

## Two Novel Types of O-Glycans on the Mugwort Pollen Allergen Art v 1 and Their Role in Antibody Binding\*

Received for publication, September 10, 2004, and in revised form, November 19, 2004  
Published, JBC Papers in Press, December 10, 2004, DOI 10.1074/jbc.M410407200

Renaud Leonard<sup>‡</sup>, Bent O. Petersen<sup>§</sup>, Martin Himly<sup>¶</sup>, Waltraud Kaar<sup>‡</sup>, Nicole Wopfner<sup>¶</sup>,  
Daniel Kolarich<sup>‡</sup>, Ronald van Ree<sup>||</sup>, Christof Ebner<sup>\*\*</sup>, Jens Ø. Duus<sup>§</sup>, Fátima Ferreira<sup>||</sup>,  
and Friedrich Altmann<sup>‡</sup> <sup>‡‡</sup>

From the <sup>‡</sup>Division of Biochemistry, Department of Chemistry, Universitaet fuer Bodenkultur Wien, Muthgasse 18, A-1190 Vienna, Austria, the <sup>§</sup>Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Denmark, the <sup>¶</sup>Division of Allergy and Immunology, Department of Molecular Biology, University of Salzburg, Hellbrunnerstrasse 34, A-5020 Salzburg, Austria, the <sup>||</sup>Department of Immunopathology, Sanquin Research at CLB, Plesmanlaan 125, 1066 CX, Amsterdam, Netherlands, and the <sup>\*\*</sup>Institute of Pathophysiology, Medical University of Vienna, Spitalgasse 23, A-1090 Wien, Austria

Art v 1, the major allergen of mugwort (*Artemisia vulgaris*) pollen contains galactose and arabinose. As the sera of some allergic patients react with natural but not with recombinant Art v 1 produced in bacteria, the glycosylation of Art v 1 may play a role in IgE binding and human allergic reactions. Chemical and enzymatic degradation, mass spectrometry, and 800 MHz <sup>1</sup>H and <sup>13</sup>C nuclear magnetic resonance spectroscopy indicated the proline-rich domain to be glycosylated in two ways. We found a large hydroxyproline-linked arabinogalactan composed of a short  $\beta$ 1,6-galactan core, which is substituted by a variable number (5–28) of  $\alpha$ -arabinofuranose residues, which form branched side chains with 5-, 2,5-, 3,5-, and 2,3,5-substituted arabinoses. Thus, the design of the Art v 1 polysaccharide differs from that of the well known type II arabinogalactans, and we suggest it be named type III arabinogalactan. The other type of glycosylation was formed by single (but adjacent)  $\beta$ -arabinofuranoses linked to hydroxyproline. In contrast to the arabinosylation of Ser-Hyp<sub>4</sub> motifs in other hydroxyproline-rich glycoproteins, such as extensins or solanaceous lectins, no oligo-arabinosides were found in Art v 1. Art v 1 and parts thereof produced by alkaline degradation, chemical deglycosylation, proteolytic degradation, and/or digestion with  $\alpha$ -arabinofuranosidase were used in enzyme-linked immunosorbent assay and immunoblot experiments with rabbit serum and with the sera of patients. Although we could not observe antibody binding by the polysaccharide, the single hydroxyproline-linked  $\beta$ -arabinose residues appeared to react with the antibodies. Mono- $\beta$ -arabinosylated hydroxyproline residues thus constitute a new, potentially cross-reactive, carbohydrate determinant in plant proteins.

Art v 1, the major allergen from the pollen of mugwort (*Artemisia vulgaris*), which constitutes a prominent allergen in the temperate climate zone, has recently been characterized and sequenced, and its cDNA has been cloned (1). It is a

two-domain protein with a defensin-like, disulfide bond-rich domain and a proline-rich domain with much of the proline being hydroxylated. The hydroxyproline (Hyp)<sup>1</sup>-rich domain carries substantial amounts of galactose and arabinose, a part of which forms a type of arabinogalactan (1). The mature polypeptide of Art v 1 has a theoretical mass of 10.8 kDa. MALDI mass spectrometry revealed it to consist of a series of isoforms spaced by the mass of an arabinose residue, with one smaller set of peaks around 13 and a second set around 15 kDa (1). Remarkably, natural and recombinant Art v 1 exhibited different IgE binding capacities, i.e. a subgroup of patients contained partly or mainly IgE binding only to the natural allergen (1). Different *in vivo* allergenic potentials were shown by skin prick and nasal provocation tests (2). Although various posttranslational modifications might be considered, we suspected the glycosylation state of the allergen to be the primary cause of these differences. O-Glycosylated plant proteins belong to the hydroxyproline-rich glycoproteins with the two main members being arabinogalactan proteins (AGP(s)) and extensins (3–5). AGPs contain large arabinogalactan polysaccharides where the glycan part often is bigger than the protein (4, 6–8). The carbohydrate part of AGPs consists of a hydroxyproline-linked  $\beta$ 1,3-D-galactan backbone having  $\beta$ 1,6-D-galactan side chains or kinks, which in turn are substituted by arabinoses and other less abundant sugar residues (8–10). These type II arabinogalactans (in contrast to the type I with  $\beta$ 1,4 linkages (8)) typically contain more galactose than arabinose. The latter may occur as a single furanose residue linked to galactose or in very short chains with mainly  $\alpha$ 1,5 and maybe  $\alpha$ 1,3 linkages (2, 3, 8, 9, 11).

Extensins and solanaceous lectins exhibit two different types of O-glycosylation, single  $\alpha$ -galactosyl residues linked to Ser residues (11, 12) and short arabinan chains linked directly to hydroxyproline (5, 8, 10, 13–15). Such arabinans have been reported to consist of 1–5 arabinofuranose residues, which are joined by  $\beta$ 1,2 and  $\beta$ 1,3 linkages (13, 15–17). The same protein

\* This work was supported by the Joint Research Project S88-MED (S8802, S8803, S8808) of the Austrian Science Fund. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>‡‡</sup> To whom correspondence should be addressed: Dept. of Chemistry, Universität für Bodenkultur Wien, Muthgasse 18, A-1190 Wien, Austria. Tel.: 43-1-36006-6062; Fax: 43-1-36006-6059; E-mail: friedrich.altmann@boku.ac.at.

<sup>1</sup> The abbreviations used are: Hyp, hydroxyproline; AGP, arabinogalactan protein; nArt v 1, natural Art v 1; rArt v 1, recombinant Art v 1; PTM, posttranslational modification; PS, polysaccharide; MALDI-TOF, matrix-assisted laser desorption time-of-flight; deglyc-nArt v 1, deglycosylated nArt v 1; dearagal-nArt v 1,  $\alpha$ -arabinofuranosidase and  $\beta$ -galactosidase-digested nArt v 1; dearagal-nArt v 1-tryp, tryptic glycopeptide of dearagal-nArt v 1; ELISA, enzyme-linked immunosorbent assay; HPLC, high pressure liquid chromatography; Fmoc, N-(9-fluorenyl)methoxycarbonyl; NOE, nuclear Overhauser effect; HSQC, heteronuclear single quantum correlation; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; HMBC, heteronuclear multiple bond correlation.

may carry both arabinosides and arabinogalactans (5, 14, 18). The type of glycan synthesized at a given Hyp residue appears to be determined by the peptide sequence (5, 15, 19). Contiguous Hyp residues form attachment sites for arabino-oligosaccharides, whereas non-contiguous, clustered Hyp residues bear arabinogalactan polysaccharides (5, 10).

Type II arabinogalactans are immunogenic, and antibodies against arabinogalactans have been employed to study the location and function of AGPs (4, 20–22). Several monoclonal antibodies have been characterized as binding to arabinose-containing epitopes (22, 23). A monoclonal antibody against linear  $\alpha$ 1,5-arabinan oligosaccharides was reported to bind with a nanomolar affinity (24). Often structural elements other than arabinose are responsible for antibody binding by sera against natural AGPs (25, 26). In the present study we report on the structure of the arabinogalactan polysaccharide and on the unexpected finding of a new type of *O*-glycosylation of the mugwort (*A. vulgaris*) pollen allergen Art v 1 as well as on the antibody binding abilities of these posttranslational modifications.

#### MATERIALS AND METHODS

**Allergens, Antibodies, and Other Materials**—Natural Art v 1 (nArt v 1) and recombinant Art v 1 (rArt v 1) were obtained as described recently (1). For immunological characterization, sera of 13 mugwort pollen-allergic patients selected by typical case history, positive skin test, and radioallergosorbent test class >3.0 were used. All sera reacted specifically with nArt v 1 in immunoblots or ELISA. Serum from a non-allergic healthy subject was used for control. Rabbit antiserum was raised against nArt v 1, and a fraction specific for the posttranslational modifications of nArt v 1 was selected (rabbit anti-PTM) (1).

$\beta$ -Glucosyl-Yariv reagent and gum arabic were obtained from Biosupplies (Victoria, Australia). Sugar beet arabinan and *Aspergillus niger*  $\alpha$ -L-arabinofuranosidase were obtained from Megazyme (Wicklow, Ireland). Potato, tomato, and *Datura stramonium* lectins were bought from Sigma.

