THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 32, pp. 24575–24583, August 6, 2010

Plant *O*-Hydroxyproline Arabinogalactans Are Composed of Repeating Trigalactosyl Subunits with Short Bifurcated Side Chains*[S]

Received for publication, January 1, 2010, and in revised form, April 14, 2010 Published, JBC Papers in Press, May 20, 2010, DOI 10.1074/jbc.M109.100149

Li Tan[‡1], **Peter Varnai**[§2], **Derek T. A. Lamport**[¶2], **Chunhua Yuan**[‖2], **Jianfeng Xu**[‡3], **Feng Qiu**[**4], **and Marcia J. Kieliszewski**[‡5]

From the [‡]*Department of Chemistry and Biochemistry, Biochemistry Research Facility, Ohio University, Athens, Ohio 45701, the* [§]*Department of Chemistry and Biochemistry and the* [¶]*School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9QG, United Kingdom, the* [‖]*Campus Chemical Instrument Center, Ohio State University, Columbus, Ohio 43210, and the* **Department of Chemistry, Temple University, Philadelphia, Pennsylvaina 19122*

Classical arabinogalactan proteins partially defined by type II *O*-Hyp-linked arabinogalactans (Hyp-AGs) are structural components of the plant extracellular matrix. Recently we described the structure of a small Hyp-AG putatively based on repetitive trigalactosyl subunits and suggested that AGs are less complex and varied than generally supposed. Here we describe three additional AGs with similar subunits. The Hyp-AGs were isolated from two different arabinogalactan protein fusion glycoproteins expressed in tobacco cells; that is, a 22-residue Hyp-AG and a 20-residue Hyp-AG, both isolated from interferon α2b-(Ser-Hyp)$_{20}$, and a 14-residue Hyp-AG isolated from (Ala-Hyp)$_{51}$-green fluorescent protein. We used NMR spectroscopy to establish the molecular structure of these Hyp-AGs, which share common features: (i) a galactan main chain composed of two 1→3 β-linked trigalactosyl blocks linked by a β-1→6 bond; (ii) bifurcated side chains with Ara, Rha, GlcUA, and a Gal 6-linked to Gal-1 and Gal-2 of the main-chain trigalactosyl repeats; (iii) a common side chain structure composed of up to six residues, the largest consisting of an α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→3)-α-L-Ara*f*-(1→3- unit and an α-L-Rha*p*-(1→4)-β-D-GlcUA*p*-(1→6)-unit, both linked to Gal. The conformational ensemble obtained by using nuclear Overhauser effect data in structure calculations revealed a galactan main chain with a reverse turn involving the β-1→6 link between the trigalactosyl blocks, yielding a moderately compact structure stabilized by H-bonds.

Hydroxyproline-rich glycoproteins of the cell surface comprise groups of related structural proteins, including the extensins that form cell wall scaffolding networks essential for cytokinesis (1) and the classical arabinogalactan proteins (2) that are largely at the membrane wall interface (3) and have diverse functions (4). *O*-Hyp[6] glycosylation characterizes the hydroxyproline-rich glycoproteins and is of much interest as it defines molecular properties and, hence, biological function. Arabinogalactan proteins are highly glycosylated mainly with *O*-Hyp-arabinogalactan polysaccharides (5, 6). Extensins are less highly glycosylated mainly with small *O*-Hyp arabinooligosaccharides (7, 8), whereas the related proline-rich proteins are minimally glycosylated also with arabinooligosaccharides (9). Peptide sequence directs *O*-Hyp glycosylation by the addition of small oligosaccharides to contiguous Hyp residues and larger acidic arabinogalactan polysaccharides to clustered noncontiguous Hyp (10, 11). For example, clustered Ala-Hyp and Ser-Hyp are typical AGP glycosylation motifs (12), whereas the Hyp residues in repetitive blocks of Ser-Hyp$_2$ orSer-Hyp$_4$ are arabinosylated. The "hyperglycosylation" of closely related AGPs complicates their purification, a problem that can be overcome by expressing single individual AGPs as GFP fusion glycoproteins, the hydrophobic GFP tag enabling chromatographic purification (13, 14). This approach also allows purification of neo-AGPs containing single repeating AGP glycosylation motifs, for example (Ala-Hyp)$_n$ or (Ser-Hyp)$_n$, for base-catalyzed peptide bond hydrolysis. Base hydrolysis releases alkali-stable Hyp-arabinogalactan glycoamino acids, designated Hyp-AGs (5, 6), that can be further purified by size-exclusion chromatography. Here we report the complete structural elucidation of three such Hyp-AGs ranging in size from 14 to 22 sugar residues. The structures were determined using multidimensional homonuclear and heteronuclear NMR spectroscopy in conjunction with molecular simulations in the presence of water.

Ala-Hyp-polysaccharide-1 isolated earlier from (Ala-Hyp)$_{51}$-GFP expressed in tobacco cells represented the first complete structure of a small Hyp-AG (15) characteristic of type II arabi-

* This work was supported, in whole or in part, by National Institutes of Health Grant 2-P41-RR05351-06 (to the Complex Carbohydrate Research Center). This work was also funded by United States Department of Agriculture Grants 2004-34490-14579), Herman Frasch Foundation Grant 526-HF02, The Ohio University Biomimetic, Nanoscience, and Nanotechnology Program Grant GC0013845), and by the Ohio University Molecular and Cellular Biology Program.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables I–V and Figs. 1–10.

[1] Present address: 315 Riverbend Rd., CCRC, University of Georgia, Athens, GA 30602.

[2] These authors contributed equally to this manuscript.

[3] Present address: Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR 7240.

[4] Present address: R&D, Bristol-Myers Squibb Co., 311 Pennington-Rocky Hill Rd., Pennington, NJ 08534.

[5] To whom correspondence should be addressed: 350 West State St., Biochemistry Research Facility, Ohio University, Athens, OH 45701. Tel.: 740-593-9466; Fax: 740-597-1772; E-mail: kielisze@helios.phy.ohiou.edu.

[6] The abbreviations used are: Hyp, hydroxyproline; AGP, arabinogalactan-protein; AG, arabinogalactan; GFP, green fluorescent protein; COSY, homonuclear correlation spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; HSQC, heteronuclear single quantum coherence; HMBC, heteronuclear multiple bond coherence.

nogalactans, which are defined by β-1,3-linked galactose residues. Ala-Hyp-polysaccharide-1 consisted of a 5-residue 1,3-β-D-Gal backbone interrupted by a single β-1,6 linkage connecting Gal-3 and Gal-4 (numbered from the reducing end). Galactose residues 1 and 2 had small-bifurcated side-chain substituents at C-6. The side chains possessed a single β-D-Gal di-substituted at C-3 with α-L-arabinose di- or trisaccharides and at C-6 with β-D-glucuronic acid or α-L-rhamnosylglucuronic acid (12, 15). We proposed that ~15-residue repetitive blocks of decorated β-(1–3) trigalactosyl subunits connected by β-1,6 linkages create the larger arabinogalactan polysaccharides of AGPs, consistent with the small blocks separated by periodate-sensitive residues suggested earlier (16, 17).

As type II arabinogalactans are often considered intractably complex, and because Ala-Hyp-polysaccharide-1 was only a single example of a Hyp-AG subunit, we determined the structure of three additional Hyp-AGs derived from two different AGP motifs, repetitive Ala-Hyp and Ser-Hyp. Two of the Hyp-AGs designated interferon-polysaccharide-1 (interferon Hyp-polysaccharide 1) and interferon-polysaccharide-2 were isolated from a fusion glycoprotein of human interferon α2b fused to a $(Ser\text{-}Hyp)_{20}$ AGP glycomodule (18). The third Hyp-AG, designated Ala-Hyp polysaccharide-2, was isolated from $(Ala\text{-}Hyp)_{51}$-GFP similar to Ala-Hyp-polysaccharide-1 described earlier (15). Here we identified the fundamental similarities between these Hyp-arabinogalactans and determined if the non-glycosylated domains (interferon α2 *versus* GFP) or AGP motifs (Ser-Hyp *versus* Ala-Hyp repeats) influenced the glycan structure. Significantly, the six-residue galactan backbone of these new Hyp-AGs consisted of two β-1,3-linked galactosyl trisaccharides connected by a β-1,6 linkage. Such "decorated" ~15-residue trisaccharide subunits likely constitute the fundamental building blocks of type II arabinogalactan polysaccharides; hence, they are far less complex than commonly supposed (4, 19). Finally, the NMR analyses and molecular modeling of the glycans revealed major conformers that include a moderately compact folded structure.

