

**Most of the Rhamnogalacturonan-I from Cultured Arabidopsis Cell Walls is
Covalently Linked to Arabinogalactan-Protein**

Li Tan^{1,2*}, Liang Zhang^{1,2,3}, Ian Black^{1,2}, John Glushka¹, Breeanna Urbanowicz^{1,2,3},
Christian Heiss^{1,2}, Parastoo Azadi^{1,2}

¹Complex Carbohydrate Research Center, ²DOE Center for Plant and Microbial Complex Carbohydrates, ³Department of Biochemistry and Molecular Biology, University of Georgia, 315 Riverbend Road, Athens, GA 30602

* Corresponding author. Tel: 1-706-542-4468, fax: 1-706-542-4412.

Email address: tan@ccrc.uga.edu

Abstract

To characterize a purified rhamnogalacturonan-I (RG-I) containing both RG-I and arabinogalactan-protein (AGP) types of glycosyl residues, an AGP-specific β -1,3-galactanase that can cleave the AG backbone and release the AG sidechain was applied to this material. Carbohydrate analysis and NMR spectroscopy verified that the galactanase-released carbohydrate consists of RG-I covalently attached to the AG sidechain, proving a covalent linkage between RG-I and AGP. Size exclusion chromatography-multiangle light scattering-refractive index detection revealed that the galactanase-released RG-I has an average molecular weight of 41.6 kDa, which, together with the percentage of pectic sugars suggests an RG-I-AGP comprising one AGP covalently linked to two RG-I glycans. Carbohydrate analysis and NMR results of the RG-I-AGP, the galactanase-released glycans, and the RG lyase-released glycans demonstrated that the attached RG-I glycans are decorated with α -1,5-arabinan, β -1,4-galactan, xylose, and 4-*O*-Me-xylose sidechains. Our measurement suggests that the covalently linked RG-I-AGP is the major component of the traditionally prepared RG-I.

Keywords: RG-I-AGPs, RG-I, AGPs, covalent linkage, RG-I size, NMR

1. Introduction

Pectins and arabinogalactan-proteins (AGPs) are the two major plant cell wall glycopolymers rich in uronic acids (UA). These acidic UA residues, including galacturonic acid (GalA) in pectins or glucuronic acid (GlcA) in AGPs, introduce negative charges to these macromolecules at physiological conditions, affecting their physiochemical properties and corresponding biological functions (Voragen *et al.*, 2009, Lamport and Varnai, 2013). This is exemplified by the Ca^{2+} binding capacity of these two types of polymers and subsequent intramolecular and intermolecular formation of calcium bridges that play important roles in wall integrity and cell-cell adhesion (Lopez-Hernandez *et al.*, 2020, Pfeifer *et al.*, 2020, Hyodo *et al.*, 2013).

AGPs are distinctive glycoproteins that have large and highly branched polysaccharides attached to hydroxyproline (Hyp) and thus are often referred to as hydroxyproline (Hyp)-rich glycoproteins (HRGPs). The Hyp-O-linked arabinogalactans have a β -1,3-linked galactan backbone that is 1,6-substituted with β -galactosyl or β -1,6-galactan sidechains, and these sidechains are further decorated with arabinose (Ara), GlcA, rhamnose (Rha), and fucose (Fuc) (Ellis *et al.*, 2010, Tan *et al.*, 2012, Classen *et al.*, 2019). This type of polysaccharide is referred to as type II arabinogalactan (AG) and is structurally distinct from type I and type III arabinogalactans, which have a β -1,4-linked or β -1,6-linked galactan backbone, respectively. The stable linkage between Hyp and arabinogalactan not only makes it difficult to release and analyze an intact arabinogalactan polysaccharide but also impedes efforts to sequence the polypeptides to identify the corresponding genes for further functional analyses.

Pectins have three different structural elements: homogalacturonan (HG), rhamnogalacturonan-I (RG-I), and rhamnogalacturonan-II (RG-II) (Mohnen, 2008, Ropartz and Ralet, 2020, Kaczmarzka, *et al.*, 2022). HG is composed of a linear α -1,4-linked GalA chain, while RG-II comprises an HG backbone decorated with six sidechains, referred to as sidechains A to F (Barnes *et al.*, 2021). In contrast, the backbone of RG-I consists of repeats of $[-2-\alpha\text{-L-Rhap}-(1\rightarrow4)-\alpha\text{-D-GalpA-1}]$ -disaccharides, substituted with different sidechains, including α -1,5-arabinan, β -1,4-galactan, and both type-I and type-II arabinogalactan (Harholt *et al.*, 2010, Nascimento *et al.*, 2017). Some pectins are weakly associated with the cell wall and can be easily

32 extracted with oxalate and carbonate solutions. However, other pectins are tightly bound
33 to the wall and can only be released from the wall with *endo*-polygalacturonase (EPG)
34 treatment, yielding RG-I, RG-II, and oligogalacturonides (uronides) that can be separated
35 with size-exclusion chromatography (SEC) (York *et al.*, 1986). Importantly, RG-II
36 always requires EPG treatment and cannot simply be extracted.

37 The existence of a pectin-AGP complex in the plant cell wall has been proposed
38 based upon numerous reports of co-purification of these macromolecules (Keegstra *et al.*,
39 1973, Yamada *et al.*, 1987, Pellerin *et al.*, 1995, Oosterveld *et al.*, 2002, McKenna *et al.*,
40 2006, Leivas *et al.*, 2016, Makarova and Shakhmatov, 2021). Previously, we provided
41 direct proof confirming that a covalent linkage is present between the reducing end GalA
42 of RG-I and the terminal rhamnosyl (Rha) residue of an AGP sidechain, -4- α -D-GalpA-
43 (1 $\rightarrow$ 2)- α -L-Rhap-, in a covalently linked arabinoxylan-pectin-arabinogalactan-protein 1
44 (APAP1) conjugate isolated from Arabidopsis suspension culture media (Tan *et al.*,
45 2013). The non-reducing-end of RG-I is further extended with repeats of HG and RG-I
46 oligomers, with the RG-I portion and AGP sidechain Ara residues substituted with
47 arabinoxylan. While we confirmed the covalent linkage between RG-I and AGP in
48 APAP1, which is secreted from the cell wall to the culture media, whether such a
49 structure is also present in the walls themselves is still an open question. To confirm this
50 hypothesis, an RG-I-AGP fraction was isolated from alcohol insoluble residues (AIR) of
51 Arabidopsis suspension cultured cells (Tan *et al.*, 2013) and subjected to extensive
52 structural analysis. The present work reveals that APAP1-like structures do exist in plant
53 cell walls and describes the detailed structure of the RG-I-AGP. Using isolation,
54 composition, linkage, and NMR analyses of the glycans released from the RG-I-AGP
55 fraction by AGP-specific β -1,3-galactanase and by RG lyase, we show that RG-I is
56 indeed linked to AGP. Quantitative analysis showed that this covalently linked RG-I-
57 AGP is the major component of the RG-I preparation released from the AIR by EPG. In
58 addition, the size of the RG-I linked to the AGPs proved to be a polysaccharide of around
59 41 KDa. This molecular weight and the calculated percentage of pectic sugars suggest an
60 RG-I-AGP is composed of one AGP covalently linked to two RG-I glycans.

62 2. Materials and Methods

2.1. Preparation of RG-I enriched fraction

Suspension cultured wild type (WT) Arabidopsis cells were harvested using a sintered glass Buchner funnel. Alcohol-insoluble residues (AIR) were prepared from the cells as described (York *et al.*, 1986). The AIR sample was resuspended in cold water, titrated to pH 12.0 with 1 M NaOH on ice, and de-esterification was carried out overnight at 4 °C, followed by titrating to pH 5.2 with 1 M HCl. De-esterified AIR was dialyzed and lyophilized. An RG-I enriched fraction was prepared by treating de-esterified AIR with *endo*-polygalacturonase (EPG) (E-PGALPC, Megazyme) in 50 mM sodium acetate buffer (pH 5.2). The solubilized materials were filtered from the residues, dialyzed and lyophilized. The dried pectic materials were fractionated on a Superdex 75 column (10 × 300 mm, GE Healthcare Biosciences) eluted with 50 mM ammonium acetate buffer, which yielded three uronic acid-containing fractions. The three fractions have traditionally been referred to as RG-I, RG-II and uronides in order of decreasing molecular sizes (York *et al.*, 1986, Barnes *et al.*, 2021). Fractions were collected and lyophilized.

2.2. Purification of RG-I-AGPs from the RG-I preparation

The above RG-I preparation was dissolved in 50 mM Tris-HCl pH 7.0 and loaded on a 100-ml DEAE (GE Healthcare Biosciences) anion exchange column, which was washed with 100 ml 50 mM Tris-HCl pH 7.0 and eluted by a linear gradient from 50 mM Tris-HCl pH 7.0 to 2 M NaCl in 120 min. Eluted fractions were pooled, dialyzed against water for 2 days, and lyophilized. These materials were further purified on a PRP-1 reverse phase column (5 µm, C-18, 4.1 × 150 mm, Hamilton) using a gradient from 0.1% TFA in H₂O to 80% (v/v) acetonitrile in 0.1% TFA in 100 min. The major fraction at 20 min was collected and lyophilized for further analysis.