**Degradations of nArt v 1**—For alkaline degradation, nArt v 1 was incubated with 0.22 M Ba(OH)<sub>2</sub> for 6 h at 100 °C (27). Glycopeptides were recovered by passage over Sephadex G50 fine in 25 mM ammonium acetate at pH 6.0, and fractions were analyzed for carbohydrate (28). Mild acid degradation of glycopeptides was achieved by incubation with 0.2 M trifluoroacetic acid at 80 °C for 60 min. Enzymatic digestion of  $\alpha$ -arabinosyl residues was performed in 50 mM ammonium acetate of pH 4.0 at 37 °C with *A. niger*  $\alpha$ -arabinofuranosidase (0.3 units/nmol of substrate).  $\beta$ -Galactosidase from *Aspergillus oryzae* (0.05 units/nmol) was subsequently added (29). Degraded glycopeptide was purified as above but with Sephadex G15.

Larger glycopeptides from nArt v 1 were obtained by digestion with trypsin in the absence or presence of dithiothreitol (28). Chemical deglycosylation was done with trifluoromethanesulfonic acid in the presence of anisole for 1 h at –20 °C (30). Precipitation of nArt v 1 and gum arabic, as a control, was performed as described (31).

**Analytical Methods**—Monosaccharides were analyzed by gas chromatography/mass spectroscopy of alditol acetates (32) and by HPLC (33) after hydrolysis with 4 M trifluoroacetic acid at 100 °C for 2 h. Hydroxyproline was determined by HPLC as its Fmoc (34) or dansyl derivative (35). Trans-4-hydroxy-L-proline and cis-4-hydroxy-L-proline (Sigma) were used as standards.

MALDI-TOF mass spectrometry of Hyp-linked polysaccharide and of nArt v 1 was carried out using a Dynamo mass spectrometer (Thermo-BioAnalysis, Ltd.) in the delayed extraction mode (36). The instrument was externally calibrated against partially hydrolyzed dextran, bee venom melittin, or porcine insulin.

Dansylated Ara-Hyp was analyzed by reversed phase HPLC coupled to electrospray ionization mass spectrometry on a Waters-Micromass Q-ToF Ultima Global instrument equipped with a 0.18  $\times$  150-mm Bio-Basic18 column (Thermo Hypersil) and a 0.35  $\times$  5-mm Opti-Pak<sup>TM</sup> C18 precolumn (Waters). A gradient from 5 to 40% acetonitrile in 0.1% formic acid was formed in 40 min at a flow rate of 2  $\mu$ l/min.

NMR spectra were recorded at 25 °C in 3- or 1.7-mm tubes in D<sub>2</sub>O in a 3-mm HCN probe at a Varian Unity Inova 800 instrument at 799.96 MHz for proton and 201.12 MHz for carbon, using acetone as reference for proton (2.225 ppm) and 1,4-dioxane for carbon (67.4 ppm). Varian standard programs tndqcoy, tnnocoy (mixing time of 100 ms), tntocoy

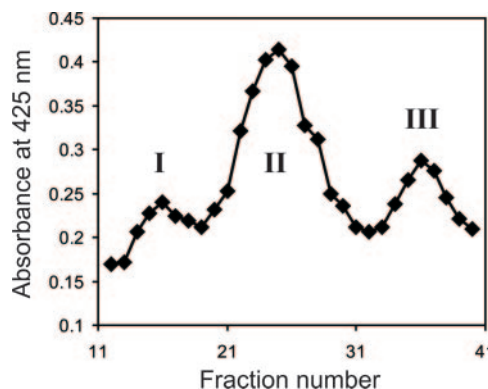


FIG. 1. Gel filtration of alkali degraded nArt v 1. The Ba(OH)<sub>2</sub> hydrolysate was separated over Sephadex G50 fine. Fractions were analyzed for carbohydrate and pooled as indicated.

(spinlock time 80 ms), gHSQC, gHSQCTOCSY (spinlock time 80 ms), gHSQCNOESY (mixing time 200 ms), and gHMBC were used with digital resolution in F2 dimension <2 Hz/pt (37). Processed spectra were assigned using the computer program PRONTO (38).

**Immunological Methods**—Immunoblots with rabbit anti-PTM were performed with 0.1  $\mu$ g of protein/lane as described (36). IgE blots were done as described previously (1). In certain cases, binding to *N*-glycans was suppressed by the addition of bovine serum albumin-conjugated bromelain glycopeptide (10  $\mu$ g/ml) to the diluted serum (39).

ELISA with the sera of the patients was performed as described previously (36) with minor modifications. Tris-buffered saline was used instead of phosphate-buffered saline, and 3% dry milk powder was used instead of bovine serum albumin. For coating, 0.6 ng of protein were applied to each well. Color development was performed between 2 and 16 h. ELISA with rabbit anti-PTM was performed in the same way but with a development time of 1 h.

In the case of inhibition experiments, the sera were preincubated with the substance to be tested for 1 h at 37 °C. All ELISA measurements were done in triplicate.

Periodate oxidation of nArt v 1 was performed with 10 mM sodium metaperiodate in 40 mM sodium acetate buffer of pH 4.2 at 4 °C. An incubation time of 7 h was found to be just sufficient to fully destroy the reactivity with rabbit anti-PTM.

#### RESULTS

**Overview on the Structural Analysis**—nArt v 1 contained Gal and Ara in a ratio of 1:11.4. To remove the protein part, the arabinogalactan protein Art v 1 was subjected to alkaline hydrolysis, which leaves intact Hyp-linked oligosaccharides while it destroys sugars linked to serine or threonine (27). Gel filtration of the alkali digest of Art v 1 yielded three peaks (Fig. 1). The first one presumably consisted of incompletely digested material. The second, most intense, peak constituted a Hyp-linked arabinogalactan, which will be referred to as Hyp-polysaccharide (Hyp-PS). A smaller arabinose-containing peak in the inclusion volume of the column will be referred to as  $\beta$ -Ara-Hyp.

Arabinogalactan proteins react with Yariv reagent (31). Indeed, nArt v 1 precipitated with  $\beta$ -glucosyl Yariv reagent as judged from dissolution and semi-quantitative SDS-PAGE of the pellet. Thus, Art v 1 can be called an AGP, although it will be shown to be an atypical one.

**Analysis of the Hyp-polysaccharide**—The glycopeptide obtained from Art v 1 by alkaline hydrolysis contained arabinose and galactose in a molar ratio of 5.5:1 (1). The Hyp-PS pool contained mainly 4-trans-Hyp accompanied by 4-cis-Hyp, an isomerization product of trans-4-Hyp (13). The ratio of Hyp to Gal was 1:2.9, which fits well to the masses reported previously where the Hyp-glycan was shown to consist of a Hyp<sub>1</sub>Gal<sub>3</sub>Ara<sub>5–28</sub> series (Fig. 2) (1). Linkage analysis by permethylation indicated the Hyp-glycan to contain six major substitution types, terminal, 5-substituted, 2,5-, or 3,5-disubstituted and even 2,3,5-trisubstituted arabinofuranosyl residues,

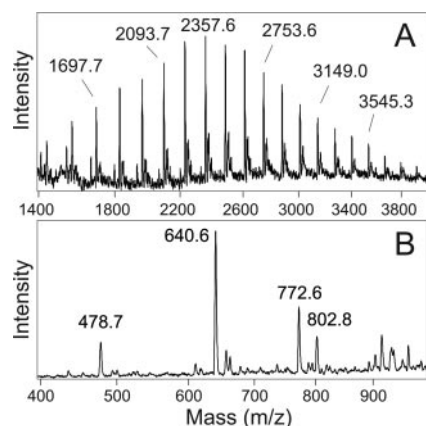


FIG. 2. MALDI mass spectrometry of the hydroxyproline-linked polysaccharide before and after mild acid hydrolysis. Spectra of complete (A) and degraded Hyp-polysaccharide (B) were acquired on the Dynamo MALDI-TOF mass spectrometer with gentisic acid as the matrix. The masses in A represent  $[M+Na]^+$  ions of mainly Hyp-Gal<sub>3</sub> plus a varying number of Ara residues (1), in B Hyp-Gal<sub>2</sub>, Hyp-Gal<sub>3</sub>-Ara, and Hyp-Gal<sub>4</sub> can be observed in addition to the major peak representing Hyp-Gal<sub>3</sub>.

on one hand and 3,6-disubstituted galactopyranosyl-residues on the other (Fig. 3). The ring forms are deduced from spectra obtained for the C1-deuterated partially methylated alditol acetates of terminal arabinose and terminal galactose. These assumptions were later corroborated by enzymatic digestions and NMR. Permethylated analysis thus indicated considerable branching of the Hyp-glycan. Remethylation and control experiments did not point at undermethylation. An approximate quantitation of the substitutions can be deduced from Fig. 3 where, however, the amount of terminal Ara is certainly underestimated due to the volatility of its derivative. Permethylated analysis of a Pronase glycopeptide of Art v 1 resulted in a considerably larger peak of terminal furanosidic arabinose. This difference later turned out not to be just a fluctuation in the yield of the derivative of terminal Ara but to result from Ara residues directly linked to Hyp (see below).