EXPERIMENTAL PROCEDURES

Gene Construction and Expression in Tobacco Cells—Genes encoding Interferon α2-$(Ser\text{-}Hyp)_{20}$ and $(Ala\text{-}Hyp)_{51}$-GFP were constructed and expressed as described in detail earlier (12, 18). Briefly, proteins were targeted for secretion using a tobacco extensin signal sequence, and gene expression was under control of the 35 S cauliflower mosaic virus promoter. The genes were subcloned into the plant transformation vector *pBI121* and expressed in tobacco Bright Yellow-2 cells selected and maintained as described earlier (12, 18).

Isolation of the Interferon α2-$(Ser\text{-}Hyp)_{20}$ and $(Ala\text{-}Hyp)_{51}$-GFP Fusion Glycoproteins—Interferon α2-$(Ser\text{-}Hyp)_{20}$ and $(Ala\text{-}Hyp)_{51}$-GFP were isolated from the medium of suspension-cultured tobacco Bright Yellow-2 cells by hydrophobic interaction chromatography on a phenyl-Sepharose column as described earlier (18).

Isolation of Hyp-arabinogalactans—Two hundred mg of Interferon α2-$(Ser\text{-}Hyp)_{20}$ were hydrolyzed in 20 ml of 0.44 N NaOH solution at 108 °C for 18 h. The cooled solution was titrated to pH 7.8 with cold 1 N HCl and then freeze-dried. Hydrolysates were fractionated on an analytical Superdex-Peptide column (HR 10/30, GE Healthcare) equilibrated in 20% acetonitrile (aqueous) and eluted at a flow rate of 0.3 ml/min (15).

Fractions (0.6 ml total volume each) were freeze-dried and analyzed for Hyp and monosaccharides colorimetrically or by gas chromatography using methods described earlier (18). The fraction containing the most Hyp (fraction 16 described in Ref. 18, interferon Hyp-polysaccharide-1) and a fraction containing later-eluting Hyp-glycans (fraction 18, Ref 18, interferon Hyp-polysaccharide-2) were rerun on the Superdex column, freeze-dried, and then used for NMR analyses. The Hyp-glycan Ala-Hyp-polysaccharide-2 from $(Ala\text{-}Hyp)_{51}$-GFP was isolated by a combination of cation exchange and gel filtration chromatography as described earlier (15, 20).

NMR Spectroscopy—A 1-mg sample of each Hyp-AG was dissolved in 0.5 ml of 99.996% D_2O (Cambridge Isotope Laboratories, Andover, MA). NMR experiments were carried out either at 55 °C on a Bruker DMX-800 equipped with a cryoprobe or at 25 °C on a Bruker DMX-600 spectrometer equipped with a triple-resonance probe and three-axis gradient coils. The parallel data sets include one-dimensional ^1H, two-dimensional ^1H-homonuclear correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY) (mixing time 60 and 90 ms), rotating frame NOE spectroscopy (ROESY) (200 ms), and nuclear Overhauser effect spectroscopy (NOESY) (mixing time 150, 300 and 500 ms), and two-dimensional ^{13}C,^1H heteronuclear single quantum coherence (HSQC) and heteronuclear multiple bond coherence (HMBC) NMR spectra. In addition, several more interferon Hyp-polysaccharide-1 experiments were recorded in an effort to resolve assignment ambiguities such as magnitude COSY with one-, two-, and three-step relay transfer and two-dimensional ^{13}C,^1H heteronuclear HSQC-TOCSY and HSQC-NOESY together with diffusion-ordered spectroscopy for measuring the diffusion constant. Water suppression was achieved by either presaturation or WATERGATE techniques. Data were processed with NMRPipe (21) and visualized using NMRView (22) Chemical shifts were referenced to an external standard: 4,4-dimethyl-4-silapentane-1-sulfonic acid.

NMR Structure Calculations—Interferon Hyp-polysaccharide-1 was constructed in an arbitrary extended conformation using the *LEaP* module of Amber 10 (23). The starting model was subjected to a restrained simulated annealing conformational search protocol to obtain an ensemble of structures consistent with the NMR data. All assigned NOESY cross-peaks were classified as strong (1.8–2.7 Å), medium (1.8–3.7 Å), weak (1.8–5.0 Å), and very weak (1.8–6.0 Å) interproton distance restraints according to their intensities. Beyond these bounds, a quadratic penalty potential was applied with a force constant of 20 kcal mol^{-1} $Å^{-2}$. A total of 49 distance restraints were used for interferon Hyp-polysaccharide-1 of which 34 were assigned non-ambiguously to protons in sequential residues. Cross-peaks that correspond to non-sequential assignments gave rise to ambiguous restraints where either more than one proton pair contributes to the NOESY volume or unambiguous assignment was not possible. Ambiguous peaks were interpreted as

an $< r^{-6} >^{-1/6}$ averaged value of the contributing interproton distances.

For a restrained molecular dynamics conformational search, initial models were energy-minimized without restraints and subjected to 200 simulated annealing cycles. A cycle started from the structure obtained in the previous cycle and included 10-ps heating from 300 to 1000 K followed by an equilibration of 10 ps at 1000 K without any restraints applied. The force constants for all restraints were then scaled gradually from 0 to the final values during 20 ps at 1000 K followed by cooling the system to 300 K over 30 ps. Atomic interactions within the system were calculated using the Glycam06 parameter set for sugars (24) and Generalized Born solvation (*igb* = 2) with monovalent salt concentration corresponding to 0.1 M. The last structures in each cycle were energy-minimized using the same parameters and restraints as described above. Ten best models for an average structure were selected based on NMR restraint violations and the potential energy of the molecule to undergo further refinement in explicit water and counterion environment. Each of these model structures was placed in a truncated octahedral box of about 5000 TIP3P water molecules and two K^+ counterions to neutralize the total charge. In one case we used Ca^{2+} to neutralize the charge of the uronic acids. Parameters related to water and counterions were taken from the standard Amber libraries. The system was energy-minimized and then heated to 300 K at constant volume during 50 ps, whereas the solute was kept under positional restraints with a force constant of 25 kcal mol^{-1} $Å^{-2}$. The positional restraints were gradually removed over 300 ps at constant pressure (1 atm) and temperature (300 K), and a production phase was initiated for 2 ns with the full set of restraints applied. The final structures were energy-minimized and used for subsequent analysis. The hydrodynamic radius was calculated for the model structures using HYDROPRO Version 7c2 (25).

RESULTS

The Structures of Interferon Hyp-polysaccharide-1, interferon Hyp-polysaccharide-2, and Ala-Hyp-polysaccharide-2 were determined based on earlier composition analyses and on the chemical shifts (15, 26) observed in one-dimensional ^{1}H NMR spectra and two-dimensional COSY, TOCSY, HSQC, and HMBC NMR spectra as follows:

Primary Structure of Interferon Hyp-polysaccharide-1—Size-fractionated base hydrolysates of Interferonα2-(Ser-Hyp)$_{20}$ yielded a single peak containing sugar and Hyp residues, described earlier (26). Two subfractions were chosen for structural analyses; one contained the major Hyp-AG species, designated Interferon Hyp-polysaccharide-1, with 22 glycosyl residues estimated by the Hyp to monosaccharide molar ratios, and a second fraction contained a smaller, less abundant Hyp-arabinogalactan of 20 glycosyl residues, designated interferon Hyp-polysaccharide-2.