2.3. Glycosyl composition and linkage analysis

Glycosyl composition analysis was performed by gas chromatography-mass spectrometry (GC-MS) of the trimethylsilyl (TMS) derivatives of the monosaccharide methyl glycosides generated from the samples by HCl methanolysis as described

previously (Santander *et al.*, 2013). In brief, each sample was heated in 1 M methanolic HCl for 18 h at 80 °C. The hydrolysate was dried, and the sample was derivatized with Tri-Sil™ HTP Reagent (Thermo Scientific) at 80 °C for 30 min. Following extraction with hexane, GC-MS analysis of the TMS methyl glycosides was performed on an Agilent 7890A GC interfaced to a 5975C MSD, using a Supelco Equity-1 fused silica capillary column (30 m × 0.25 mm ID). GC temperature program: 80 °C 2 min; 80 to 140 °C at 20 °C/min, hold for 2 min; 140 to 200 °C at 2 °C/min; 200 to 250 °C at 30 °C/min, hold for 5 min. Split ratio: 100 to 1. MS source temperature: 230 °C; Quadrupole temperature: 150 °C. Mass range: 50 – 550 Da.

Glycosyl linkage analysis was performed by GC-MS of the partially methylated alditol acetate (PMAA) derivatives of the sample (Willis *et al.*, 2013). Briefly, each sample was suspended in DMSO and methylated using dimethyl potassium and iodomethane (MeI) (Phillips and Fraser, 1981). The carboxylic acid methyl esters were reduced using lithium aluminum deuteride in THF (80 °C, 8 h). The sample was re-methylated using MeI and hydrolyzed using 2 M TFA for 2 h at 120 °C. The hydrolysate was reduced with NaBD₄, followed by acetylation using acetic anhydride and TFA. The resulting PMAAs were analyzed on an Agilent 7890A GC interfaced to a 5975C MSD; separation was performed on a 30 m Supelco SP-2331 bonded phase fused silica capillary column. GC temperature program: 60 °C 1 min; 60 to 170 °C at 27.5 °C/min; 170 to 235 °C at 4 °C/min, hold for 2 min; 235 to 240 °C at 3 °C/min, hold for 12 min. Split ratio: 100 to 1. MS parameters were the same as above.

2.4. RG lyase expression and purification

A pET28a plasmid containing the *BT4175* gene was a kind gift from Dr. Harry J. Gilbert (Luis *et al.*, 2018). *Escherichia coli* strain TUNER was transformed with the plasmid and grown to mid-exponential phase before induction with 0.1 mM isopropyl β-D-galactopyranoside (IPTG), and the culture was grown for overnight at 16 °C and 160 rpm. The recombinant proteins were purified by using HisPur™ cobalt resin (ThermoFisher) and dialyzed in HEPES sodium buffer (75 mM, pH 7.0). After dialysis, the RG lyase was concentrated to 0.5 mg/ml using Amicon Ultra centrifugal filter device

(MWCO: 30 kDa, EMD Millipore). Protein concentration was measured with a NanoDrop 1000 spectrophotometer (ThermoFisher).

2.5. Release of glycans from the RG-I-AGPs with β -1,3-galactanase or RG lyase

The RG-I-AGPs (2 mg) were dissolved in 400 μ l of 100 mM NaOAc, pH 5.2. Five microliters of *exo*- β -1,3-galactanase (*C. thermocellum*, Gift from Megazyme) were added to the sample solution and the mixture was incubated at 37 °C for 24 h on a rotary shaker at 90 rpm.

RG-I-AGPs (2 mg) were dissolved in 400 μ l of 100 mM HEPES, pH 7.0. Ten μ l of the RG lyase and 4.1 μ l of 1 M $MnCl_2$ were mixed with the sample. The digestion was carried on at 37 °C for 24 h on a rotary shaker at 90 rpm.

For each digestion, the released glycans voided on a PRP-1 C18 reverse-phase column and separated from the remaining RG-I-AGPs using the above conditions for RG-I-AGP isolation.

2.6. Assay of xylan synthesis

Enzyme reactions (20 μ l) consisted of 3 mM UDP-xylose (Carbosource) and 4 μ g xylan synthase *KfXYS1* (Jensen *et al.*, 2018) were proceeded in HEPES sodium buffer (50 mM, pH 7.0) at room temperature for 30 min. RG-I-AGP (10 μ g), APAP1 (10 μ g), xylopentaose (Megazyme, 50 μ M), and cellopentaose (Megazyme, 50 μ M) were used as the acceptor substrates. After 30 min, the production of UDP was measured by using UDP-Glo™ Glycosyltransferase Assay kit (Promega). Briefly, 5 μ l of each reaction was mixed with 5 μ l UDP detection reagent in a 384-well plate (Greiner Bio-one). After 1 h incubation at room temperature, the luminescence was read in Synergy LX Microplate Spectrometer (BioTek).

2.7. Size determination of glycans by size exclusion chromatography with multi-angle light scattering and refractive index detection (SEC-MALS-RI)

Glycans (0.8 mg) were dissolved in 400 μ l nanopure water and fractionated on an analytical Superdex 75 size exclusion column (10 $\times$ 300 mm, GE Healthcare Biosciences). The column was eluted with 50 mM ammonium formate at a flow rate of

0.45 ml/min and monitored with an RI detector (Agilent). The molecular weight of each fraction was estimated with a DAWN MALS detector (Wyatt technology). A series of Dextran standards, of average sizes 167, 40, 10, and 1 kDa, were also separated on the same system to confirm the glycan sizes. The SEC-purified galactanase-released or RG lyase-released glycans were lyophilized for NMR analysis.

2.8. Nuclear magnetic resonance (NMR) spectroscopy analysis

RG-I-AGPs (2 mg) or released glycans (500 µg) were dissolved in 150 µl of deuterium oxide. One microliter of 50 mM sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) in D₂O was added as internal standard. ¹H and ¹³C NMR spectra were obtained at 25 °C with a Varian Inova 600 equipped with a 3 mm cryoprobe at the Complex Carbohydrate Research Center. One- and two-dimensional NMR spectra were collected as described previously (Tan *et al.*, 2010).

3. Results

3.1. Structural analysis of the RG-I-AGPs derived from AIR

Our previous glycosyl composition and linkage analyses showed that the purified RG-I-AGPs released from AIR of Arabidopsis suspension cultured cells by EPG consisted of 64.8 mol% Ara, 11.4 mol% GalA, 10.4 mol% Rha, and 13.4 mol% Gal, and trace amount of Xyl, which were mostly pectic and AGP types of sugar residues (Tan *et al.*, 2013). To fully characterize their structures and understand their function in the cell wall, NMR experiments were carried out analyzing the structural elements of the purified RG-I-AGPs (Figure 1 and Table A.1). Chemical shift assignments were based on a heteronuclear single quantum coherence (HSQC) spectrum and published literature reports (Zheng & Mort, 2008, Sun *et al.*, 2010, Tan *et al.*, 2004, 2010, 2013). Characteristic signals are summarized as follows: The ¹³C/¹H chemical shifts at 99.0/5.20 and 97.7/5.03 ppm represent the anomeric C/H of α-2-Rhap and α-GalpA residues, respectively. In addition, the ¹³C/¹H signals at 77.7/4.41 (GalpA-C/H4) and 70.4/4.88 ppm (GalpA-C/H5) corroborate that GalpA is 4-substituted, while signals at 77.5/4.14 (Rhap-C/H2) & 72.8/3.35 (Rhap-C/H4) ppm, and at 76.6/4.11 (Rhap-C/H2) & 81.5/3.84

(Rhap-C/H4) ppm substantiate the presence of 2-Rhap and 2,4-Rhap residues, respectively. These results confirm the existence of branched RG-I in this RG-I-AGP preparation. The intense $^{13}\text{C}/^1\text{H}$ signals C/H5 of t-Araf and 3-Araf at 61.0/3.81, 3.69 ppm and C/H5 of (2, and/or 3),5-Araf signals at 66.5/3.86, 3.77 ppm, as well as the presence of 3,5-Araf, 2,5-Araf, & 2,3,5-Araf in the glycosyl linkage data, suggest that the RG-I is substituted with branched 5- α -arabinan sidechains.

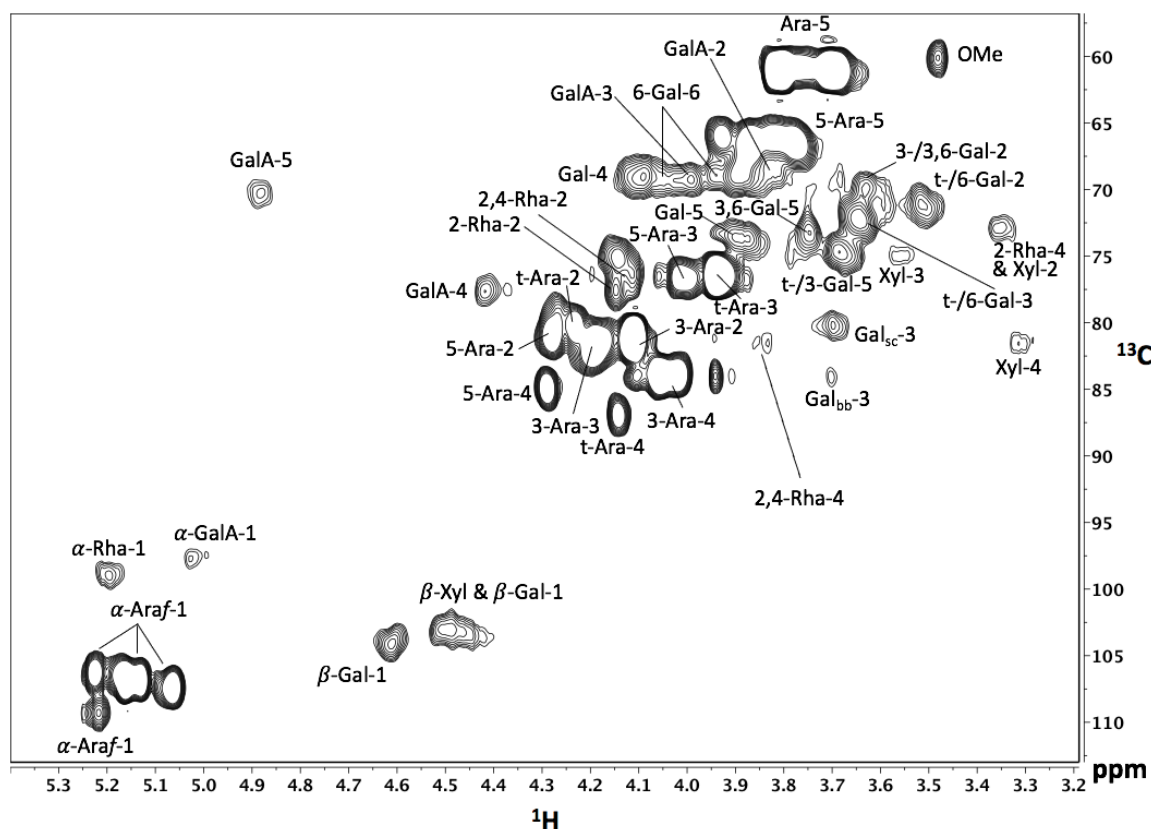


Figure 1. HSQC NMR spectrum of the intact RG-I-AGPs.