Limited acid hydrolysis of the highly disperse polysaccharide (1) led to the disappearance of large glycans and to an accumulation of a peak of mass 640 Da which corresponds to trigalactosyl-Hyp (Fig. 2). Smaller peaks at 478 and 802 Da indicated the occurrence of additional di- and tetragalactosyl series. Similar results were obtained when Hyp-PS was incubated with  $\alpha$ -L-arabinofuranosidase. These data imply the Art v 1 Hyp-PS consists of an inner galactosyl-core to which arabinan chains composed of mainly 3–5  $\alpha$ -L-arabinofuranosyl residues are attached.

The galactan was sensitive to fungal  $\beta$ -D-galactosidase, which, notably, is essentially inactive toward  $\beta$ 1,3-linked galactosides (29). The formation of free Hyp as found by HPLC after Fmoc derivatization indicated that the innermost galactose residues also had been of the  $\beta$ -configuration. NMR and linkage analyses (Fig. 3) indicated the D-galactoses to be  $\beta$ 1,6 linked and not  $\beta$ 1,3 as found in type II arabinogalactans.

**NMR Spectroscopy of Hyp-polysaccharide from nArt v 1**—Extensive two-dimensional NMR spectroscopy of Hyp-PS was carried out at high field (800 MHz for  $^1H$ ). State of the art two-dimensional approaches, such as HSQC-TOCSY and HSQC-NOESY, were used, and assignments were based on well established methods as reviewed by Duus *et al.* (37). Information from all the available two-dimensional data were combined with published spectral data. The assignments described were initially performed on the isolated complex mixture of Hyp-PS but subsequently confirmed for the intact glycoprotein. The data given in Table I are all for the Hyp-PS, but

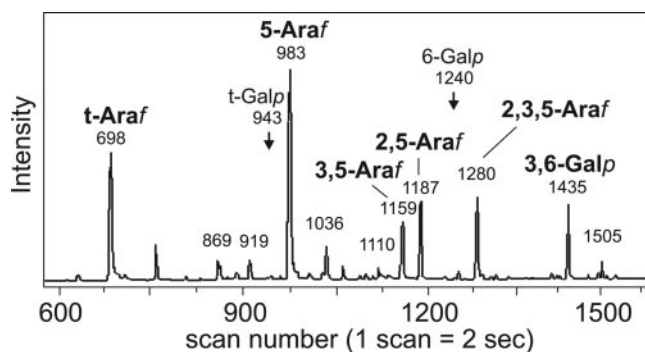


FIG. 3. Linkage analysis of the Hyp-linked polysaccharide from Art v 1. Total ion current of the gas chromatogram of the partially methylated alditol acetates from Hyp-glycan showing peaks for terminal (*t*-), 5-substituted, and disubstituted arabinofuranosidase as implied by the mass spectra. The arrows indicate the positions and nature of the two peaks in the chromatogram obtained with dearabinosylated Hyp-glycan (free galactan backbone).

very similar chemical shifts were observed for the PS linked to the intact protein.

The structure of the backbone of the PS could clearly be assigned to be  $\beta$ -galactosidase-(1–6)- $\beta$ -galactosidase-(1–6)- $\beta$ -galactosidase-(1-*O*)-Hyp. The anomeric  $^1H$  signals at 4.65, 4.47, 4.52, and 4.45 ppm in an approximate ratio of 1:1:0.7:0.3, respectively, could be assigned to  $\beta$ -galactosidase from the chemical shifts, from the  $^1H$ - $^1H$  three bond coupling constants ( $J_{1,2} \cong 7$  Hz,  $J_{2,3} \cong 10$  Hz,  $J_{3,4} \cong 3$  Hz) and from NOE correlations between H1–H3 and H1–H5 (37, 40). The clear NOEs to H- $\gamma$  of Hyp, 4.74 and 4.66 ppm, respectively, indicated the  $\beta$ -galactosidase with anomeric signals at 4.52 and 4.45 ppm was linked to two types of Hyp. The chemical shifts of the two types of Hyp could be fully assigned, but the configuration (*cis* or *trans*) could not be determined. Normally in peptide NMR (41) the *cis*- or *trans*-configuration is assigned based on NOE to the adjacent peptide H $\alpha$ , which is not present in an isolated Hyp.

The two remaining  $\beta$ -galactosidase residues appeared to be linked 1–6 to another galactosidase by NOE connectivities. The chemical shifts assigned for all the  $\beta$ -galactosidase moieties showed these to be substituted both at the 3 and 6 positions, and the residues attached to those positions could be identified by HMBC and NOE correlations.

The assignment of the arabinofuranose residues was hampered by overlap because of the high heterogeneity of the sample, which contained a series of structures ranging from 5 to 28 Ara residues at the extreme ranges, as judged by mass spectrometry (Fig. 3). The anomeric signals for a series of different residues also could be identified and many of these were assigned to specific linkage positions by NOE and HMBC. For example, the residues named b1 and b2 could be linked to galactosidase-3 and b3 to galactosidase-6. These residues could not all be fully assigned, but based on the  $^{13}C$  chemical shifts compared with data in the literature (42–45) of *e.g.* the C-2 (82.1 ppm), C-3 (83.1 ppm), and C-5 (67.2 ppm) of residue b1, it can be deduced that this Ara residue is substituted at C-3 and C-5 and not at C-2. In contrast, the residue b2 with an anomeric signal at 5.19 ppm can only be partly assigned, but based on the chemical shift of C-2 (85.9 ppm) b2 can be deduced to be substituted at C-2 and C-3. However, it could not be determined if this residue is substituted at C-5. therefore, it is given as (2,3,?) $\alpha$ -Ara-(1–3)- $\beta$ -galactosidase, representing both (2,3)- $\alpha$ -Ara-(1–3)- $\beta$ -galactosidase and (2,3,5)- $\alpha$ -Ara-(1–3)- $\beta$ -galactosidase. An important tool in the assignment of the substitution pattern of different plant arabinans is the  $^{13}C$  chemical shift data published by several groups (43–45) (added to Table I).

