degree of polymerisation of the oligosaccharides influences the ability of the enzymes to hydrolyse the glycosidic linkages, so it would be difficult to distinguish whether the reason for the inability of the enzyme to hydrolyse the structure is its degree of polymerisation or the structure of the oligomers.

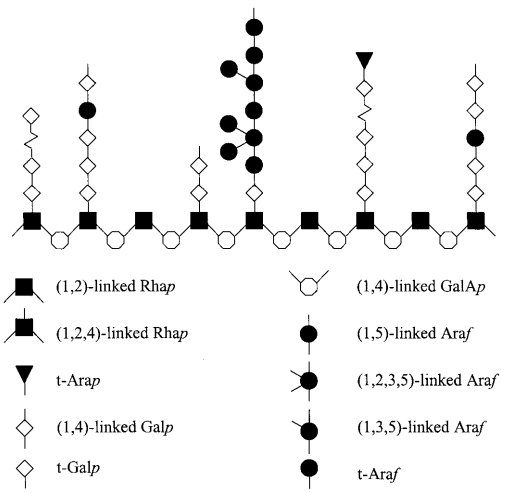

Fig. 8. Hypothetical structure of the rhamnogalacturonan regions of CDTA-extractable soybean pectic substances. The distribution of the acetyl groups is not shown. A strictly alternating sequence of Rha and GalA, termed RG-I,[28] is presented here, but the structure of the soybean rhamnogalacturonan backbone is as yet unknown. Only short (arabino)galactan side chains are presented in this figure, but the true average length of the side chains is 45–50 residues.

To illustrate the difficulty of carrying out enzymatic degradation studies on these mixtures, Bio-Gel P-2 Pool 6 was incubated with endo-galactanase. The resulting HPAE chromatogram obtained from the digest was indeed very complex. Peaks corresponding to fractions 6.3, 6.4 and 6.5 were completely and very rapidly lost, whereas peak 6.2 was removed more slowly. This may indicate that the enzyme has different affinities for the isomers in the mixture or that all these isomers are slowly degraded because of the presence of the internal Ara*f*. The reaction products are very diverse and it is impossible to determine from which oligosaccharide in the pool a specific product peak is derived. It is therefore not possible to draw any conclusions concerning the structures of the components giving rise to the product peaks or to identify from which parental oligosaccharides they were produced.

This study shows the need to isolate enzymes able to hydrolyse the glycosidic linkage between a (terminal or (1→4)-substituted) galactose residue and a (1→5)-substituted arabinose residue, and to hydrolyse the glycosidic linkage between a (terminal or (1→5)-substituted) arabinose residue and a (1→4)-substituted galactose residue, together with the need for arabinopyranosidases. These enzymes could be helpful in the determination of the structure of the purified oligosaccharides, and they could be helpful in the elucidation of the structure of the polymeric pectic arabinogalactan side chains from soybean meal.

A hypothetical structure of the pectic arabinogalactan side chains in soybean meal.—The pectic arabinogalactan side chains present in soybean pectic substances have been shown to be more complex than had been suggested by previous studies.[8–11] Based on our results, a hypothetical structure of the pectic arabinogalactan side chains is proposed (Fig. 8). The ratio of (1, 2)- to (1, 2, 4)-substituted Rha in the ChSS extract is 1:2.[27] From the monosaccharide composition of the ChSS fraction,[6] it can be concluded that the average length of the (arabino)galactan side chains is in the

range of 45–50 residues. It is not clear how the arabinan side chains were attached to the rhamnogalacturonan backbone. Although side chains commencing with an arabinosyl residue might be present in soybean, all side chains in Fig. 8 start with a galactosyl residue. This is mainly based on the results of previous studies, in which NMR analysis showed that terminal or (1→4)-substituted Gal was attached to the rhamnose residues after enzymatic and chemical removal of large stretches of the chains leaving a short stub behind.[27]

Acknowledgements

This research was supported by the Dutch Technology Foundation (STW), Gist-Brocades and the Product Board for Feeding Stuffs (VVR).