The chemical shifts confirmed the presence of RG-I and AGP types of structural elements. The spectrum was collected at 25 °C and referenced to DSS internal standard.

The distinctive $^{13}\text{C}/^1\text{H}$ chemical shifts at 84.2/3.70 ppm and 80.3/3.69, representing C/H3 of the branching 3,6-Galp residues on the 1,3-galactan backbone (Gal_{bb}) and on the 1,3-galactosyl sidechains (Gal_{sc}), respectively, confirm the presence of type-II AG (Tan *et al.*, 2004, 2010). In addition, the unique $^{13}\text{C}/^1\text{H}$ signals of Xylp residues at 81.6/3.31 ppm suggest that some of the Xylp residues are 4-O-methylated (Tan *et al.*, 2013). NMR analysis of the sample verified the presence of RG-I and type II

AG structural elements in the RG-I-AGP preparation, consistent with the structure depicted in Figure 2. However, direct evidence of the linkages between the RG-I and AG elements is still lacking since heteronuclear multiple bond correlation (HMBC) spectrum of polysaccharides as large as the RG-I-AGPs suffers from insufficient resolution and sensitivity. Data resolving these linkages are described further below.

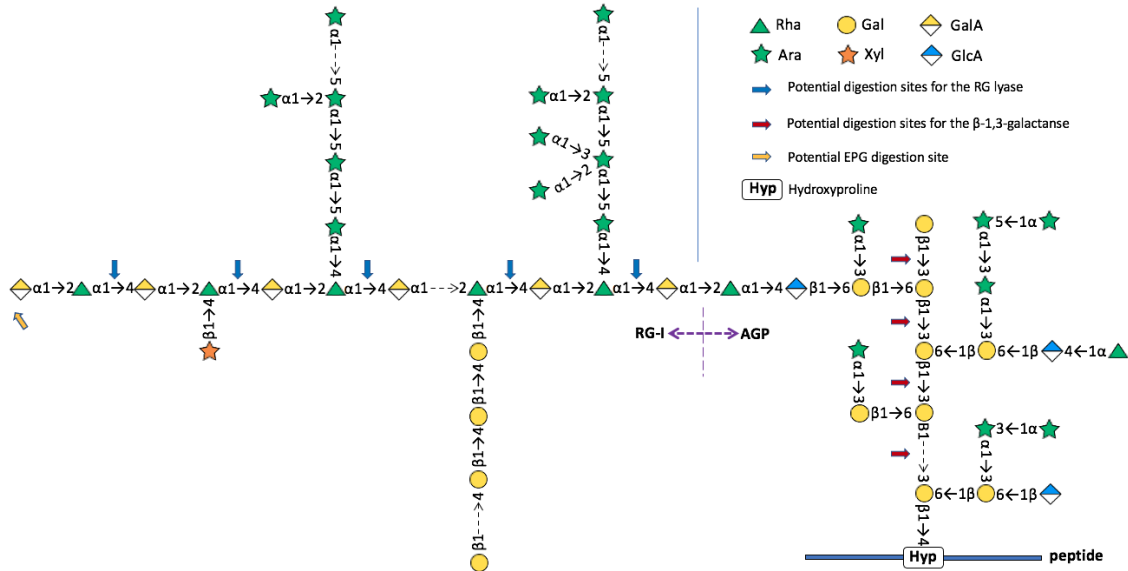


Figure 2. Scheme of the RG-I-AGPs isolated from AIR of Arabidopsis suspension cultured cells.

The β -1,3-galactanase can release a whole RG-I glycan with the AG sidechain residues, while the RG lyase can only produce RG-I fragments from the RG-I glycan.

3.2. Structure of glycans released by an *exo*- β -1,3-galactanase

To confirm the linkage between the RG-I and AGP in the RG-I-AGP preparation, an AGP-specific *exo*- β -1,3-galactanase was used since this enzyme can sequentially liberate galactosyl residues (and the attached sidechains) from the non-reducing end of the AGP β -1,3-galactan backbone (Figure 2) (Tsumuraya *et al.*, 1990, Tryfona *et al.*, 2010, 2012). If the RG-I is covalently linked to AGP, digestion with this galactanase should generate fragments consisting of a reducing-end Gal (produced from the galactan backbone) and the attached AG sidechain with the covalently linked RG-I (see Figure 2).

Our analysis proved that such a fragment structure in fact exists in the galactanase-generated glycans.

Specifically, after enzyme treatment, the released glycans were separated from the residual AGP conjugates by reverse phase chromatography (Figure A.1C). Glycosyl linkage analysis of the galactanase-released glycans gave significant percentages of partially methylated alditol acetates (PMAAs) corresponding to 2-linked and 2,4-linked Rhap and 4-linked GalpA (Table 1), which are characteristic residues found in RG-I, in addition to the PMAAs coming from 3-Galp, 6-Galp, and 3,6-Galp, as well as the different AraF linkages, which are all typical AG residues. This confirms the existence of a highly branched RG-I structure in these β -1,3-galactanase-released glycans. This result supports the hypothesis that RG-I is covalently linked to AG in the RG-I-AGP preparation, as the AGP-specific enzyme can liberate RG-I from the preparation.

NMR analysis of the released glycans confirms these linkage results (Figure 3 and Table A.2). A comprehensive chemical shift assignment was achieved based on a suite of two-dimensional (2D) NMR experiments summarized as follows: The proton resonances of each glycosyl residue were identified in a correlated spectroscopy (COSY) spectrum (Figure 3A) and a total correlation spectroscopy (TOCSY), and the corresponding carbon resonances were determined in an HSQC spectrum (Figure 3B). The glycosidic linkages of these glycosides, described in the following paragraph, were determined using an HMBC spectrum (Figure 3C). Specifically, like in intact RG-I-AGPs, the $^{13}\text{C}/^1\text{H}$ signals at 98.8/5.20 ppm [2,(4)- α -Rhap-C/H1] and 97.6-97.7/5.01-5.03 (α -GalpA-C/H1) ppm confirm the presence of 2,(4)- α -Rhap and α -GalpA (Figure 3B), respectively. Moreover, chemical shifts at 77.7/4.42 ppm (α -GalpA-C/H4) in the HSQC spectrum suggest α -GalpA residues are 4-linked, while those at 76.4/4.1 ppm (α -Rhap-C/H2) and 80.1/3.71 ppm (α -Rhap-C/H4) verify that some α -Rhap residues are 2,4-linked. These results are consistent with the linkage analysis, which showed that the RG-I glycans released by the AGP-specific galactanase are highly branched. H-H correlation between 4.53 and 3.63 ppm in Figure 3A also confirms the presence of β -GlcA, which was absent in the parent RG-I-AGP because of its low abundance in the whole large molecule (Figure 1). The corresponding anomeric C signal (103.0 ppm) is identified in the HSQC spectrum (Figure 3B). The chemical shifts at 92.3/5.25 & 96.3/4.58 ppm (Galp-C/H1) represent the α - and

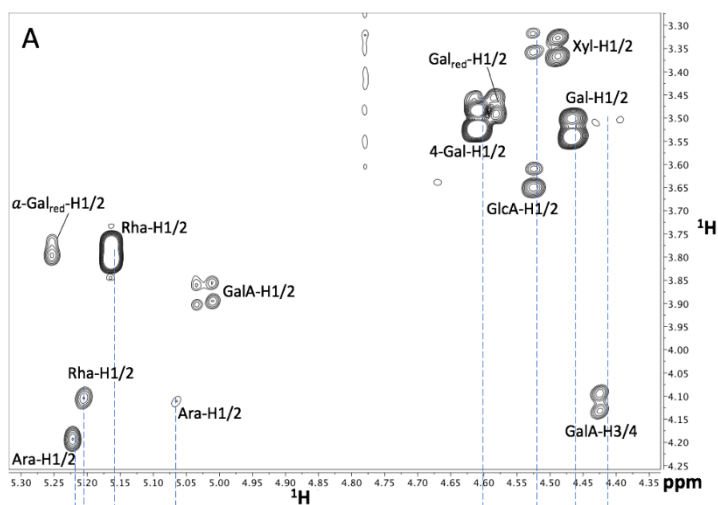
β -anomers of the reducing end Galp residue (Figure 3A), produced by the 1,3-galactanase cleavage.

Table 1. Glycosyl linkage analyses of β -1,3-galactanase and RG lyase released glycans from the RG-I-AGPs.