Sugar Composition and Configuration—The integrated areas in the anomeric region in a one-dimensional ^{1}H NMR spectrum (Fig. 1) corroborated the molar ratios Hyp Gal_{10} Ara_8 $GlcU_2$ Rha_2 of interferon Hyp-polysaccharide-1 (26), confirming interferon Hyp-polysaccharide-1 as a 22-residue arabinogalactan attached to Hyp. The anomeric configurations were confirmed by the corresponding anomeric carbon signals in the HSQC spectrum (Fig. 2, supplemental Table I). Interferon Hyp-polysaccharide-1 contained a total of eight α-L-Ara*f* residues; 6 of the 8 α-Ara*f* residues were either terminal, 1,5-linked, or 1,3-linked (~5.25 ppm), and 2 gave signals that were consistent with being both terminal and 1→ 5-linked to L-Ara*f* residues (5.087 ppm). Interferon Hyp-polysaccharide-1 also contained 2 α-L-Rha*p* residues (4.797 ppm), 2 β-D-GlcUA*p* residues (~4.5 ppm), and a total of 10 β-D-Gal*p* residues; 1 was *O*-linked to Hyp (4.57 ppm), 5 were consistent with a galactan backbone (~ 4.70 ppm) (15), and 4 others that produced signals typical of side-chain β-D-Gal*p* residues (4.40 – 4.54 ppm). The chemical shift at 4.786 ppm identified 4-H of Hyp.

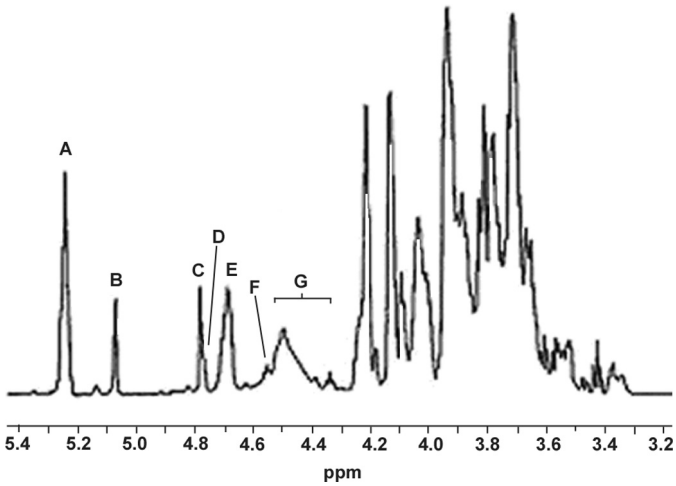

FIGURE 1. **^{1}H NMR spectrum of Hyp-AG interferon Hyp-polysaccharide-1 at 55 °C.** Signals *A* and *B* were assigned respectively to H-1 of six and two α-L-Ara residues, signal *C* to H-1 of two α-L-Rha residues, signal *D* (shoulder peak of C) to H-4 of Hyp, and signal *E* to the five β-D-Gal residues of the galactan backbone. Signal *F* was assigned to H-1 of β-D-Gal linked to Hyp and signal *G* to H-1 of the four side chain β-D-Gal residues and two β-D-GlcUA residues.

The Interferon Hyp-polysaccharide-1 Hyp-galactose Linkage—Signals arising from the Hyp residue were also identified in the TOCSY (supplemental Fig. 1) and HSQC spectra (Fig. 2). The chemical shifts are shown in supplemental Table I. The H-4 and C-4 resonances characteristic of non-glycosylated Hyp shifted downfield from 4.62 and 70.6 to 4.786 and 78.02 ppm, respectively, judging from the HSQC spectrum. This indicated that the hydroxyl group of Hyp was galactosylated (15, 18). The HMBC spectrum (Fig. 3) confirmed this in cross-peak G, which arose from C-4 of Hyp (78.02 ppm) and H-1 of a β-D-Gal*p* residue (4.57) ppm, designated G_1 in Fig. 4 and supplemental Table I.

Galactan Backbone and Side Chain Gal Residues—In addition to G_1 linked to Hyp, there were another five β-D-Gal*p* residues in the interferon Hyp-polysaccharide-1 galactan backbone (designated G_{1-6} in supplemental Table I and Figs. 4 and 5) judging by H-1 resonances at ~4.70 ppm in the one-dimensional ^{1}H spectrum. Four of the five backbone Gal residues were 3-linked to each other and to G_1, as deduced from cross-peaks E in the HMBC spectrum (Fig. 3), which correlated backbone Gal H-1 signals with backbone Gal 3-C signals. A fifth backbone Gal participated in a 1→ 6 linkage to another backbone Gal residue deduced from cross-peak F in the HMBC spectrum. Thus, based

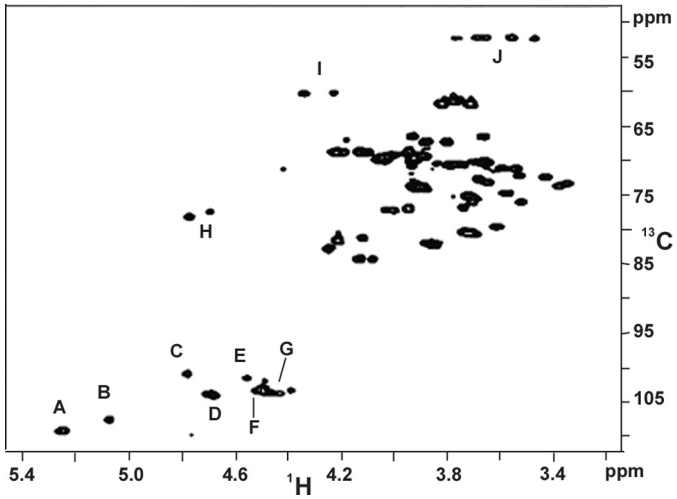

FIGURE 2. **HSQC spectrum of Hyp-AG Interferon Hyp-polysaccharide-1 at 55 °C.** Cross-peaks *A* and *B* were assigned to H-1/C-1 of α-L-Ara, *C* to H-1/C-1 of α-L-Rha, *D* to the H-1/C-1 of β-D-Gal residues of the galactan backbone, and *E* to H-1/C-1 of β-D-Gal linked to Hyp. Cross-peak *F* was assigned to H-1/C-1 of β-D-GlcUA residues, *G* to H-1/C-1 of side-chain β-D-Gal, *H* to H-4/C-4 of Hyp, *I* to H-2/C-2 of Hyp, and *J* to H-5/C-5 of Hyp.

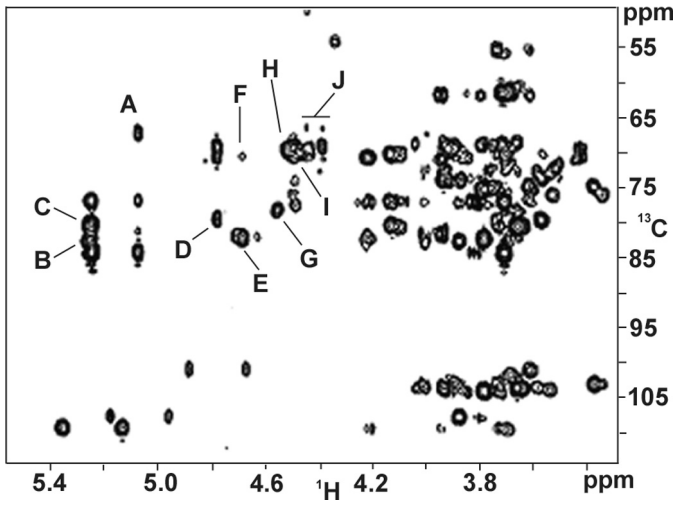

FIGURE 3. **HMBC spectrum of Hyp-AG interferon Hyp-polysaccharide-1 at 55 °C.** This helped identify the interferon Hyp-polysaccharide-1 monosaccharide sequence. Cross-peak *A* correlated Ara H-1 with Ara C-5, cross-peak *B* correlated Ara H-1 with Ara C-3, cross-peak *C* correlated Ara H-1 with side-chain Gal C-3, cross-peak *D* correlated Rha H-1 with GlcUA C-4, cross-peak *E* correlated backbone Gal H-1 with backbone Gal C-3, cross-peak *F* correlated backbone Gal H-1 with backbone Gal C-6, cross-peak *G* correlated G_1 H-1 with Hyp C-4, cross-peak *H* correlated GlcUA H-1 with side-chain Gal C-6, and cross-peaks *I* and *J* correlated side-chain Gal H-1 to backbone Gal C-6.

on our earlier work (15) and the NMR spectra presented here, the interferon Hyp-polysaccharide-1 galactan backbone had the structure β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-*O*-Hyp.