References

1. Stephen, A. M. In *The Polysaccharides*; Aspinall, G. O., Ed.; Academic: New York, 1983; Vol. 2, pp. 97–193.
2. Clarke, A. E.; Anderson, R. L.; Stone, B. A. *Phytochemistry* **1979**, *18*, 521–540.
3. McNeil, M.; Darvill, A. G.; Fry, S. C.; Albersheim, P. *Ann. Rev. Biochem.* **1984**, *53*, 625–663.
4. Selvendran, R. R. *J. Cell Sci. Suppl.* **1985**, *2*, 51–88.
5. Kiyohara, H.; Yamada, H. *Carbohydr. Res.* **1989**, *193*, 173–192.
6. Huisman, M. M. H.; Schols, H. A.; Voragen, A. G. J. *Carbohydr. Polym.* **1998**, *37*, 87–95.
7. Huisman, M. M. H.; Schols, H. A.; Voragen, A. G. J. *Carbohydr. Polym.* **1999**, *38*, 299–307.
8. Morita, M. *Agric. Biol. Chem.* **1965**, *29*, 564–573.
9. Morita, M. *Agric. Biol. Chem.* **1965**, *29*, 626–630.
10. Morita, M.; Okuhara, M.; Kikuchi, T.; Sakurai, Y. *Agric. Biol. Chem.* **1976**, *31*, 314–318.
11. Aspinall, G. O.; Begbie, R.; Hamilton, A.; Whyte, J. N. C. *J. Chem. Soc. C* **1967**, *170*, 1065–1070.
12. Labavitch, J. M.; Freeman, L. E.; Albersheim, P. *J. Biol. Chem.* **1976**, *251*, 5904–5910.
13. Schols, H. A.; Voragen, A. G. J.; Colquhoun, I. J. *Carbohydr. Res.* **1994**, *256*, 97–111.
14. Kerek, C.; Kerek, F. *Carbohydr. Res.* **1984**, *131*, 209–217.
15. deRuiter, G. A.; Schols, H. A.; Voragen, A. G. J.; Rombouts, F. M. *Anal. Biochem.* **1992**, *207*, 176–185.
16. deLederkremer, R. M.; Lima, C.; Ramirez, M. I.; Ferguson, M. A. J.; Homans, S. W.; Thomas-Oates, J. *J. Biol. Chem.* **1991**, *266*, 23670.
17. Cooper, J. M.; McCleary, B. V.; Morris, E. R.; Richardson, R. K.; Marrs, W. M.; Hart, R. J. *Gums Stabil. Food Ind.* **1992**, *6*, 451–460.
18. Brüll, L.; Huisman, M.; Schols, H.; Voragen, F.; Critchley, G.; Thomas-Oates, J.; Haverkamp, J. *J. Mass Spectrom.* **1998**, *33*, 713–720.
19. Lee, Y. C. *Anal. Biochem.* **1990**, *189*, 151–162.
20. Kiyohara, H.; Yamada, H.; Otsuka, Y. *Carbohydr. Res.* **1987**, *167*, 221–237.
21. Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397–409.
22. Bouveng, H. O.; Lindberg, B. *Acta Chem. Scand.* **1958**, *12*, 1977.
23. Young, R. A.; Sarkanen, K. V. *Carbohydr. Res.* **1977**, *59*, 193–201.
24. Brüll, L. P.; Kovácik, V.; Thomas-Oates, J. E.; Heerma, W.; Haverkamp, J. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1520–1532.
25. Dey, P. M. *Biochim. Biophys. Acta* **1973**, *302*, 393–398.
26. Dey, P. M. *Biochim. Biophys. Acta* **1983**, *746*, 8–13.
27. Huisman, M.M.H.; Fransen, C.T.M.; Kamerling, J.P.; Vliegenthart, J.F.G.; Schols, H.A.; Voragen, A.G.J. *Biopolymers* **2001**, in press.
28. O'Neill, M. A.; Albersheim, P.; Darvill, A. G. In *Carbohydrates*; Dey, P. M., Ed. Methods in plant biochemistry. Academic: London, 1990; Vol. 2, pp. 415–441.