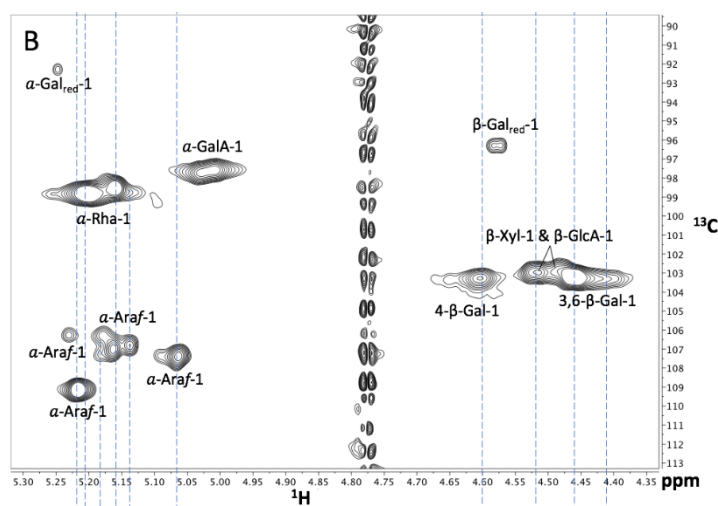
After digestion, the oligosaccharides were desalted on a LH20 column and separated from the rest of glycoproteins on a reverse-phase column as a void fraction.

Glycans released with β -1,3-galactanase		Glycans released with RG lyase	
Glycosyl Residue	Mol%	Glycosyl Residue	Mol%
		t-Rhap	0.8
2-Rhap	3.1	2-Rhap	10.0
4-Rhap	3.5	4-Rhap	0.7
2,4-Rhap	4.9	2,4-Rhap	3.8
		2,3,4-Rhap	24.3
		t-GalpA	0.8
4-GalpA	13	4-GalpA	13.2
t-Araf	11.5	t-Araf	1.4
3-Araf	1.0		
5-Araf	5.5	5-Araf	0.9
3,5-Araf	0.9		
2,5-Araf	2.3		
2,3,5-Araf	1.3		
		2,3,5-Araf	10.4
t-Galp	8.7	t-Galp	3.3
3-Galp	7.5		
6-Galp	13.8		
3,6-Galp	13.9		
4,6-Galp	1.2		
3,4,6-Galp	1.5		
		2,3,4,6-Galp	20.9
4-GlcpA	2.9		
3,4-GlcpA	1.6		
t-Xylp	1.4		
		4-Xylp	0.9
		2,3,4-Xylp	1.2
2-Manp	0.6		
		t-Glcp	0.9
		4-Glcp	4.3
		2,3,4,6-Glcp	2.2

263



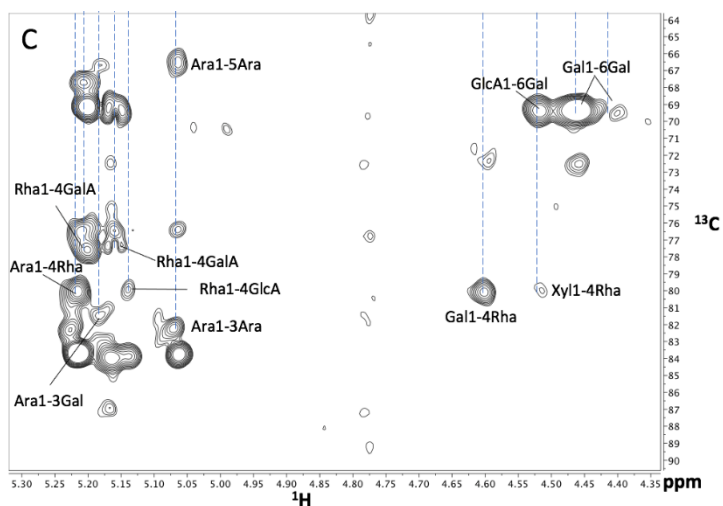
264



265

266 Figure 3. NMR spectra of glycans released from the RG-I-AGPs with the β -1,3-

267 galactanase.



A. COSY spectrum, B. HSQC spectrum, C. HMBC spectrum. The structural elements in the spectra confirm the released glycans are RG-I with different sidechains. The spectra were collected at 25 °C and referenced to DSS internal standard.

The HMBC spectrum provided linkage information between different structural elements in the AGP-specific galactanase-released glycans (Figure 3C). The β -3,6-Gal residues with anomeric C/H at 103.0/4.4 ppm are from the sidechain β -Galp residues substituted at position-3 by α -Araf residues, as evidenced by the correlation between α -Araf-H1 (5.19 ppm) and β -Galp-C3 (81.8 ppm). Correlations between sidechain β -Galp-H1 (4.46 or 4.39 ppm) and β -Galp-C6 at 69.3 ppm support the β -Galp-(1 $\rightarrow$ 6)- β -Galp linkage, while those between β -GlcA-H1 (4.53 ppm) and β -Galp-C6 (69.3 ppm) confirm the β -GlcA-(1 $\rightarrow$ 6)- β -Galp substitution. These structural elements are consistent with a structural element β -GlcA-(1 $\rightarrow$ 6)-[α -Araf-(1 $\rightarrow$ 3)]- β -Galp-(1 $\rightarrow$ 6)-Gal_{reducing} produced from typical AGPs by the galactanase (Figure 2) (Tan *et al.*, 2014, 2010, Tryfona *et al.*, 2010).

The linkages within the RG-I structure are confirmed by the correlations in the HMBC spectrum (Figure 3B). The correlations of α -Rhap-H1 (5.20 or 5.16 ppm) to α -GalpA-C4 (77.7 ppm) support the α -Rhap-(1 $\rightarrow$ 4)- α -GalpA linkages. The arabinan is corroborated by the correlations between α -Araf-H1 (5.06 ppm) & α -Araf-C5 (67.8 ppm) and α -Araf-H1 (5.06 ppm) & α -Araf-C3 (83.7 ppm). The RG-I sidechain substitutions are supported by correlations between α -Araf-H1 (5.22 ppm) & α -Rhap-C4 (80.1 ppm), β -Galp-H1 (4.61 ppm) & α -Rhap-C4 (80.1 ppm), and β -Xylp-H1 (4.53 ppm) & α -Rhap-C4 (80.1 ppm), which establish the attachment of 1,5-arabinan, 1,4-galactan, and t-4-O-Me- β -Xylp to the RG-I Rhap residues, respectively. We did not observe 4-linked Xylp in the NMR spectra and in the glycosyl linkage analysis (Table 1), and together with the HMBC correlation between Xyl H1 and Rha-C4, this suggests that the xylosyl sidechains consist of a single residue. These results affirm that the RG-I portion of the RG-I-AGP preparation is highly decorated with different substituents.

In general, 2-O-glycosylated α -Rhap residues show a chemical shift displacement relative to the α -Rhap anomeric C/H from 102/4.80 ppm (t- α -Rhap) to ~99/~5.1 ppm (2-

α -Rhap) (Tan *et al.*, 2004, 2010, 2013). Thus, the signal at 5.14/99.6 ppm is likely from a 2- α -Rhap. An HMBC correlation between its anomeric proton and a carbon signal at 79.9 ppm (C4 of β -Glc pA) demonstrates the presence of a 2- α -Rhap-(1 $\rightarrow$ 4)- β -Glc pA linkage, which further supports that the AGP sidechain Rha residues were 2-O-substituted. Considering that the glycans, with RG-I and the -2- α -Rhap-(1 $\rightarrow$ 4)- β -Glc pA- structural elements, were released from the RG-I-AGP complex by the galactanase and enriched by SEC separation (Figure 4A), the results support the attachment of RG-I to the O2 of the sidechain Rhap of AGP (Figure 3C) (Tan *et al.*, 2013), further confirming our hypothesis that RG-I is covalently linked to AGP.

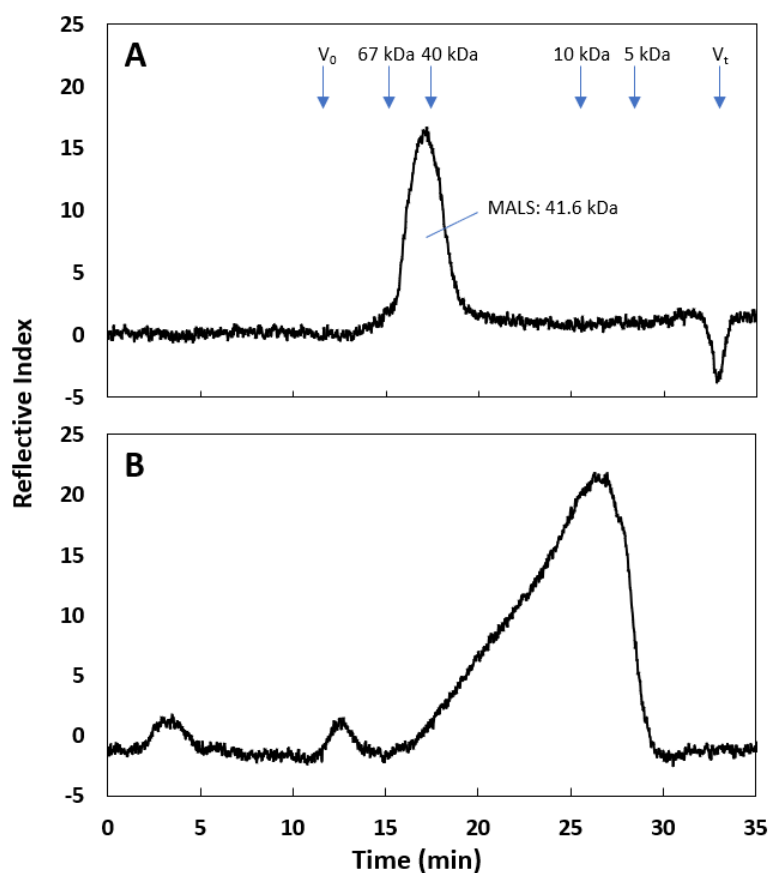


Figure 4. HPLC chromatograms of glycans released from the RG-I-AGPs on SEC-MALS-RI. Glycans released with β -1,3-galactanase (A) and with RG lyase (B) were separated on a Superdex 75 size-exclusion column and monitored with MALS-RI detectors. The

molecular weights of all dextran standards were labeled at the corresponding retention times in panel A.