The final 4 of the 10 Gal residues of interferon Hyp-polysaccharide-1 occurred in side chains (designated G_a, G_b, G_c, G_d in supplemental Table I and Figs. 4 and 5) linked 1→6 to the backbone Gal residues. This was deduced from one-dimensional ^{1}H and two-dimensional ^{13}C,^{1}H HMBC spectra. Resonances at 4.40–4.54 ppm in the one-dimensional ^{1}H NMR spectrum were consistent with four Gal side chains attached to the galactan backbone (15), and cross-peaks I and J in the HMBC spectrum indicated that the four side-chain Gal residues were 1→6 linked to backbone Gal residues.

Interferon Hyp-polysaccharide-1 Side-chain Composition and Linkages—The two α-L-Rha residues, designated R_1 and R_2, were terminal, deduced by a comparison of their assigned chemical shifts (supplemental Table I) obtained from TOCSY (supplemental Fig. 1) and HSQC spectra (Fig. 2) with those of earlier characterized Ala-Hyp-polysaccharide-1 (15). Rhamnose residues R_1 and R_2 were linked to *O*-4 of β-D-GlcUA*p* (UA_1 and UA_2 of supplemental Table I), deduced from cross-peak D in the HMBC spectrum (Fig. 3). Cross-peak H in the HMBC spectrum indicated side-chain Gal residues were substituted at *O*-6 with glucuronic acid residues UA_1 and UA_2. The same chemical shifts arising from α-L-Rha*p*, β-D-GlcUA*p*, and side-chain β-D-Gal*p* residues were identified earlier on side-chain Gal residues in Ala-Hyp-polysaccharide-1 (15). This indicated the side chains were attached to backbone Gal residues nearest Hyp; therefore, we assigned the two Rha-(1→4)-GlcUA subunits to side-chain Gal residues closest to Hyp, G_a, and G_b (supplemental Table I and Figs. 4 and 5) to give two side chains: α-L-Rha*p*-(1→4)-β-D-GlcUA*p*-(1→6)-β-D-G_a and $_{-Gb}$.

The Ara residues occurred in small side chains that were 3-linked to the side-chain Gal residues. The HMBC spectrum cross-peak A identified Ara 5-linked to another Ara (Fig. 3). It arose from H-1 of α-L-Ara*f* residues (5.087 ppm, A_1 and A_4 in Fig. 4*a*) and 5-C of other Ara residues (67.1 ppm, A_2 and A_5 in Fig. 4*a*) This was consistent with the one-dimensional ^{1}H NMR spectrum that indicated there were two Ara-(1→5)-Ara linkages in interferon Hyp-polysaccharide-1. HMBC signals arising from the ring carbon atoms of A_1 and A_4 (C-2, C-3, and C-4) showed that they were terminal residues (15); hence, two diarabinosyl structures occurred having the structure α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→).

The HSQC C-1/H-1 signals at 109.2 ppm and ~5.25 ppm arose from six Ara residues (A_2, A_3, A_5, A_6, A_7, A_8 in supplemental Table I, Figs. 4 and 5) and indicated two other types of linkages in the arabinose side chains. Cross-peak B in the HMBC spectrum indicated linkage of α-L-Ara*f* to *O*-3 of another α-L-Ara*f* residue (3-C signal at 83.88 ppm) as in the structure α-L-Ara*f*-(1→3)-α-L-Ara*f* (*i.e.* A_2 to A_3 and A_5 to A_6), whereas cross-peak C indicated α-L-Ara*f*-(1→3)-β-D-Gal linkages (side-chain Gal residues G_{a-d} in supplemental Table I). The side-chain Gal H-3 chemical shifts (~3.73 ppm) (15) present in the TOCSY spectrum (supplemental Fig. 1) indicated each of the four side-chain Gal residues was *O*-3 substituted. As four Ara residues were attached to Gal, and the other four occurred in two α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→ units, we concluded that two of the arabinosyl side chains had the structure α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→3)-α-L-Ara*f*-(1→3)-Gal_{sc}, and two had the structure α-L-Ara*f*-(1→3)-Gal_{sc}.

A comparison of the interferon Hyp-polysaccharide-1 HSQC and HMBC spectra with those of Ala-Hyp-polysaccharide-1 (15) showed that the interferon Hyp-polysaccharide-1 anomeric proton/carbon chemical shifts arising from G_c and G_d (~103.6/4.447 and 103.2/4.41 ppm), the side-chain Gal residues furthest from Hyp in Fig. 4, differed from the other side-chain Gal residues, G_a and G_b (both ~103.4/4.50). Furthermore, the signals from G_a and G_b in Interferon Hyp-

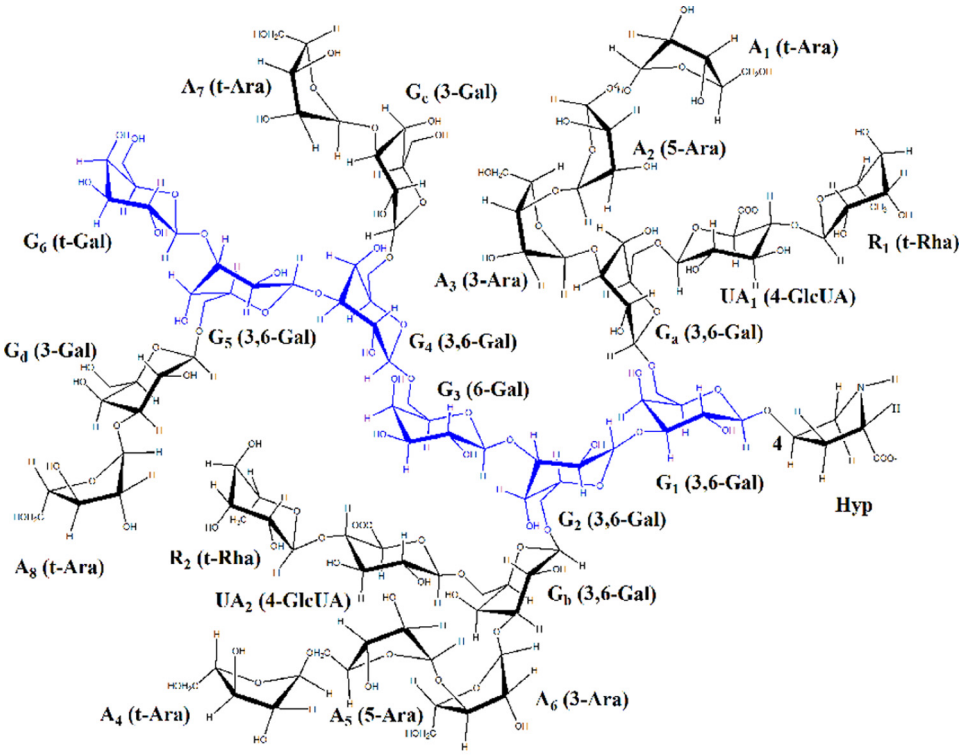

FIGURE 4. **Interferon Hyp-polysaccharide-1 and its canonical side chain.** *a*, the primary structure of the 22-residue Hyp-arabinogalactan Interferon Hyp-polysaccharide-1 is shown. Residue labeling includes the linkage of each residue and corresponds to that featured in supplemental Table I. The Hyp residue is positioned to the *right*, beneath rhamnose residue R_1. *b*, the canonical bifurcated AG side chain is shown.