TABLE I  
Assigned NMR chemical shift for residues in the oligosaccharide

| Residue <sup>a</sup>                   | Position                 |          |           |           |          |            |                   | NOE<br>from H1    | HMBC<br>from<br>H1 | App.<br>Integral |
|----------------------------------------|--------------------------|----------|-----------|-----------|----------|------------|-------------------|-------------------|--------------------|------------------|
|                                        |                          | $\alpha$ | $\beta$ 1 | $\beta$ 2 | $\gamma$ | $\delta$ 1 | $\delta$ 2        |                   |                    |                  |
| <b>Hyp</b>                             | <sup>1</sup> H           | 4.30     | 2.58      | 2.13      | 4.74     | 3.64       | 3.51              |                   |                    | 0.7              |
|                                        | <sup>13</sup> C          | 60.7     | 36.0      |           | 79.0     | 52.7       |                   |                   |                    |                  |
| <b>Hyp</b>                             | <sup>1</sup> H           | 4.20     | 2.47      | 2.40      | 4.66     | 3.75       | 3.41              |                   |                    | 0.3              |
|                                        | <sup>13</sup> C          | 60.7     | 35.8      |           | 78.3     | 52.9       |                   |                   |                    |                  |
|                                        |                          | 1        | 2         | 3         | 4        | 5a         | 5b(6a)            | 6b                |                    |                  |
| 6) $\beta$ -Gal(1-O)Hyp                | <b>h</b> <sup>1</sup> H  | 4.52     | 3.64      | 3.81      | 4.15     | 3.88       | 3.86 <sup>b</sup> | 3.86 <sup>b</sup> | 4.74               | 0.7              |
|                                        | <sup>13</sup> C          | 102.1    | 70.6      | 82.4      | 69.4     | 74.4       | 67.4 <sup>c</sup> |                   |                    |                  |
| 6) $\beta$ -Gal(1-O)Hyp                | <b>h</b> <sup>1</sup> H  | 4.45     | 3.61      | 3.78      | 4.12     | 3.86       | 3.86 <sup>b</sup> | 3.86 <sup>b</sup> | 4.66               | 0.3              |
|                                        | <sup>13</sup> C          | 102.6    | 70.6      | 82.4      | 69.4     | 74.4       | 67.4 <sup>c</sup> |                   |                    |                  |
| 6) $\beta$ -Gal(1-6) $\beta$ -Gal      | <b>i</b> <sup>1</sup> H  | 4.47     | 3.61      | 3.68      | 4.06     | 3.88       | 3.80*             | 3.80*             | 3.86               | 1                |
|                                        | <sup>13</sup> C          | 104.3    | 70.7      | 81.1      | 69.4     | 74.4       | 67.1 <sup>c</sup> |                   |                    |                  |
| 6) $\beta$ -Gal(1-6) $\beta$ -Gal      | <b>g</b> <sup>1</sup> H  | 4.65     | 3.68      | 3.67      | 4.06     | 3.80       | 3.84 <sup>b</sup> | 3.84 <sup>b</sup> | 3.80               | 1                |
|                                        | <sup>13</sup> C          | 104.8    | 71.0      | 81.1      | 69.4     | 74.4       | 67.1 <sup>c</sup> |                   |                    |                  |
| 3,5) $\alpha$ -Ara(1-3) $\beta$ -Gal   | <b>b1</b> <sup>1</sup> H | 5.20     | 4.17      | 4.20      | 4.26     | 3.90       | 3.80              |                   | 3.68               | 81.1             |
|                                        | <sup>13</sup> C          | 110.2    | 82.1      | 83.1      | 82.2     | 67.1       |                   |                   |                    | 3.7 <sup>e</sup> |
| 2,3,?) $\alpha$ -Ara(1-3) $\beta$ -Gal | <b>b2</b> <sup>1</sup> H | 5.19     | 4.26      |           |          |            |                   |                   | 3.67               | 81.1             |
|                                        | <sup>13</sup> C          | 110.2    | 85.9      |           |          |            |                   |                   |                    | <sup>c</sup>     |
| ?) $\alpha$ -Ara(1-6) $\beta$ -Gal     | <b>b3</b> <sup>1</sup> H | 5.20     |           |           |          |            |                   |                   | 3.84               | 67.1             |
|                                        | <sup>13</sup> C          | 107.1    |           |           |          |            |                   |                   |                    | <sup>c</sup>     |
| 3) $\alpha$ -Ara(1-?)                  | <b>e</b> <sup>1</sup> H  | 5.06     | 4.24      | 4.00      |          |            |                   |                   | 82.4               | 0.8 <sup>f</sup> |
|                                        | <sup>13</sup> C          | 108.1    | 80.0      | 84.8      |          |            |                   |                   |                    |                  |
| 3) $\alpha$ -Ara(1-?)                  | <b>e</b> <sup>1</sup> H  | 5.07     | 4.24      | 4.06      |          |            |                   |                   | 83.8               | <sup>f</sup>     |
|                                        | <sup>13</sup> C          | 108.1    | 80.0      | 84.8      |          |            |                   |                   |                    |                  |
| 5) $\alpha$ -Ara(1-?)                  | <b>m</b> <sup>1</sup> H  |          |           | 3.98      | 4.17     | 3.84       | 3.75              |                   |                    | 7 <sup>d</sup>   |
|                                        | <sup>13</sup> C          |          |           | 77.5      | 83.3     | 67.7       |                   |                   |                    |                  |
| 3,5) $\alpha$ -Ara(1-3) $\alpha$ -Ara  | <b>c1</b> <sup>1</sup> H | 5.14     | 4.08      | 3.99      | 4.15     | 3.85       | 3.77              |                   | 4.20               | 83.1             |
|                                        | <sup>13</sup> C          | 107.3    | 81.9      | 84.7      | 83.3     | 67.4       |                   |                   |                    | 3.4 <sup>g</sup> |
| 2) $\alpha$ -Ara(1-5) $\alpha$ -Ara    | <b>c2</b> <sup>1</sup> H | 5.16     | 4.12      |           |          |            |                   |                   | 3.84/3.75          | <sup>g</sup>     |
|                                        | <sup>13</sup> C          | 107.2    | 87.8      |           |          |            |                   |                   |                    |                  |
| $\alpha$ -Ara(1-2) $\alpha$ -Ara       | <b>c3</b> <sup>1</sup> H | 5.13     | 4.08      | 3.92      | 4.03     | 3.78       | 3.67              |                   | 85.9               | <sup>g</sup>     |
|                                        | <sup>13</sup> C          | 107.8    | 81.9      | 77.4      | 84.7     | 61.9       |                   |                   |                    |                  |
| $\alpha$ -Ara(1-5) $\alpha$ -Ara       | <b>e</b> <sup>1</sup> H  | 5.04     | 4.09      | 3.91      | 4.05     | 3.79       | 3.67              |                   | 67.4               | 5.9              |
|                                        | <sup>13</sup> C          | 108.5    | 81.9      | 77.2      | 84.8     | 61.9       |                   |                   |                    |                  |
| $\alpha$ -Ara(1-3) $\alpha$ -Ara       | <b>d</b> <sup>1</sup> H  | 5.11     | 4.09      | 3.91      | 4.00     | 3.79       | 3.67              |                   | 4.20               | 2.8              |
|                                        | <sup>13</sup> C          | 107.7    | 81.9      | 77.4      | 84.8     | 61.9       |                   |                   |                    |                  |
| $\alpha$ -Ara(1-2)[1-3] $\alpha$ -Ara  | <b>a</b> <sup>1</sup> H  | 5.39     | 4.14      | 3.88      | 4.04     | 3.77       | 3.66              |                   | 4.12               | 0.4              |
|                                        | <sup>13</sup> C          | 109.3    | 82.3      | 77.6      | 84.8     | 61.9       |                   |                   |                    |                  |
| Arabinan data (ref. 44).               |                          |          |           |           |          |            |                   |                   |                    |                  |
| $\alpha$ -Ara                          | <sup>13</sup> C          | 108.3    | 81.7      | 77.3      | 84.7     | 62.0       |                   |                   |                    |                  |
| 5) $\alpha$ -Ara                       | <sup>13</sup> C          | 108.3    | 81.6      | 77.5      | 83.1     | 67.7       |                   |                   |                    |                  |
| 3,5) $\alpha$ -Ara                     | <sup>13</sup> C          | 108.0    | 81.9      | 83.2      | 82.4     | 67.5       |                   |                   |                    |                  |
| 2,5) $\alpha$ -Ara                     | <sup>13</sup> C          | 107.9    | 87.7      | 76.1      | 83.0     | 67.7       |                   |                   |                    |                  |
| 2,3,5) $\alpha$ -Ara                   | <sup>13</sup> C          | 107.7    | 84.7      | 83.2      | 82.1     | 67.4       |                   |                   |                    |                  |

<sup>a</sup> The chemical shifts are given for the residue in bold.

<sup>b, c</sup> The indicated chemical shifts might be interchanged.

<sup>d</sup> The integral is the sum for all 5-substituted residues.

<sup>e, f, g</sup> Within each group the integral is the sum for these residues.

In the same manner several other arabinoside residues could be identified and linked together even if the full assignment could not be obtained. For a quantitative estimation of the representative structures, the anomeric <sup>1</sup>H signals were integrated, and when these overlapped the integral represents several fragments. For a subfragment m the signals for positions 3, 4, and 5 could be identified, but it was not possible to identify the anomeric signal. Fragment m can be deduced to be 5- but clearly not 3-substituted by the <sup>13</sup>C chemical shifts of C-3 and C-5. Most likely, it is not 2-substituted either, having a C-3 <sup>13</sup>C chemical shift of 77.5 ppm. A 2 substitution is expected to bring the chemical shift up to ~76 ppm by a  $\beta$ -glycosylation shift. This 5-monosubstituted  $\alpha$ -arabinoside is integrated to ~7 residues/mol, which is in good agreement with the intense peak seen in Fig. 3 from the linkage analysis.

The chemical shifts in Table I should be regarded as description of building blocks present in the PS and not as an assignment of a specific structure. Based on the NMR data and the

information from MS and linkage analysis, a representative “average” structure can be devised as depicted in Fig. 4. The model must be considered to be tentative, because the parameters obtained by NMR and linkage analysis are averaged over a wide molecular size range. Nevertheless, we assume that this structure represents an existing, albeit small fraction of the polysaccharide, among numerous other isomers, which can arise from other combinations of the building blocks given in Table I.

**Partial Degradation of Art v 1**—Our initial hypothesis about the global structure of Art v 1 supposed it to contain either one or two O-linked polysaccharides leading to two groups of peaks at ~13 and 15 kDa (1) or, in the case of tryptic glycopeptides (nArt v 1-trypt), at ~7 and 9 kDa (Fig. 5). However,  $\alpha$ -arabinofuranosidase treatment of nArt v 1-trypt resulted in peaks ~7.3 kDa, and although subsequent  $\beta$ -galactosidase further reduced the size (Fig. 5), the product “dearagal-nArt v 1-trypt” was much larger than peptide 56–108, which has a calculated mass of

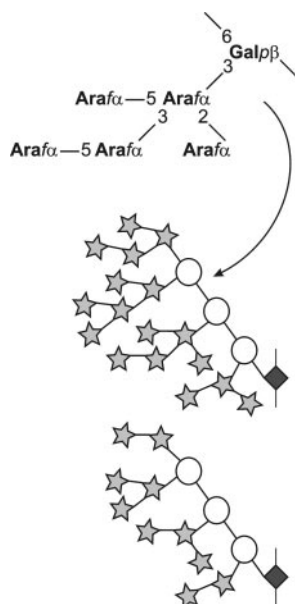


FIG. 4. **Tentative structural model of the Art v 1 polysaccharide.** At the top, an arabinan branch is depicted, which shall be understood as one of many possible combinations of the building elements detected by NMR and other techniques. Ara and Gal stand for L-arabinose and D-galactose, respectively. The schematics above arbitrarily depict polysaccharides with 17 and 10 arabinose residues. Gray stars, white circles, and the rhombus represent arabinofuranose, galactopyranose, and hydroxyproline residues, respectively.

4,846 Da if 16 Pro residues are oxidized to Hyp. Thus only ~2 kDa of glycan mass had been removed by this enzymatic treatment, which corresponds to one polysaccharide chain of the kind described above. In contrast, chemically deglycosylated nArt v 1 exhibited a peak at 4,854 Da (Fig. 5). The dominant peak at 4,704 Da may arise from loss of the C-terminal His. In other words, there was ~2.5 kDa of dark matter, *i.e.* unexplained mass, in the arabinosidase and galactosidase treated tryptic peptide of nArt v 1. Taking into account the findings described later, the difference between the peaks in Fig. 5, C and D can be interpreted as 16–17 arabinose residues within the limits of accuracy set by multiple alkali metal adducts and the probably disperse number of Hyp residues. These Ara residues must occur in a structure different from the Hyp-polysaccharide with its terminal glycosidase-sensitive  $\alpha$ -Ara residues.