3.3. The structure of glycans released by rhamnogalacturonan lyase

To obtain more structural information of the RG-I portion of the complex, the RG-I-AGP preparation was treated with rhamnogalacturonan (RG) lyase BT4175 (Luis *et al.*, 2018), and the released glycans were separated from the remaining RG-I-AGPs using a C18-reverse phase column (Figure A.1D). Glycosyl linkage analysis of these released glycans, which were eluted from the column in the void volume, showed the presence of typical RG-I type glycosyl linkages (Table 1), although certain glycosyl residues were under-methylated. Further NMR analysis was carried out to secure the RG-I structure of these glycans (Figure 5 and Table A.3). The chemical shift assignment was based on a total correlation spectroscopy (TOCSY) spectrum and an HSQC spectrum, in addition to the NMR results in above section 3.2, and assignments from the literature (Mutter *et al.*, 1998, Deng *et al.*, 2006). In contrast to the galactanase digest, no residues originating from AGP were detected. The HSQC spectrum confirmed the RG-I type glycosyl residues in the sample, including the backbone α -2-Rhap, α -2,4-Rhap, and α -4-GalpA residues, as well as the α -t-Araf, α -3-Araf, and α -5-Araf residues, β -4-Gal residues, and β -t-Xyl residues on the sidechains. The characteristic $^{13}\text{C}/^1\text{H}$ chemical shifts at 107.5/5.79 ($^{4,5}\Delta$ -GalpA-C/H4) and 65.8/4.31 ($^{4,5}\Delta$ -GalpA-C/H3) ppm verified the non-reducing ends of the released glycans are $^{4,5}\Delta$ -GalpA (Mutter *et al.*, 1998, Deng *et al.*, 2006), generated by the RG lyase. No AGP linkages such as β -3-Galp, β -3,6-Galp, and β -4-GlcpA were identified in this sample.

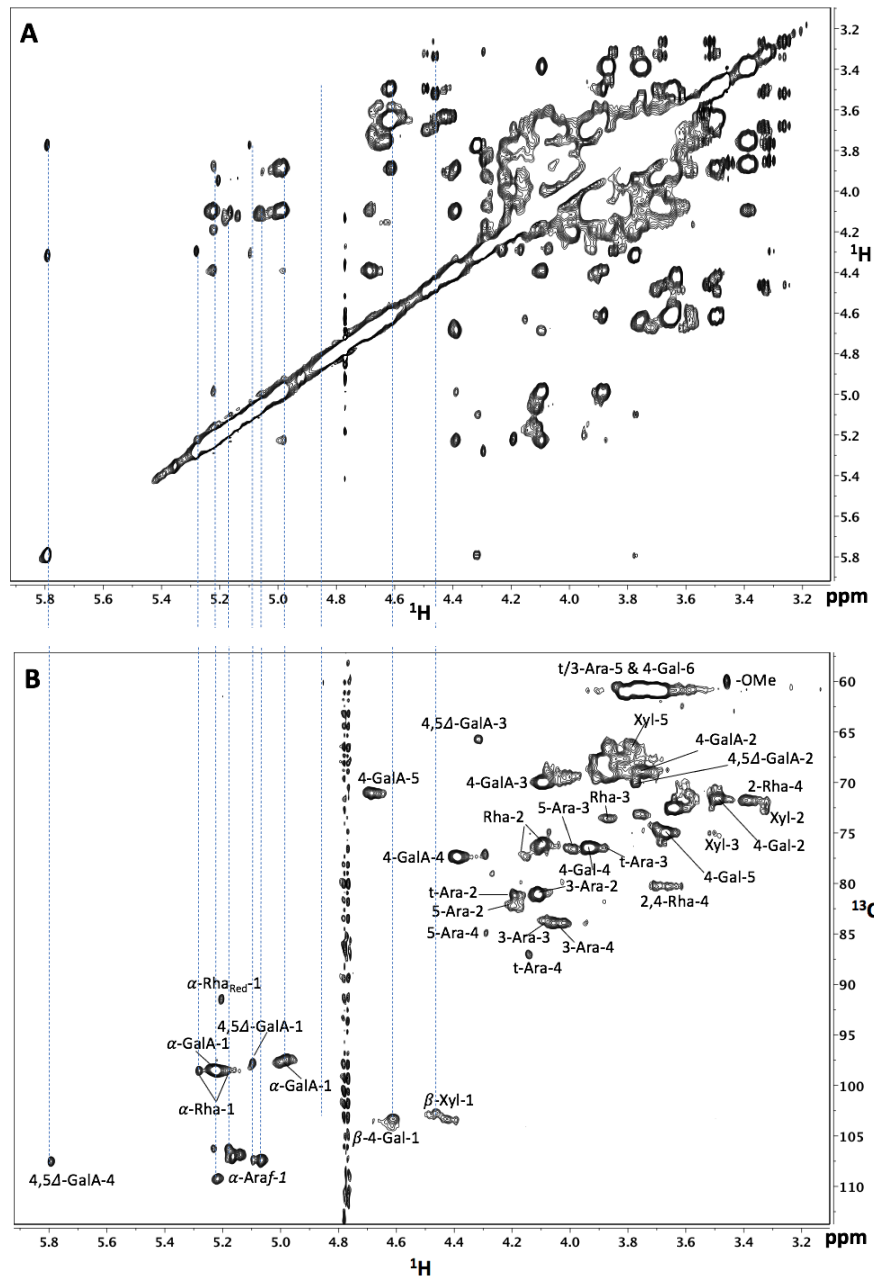


Figure 5. NMR spectra of glycans released from the RG-I-AGPs with the RG lyase. A. TOCSY spectrum, B. HSQC spectrum. The chemical shifts confirm the enzyme released glycans are RG-I fragments with different sidechains. The spectra were collected at 25 °C and referenced to DSS internal standard.

3.4. SEC-MALS-RI reveals size of the RG-I glycans appended to the RG-I-AGPs

The molecular weight of the RG-I glycans attached to the AGPs was determined by SEC-MALS-RI. Specifically, Superdex 75 SEC-MALS-RI was used to measure the

molecular weight using a MALS detector directly and an RI detector (Figure 4A) with dextran standards. The released pectic glycans showed a short retention time (16 to 18.5 min, centered at 17.6 min) on the SEC column. The MALS detector measured the average molecular weight as 41.6 kDa, while comparison to the dextran standards suggested the molecular weight is close to 40 kDa in good agreement with the MALS measurement. Thus, the average molecular weight of the RG-I glycans in the RG-I-AGP complexes is about 41 kDa.

3.5. Measuring the size of glycans released by a rhamnogalacturonan lyase

The branching Rha residues, 2,4-Rha, account for 42.6% of the total Rha in the galactanase-released RG-I glycans (Table 1), indicating the RG-I is highly branched. Despite this high degree of branching, one treatment with the RG lyase can remove RG-I glycans corresponding to one third of the starting weight of RG-I-AGP, confirming that the RG lyase can cleave the RG-I backbone even in the presence of sidechains. To measure the molecular weight of RG-I glycans released by the RG lyase, we used the Superdex 75 SEC-MALS-RI system (Figure 4B). The SEC chromatogram showed a broad profile with retention time ranging from 16 to 29 min, with measured sizes ranging from 5 to 41 kDa by the MALS detector and RI dextran standards. The SEC profile also showed an asymmetric peak distribution with higher abundance of materials with lower molecular weight, correspondingly producing a range of short RG-I oligosaccharides to long RG-I glycans (41 kDa) that are close to the full-length RG-I glycans (41 kDa) released by the galactanase from the RG-I-AGPs. This result confirms the accuracy of the measured sizes of the RG-I in the RG-I-AGPs.

3.6. RG-I-AGP is the major component of the RG-I preparation

To measure how much of the traditional RG-I preparation (Figure A.1A) consisted of RG-I-AGP conjugates, we chose the following experimental approach (Figure 6): A one-step EPG treatment released about 245.8 mg [9.8% (w/w)] of materials from 2.5 g of the AIR residues, in which 95.9 mg of the materials [39% (w/w)] became insoluble after dialysis and lyophilization. After separation of the soluble portion (149.9 mg) of the EPG-released materials on a Superdex 75 size exclusion column (Fig A.1A),

the largest molecular weight fraction (retention time 16-20 min), which has been traditionally named RG-I, was collected and freeze-dried (105.7 mg). The yield of traditional RG-I prep from the AIR was 4.2% (w/w).

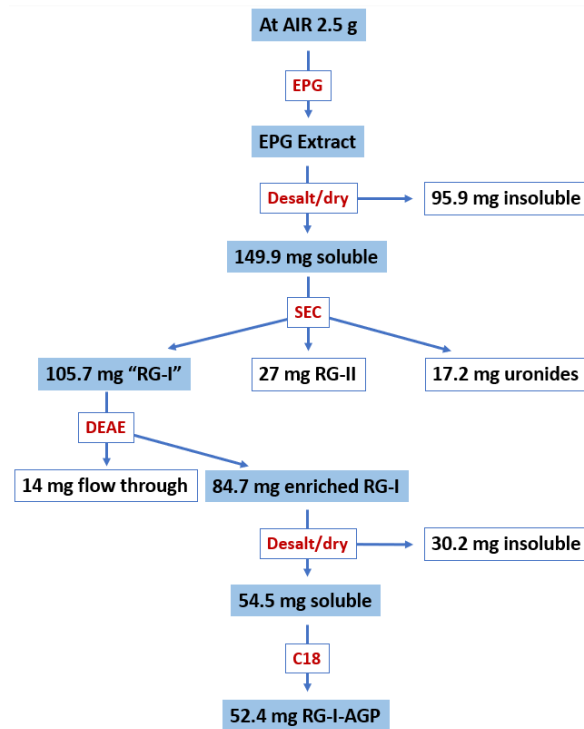


Figure 6. Yields of materials during the steps of RG-I-AGPs purification.