polysaccharide-1 were identical to those of Ala-Hyp-polysaccharide-1 characterized earlier (15). Therefore, we designated the specific side-chain Gal residues in the two α-L-Ara*f*-(1→5)-α-L-Ara*f*-(1→3)-α-L-Ara*f*-(1→3)-Gal units as G_a and G_b. They were part of the two bifurcated six-residue side chains of Interferon Hyp-polysaccharide-1. The Gal residues in the two α-L-Ara*f*-(1→3)-side-chain Gal units were designated G_c and G_d (supplemental Table I, Figs. 4 and 5).

Interferon Hyp-polysaccharide-1 Long-range Interactions—Lowering the temperature of NMR analyses from 55 to 25 °C (see the chemical shift assignments in supplemental Table II) provided the following lines of evidence for a folded Interferon Hyp-polysaccharide-1 conformer (Fig. 6).

A diffusion-ordered spectroscopy spectrum (supplemental Fig. 2) gave a diffusion coefficient of (1.58 ± 0.1) $\times 10^{-10}$ m^2/s. Using the Stokes-Einstein equation, this value corresponds to a hydrodynamic radius of 15.6 ± 1.0 Å, which is consistent with a globular folded structure and close to the calculated value for model structures of 13.9 ± 0.3 Å (see below).

The intensity of NOEs arising from protons between sequential *O*-linked sugar residues, for example H-1 of R_1 and H-4 of glucuronic acid residue UA_1 (supplemental Table III and Fig. 3) suggested a somewhat restricted conformation around glycosidic links rather than free rotation. Furthermore, the single set of chemical shifts rules out the existence of several stable conformations.

Although the major spectral region was not amenable to unequivocal analysis due to considerable resonance overlap, a region possessing unique chemical shifts showed two NOE clusters at ~5.25/4.50 and ~5.25/4.70 ppm. The first cluster included three NOEs indicating long-range interactions; (i) δ 5.241/δ 4.526 attributed to H-1 of Ara residue A_7 or A_8 and H-1 of Gal residue G_a; (ii) δ 5.233/δ 4.500 attributed to H-1 of A_7 or A_8 and H-1 of G_b, UA_1, or UA_2; (iii) δ 5.262/δ 4.511 attributed to H-1 of Ara A_3 or A_6 and H-1 of UA_1 or UA_2. NOEs (i) and (ii) indicate a side chain Ara of the second trigalactosyl unit is close to the side-chain Gal/UA residues of the first trigalactosyl unit; this suggests that the β-1–6 linkage between two trigalactosyl units allows the main chain to fold. The second cluster also included diagnostic NOEs; (i) 5.237/4.694 ppm attributed to H-1 of Ara A_7 or A_8 and H-1of Gal G_2; (ii) 5.265/4.700 ppm attributed to H-1 of A_3 or A_6 and H-1 of G_4. The intensity of these NOEs indicates a distance of 6 Å between the respective protons that suggests the molecule is folded. The possible effect of spin diffusion can be excluded because these NOEs could be observed with a short mixing time of 150 ms.

Primary Structure of Interferon Hyp-polysaccharide-2—Neutral sugar, uronic acid, and Hyp analyses of interferon Hyp-polysaccharide-2 gave a molar ratio of Hyp Gal_{10} Ara_5 $GlcUA_4$ Rha. The interferon Hyp-polysaccharide-2 ^1H NMR spectrum (supplemental Fig. 4) gave a very similar signal pattern in the anomeric proton region as interferon Hyp-polysaccharide-1, except the peak area ratios differed, as interferon Hyp-polysaccharide-2 contained two more GlcUA and one less Rha and only five Ara residues. Of the five α-Ara*f* residues, only one (anomeric proton signal at ~5.08 ppm) was 5-linked to another Ara; the other four (anomeric proton signal at ~5.24 ppm) were terminal, 1,3-linked, or 1,5-linked. The α-Rha*p* residue was ter-

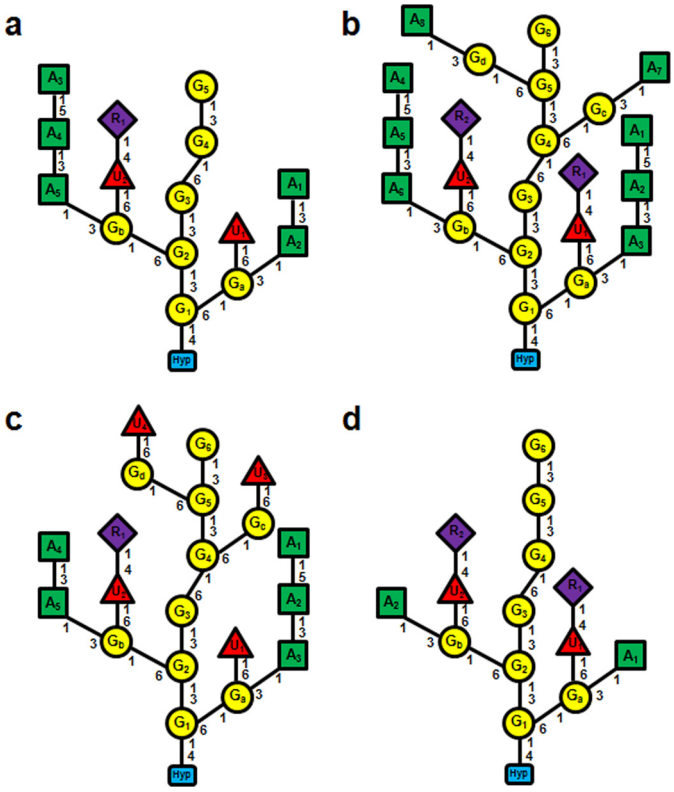

FIGURE 5. **Comparison of the primary structures of Hyp-AGs.** *a*, Ala-Hyp-polysaccharide-1 (15). *b*, interferon Hyp-polysaccharide-1. *c*, interferon Hyp-polysaccharide-2. *d*, Ala-Hyp-polysaccharide-2. The *yellow circles* indicate β-D-Gal*p* residues, the *green squares* indicate α-L-Ara*f*, the *purple diamonds* indicate α-L-Rha*p*, the *red triangles* indicate β-D-GlcUA*p*, and the *blue blocks* indicate Hyp. The residue labels inside the shapes (A_3, G_1, etc.) correspond to the residues described under "Results," Fig. 4, and the supplemental Tables I–V.