**NMR Spectroscopy of the Intact nArt v 1 Glycoprotein**—In the NMR spectra of the isolated Hyp-PS described above, a set of resonances typical for dextran were present with an anomeric signal at 4.90 ppm (40). The dextran was suspected to originate from a Sephadex column used during fractionation. To prove that the dextran was not a part of nArt v 1, NMR data of the intact glycoprotein was obtained. In this case, the results showed the dextran was clearly not present but surprisingly another pronounced anomeric  $^1\text{H}$  signal showed up at ~5.12 ppm. Extensive two-dimensional  $^1\text{H}$ - $^1\text{H}$  and  $^1\text{H}$ - $^{13}\text{C}$  data were acquired and these results indicated that this signal originated from a single  $\beta$ -arabinofuranoside linked to Hyp. Overlap with the  $\alpha$ -arabinoside made the assignment somewhat difficult. Thus, to further substantiate this observation, the nArt v 1 glycoprotein was treated with  $\alpha$ -arabinofuranosidase to remove the  $\alpha$ -arabinoside residues from the polysaccharide. This made the  $\beta$ -Ara anomeric signals clearly visible even in the one-dimensional  $^1\text{H}$  spectrum (Fig. 6). Based on two-dimensional data,  $^1\text{H}$  signals could be assigned corresponding to three very similar  $\beta$ -arabinoside residues (Table II). The  $^{13}\text{C}$  chemical shifts could not be distinguished for each of the three types, and only one set of signals is given in the table. The comparison of

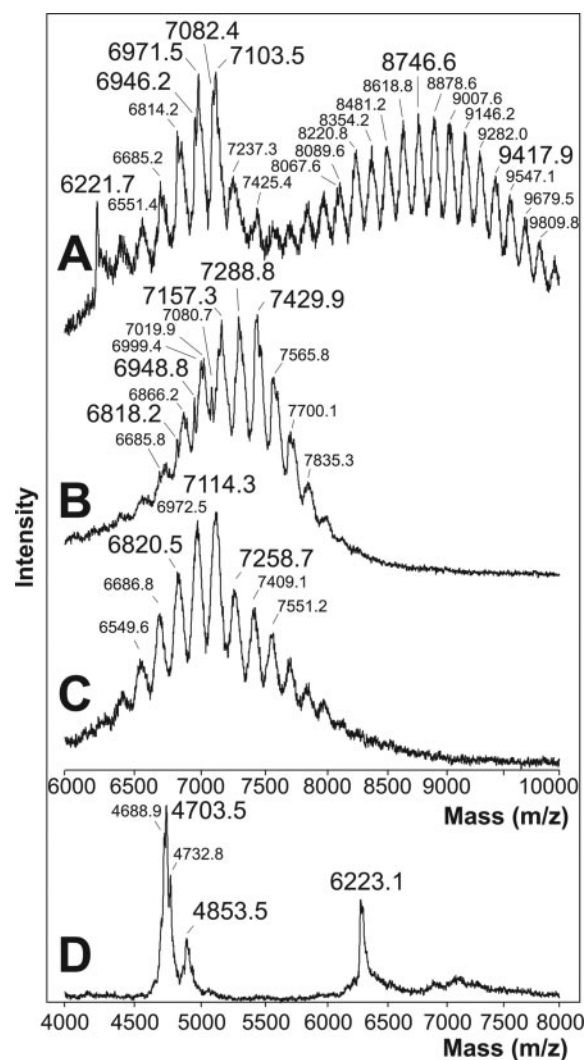


FIG. 5. **MALDI mass spectrometry of glycoforms of the Hyp-rich domain.** Tryptic glycopeptides from nArt v 1 were obtained under non-denaturing conditions and purified by gel filtration to remove most of the globular domain ( $M_r = 6220$  with 4 disulfide bridges). Fully glycosylated tryptic glycopeptide (A) is shown, after treatment with  $\alpha$ -arabinofuranosidase (B), and after additional  $\beta$ -galactosidase treatment (dearagal-nArt v 1-trypt) (C). D, tryptic peptides obtained from deglyc-nArt v 1. Mass values may be erroneous by  $\pm 0.2\%$ . Spectra A–C were acquired with gentisic acid as the matrix.

the  $^{13}\text{C}$  chemical shifts with literature values (42, 45) clearly proves that the configuration is  $\beta$ , not  $\alpha$ , like the residues present in the PS. In particular, the anomeric  $^{13}\text{C}$  signal at 100.9 ppm contrasted significantly with the values ~107–109 ppm for the  $\alpha$ -form, and the difference between  $\alpha$  and  $\beta$  is in good agreement with the general trend in furanositides (40, 42). A comparison of the  $^{13}\text{C}$  chemical shifts for positions 2, 3, and 5 with published data for  $\beta$ -arabinose (45) and with data for tri- or tetra-arabinosides linked to Hyp (16) clearly proves the  $\beta$ -Ara to be unsubstituted. The linkage between  $\beta$ -Ara and Hyp could be assigned by the NOE between the anomeric proton of  $\beta$ -Ara and the  $\gamma$ -proton of the Hyp and HMBC correlation from the anomeric proton of  $\beta$ -Ara to the  $\gamma$ -carbon of the Hyp. (Table II)

Likewise, three sets of  $^1\text{H}$  data could be assigned for the Hyp residues. As the chemical shifts of the three assigned Hyp residues are very similar, it can be deduced that these most likely have the same configuration. Hence, the difference between the three different  $\beta$ -arabinoside residues is most likely because of slight differences in their environment, *i.e.* the na-

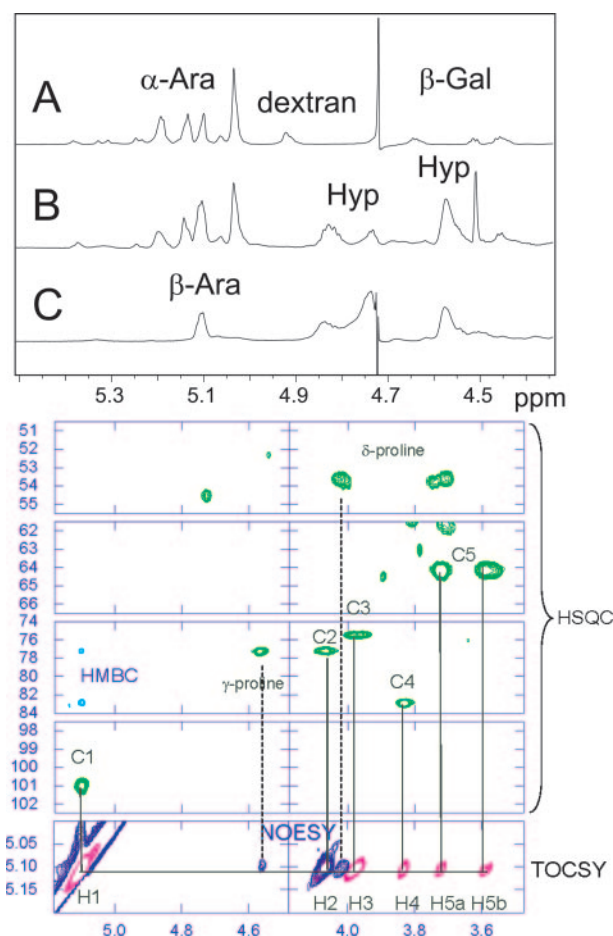


FIG. 6. NMR spectroscopic identification of  $\beta$ -arabinose attachment to nArt v 1. Top, sections of one-dimensional  $^1\text{H}$  spectra of isolated Hyp-polysaccharide (A), intact Art v 1 before (B) and after treatment with  $\alpha$ -arabinofuranosidase (C). The dextran peak in spectrum A is a contamination introduced by gel filtration. Bottom, parts of two-dimensional  $^1\text{H}$ - $^1\text{H}$  NOESY and TOCSY spectra and of  $^1\text{H}$ - $^{13}\text{C}$  correlated HSQC and HMBC spectra of  $\alpha$ -arabinofuranosidase-treated nArt v 1 protein showing the assignment of the  $\beta$ -Ara attached to Hyp.

ture and glycosylation status (if Hyp) of adjacent amino acid residues. Craik and co-workers (41, 46) have demonstrated that the difference between the  $\gamma$ - and  $\beta$ - $^{13}\text{C}$  chemical shift can be used to determine the configuration of Hyp in proteins. They observed that the difference between  $\gamma$  and  $\beta$  would be  $\sim 33$  ppm for trans and 30 ppm for cis. Apparently, the glycosylation shift overrules this small difference and the  $\Delta\gamma\text{-}\beta$  is  $\sim 42$  ppm, so the configuration cannot be determined here.