At AIR: alcohol insoluble residues prepared from Arabidopsis suspension cultured cells;
SEC: size-exclusion chromatography.

This traditional RG-I prep was further purified on a DEAE anion exchange column to remove neutral glycan components (Fig A.1B) and reverse-phase column (Tan *et al.*, 2013) to enrich the RG-I-AGPs. The results showed that the DEAE purification enriched 84.7 mg [80.1% (w/w)] of RG-I prep from the 105.7 mg of traditional RG-I prep (Figure 6), in which 30.2 mg of materials, accounting for 36% w/w of the enriched RG-I prep, became insoluble after lyophilization. Further reverse-phase purification yielded 52.4 mg of the final RG-I-AGP, which accounted for 96% (w/w) of the soluble DEAE-purified RG-I prep (54.5 mg), 69.4% (w/w) of the soluble traditional RG-I [75.5 mg (105.7-30.2 mg)], and 2.1% (w/w) of the AIR. Therefore, the traditional RG-I prep (105.7

mg) from the AIR of Arabidopsis suspension cultured cells mainly comprises covalently linked RG-I and AGPs [52.4 mg (49.6%, w/w)].

4. Discussion

In this work, we showed that the EPG-released RG-I from the AIR preparation of cultured wild-type Arabidopsis cells consisted largely of RG-I that was covalently attached to AGP, i.e. RG-I-AGP. Structural evidence for this RG-I-AGP attachment was provided by the AGP-specific β -1,3 galactanase release of RG-I glycans and the identification of the linkage between RG-I and AGP sidechain Rha in the HMBC NMR spectrum of the galactanase-released glycans. Furthermore, it was demonstrated that RG-I-AGP was the major molecule present in the RG-I preparation, and that the RG-I polysaccharide was relatively large in size, i.e. 41 kDa.

The yield of the RG-I-AGPs (52.4 mg) from the soluble RG-I prep (54.5 mg) suggests that about 96% (w/w) of the soluble RG-I is covalently attached RG-I-AGPs, and the purified RG-I-AGPs (52.4 mg) account for 69.4% (w/w) of the soluble traditional RG-I preparation [75.5 mg (105.7-30.2 mg)] (Figure A.1A). Moreover, glycosyl composition analysis of the water-insoluble RG-I preparation (30.2 mg) after DEAE column showed that these materials are also mainly composed of RG-I and AGP types of glycosyl residues (Table A.4). However, the structure of this insoluble fraction still needs to be elucidated, and we have recently developed methods (Black *et al.*, 2019, Ndukwe *et al.*, 2020) for the structural characterization of insoluble polysaccharides that we plan to exploit for this purpose. Traditionally, we designate the first or the largest-sized SEC fraction of the EPG released cell wall materials as RG-I (Figure A.1A). Here, we characterized the majority of this traditional RG-I as covalently linked RG-I-AGP.

In recent years, many cell wall specific carbohydrate degrading enzymes have been discovered (Ndeh *et al.* 2017, Cartmell *et al.*, 2018, Ndeh and Gilbert 2018, McKee *et al.*, 2021). This provides powerful analytical tools for the selective cleavage of individual linkages, as well as structural information of targeted wall components. Dupree, Tsumuraya, and colleagues (Tryfona *et al.*, 2010, 2012, Shimoda *et al.*, 2014) successfully characterized some AGP structures by making use of different AGP-specific enzymes, including *exo*- β -1,3-galactanase, α -L-arabinofuranosidase, β -glucuronidase,

and *endo*- β -1,6-galactanase and by analyzing the products of these enzymes with gel electrophoresis and mass spectrometry (MS) techniques. These enzymes, especially the *exo*- β -1,3-galactanase that can notably trim the core AG structures, enable the analysis of AG sidechains of low abundances that cannot be detected by traditional MS and NMR analyses.

Here, the release of RG-I glycans from the RG-I-AGPs by the AGP-specific β -1,3-galactanase provided strong evidence supporting that RG-I is covalently attached to the sidechains of AGPs. The galactanase-catalyzed cleavage also significantly simplified the RG-I structural analysis by removing most of the AG portion of the RG-I-AGPs from the released RG-I glycans. This feature also makes it possible to apply this enzyme directly on cell walls or AIR to release RG-I or pectin as a complementary approach in place of EPG treatment. Both glycosyl composition and linkage analyses as well as NMR results demonstrate that the RG-I backbone is decorated with α -1,5-arabinan, β -1,4-galactan, t- β -Xyl, and t-4-*O*-Me- β -Xyl sidechains. The relative molar ratio of 2-Rha and 2,4-Rha (1:1.6) suggest that the Rha residues on the RG-I backbones are highly substituted with those sidechains. It is noteworthy that the released RG-I did not contain type II AG sidechains, raising the question whether the type II AG that has traditionally been viewed as RG-I sidechains in reality originates from the AG portion of RG-I-AGPs also in other plants. We will need to isolate and characterize RG-I-AGPs from several additional species to evaluate whether this is generally the case.

APAP1 and RG-I-AGPs are pectic AGPs with reducing ends of pectin and RG-I covalently linked to AGPs, respectively. In the case of APAP1, which was secreted into the culture media from the walls, the calculated pectin length is short with an average 4-GalA/4-GlcA ratio of $\sim$ 6:1, i.e. one pectin glycan has $\sim$ 6 GalA residues if all GlcA-containing AG sidechains are substituted with pectin glycans (Tan *et al.*, 2013). Here, the RG-I-AGPs were released from the AIR by EPG digestion, which would cleave the HG portions of pectin molecules and liberate the covalently attached RG-II, RG-I (Harholt *et al.*, 2010), and RG-I-AGP located at the reducing end of pectin/RG-I. The long RG-I glycans ($\sim$ 41 kDa) of the RG-I-AGP consisted of several hundred glycosyl residues. We found that the glycoconjugates of which the RG-I-AGPs are part, and which may also comprise the covalently linked HG and RG-II (and even covalently attached xylan chains

as presented in APAP1) cannot be extracted from the cell wall but require digestion by EPG. This indicates that the RG-I-AGPs in their context as part of a larger glycoconjugate are tightly associated in the cell wall. This feature fits our hypothesis that was stated in the previous paper; namely that “APAP1 with longer pectin and arabinoxylan chains may exist in the cell walls and that such macromolecules are highly crosslinked and interwoven in the walls and thus cannot be released into the culture media” (Tan *et al.*, 2013).

Arabinosylation is a major modification of both RG-I and AGP. The arabinan sidechain of RG-I is composed of a 1,5-linked α -L-arabinan backbone, which is usually further 2-O and/or 3-O substituted with α -L-arabinosyl residues (Mohnen, 2008, Luis *et al.*, 2018). AGP galactan arabinosylation includes formation of t- α -L-Araf-(1 $\rightarrow$ 5)- α -L-Araf-(1 $\rightarrow$ 3)- α -L-Araf-(1, α -L-Araf-(1 $\rightarrow$ 3)- α -L-Araf-(1, and t- α -L-Araf-(1 sidechains that are 3-O linked to AG sidechain Gal (Defaye and Wong, 1986, Gane *et al.*, 1995, Tan *et al.*, 2004, 2010, Knoch *et al.*, 2014). Therefore, 5-Araf, as well as 2,5-Araf and 3,5-Araf, are the main Ara linkages in RG-I, while 5-Araf is a minor Ara linkage in AGP and even absent in some AGPs (Tryfona *et al.*, 2010, Cartmell *et al.*, 2018, Ndeh and Gilbert, 2018, Munoz-Munoz *et al.*, 2021). In the case of the present RG-I-AGP, which is composed of 64.8 mol% Ara, 11.4 mol% GalA, 10.4 mol% Rha, and 13.4 mol% Gal (Tan *et al.*, 2013), 5-Araf in the AGP portion should be no more than the mol% of 3-Araf, which is 0.4%. Most 5-Araf [$>8.0\%$ (8.4-0.4 mol%)] is from the RG-I portion. With other types of RG-I Araf linkages such as 1.9% 3,5-Araf, 4.2% 2,5-Araf, & branching t-Araf [substituting 3-O (1.9%) & 2-O (4.2%) on the 3,5- & 2,5-Araf, respectively], at least 20.2% (8.0+1.9+4.2+1.9+4.2) out of all 35.3% Ara linkages is on RG-I, i.e. at least 57.2% (20.2/35.3) of Ara residues (37.1 mol%) are from RG-I sidechain arabinan (Tan *et al.*, 2013).

Almost all of the Rha (10.2 mol%) is pectic 2-, 2,3-, & 2,4-Rha (Tan *et al.*, 2013). In addition, based on a 3:1 relative ratio of anomeric signals of Araf:Galp in the HSQC spectrum of the RG lyase released RG-I glycans (Figure 5B), the 4-Galp attributes to about 10.8 mol% (one third of 32.4 mol% of RG-I arabinosyl residues) of total sugars in the RG-I-AGP. Thus, the RG-I accounts for at least 69.5 mol% (11.4% GalA + 10.2% Rha + 37.1% Ara + 10.8% Gal) of carbohydrates in the RG-I-AGP, and the other 30.5

mol% of total carbohydrates (27.7 mol% Ara + 2.6 mol% Gal + 0.2 mol% Rha) are from the AG portion. Considering the molecular weight of the RG-I-AGP (~120 kDa), the attached RG-I (≤ 41 kDa), and the matured polypeptide of the AGP (10.8 kDa), this suggests that one RG-I-AGP molecule consists of one molecule of AGP and two molecules of RG-I (the calculation is presented in Figure A.3), i.e. one AGP can covalently link two individual RG-I. It indicates that *in vivo* one AGP can crosslink two pectin glycans in the cell wall. Therefore, AGPs involved in such complex structures play an important role in crosslinking pectin glycans for a functional wall structure.