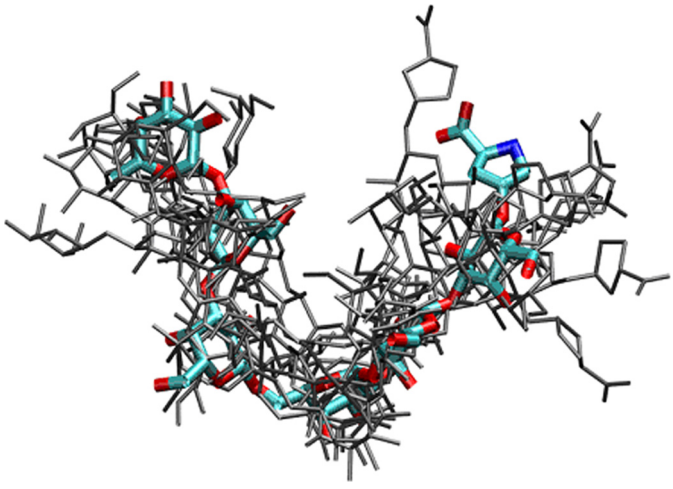

FIGURE 6. **Molecular model of Interferon Hyp-polysaccharide-1.** A range of galactan backbone conformations (in *gray*) is consistent with the experimental NOE data. A representative conformation is shown in color with the terminal Hyp residue *upper right*.

minal, and of the 10 β-Gal*p* residues, 6 were backbone Gal residues. Five of the backbone Gal residues gave H-1 signals at 4.68–4.71 ppm, and the sixth, G_1 in supplemental Table IV, was linked to Hyp. The other four Gal residues were part of the side chains (4.39–4.49 ppm). The H-1 signals of β-GlcUA*p* were not resolved from those of the side-chain Gal; however, the signal peak integral combined with chemical analyses of interferon Hyp-polysaccharide-2 indicated it had four GlcUA residues. Together, the ^{1}H NMR spectrum and the sugar analyses indicated interferon Hyp-polysaccharide-2 was a 20-sugar residue Hyp-arabinogalactan (Fig. 5*c*). We assigned the chemical shifts of interferon Hyp-polysaccharide-2 (supplemental Table IV) using two-dimensional TOCSY (supplemental Fig. 5), HSQC (supplemental Fig. 6), and HMBC (supplemental Fig. 7) spectra discussed below.

Hyp-Gal Linkage and Galactan Backbone—Interferon Hyp-polysaccharide-2 had the same galactan backbone structure as interferon Hyp-polysaccharide-1. The Hyp-Gal linkage was established by cross-peak G in the HMBC spectrum (supplemental Fig. 7). Like interferon Hyp-polysaccharide-1, the galactan backbone of interferon Hyp-polysaccharide-2 was composed of five β-D-Gal*p* residues (G_2-G_6 in supplemental Table IV) and six Gal linked to Hyp (G_1). Cross-peak E indicated the backbone Gal residues were mainly 1→3-linked, although a 1→6 link occurred between G_3 and G_4 (cross-peak F in supplemental Fig. 7 and Fig. 5*c*).

Interferon Hyp-polysaccharide-2 Side Chains—The side-chain structures were similar to those of interferon Hyp-polysaccharide-1, but generally smaller. The four side-chain Gal residues evident in the one-dimensional ^{1}H NMR spectrum were attached to backbone Gal residues through the *O*-6 position, deduced from cross-peaks I and J in the HMBC spectrum (supplemental Fig. 7) (side-chain Gal H-1 at 4.39–4.49 ppm to backbone Gal 6-C at ~70.0 ppm). Interferon Hyp-polysaccharide-2 had only one terminal α-L-Rha*p* but four β-D-GlcUA*p* residues. Cross-peak D in the HMBC spectrum indicated an α-L-Rha*p*-(1→4)-β-D-GlcUA*p* unit, and the chemical shifts of the other GlcUA residues (supplemental Table IV) indicated they were unsubstituted and, therefore, terminal. Cross-peak H showed that all four GlcUA residues were β-1→ 6-linked to side-chain Gal residues. In Fig. 5*c*, we assigned the Rha-(1→4)-GlcUA unit to one of the two side-chain Gal residues closest to Hyp, G_b, based on earlier work with Ala-Hyp-polysaccharide-1 (15); however, we had no direct evidence for this assignment, and any one of the four side-chain Gal residues were candidates.

HMBC cross-peaks A, B, and C (supplemental Fig. 7) corresponded to the following side chain arabinosyl units: α-L-Ara-(1→5)-α-L-Ara, α-L-Ara*f*-(1→3)-α-L-Ara*f*, and α-L-Ara*f*-(1→3)-Gal (15). These units were part of two arabinosyl side chains, α-L-Ara-(1→5)-α-L-Ara-(1→3)-α-L-Ara*f*-(1→3 and α-L-Ara*f*-(1→3)-α-L-Ara*f*-(1→3, that we assigned to side-chain Gal residues G_a and G_b rather than G_c and G_d because the H-3 resonances (supplemental Table IV, Fig. 5*c*) in the TOCSY spectrum (supplemental Fig. 5) indicated residues G_c and G_d were unsubstituted at *O*-3. However, we could not discern the precise distribution of the diarabinosyl and triarabinosyl units between G_a and G_b.

Primary Structure of Ala-Hyp-polysaccharide-2—Ala-Hyp-polysaccharide-2 derived from base-catalyzed hydrolysates of $(Ala\text{-}Hyp)_{51}$-enhanced green fluorescence protein was a 14-sugar residue Hyp-AG determined from the peak areas of its one-dimensional ^{1}H NMR spectrum (not shown) and sugar/Hyp analyses corresponding to Hyp_1 Gal_8 Ara_2 $GlcUA_2$ Rha_2, one residue less than Hyp_1 Gal_7 Ara_5 $GlcUA_2$ Rha of Ala-Hyp-

polysaccharide-1 (15). The chemical shifts of Ala-Hyp-polysaccharide-2 (supplemental Table V) were assigned based on HSQC (supplemental Fig. 8) and HMBC (supplemental Fig. 9) spectra compared with those of Ala-Hyp-polysaccharide-1.

Ala-Hyp-polysaccharide-2 Hyp-Gal Linkage and Galactan Backbone—Cross-peak G of Ala-Hyp-polysaccharide-2 in the HMBC spectrum established the linkage of β-D-Gal*p* (G_1) to position 4 of Hyp (Fig. 5*c*). G_1 with five other backbone β-D-Gal*p* residues comprised the galactan backbone, confirmed by the four β-D-Gal*p*-(1→3)-Gal linkages evident in cross-peak E and by a single β-D-Gal*p*-(1→6)-Gal linkage (cross-peak F). Thus, Ala-Hyp-polysaccharide-2 had one more backbone Gal residue than Ala-Hyp-polysaccharide-1.

Ala-Hyp-polysaccharide-2 Side Chains—Like Ala-Hyp-polysaccharide-1, interferon Hyp-polysaccharide-1, and interferon Hyp-polysaccharide-2 side chain Gal residues (G_a and G_b) of Ala-Hyp-polysaccharide-2 were attached to the galactan backbone through 1–6 linkages, deduced from cross-peak I. These two Gal side chains were assigned to G_1 and G_2 (Fig. 5*c*, supplemental Table V), as the chemical shifts were like those already characterized in Ala-Hyp-polysaccharide-1 (15).

Like the other Hyp-AGs, the β-D-GlcUA*p* residues of Ala-Hyp-polysaccharide-2 *O*-4 substituted with terminal α-L-Rha*p* residues (cross-peak D in supplemental Fig. 9 and Table V). Furthermore, the GlcUA residues were *O*-6-linked to Gal_{sc} deduced from cross-peak H. Therefore, Ala-Hyp-polysaccharide-2 contained two side-chain units of α-L-Rha*p*-(1→4)-β-D-GlcUA*p*-(1→6)-β-D-Gal.

Judging by the C/H chemical shifts at δ 82.3/4.21 ppm in the HSQC spectrum (supplemental Fig. 8 and Table V), the two α-L-Ara*f* residues were terminal residues. Cross-peak C in the HMBC spectrum indicated two side-chain units of α-L-Ara*f*-(1→3)-β-D-Gal.