**Detection of  $\beta$ -Ara-Hyp**— $\beta$ -Ara-Hyp might form peak III in the gel filtration of alkali-digested nArt v 1. A part of this pool was dansylated and subjected to reversed phase HPLC coupled to an electrospray ionization mass spectrometer. The chromatogram revealed distinct peaks for the masses of dansylated Hyp, Pro and some other amino acids. Two major peaks displayed the mass of dansylated  $\beta$ -Ara-Hyp and are interpreted as arabinosylated trans- and cis-4-Hyp (Fig. 7). The data did not contain a hint to the occurrence of di-, tri-, or tetra-arabinosides as found in extensin and solanaceous lectins. However, another set of masses with the revealing spacing of 132 Da can be interpreted as dansyl-(Ara-)Hyp-(Ara-)Hyp which indicates the presence of contiguous  $\beta$ -Ara-Hyp pairs in Art v 1. Indeed, as nArt v 1 appears to contain  $\sim 16$ – $17$   $\beta$ -Ara residues, most of the (hydroxy)proline residues must be substituted with arabinose as depicted in a schematic model of nArt v 1 and its glycoforms (Fig. 8). We suggest that natural Art v 1 occurs in

two glycoforms, one containing only  $\beta$ -Ara residues, the other one bearing in addition a polysaccharide consisting of  $\beta$ -galactosidase and  $\alpha$ -Ara residues.

**Binding of Antibodies to the Carbohydrate Moieties of Art v 1**—The differential binding of the IgE of the patients by nArt v 1 and rArt v 1 had implied a significant role for posttranslational modifications of Art v 1 in its allergenic properties (1). Therefore, experiments were performed to compare the binding capacities of variously deglycosylated Art v 1 preparations and the natural protein. Periodate oxidation of nArt v 1 resulted in a loss of IgE binding; however, this effect was also observed with sera, which by other means had been shown to bind to the peptide moiety. Hence, periodate experiments were judged unreliable and were not pursued.

Initially, we tried to substantiate the involvement of the Hyp-polysaccharide in antibody binding, and this was done in two ways following recent work on plant *N*-glycans (39). The first strategy used isolated Hyp-polysaccharide as an inhibitor of antibody binding to nArt v 1. In the second, this Hyp-polysaccharide was covalently coupled to bovine serum albumin, and this neo-glycoprotein was used in ELISA and immunoblot. This latter method failed to reveal any binding of either rabbit serum or the IgE of the patients to Hyp-polysaccharide. In contrast, inhibition of IgE binding could be achieved but only with a fairly high concentration of inhibitor. The  $\text{IC}_{50}$  for the Hyp-PS was estimated to be in the range of 12 mM in terms of Ara or  $\sim 0.7$  mM in terms of polysaccharide. In retrospect we assume that even this weak inhibitory potency was not because of the polysaccharide itself but because of residual  $\beta$ -Ara-containing peptides. In line with this assumption, gum arabic consisting to a large part of an arabinogalactan protein (14) and sugar beet arabinan being a branched arabinan with mainly  $\alpha 1,5$  linkages (47) inhibited even less efficiently (supplier's information).

Treatment of nArt v 1 with  $\alpha$ -arabinofuranosidase did not affect binding of the rabbit anti-PTM or of the sera of patients 10–13. Thus,  $\alpha$ -Araf residues on the Hyp-polysaccharide were not involved in antibody binding.

On immunoblots, rabbit anti-PTM and the sera of the patients stained both bands in nArt v 1 (Fig. 9). The single band of dearagal-Art v 1 was stained even stronger due to conflation of two bands into one (Fig. 9). This result was obtained with all sera, even with those reacting with rArt v 1, i.e. with the peptide moiety alone, showing that the epitopes for both the PTM- and the peptide-reactive sera had persisted through the glycosidase treatment. The rArt v 1-binding sera even bound to chemically deglycosylated nArt v 1. On the contrary, this deglyc-nArt v 1 was no longer recognized by those sera that were unable to bind to rArt v 1 (Fig. 9, patients 10–13). We conclude that these sera exclusively contained antibodies binding to a carbohydrate determinant. This epitope, however, is not part of the polysaccharide, which explains that these sera also stained the lower mass band that is devoid of the polysaccharide (see Fig. 8). Only four sera bound exclusively to nArt v 1. However, we assume that more sera contain anti-arabinose IgE in a mixture with anti-peptide IgE.

In another experiment, the  $\alpha$ -arabinosidase and  $\beta$ -galactosidase-treated tryptic peptide dearagal-nArt v 1-trypt (Fig. 10C) resembling the Hyp-rich region of Art v 1 was used as inhibitor of antibody binding to nArt v 1. At a concentration of 100  $\mu\text{M}$  (in terms of Ara, i.e. 6  $\mu\text{M}$  in terms of glycopeptide) the binding to nArt v 1 in several PTM sera was strongly inhibited by dearagal-nArt v 1-trypt (Fig. 10). For sera 9 and 11 inhibition was also performed at 10  $\mu\text{M}$ . The estimated  $\text{IC}_{50}$  of  $0.5 (\pm 0.3)$   $\mu\text{M}$  indicated that this tryptic glycopeptide contained the actual antibody epitope. Taken together, we conclude the  $\beta$ -Araf residues

TABLE II  
Assigned  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of Hyp-linked  $\alpha$ -Ara in  $\alpha$ -arabinofuranosidase-treated nArt v 1

| Residue <sup>a</sup>                   |              |                 | Position |           |           |          |            |            | NOE from H1 | HMBC from H1 |
|----------------------------------------|--------------|-----------------|----------|-----------|-----------|----------|------------|------------|-------------|--------------|
|                                        |              |                 | $\alpha$ | $\beta 1$ | $\beta 2$ | $\gamma$ | $\sigma 1$ | $\sigma 2$ |             |              |
| <b>Hyp</b>                             | Type 1       | $^1\text{H}$    | 4.837    | 2.547     | 2.000     | 4.581    | 4.035      | 3.729      |             |              |
|                                        | Type 2       | $^1\text{H}$    | 4.844    | 2.552     | 2.005     | 4.574    | 4.012      | 3.758      |             |              |
|                                        | Type 3       | $^1\text{H}$    | 4.831    | 2.531     | 2.001     | 4.580    | 4.028      | 3.722      |             |              |
|                                        | Type 1, 2, 3 | $^{13}\text{C}$ | 58.6     | 35.7      |           | 77.5     | 53.8       |            |             |              |
| <b><math>\beta</math>-Ara(1-O)Hyp</b>  | Type 1       | $^1\text{H}$    | 5.120    | 4.10      | 4.012     | 3.848    | 3.744      | 3.600      | 4.581       | 77.5         |
|                                        | Type 2       | $^1\text{H}$    | 5.108    | 4.069     | 3.991     | 3.840    | 3.734      | 3.577      | 4.574       | 77.5         |
|                                        | Type 3       | $^1\text{H}$    | 5.105    | 4.089     | 3.972     | 3.840    | 3.734      | 3.572      | 4.580       | 77.5         |
|                                        | Type 1, 2, 3 | $^{13}\text{C}$ | 100.9    | 77.2      | 75.3      | 82.8     | 64.1       |            |             |              |
| Literature data (Ref. 45) $\beta$ -Ara |              | $^{13}\text{C}$ | 102.3    | 77.2      | 75.1      | 83.2     | 63.9       |            |             |              |

<sup>a</sup> The chemical shifts are given for the residue in bold.

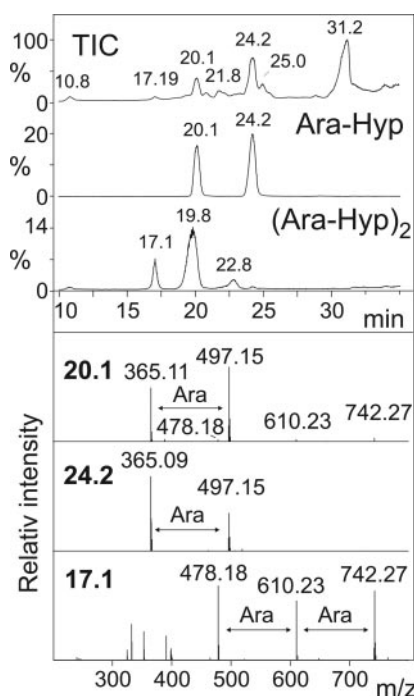


FIG. 7. **Detection of Ara-Hyp by liquid chromatography-MS.** Dansyl-amino acids from pool III of the alkaline digest of nArt v 1 were separated by reversed phase HPLC with electrospray ionization mass spectrometric detection. In addition to the total ion current (TIC), reconstructed ion currents for mass 497 Da (*Ara-Hyp*) and 742 Da (*(Ara-Hyp)<sub>2</sub>*) are shown. In the MS spectra of selected peaks, loss of arabinose by in source fragmentation is indicated.

in the Hyp-rich region of Art v 1 to form the core of an IgE epitope.

**Cross-reactivity of Antibodies with Solanaceous Lectins**—The amino acid sequence and the sugar composition of nArt v 1 pointed to a structural similarity with extensin-like proteins or with domains such as those found in the lectins from potato, tomato, and thorn-apple. On immunoblots, rabbit anti-PTM showed no reaction with tomato and thorn-apple lectin. The binding to several high molecular mass proteins in the potato lectin preparation could be inhibited by bromelain glycopeptide coupled to bovine serum albumin (39). The affinity of IgE from serum 10 of a patient to the lectins was tested with ELISA. Although the reaction with thorn-apple lectin was extremely intense and could not have possibly been due to Art v 1-specific antibodies, no binding to the other two lectins could be observed. Taken together, the arabinan chains on contiguous Hyp-motifs of solanaceous lectins are not able to bind antibodies with specificity for mono-arabinosylated Hyp residues as they occur in the mugwort pollen allergen Art v 1.