The attachment of RG-I to AGP raises an important question regarding the biosynthesis of such a complex structure. All identified glycosyltransferases (GTs) involved in pectin and AG biosynthesis are Golgi-localized. Classic AGPs, such as AtAGP57, are anchored to the Golgi membrane via a C-terminal GPI anchor during glycosylation. Since both GTs and AGPs are Golgi membrane-anchored, it is likely that RG-I-AGP synthesis occurs inside the Golgi. In this context, t-Rha on AGP would serve as the initiation site for addition of RG-I. This is accomplished by the Golgi-localized RG-I specific enzymes rhamnosyl transferase and galacturonic acid transferase (Figure A.2A).

According to our previous work (Tan *et al.*, 2013), APAP1, which is isolated from the culture medium, is decorated with xylan sidechains with an average degree of polymerization (DP) of 11. However, we have observed that the RG-I of RG-I-AGP from the cell wall characterized here has only monomeric Xyl and 4-O-methylated Xyl residues as sidechains. Some other reported pectin-AGP complexes also contained a small amount of Xyl (Saulnier & Brillouet, 1988, Ryden & Selvendran, 1990, Pellerin *et al.*, 1995, Oosterveld *et al.*, 2002, Duan *et al.*, 2004, Nergard *et al.*, 2005). The xylan DP differences between the reported pectic AGPs and the common presence of Xyl in the pectic AGPs raise the following questions: Are these minor Xyl moieties on RG-I the initiation sites for xylan elongation or addition? What is the function of 4-O-methylation of Xyl in RG-I-AGPs? Does it serve as the terminating signal for xylan elongation? Although xylan synthase and UDP-Xyl were unable to elongate the xylosyl sidechains on the RG-I-AGP (Figure A.4) *in vitro*, it is reasonable to expect that the xylan endotransglycosylase from plant cell wall (Johnston *et al.*, 2013, Frankova and Fry, 2020)

will be able to accomplish this chain elongation by adding xylan fragments hydrolyzed from wall xylan onto the 4-O position of the t-Xyl (Figure A.2B). This xylan elongation process on structures like RG-I-AGPs catalyzed by wall endotransglycosylation may allow the formation of crosslinks between pectin and xylan in the wall *in vivo*.

Undoubtedly with more carbohydrate-specific enzymes, such as enzymes that can cleave 4- β -GlcA with the directly linked RG-I from the RG-I-AGPs, and more powerful analytical probes and tools, such as an antibody that recognize the specific - α -GalpA-(1 $\rightarrow$ 2)- α -Rhap-(1 $\rightarrow$ 4)- β -GlcA-(1 $\rightarrow$ 6)- β -Galp- linkage between RG-I and AGPs, we will be able to track when and where RG-I and AGPs start to covalently link *in planta*. When we understand how these macromolecules are synthesized, the role of such complex structures in plant cell wall architecture will be revealed.

5. Conclusions

The RG-I preparation released from the AIR of Arabidopsis suspension cultured cells are mainly composed of covalently linked RG-I and AGPs. The RG-I portion is highly substituted with branched α -1,5-arabinan, β -1,4-galactan, and single β -Xyl or 4-O-Me- β -Xyl residues. The RG-I glycans have an average molecular size of 41 kDa, suggesting one AGP can crosslink two RG-I glycans.

CRedit Authorship Contribution Statement

LZ, BU, IB, JG, LT: Formal analysis; LT: conceptualization, investigation, original draft; LT, CH, & PA: writing, review, funding acquisition. All authors: revision.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgements

LT thanks Dr. Vincent McKie from Megazyme for kindly providing the β -1,3-galactanase from *C thermocellum*. The authors acknowledge Dr. Russell Carlson for critical review of this manuscript. This work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Grant number DE-SC0015662.

551

552 **References**

553

554 Barnes, W.J., Koj, S., Black, I.M., Archer-Hartmann, S.A., Azadi, P., Urbanowicz, B.R.,

555 Pena, M.J., O'Neill, M.A. (2021) Protocols for isolating and characterizing

556 polysaccharides from plant cell walls: a case study using rhamnogalacturonan-II.

557 Biotechnol. Biofuels. 14: 142.

558 Black, I., Heiss, C., Azadi, P. (2019) Comprehensive monosaccharide composition

559 analysis of insoluble polysaccharides by permethylation to produce methyl alditol

560 derivatives for gas chromatography/mass spectrometry. Anal. Chem. 91: 13787-13791.

561 Cartmell, A., Munoz-Munoz, J., Briggs, J.A., Ndeh, D.A., Lowe, E.C., Basle, A.,

562 Terrapon, N., Stott, K., Heunis, T., Gray, J., Yu, L., Dupree, P., Fernandes, P.Z., Shah, S.,

563 Williams, S.J., Labourel, A., Trost, M., Henrissat, B., Gilbert, H.J. (2018) A surface

564 endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for

565 arabinogalactan degradation. Nat. Microbiol. 3: 1314-1326.

566 Classen, B., Baumann, A., Utermohlen, J. (2019) Arabinogalactan-proteins in spore-

567 producing land plants. Carbohydr. Polym. 210: 215-224.

568 Defaye, J., & Wong, E. (1986) Structural studies of gum arabic, the exudate

569 polysaccharide from *Acacia senegal*. Carbohydr. Res. 150: 221-231.

570 Deng, C., O'Neill, M.A., York, W.S. (2006) Selective chemical depolymerization of

571 rhamnogalacturonans. Carbohydr. Res. 341: 474-484.

572 Duan, J., Zheng, Y., Dong, Q., Fang, J.N. (2004) Structural analysis of a pectic

573 polysaccharide from the leaves of *Diospyros kaki*. Phytochemistry 65: 609-615.

574 Ellis, M., Egelund, J., Schultz, C.J., Bacic, A. (2010) Arabinogalactan-proteins: key

575 regulators at the cell surface? Plant Physiol. 153: 403-419.

576 Frankova, L., & Fry, S.C. (2020) Activity and action of cell-wall transglycanases.

577 Methods Mol. Biol. 2149: 165-192.

578 Gane, A.M., Craik, D., Munro, S.L., Howlett, G.J., Clarke, A.E., Bacic, A. (1995)

579 Structural analysis of the carbohydrate moiety of arabinogalactan-proteins from stigma

580 and styles of *Nicotiana glauca*. Carbohydr. Res. 277: 67-85.

581 Harholt, J., Suttangkakul, A., Scheller, H.V. (2010) Biosynthesis of pectin. Plant Physiol.

582 153: 384-395.

583 Hyodo, H., Terao, A., Furukawa, J., Sokamoto, N., Yurimoto, H., Satoh, S., Iwai, H.
584 (2013) Tissue specific localization of pectin-Ca²⁺ cross-linkages and pectin methyl-
585 esterification during fruit ripening in tomato (*Solanum lycopersicum*) PLOS One 8:
586 e78949.

587 Johnston, S.L., Prakash, R., Chen, N.J., Kumagai, M.H., Turano, H.M., Cooney, J.M.,
588 Atkinson, R.G., Paull, R.E., Cheetamun, R., Bacic, A., Brummell, D.A., Schroder, R.
589 (2013) An enzyme activity capable of endotransglycosylation of heteroxylan
590 polysaccharides is present in plant primary cell walls. *Planta* 237: 173-187.

591 Kaczmarek, A., Pieczywek, P.M., Cybulska J., Zdunek, A. (2022) Structure and
592 functionality of rhamnogalacturonan I in the cell wall and in solution: a review.
593 *Carbohydr. Polym.* 278: 118909.

594 Keegstra, K., Talmadge, K.W., Bauer, W.D., Albersheim, P. (1973) The structure of
595 plant cell walls III. A model of the walls of suspension-cultured sycamore cells based on
596 the interconnections of the macromolecular components. *Plant Physiol.* 51: 188-197.

597 Knoch, E., Dilokpimol, A., Geshi, N. (2014) Arabinogalactan proteins: focus on
598 carbohydrate active enzymes. *Front. Plant Sci.* 5: 198.

599 Lamport, D.T.A., & Varnai, P. (2013) Periplasmic arabinogalactan glycoproteins act as
600 calcium capacitor that regulates plant growth and development. *New Phytol.* 197: 58-64.

601 Leivas, C.L., Iacomini M., Cordeiro L.M.C. (2016) Pectic type II arabinogalactans from
602 starfruit (*Averrhoa carambola* L.). *Food Chem.* 199: 252-257.

603 Lemaire, A., Garzon, C.D., Perrin, A., Habrylo, O., Trezel, P., Bassard, S., Lefebvre, V.,
604 Wuytswinkel, O.V., Guillaume, A., Pau-Roblot, C., Pelloux, J. (2020) Three novel
605 rhamnogalacturonan I-pectins degrading enzymes from *Aspergillus aculeatinus*:
606 biochemical characterization and application potential. *Carbohydr. Polym.* 248: 116752.

607 Lopez-Hernandez, F., Tryfona, T., Rizza, A., Yu, X.L., Harris, M.O.B., Webb, A.A.R.,
608 Kotake, T., Dupree, P. (2020) Calcium binding by arabinogalactan polysaccharides is
609 important for normal plant development. *Plant Cell* 32: 3346-3369.

610 Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., Basle, A., Cartmell,
611 A., Terrapon, N., Scott, K., Lowel, E.C., McLean, R., Shearer, K., Schuckel, J., Venditto,
612 I., Ralet, M., Henrissat, B., Martens, E.C., Mosimann, S.C., Abbott, D.W., Gilbert, H.J.