NMR Structure Calculations of the Hyp-AGs—Model structures of interferon Hyp-polysaccharide-1 were generated by simulated annealing, and the 10 best models were further refined in explicit water and ion environment. These models, consistent with NOE distance information obtained at 25 °C (supplemental Table III), depict an overall folded structure with backbone residues Gal-6 and Gal-1 in proximity forming a sharp bend or "reverse turn" formed by the β-1,6-link between repetitive β-1,3-linked trigalactosyl subunits. We note that some conformational flexibility was seen in the NMR ensemble without violating the experimental distance information (Fig. 6). Significantly, ~10 intramolecular H-bonds stabilized interactions in which uronic acid carboxyls appeared close enough to chelate Ca^{2+}, a possibility supported by NOE-restrained molecular dynamics simulation of interferon Hyp-polysaccharide-1 in explicit water and Ca^{2+} (supplemental Fig. 10). The chelated ion remained strongly bound to the uronic acids without violating the available NOE distance data throughout the simulation. As an additional test of the global properties of the structural models, we calculated the hydrodynamic radius of the structures as 13.9 ± 0.3 Å.

DISCUSSION

Complete elucidation of type II arabinogalactan structure has been a major goal since the isolation of Hyp-AGs more than 30 years ago (5, 6). Churms *et al.* (16, 27, 28) notably suggested a repetitive structure with "AG substituents showing blocks of 1,3–1inked galactan backbone interrupted by periodate susceptible residues (kinked region)." The structure of Ala-Hyp-polysaccharide-1 supported the repetitive subunit hypothesis and also suggests that the arabinogalactan structure is highly conserved in both classical AGPs and the less abundant chimeric glycoproteins of the cell surface (29). As Ala-Hyp-polysaccharide-1 was only the first Hyp-AG (15), we characterized additional Hyp-AGs to test the possibility that AGs are highly varied structures dictated by regional peptide sequence and too complex for structural elucidation. We expressed the two most frequently occurring AGP motifs Ser-Hyp and Ala-Hyp in tobacco Bright Yellow-2 cells as chimeric fusion glycoproteins with non-glycosylated partners interferon α2b and green fluorescence protein, respectively. This allowed comparison of Hyp-AGs isolated from Ser-Hyp and Ala-Hyp repeats in the fusion proteins interferon α2-$(Ser\text{-}Hyp)_{20}$ (15, 30) and $(Ala\text{-}Hyp)_{51}$-GFP.

Interferon α2-$(Ser\text{-}Hyp)_{20}$ yielded Hyp-AGs that ranged from 18 to 26 residues. Interferon Hyp-polysaccharide-1 was the largest characterized and contained 22 residues corresponding to the composition Gal_{10} Ara_8 $GlcUA_2$ Rha_2. Hyp-AGs isolated from $(Ala\text{-}Hyp)_{51}$-GFP were marginally smaller and somewhat more disperse (14–25 residues). Significantly, all the newly characterized Hyp-AGs (Fig. 5) had a six-residue galactan main chain consisting of two β-1,3-linked trigalactosyl blocks linked by a β-1,6 linkage. Ala-Hyp-polysaccharide-1 also shared this main-chain structure except that it lacked the sixth (terminal) Gal residue (15).

The side-chain substituents linked to the galactan main chain were similar in position and composition. The small Hyp-AGs, Ala-Hyp-polysaccharide-1, and Ala-Hyp-polysaccharide-2 contained only two side chains attached to main-chain Gal residues G_1 and G_2 (15) and differed from each other only in their Rha and Ara content (Fig. 5). In contrast, the larger Hyp-AGs, interferon Hyp-polysaccharide-1 and interferon Hyp-polysaccharide-2 each had four side chains ranging from two to six glycosyl residues linked to G_1, G_2, G_4, and G_5 (Figs. 4 and 5). These three new Hyp-AGs and Ala-Hyp-polysaccharide-1 shared an ~15-residue subunit consisting of a repetitive trisaccharide with two bifurcated acidic side chains (Figs. 4–6), each with a maximum of six residues; these side chains are attached to C-6 of G_1 and G_2, *i.e.* the first and second Gal residues, numbered from the reducing end of the galactan backbone.

This relatively invariant Hyp-AGs structure with bifurcated side chains is apparently widespread; it is consistent with compositional data from diverse species (16, 27, 28) and with the fact that AGPs from diverse species selectively co-precipitate with the β-Yariv reagent. The six-residue side chain of interferon Hyp-polysaccharide-1 is identical with the gum arabic side chain in the legume *Acacia senegal* (31) except for the addition of terminal 5-linked Ara*f* in interferon Hyp-polysaccharide-1. Furthermore, neither the type of Hyp-AGs peptide motif (Ser-Hyp *versus* Ala-Hyp) nor the attached non-glycosylated domain (interferon *versus* GFP) affected the composition or structure of this 15-residue glycan subunit, again consistent with a highly conserved structure. The "type II" (*i.e.* β-1,3-linked) Hyp-AGs represented here by interferon Hyp-polysac-

charide-1, interferon Hyp-polysaccharide-2, Ala-Hyp-polysaccharide-2 (Figs. 4 and 5), and Ala-Hyp-polysaccharide-1 (15) confirm that tobacco Bright Yellow-2 cell Hyp-AGs consist of small ~15-residue subunit repeats with a common composition and linkage pattern differing mainly in the number of trigalactosyl subunits decorated with two side chains each composed of up to 6 residues. Thus, larger Hyp-AGs of up to 150 sugar residues (5, 6) may consist of ~10 repetitive subunits. There are variations on the major 15-residue theme. For example, some *Arabidopsis* Hyp-AGs lack rhamnose (32) and likely contain fucose; furthermore, some monosaccharide residues may be modified by acetylation (33) or 4-*O*-methylation (34).

Although interferon Hyp-polysaccharide-1, interferon Hyp-polysaccharide-2, and Ala-Hyp-polysaccharide-2 each had a complete 15-residue subunit, *i.e.* a galactose trisaccharide with 6-residue bifurcated side chains, these Hyp-AGs also contained incomplete subunits with truncated side chains, for example (α-L-Ara-(1→3)-Gal) in interferon Hyp-polysaccharide-1 and (β-D-GlcUA-(1→6)-Gal) in interferon Hyp-polysaccharide-2. These variations, particularly the incomplete Ala-Hyp-polysaccharide-1 Gal backbone, argue for biosynthesis of Hyp-AGs via stepwise saccharide addition to an AGP polypeptide or alternatively *en bloc* transfer of incomplete lipid-linked arabinogalactan intermediates (15) or even sugar trimming after transfer, although compelling evidence for degradative turnover of AGPs has not yet been observed. The existence of AGP microheterogeneity like that shown here may account for notions of Hyp-AGs as intractably complex. However, a small Hyp-arabinogalactan containing only 4 different sugars and 7 different glycosidic linkages is a relatively simple structure compared, for example, with the complex rhamnogalacturonan-II pectic polysaccharide with 12 different sugar residues linked by more than 20 different glycosidic linkages (35).

Signaling functions and roles as determinants of cell fate dominate current AGP discussion. However, the location of classical AGPs at the cell surface and their sheer physical abundance (3) predict primarily structural functions. Structural conservation implies that Hyp-AGs play similar conserved roles in both classical AGPs and AGP chimeras (3, 29). Nevertheless, despite considerable speculation (4, 36), specific biological roles of classical AGPs and their Hyp-AGs remain to be elucidated. The ubiquity of AGPs at the cell surface and involvement of classical AGPs in numerous fundamental developmental processes is clear (3, 37–41), yet AGP function at molecular levels is relatively unexplored. The current structural elucidation including computer simulations support structural roles for classical AGPs as follows.

Exclusively β-1–3-linked galactans belong to the compact hollow helix polysaccharide family (42). However, interspersed β-1–6-linkages form "kinks" (17) that may be analogous to the classical reverse β-turns of polypeptides. A molecular model (Fig. 6) depicts interferon Hyp-polysaccharide-1 as a folded polysaccharide that forms a moderately compact spheroid consistent with both the NOE data and the hydrodynamic radius determined experimentally.

Modeling also revealed other conserved features of interferon Hyp-polysaccharide-1 that may contribute to its stability and molecular function. They include the general stabilizing role of ~10 intramolecular H-bonds and shrinkage of conformational space (43) by the bulky bifurcated AG side chains; these limit the conformational AG landscape particularly when restrained even further *in vivo* by their attachment to the polypeptide backbone.