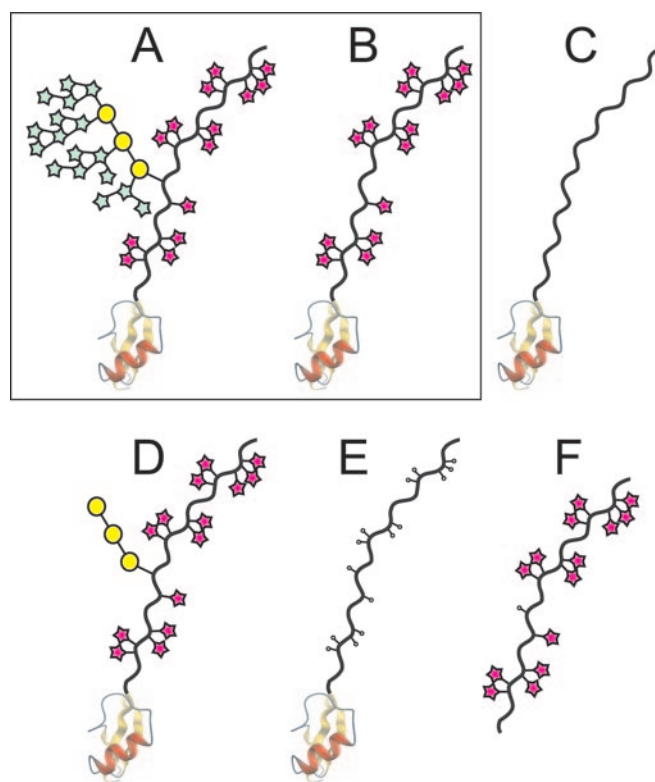


FIG. 8. **Models of natural and artificial glycoforms of Art v 1.** The gray line represents the protein backbone with the globular domain at the bottom. In addition to the Hyp-polysaccharide (yellow circles are  $\beta$ -galactosidase, emerald stars are  $\alpha$ -Ara residues arranged as in Fig. 4), the Hyp-linked  $\beta$ -arabinofuranosyl residues are shown as red stars. The square encloses the two naturally occurring glycoforms. Schematic A shows the larger nArt v 1 glycoform of 15 kDa, B shows the one of 13 kDa, and C shows rArt v 1. Schematic D depicts Art v 1 after de- $\alpha$ -arabinosylation and E after chemical deglycosylation is deglyc-nArt v 1, where the small circles symbolize OH groups of Hyp. The tryptic peptide dearagal-nArt v 1-trypt is shown in Schematic F.

## DISCUSSION

At first glance, Art v 1 is yet another arabinogalactan protein and seemingly it is not worth dealing with it in any detail. However at a closer look, this mugwort pollen allergen holds a few surprises. Firstly, for a plant proteomics expert, Art v 1 merely constitutes a chimeric protein having a Hyp-rich domain that contains a number of contiguous Hyp residues e.g. in the sequence Ser-Hyp-Hyp-Hyp. Extrapolating from work done on various Hyp-rich glycoproteins (5, 15, 18) one would expect this extensin-like domain to bear short Hyp-linked arabinan chains. Thus, Art v 1 would highly resemble the lectins from solanaceous lectins except that its globular domain has a different function. However, Art v 1 bears only single  $\beta$ -Ara

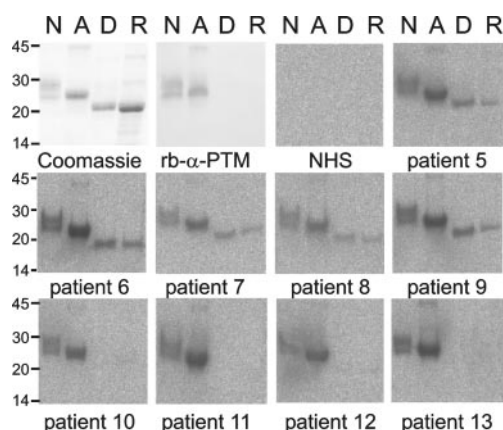


FIG. 9. **Antibody binding to Art v 1 glycoforms.** Similar amounts of nArt v 1 (lanes N), dearagal-nArt v 1 (lanes A), deglyc-nArt v 1 (lanes D) and rArt v 1 (lanes R) were subjected to SDS-PAGE and immunoblotting. One gel was stained with Coomassie Brilliant Blue, one blot was developed with rabbit anti-PTM (*rb-α-PTM*), all other blots show the IgE binding from human sera. *NHS*, non-allergic healthy subject.

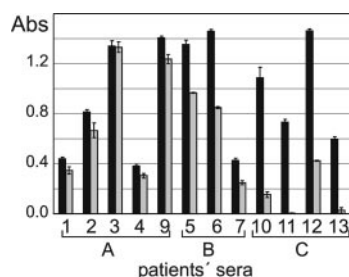


FIG. 10. **Inhibition of IgE binding to nArt v 1 by truncated glycopeptide.** Black bars, binding of the IgE of the patients to nArt v 1 in the absence of inhibitor. Gray bars, show the binding in the presence of 10  $\mu$ M trypsin-treated dearagal-nArt v 1, which lacks the polysaccharide but still contains  $\beta$ -arabinose (see also Fig. 8).

residues attached to, often adjacent, hydroxyprolines. Surprisingly though, there was no immunological cross-reactivity between the mugwort allergen and solanaceous lectins.

Secondly, taking its ability to precipitate with Yariv reagent, Art v 1 constitutes an arabinogalactan protein. Indeed, it does contain a Hyp-linked polysaccharide consisting of arabinose and galactose. However, with its small size (which makes it amenable to MALDI mass spectrometry), lack of other monosaccharides (1), and shortness of the galactan main chain, the  $\beta$ 1,6 linkages between the galactoses and the relatively large branched arabinan side chains this polysaccharide differs from "typical" type II arabinogalactans (3, 8). Assuming that similar polysaccharides will be found on other proteins (see below), we suggest to call these glycans type III arabinogalactans (Fig. 4). Given the pronounced size heterogeneity and the likely presence of various isomers within each size class, a definitive structure of this polysaccharide cannot be given. Even though the structures depicted in Fig. 4 represent a highly likely consensus model that combines the results of NMR and linkage analysis and enzymatic degradations.

Finally, considering the immunogenicity of arabinogalactans and arabinans (4, 24–27), we initially assumed the Hyp-polysaccharide to be responsible for antibody binding in nArt v 1. This hypothesis, however, was found to be inappropriate. In a publication from Haavik *et al.* (48), the carbohydrate moiety of a basic glycoprotein allergen from *Phleum pratense* exhibited striking similarities with the Hyp-polysaccharide of Art v 1 (48). In keeping with the present study, the same authors could not find any evidence for binding of IgE of the patients to this arabinogalactan (49).

The results of the present paper indicate that the Hyp-linked  $\beta$ -arabinose residues constitute the IgG- and IgE-binding determinants. We assume that the epitope comprises more than just one such residue. It may take two or more adjacent arabinosides or even the flanking peptide region. The most convincing proofs for the role of the  $\beta$ -arabinoses would be x-ray diffraction of a glycopeptide-antibody complex or biosynthesis of the epitope akin to that recently performed with plant *N*-glycan structures (36). Both approaches have to await the availability of the relevant reagents, *e.g.* of the arabinosyl transferase.

Taken together, the mugwort pollen allergen Art v 1 appears as an unusual glycoprotein partly akin to well known examples but with several unique aspects. We cannot at the moment locate the Hyp-polysaccharide and the many  $\beta$ -arabinose residues along the C-terminal domain of Art v 1. But we attempted to draft a model for its structure (Fig. 8), which may resemble what Qi *et al.* (31) termed a twisted, hairy rope. As the rather stiff nature of the extended Hyp-rich domain, which also causes the huge difference between true and apparent mass on SDS-PAGE, may not be well described by the term "rope," which implies flexibility, we would rather call it a "bottle-brush." Mugwort pollen thus provides us with two models of bottle-brushes, the simple version having solely  $\beta$ -arabinose residues, which give it a mass of  $\sim$ 13 kDa, and the complex version, equipped additionally with an arabinogalactan polysaccharide resulting in a mass of  $\sim$ 15 kDa (Fig. 8). When it comes to antibody binding the simple version, however, has all that it takes. It remains to be answered whether the only partial decoration with a Hyp-polysaccharide is merely accidental underglycosylation or a consequence of the existence of isoforms of Art v 1 as described for several other pollen allergens (50).