613 (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human
 614 colonic Bacteroides. Nat. Microbiol. 3: 210-219.

615 Makarova, E.N., & Shakhmatov, E.G. (2021) Characterization of pectin-xylan-glucan-
 616 arabinogalactan proteins complex from Siberian fir *Abies sibirica* Ledeb. Carbohydr.
 617 Polym. 260: 117825.

618 McKee, L.S., Rosa, S.L.L., Westereng, B., Eijsink, V.G., Pope P.B., Larsbrink, J. (2021)
 619 Polysaccharide degradation by the Bacteroidetes: mechanisms and nomenclature.
 620 Environ. Microbiol. Rep. 13: 559-581.

621 McKenna, C., Al-Assaf, S., Phillips, G.O., Funami, T. (2006) The protein component in
 622 pectin: Is it an AGP? FFI J. 211: 264-274.

623 Mohnen, D. (2008) Pectin structure and biosynthesis. Curr. Opin. Plant Biol. 11: 266-
 624 277.

625 Munoz-Munoz, J., Ndeh, D., Fernandez-Julia, P., Walton, G., Henrissat, B., Gilbert, H.J.
 626 (2021) Sulfation of arabinogalactan proteins confers privileged nutrient status to
 627 Bacteroides plebeius. mBio 12: e01368-21.

628 Mutter, M., Colquhoun, I.J., Beldman, G., Schols, H.A., Bakx, E.J., Voragen, A.G.J.
 629 (1998) Characterization of recombinant rhamnogalacturonan α -L-rhamnopyranosyl-1-
 630 (1,4)- α -D-galactopyranosyluronide lyase from *Aspergillus aculeatus*. Plant Physiol. 117:
 631 141-152.

632 Nascimento, G.E.d., Iacomini, M., Cordeiro, L.M.C. (2017) New findings on green sweet
 633 pepper (*Capsicum annum*) pectins: rhamnogalacturonan and type I and II
 634 arabinogalactans. Carbohydr. Polym. 171: 292-299.

635 Ndeh, D., & Gilbert, H.J. (2018) Biochemistry of complex glycan depolymerization by
 636 the human gut microbiota. FEMS Microbiol. Rev. 42: 146-164.

637 Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Basle, A., Gray, J., Venditto, I.,
 638 Briggs, J., Zhang, X., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y.,
 639 Field, R.A., Zhu, Y., O'Neill, M.A., Urbanowicz, B.R., York, W.S., Davies, G.J., Abbott,
 640 D.W., Ralet, M.C., Martens, E.C., Henrissat, B., Gilbert, H.J. (2017) Complex pectin
 641 metabolism by gut bacteria reveals novel catalytic functions. Nature 544: 65-70.

642 Ndukwe, I.E., Black, I., Heiss, C., Azadi P. (2020) Evaluating the utility of permethylated
 643 polysaccharide solution NMR data for characterization of insoluble plant cell wall
 644 polysaccharides. *Anal. Chem.* 92: 13221-13228
 645 Nergard, C.S., Matsumoto, T., Inngjerdigen, M., Inngjerdigen, K., Hokputsa, S.,
 646 Harding, S.E., Michaelsen, T.E., Diallo, D., Kiyohara, H., Paulsen, B.S., Yamada, H.
 647 (2005) Structural and immunological studies of a pectin and a pectic arabinogalactan
 648 from *Vernonia kotschyana* Sch. Bip. Ex Walp. (Asteraceae). *Carbohydr. Res.* 340: 115-
 649 130.
 650 Oosterveld, A., Voragen, A.G.J., Schols, H.A. (2002) Characterization of hop pectins
 651 shows the presence of an arabinogalactan-protein. *Carbohydr. Polym.* 49: 407-413.
 652 Pellerin, P., Vidal, S., Williams, P., Brillouet, J.M. (1995) Characterization of five type II
 653 arabinogalactan-protein fractions from red wine of increasing uronic acid content.
 654 *Carbohydr. Res.* 277: 135-143.
 655 Pfeifer, L., Shafee, T., Johnson, K., Bacic, A., Classen, B. (2020) Arabinogalactan-
 656 proteins of *Zostera marina* L. contain unique glycan structures and provide insight into
 657 adaption processes to saline environments. *Sci. Rep.* 10: 8232.
 658 Phillips, L.R., & Fraser, B.A. (1981) Methylation of carbohydrates with dimsyl
 659 potassium in dimethyl sulfoxide. *Carbohydr. Res.* 90: 149-152.
 660 Ropartz, D., & Ralet, M.-C. (2020) Pectin structure. In V. Kontogiorgos (Ed.) *Pectin:*
 661 *Technological and physiological properties* (pp. 17-36). Switzerland: Springer Nature.
 662 Ryden, P., & Selvendran, R.R. (1990) Cell-wall polysaccharides and glycoproteins of
 663 parenchymatous tissues of runner bean (*Phaseolus coccineus*). *Biochem. J.* 269: 393-402.
 664 Santander, J., Martin, T., Loh, A., Pohlenz, C., Gatlin III, D.M., Curtiss III, R. (2013)
 665 Mechanisms of intrinsic resistance to antimicrobial peptides of *Edwardsiella ictaluri* and
 666 its influence on fish gut inflammation and virulence. *Microbiology* 159: 1471-1486.
 667 Saulnier, L., & Brillouet, J.M. (1988) Structural studies of pectic substances from the
 668 pulp of grape berries. 182: 63-78.
 669 Shimoda, R., Okabe, K., Kotake, T., Matsuoka, K., Koyama, T., Tryfona, T., Liang,
 670 H.C., Dupree, P., Tsumuraya, Y. (2014) Enzymatic fragmentation of carbohydrate
 671 moieties of radish arabinogalactan-protein and elucidation of the structures. *Biosci.*
 672 *Biotechnol. Biochem.* 78: 818-831.

673 Sun, Y., Cui, S.W., Tang, J., Gu, X. (2010) Structural features of pectic polysaccharide
 674 from *Angelica sinensis* (Oliv.) Diels. Carbohydr. Polym. 80: 544-550.

675 Tan, L., Leykam, J.F., Kieliszewski, M.J. (2003) Glycosylation motifs that direct
 676 arabinogalactan addition to arabinogalactan-proteins. Plant Physiol. 132: 1362-1369.

677 Tan, L., Qiu, F., Lamport, D.T.A., Kieliszewski, M.J. (2004) Structure of a
 678 hydroxyproline (Hyp)-arabinogalactan polysaccharide from repetitive Ala-Hyp expressed
 679 in transgenic *Nicotiana tabacum*. J. Bio. Chem. 279: 13156-13165.

680 Tan, L., Varnai, P., Lamport, D.T.A., Yuan, C., Xu, J., Qiu, F., Kieliszewski, M.J. (2010)
 681 Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl
 682 subunits with short bifurcated side chains. J. Biol. Chem. 285: 24575-24583.

683 Tan, L., Showalter, A.M., Egelund, J., Hernandez-Sanchez, A., Doblin, M.S., Bacic, A.
 684 (2012) Arabinogalactan-proteins and the research challenges for these enigmatic plant
 685 cell surface proteoglycans. Front. Plant Sci. 3: 140.

686 Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., Hao, Z., Zhu, X.,
 687 Avci, U., Miller, J.S., Baldwin, D., Pham, C., Orlando, R., Darvill, A., Hahn, M.G.,
 688 Kieliszewski, M.J., Mohnen, D. (2013) An Arabidopsis cell wall proteoglycan consists of
 689 pectin and arabinoxylan covalently linked to an arabinogalactan protein. Plant Cell 25:
 690 270-287.

691 Tryfona, T., Liang, H.C., Kotake, T., Kaneko, S., Marsh, J., Ichinose, H., Lovegrove, A.,
 692 Tsumuraya, Y., Shewry, P.R., Stephens, E., Dupree, P. (2010) Carbohydrate structural
 693 analysis of wheat flour arabinogalactan protein. Carbohydr. Res. 345: 2648-2656.

694 Tryfona, T., Liang, H.C., Kotake, T., Tsumuraya, Y., Stephens, E., Dupree, P. (2012)
 695 Structural characterization of Arabidopsis leaf arabinogalactan polysaccharides. Plant
 696 Physiol. 160: 653-666.

697 Tsumuraya, Y., Mochizuki, N., Hashimoto, Y., Kovac, P. (1990) Purification of an exo-
 698 β -(1,3)-D-galactanase of *Irpex lacteus* (*Polyporus tulipiferae*) and its action on
 699 arabinogalctan-proteins. J. Biol. Chem. 265: 7207-7215.

700 Voragen, A.G.J., Coenen, G.-J., Verhoef, R.P., Schols, H.A. (2009) Pectin, a versatile
 701 polysaccharide present in plant cell walls. Struct. Chem. 20: 263.

702 Willis, L.M., Stupak, J., Richards, M.R., Lowary, T.L., Li, J., Whitfield, C. (2013)
 703 Conserved glycolipid termini in capsular polysaccharides synthesized by ATP-binding

704 cassette transporter-dependent pathways in Gram-negative pathogens. Proc. Natl. Acad.
705 Sci. U.S.A. 110: 7868-7873.

706 Yamada, H., Yanahira, S., Kiyohara, H., Cyong, J.C., Otsuka, Y. (1987) Characterization
707 of anti-complimentary acidic heteroglycans from the seeds of *Coix lacryma-jobivar* Ma-
708 yuen. *Phytochemistry* 26: 3269-3275.

709 York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T., Albersheim, P. (1986) Isolation
710 and characterization of plant cell walls and cell wall components. *Methods Enzymol.*
711 118: 3-40.

712 Zheng, Y., & Mort, A. (2008) Isolation and structural characterization of a novel
713 oligosaccharide from the rhamnogalacturonan of *Gossypium hirsutum* L. *Carbohydr.*
714 *Res.* 343: 1041-1049.