A conserved core structure for type II arabinogalactans suggests a conserved function and some speculations about the precise roles played by the Hyp-arabinogalactans that decorate naturally occurring AGPs and the numerous AGP chimeras that populate the plasma membrane-cell wall interface. Indeed, in tobacco cells complete coverage of the plasma membrane by classical AGPs is likely (3).

Although the sheer abundance of classical AGPs at the cell surface suggests they (and their glycans) are structural molecules, roles for AGPs in signal transduction might arise from the composition and presentation of residues at the periphery of the glycans, the regions that also exhibit microheterogeneity. The compact folded structure deduced here for interferon Hyp-polysaccharide-1 indicates that the side chains are readily available for homophilic and heterophilic interactions and the galactan backbone somewhat less so, especially near the reducing end of the polysaccharide where the larger side chains shield the galactan backbone.

For example, although all Hyp-arabinogalactans possess a galactan backbone with arabinosyl side chains, some lack the abundant uronic acid residues prevalent in the Hyp-arabinogalactans of tobacco (44, 45) where charge repulsions of the uronic acid residues may form a compression buffer at the cell surface analogous to animal proteoglycan compression buffers (15). On the other hand, the close proximity of glucuronic acid residues may favor Ca^{2+} chelation (supplemental Fig. 10), which may be relevant to processes involving calcium signaling (45). Finally, wall AGPs may play a role in regulating cell extension, perhaps by acting as pectic plasticizers (3). The structures described here may help test these hypotheses.

Acknowledgments—We gratefully acknowledge the School of Life Sciences and the University of Sussex. We are also grateful to Charles Cottrell of the Ohio State University Campus Chemical Instrument Center for contributions to the NMR analyses.

REFERENCES

1. Cannon, M. C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B. L., Chen, L., Lamport, D. T., Chen, Y., and Kieliszewski, M. J. (2008) *Proc. Natl. Acad Sci.* **105,** 2226–2231
2. Lamport, D. T. (1970) *Annu. Rev. Plant Physiol.* **21,** 235–270
3. Lamport, D. T., Kieliszewski, M. J., and Showalter, A. M. (2006) *New Phytol.* **169,** 479–492
4. Seifert, G. J., and Roberts, K. (2007) *Annu. Rev. Plant Biol.* **58,** 137–161
5. Pope, D. G. (1977) *Plant Physiol.* **59,** 894–900
6. Lamport, D. T. A. (1977) in *Recent Advances in Phytochemistry* (Loewus, F. A., and Runeckles, V. C., eds) pp. 79–115, Plenum Publishing Corp., New York
7. Lamport, D. T. (1967) *Nature* **216,** 1322–1324
8. Lamport, D. T., and Miller, D. H. (1971) *Plant Physiol.* **48,** 454–456
9. Kieliszewski, M., de Zacks, R., Leykam, J. F., and Lamport, D. T. A. (1992) *Plant Physiol.* **98,** 919–926
10. Kieliszewski, M. J., and Lamport, D. T. A. (1994) *Plant J.* **5,** 157–172
11. Kieliszewski, M. J. (2001) *Phytochemistry* **57,** 319–323
12. Tan, L., Leykam, J. F., and Kieliszewski, M. J. (2003) *Plant Physiol.* **132,** 1362–1369

13. Shpak, E., Leykam, J. F., and Kieliszewski, M. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96,** 14736–14741
14. Zhao, Z. D., Tan, L., Showalter, A. M., Lamport, D. T., and Kieliszewski, M. J. (2002) *Plant J.* **31,** 431–444
15. Tan, L., Qiu, F., Lamport, D. T., and Kieliszewski, M. J. (2004) *J. Biol. Chem.* **279,** 13156–13165
16. Churms, S. C., Stephen, A. M., and Siddiqui, I. R. (1981) *Carbohydr. Res.* **94,** 119–122
17. Fincher, G. B., Stone, B. A., and Clarke, A. E. (1983) *Annu. Rev. Plant Physiol.* **34,** 47–70
18. Xu, J., Tan, L., Goodrum, K. J., and Kieliszewski, M. J. (2007) *Biotechnol. Bioeng.* **97,** 997–1008
19. Serpe, M. D., and Nothnagel, E. A. (1999) *Adv. Bot. Res.* **30,** 207–289
20. Tan, L. (2003) *O-Glycosylation Motifs in Arabinogalactan-Proteins*, Ph.D. dissertation, Ohio University
21. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
22. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR* **4,** 603–614
23. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., and Kollman, P. A. (2008) *AMBER 10*, University of California, San Francisco., University of California, San Francisco
24. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) *J. Comput. Chem.* **29,** 622–655
25. García De La Torre, J., Huertas, M. L., and Carrasco, B. (2000) *Biophys. J.* **78,** 719–730
26. Xu, J., Shpak, E., Gu, T., Moo-Young, M., and Kieliszewski, M. J. (2005) *Biotechnol. Bioeng.* **90,** 578–588
27. Churms, S. C., Merrifield, E. H., and Stephen, A. M. (1983) *Carbohydr. Res.* **123,** 267–279
28. Churms, S. C., and Stephen, A. M. (1984) *Carbohydr. Res.* **133,** 105–123
29. Borner, G. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003) *Plant Physiol.* **132,** 568–577
30. Hill, J. M., Alewood, P. F., and Craik, D. J. (2000) *Eur. J. Biochem.* **267,** 4649–4657
31. Defaye, J., and Wong, E. (1986) *Carbohydr. Res.* **150,** 221–231
32. Xu, J., Tan, L., Lamport, D. T., Showalter, A. M., and Kieliszewski, M. J. (2008) *Phytochemistry* **69,** 1631–1640
33. Serpe, M. D., and Nothnagel, E. A. (1995) *Plant Physiol.* **109,** 1007–1016
34. Tsumuraya, Y., Ogura, K., Hashimoto, Y., Mukoyama, H., and Yamamoto, S. (1988) *Plant Physiol.* **86,** 155–160
35. O'Neill, M. A., Ishii, T., Albersheim, P., and Darvill, A. G. (2004) *Annu. Rev. Plant Biol* ***55,*** 109–139
36. Gaspar, Y., Johnson, K. L., McKenna, J. A., Bacic, A., and Schultz, C. J. (2001) *Plant Mol. Biol.* **47,** 161–176
37. Chaves, I., Regalado, A. P., Chen, M., Ricardo, C. P., and Showalter, A. M. (2002) *Physiol. Plant.* **116,** 548–553
38. Guan, Y., and Nothnagel, E. A. (2004) *Plant Physiol.* **135,** 1346–1366
39. Takahata, K., Takeuchi, M., Fujita, M., Azuma, J., Kamada, H., and Sato, F. (2004) *Plant Cell Physiol.* **45,** 1658–1668
40. Gens, J. S., Fujiki, M., and Pickard, B. G. (2000) *Protoplasma* **212,** 115–134
41. Pennell, R. I., Knox, J. P., Scofield, G. N., Selvendran, R. R., and Roberts, K. (1989) *J. Cell Biol.* **108,** 1967–1977
42. Rees, D. A. (1977) *Polysaccharide Shapes*, pp. 1–80, Chapman & Hall, London
43. Xia, J., Daly, R. P., Chuang, F. C., Parker, L., Jensen, J. H., and Margulis, C. J. (2007) *J. Chem. Theory Comput.* **3,** 1629–1643
44. Bacic, A., Churms, S. C., Stephen, A. M., Cohen, P. B., and Fincher, G. B. (1987) *Carbohydr. Res.* **162,** 85–93
45. Gane, A. M., Craik, D., Munro, S. L., Howlett, G. J., Clarke, A. E., and Bacic, A. (1995) *Carbohydr. Res.* **277,** 67–85