Although, to the best of the knowledge of these authors, this type of glycosylation has not yet been explicitly described, we do not expect it to be unique and confined to Art v 1. Glycoproteins with a homologous glycosylation might be present in related plant species. The asteraceae (or compositae) constitute the second largest family of dicotyledonous plants, and they comprise among many other species potent producers of allergenic pollen such as *Ambrosia artemisiifolia* (common or small ragweed), *Ambrosia trifida* (giant ragweed), *Helianthus annuus* (sunflower), *Parthenium hysterophorus* (feverfew), and of course *A. vulgaris*. Protein sequences homologous to Art v 1 can be found in pollen from *H. annuus* (Swiss prot P22357) and *P. hysterophorus* (51). The latter allergen was described as containing a considerable amount of carbohydrate, which was essential for antibody binding (51). Unfortunately, its structure has not yet been elucidated.

The work reported here leads to a number of intriguing questions. For example, does the amino acid sequence provide the key to this mono-arabinosylation in contrast to the attachment of arabinans? Or do asteraceae simply not express the elongating arabinosyl transferases? How would Art v 1 be glycosylated when produced in tobacco or other non-asteracean plants? How are asteraceae extensins glycosylated?

Does the more restricted occurrence of mono-arabinosylated allergens correspond with a higher allergological impact as compared with *N*-glycans, which for many patients have no clinical relevance? The recent observation (2) that with 6 of 32 patients nArt v 1 gave drastically stronger responses in skin prick test than rArt v 1 indicates that for several patients the carbohydrate does contribute to the allergic reaction.

**Acknowledgments**—We thank Wolfgang Hemmer for the sera of patients, Thomas Dalik for amino acid analysis, and George Lomonosoff and Don Jarvis for reading the manuscript. The assistance of Markus Simmerstatter and Jin Chunsheng in the expression of recombinant allergen is gratefully acknowledged. NMR spectra were obtained using the 800 MHz Varian Unity Inova spectrometer at the Danish

Instrument Center for NMR Spectroscopy of Biological Macromolecules in Valby.

## REFERENCES

- Himly, M., Jahn-Schmid, B., Dedic, A., Kelemen, P., Wopfner, N., Altmann, F., van Ree, R., Briza, P., Richter, K., Ebner, C., and Ferreira, F. (2003) *FASEB J.* **17**, 106–108
- Schmid-Grendelmeier P., Holzmann D., Himly M., Weichel M., Ferreira F., Tresch S., Rückert B., Menz G., Blaser K., Wüthrich B., and Cramer, R. (2003) *J. Allergy Clin. Immunol.* **111**, 1328–1336
- Klis, F. M. (1995) in *Glycoproteins (Comprehensive Biochemistry 29a)* (Montreuil, J., Vliegthart, J. F. G., and Schachter, H., eds) pp. 511–520, Elsevier, Amsterdam
- Showalter, A. M. (2001) *Cell. Mol. Life Sci.* **58**, 1399–1417
- Tan, L., Leykam, J. F., and Kieliszewski, M. J. (2003) *Plant Physiol.* **132**, 1362–1369
- Majewska-Sawka, A., and Nothnagel, E. A. (2000) *Plant Physiol.* **122**, 3–10
- Gaspar, Y., Johnson, K. L., McKenna, J. A., Bacic, A., and Schultz, C. J. (2001) *Plant Mol. Biol.* **47**, 161–176
- Gane, A., Craik, D., Munro, S.L.A., Howlett, G. J., Clarke, A. E., and Bacic A. (1995) *Carbohydr. Res.* **277**, 67–85
- Classen, B., Witthohn, K., and Blaschek, W. (2000) *Carbohydr. Res.* **327**, 497–504
- Tan, L., Qiu, F., Lampion, D. T. A., and Kieliszewski, M. J. (2004) *J. Biol. Chem.* **279**, 13156–13165
- Lampion, A., Katona, L., and Roerig, S. (1973) *Biochem. J.* **133**, 125–131
- Allen, A. K., Desai, N. N., Neuberger, A., and Creeth, M. (1978) *Biochem. J.* **171**, 665–674
- Ashford, D., Desai, N. N., Allen, A. K., Neuberger, A., O'Neill, M. A., and Selvendran, R. R. (1982) *Biochem. J.* **201**, 199–208
- Goodrum, L. J., Patel, A., Leykam, J. F., and Kieliszewski, M. J. (2000) *Phytochemistry* **54**, 99–106
- Kieliszewski, M. J. (2001) *Phytochemistry* **57**, 319–323
- Akiyama, Y., Mori, M., and Kato, K. (1980) *Agric. Biol. Chem.* **44**, 2487–2489
- Campargue, C., Lafitte, C., Esquerre-Tugay, M. T., and Mazau, D. (1998) *Anal. Biochem.* **257**, 20–25
- Zhao, Z. D., Tan, L., Showalter, A. M., Lampion, D. T., and Kieliszewski, M. J. (2002) *Plant J.* **31**, 431–444
- Shpak, E., Barbar, E., Leykam, J. F., and Kieliszewski, M. J. (2001) *J. Biol. Chem.* **276**, 11272–11278
- Knox, J. P., Linstead, P. J., Peart, J., Cooper, C., and Roberts, S. K. (1991) *Plant J.* **1**, 317–326
- Pennell, R. I., Janniche, L., Scofield, G. N., Booi, H., de Vries, S. C., and Roberts, K. (1992) *J. Cell Biol.* **119**, 1371–1380
- Steffan, W., Kovac, P., Albersheim, P., Darvill, A. G., and Hahn, M. G. (1995) *Carbohydr. Res.* **275**, 295–307
- Anderson, M. A., Sandrin, M. S., and Clarke, A. E. (1984) *Plant Physiol.* **75**, 1013–1016
- Willats, W. G., Marcus, S. E., and Knox, J. P. (1998) *Carbohydr. Res.* **308**, 149–152
- Yates, E. A., Valdor, J. F., Haslam, S. M., Morris, H. R., Dell, A., Mackie, W., and Knox, J. P. (1996) *Glycobiology* **6**, 131–139
- Lind, J. L., Bacic, A., Clarke, A. E., and Anderson, M. A. (1994) *Plant J.* **6**, 491–502
- Lampion, D. T. (1969) *Biochemistry* **8**, 1155–1163
- Altmann, F., Schweiszer, S., and Weber, C. (1995) *Glycoconj. J.* **12**, 84–93
- Zeleny, R., Altmann, F., and Praznik, W. (1997) *Anal. Biochem.* **246**, 96–101
- Edge, A. S. (2003) *Biochem. J.* **376**, 339–350
- Qi, W., Fong, C., and Lampion, D. T. A. (1981) *Plant Physiol.* **96**, 848–855
- Altmann, F., Kubelka, V., Staudacher, E., Uhl, K., and März, L. (1991) *Insect. Biochem.* **21**, 467–472
- Altmann, F., and Lomonosoff, G. P. (2000) *J. Gen. Virol.* **81**, 1111–1114
- Schuster, R. (1988) *J. Chromatogr.* **431**, 271–284
- Tapuhi, Y., Schmidt, D. E., Lindner, W., and Karger, B. L. (1981) *Anal. Biochem.* **115**, 123–129
- Bencurova, M., Hemmer, W., Focke-Tejkl, M., Wilson, I. B. H., and Altmann, F. (2004) *Glycobiology* **14**, 457–466
- Duus, J. Ø., Gotfredsen, C. H., and Bock, K. (2000) *Chem. Rev.* **100**, 4589–4614
- Kjaer, M., Andersen, K. V., and Poulsen, F. M. (1994) *Methods Enzymol.* **239**, 288–308
- Wilson, I. B. H., Harthill, J. E., Mullin, N., Ashford, D., and Altmann, F. (1998) *Glycobiology* **8**, 651–661
- Bock, K., and Pedersen, C. (1983) *Adv. Carbohydr. Chem. Biochem.* **41**, 27–66
- Hill, J. M., Alewood, P. F., and Craik, D. J. (1996) *Biochemistry* **35**, 8824–8835
- Gorin, P. A. J., and Mazurek, M. (1976) *Carbohydr. Res.* **48**, 171–186
- Capek, P., Toman, R., Kardosova, A., and Rosik, J. (1983) *Carbohydr. Res.* **117**, 133–140
- Swamy, N. R., and Salimath, P. V. (1991) *Phytochemistry* **30**, 263–265
- Cardoso, S. M., Silva, A. M. S., and Coimbra, M. A. (2002) *Carbohydr. Res.* **337**, 917–924
- Hill, J. M., Alewood, P. F., and Craik, D. J. (2000) *Eur. J. Biochem.* **267**, 4649–4957
- Guillon, F., and Thibault, J.-F. (1989) *Carbohydr. Res.* **190**, 85–96
- Haavik, S., Smestad Paulsen, B., and Wold, J. K. (1987) *Int. Arch. Allergy Appl. Immunol.* **83**, 225–230
- Haavik, S., Smestad Paulsen, B., and Wold, J. K. (1987) *Int. Arch. Allergy Appl. Immunol.* **83**, 231–237
- Svoboda, I., Jilek, A., Ferreira, F., Engel, E., Hoffmann-Sommergruber, K., Scheiner, O., Kraft, D., Breiteneder, H., Pittauer, E., Schmidt, E., Vicente, O., Heberle-Bors, E., Ahorn, H., and Breitenbach, M. (1995) *J. Biol. Chem.* **270**, 2607–2613
- Gupta, N., Martin, B. M., Metcalfe, D. D., and Rao, P. V. (1996) *J. Allergy Clin. Immunol.* **98**, 903–912