Structural studies of the carbohydrate moieties of lectins from potato (*Solanum tuberosum*) tubers and thorn-apple (*Datura stramonium*) seeds

David ASHFORD, Nila N. DESAI, Anthony K. ALLEN and Albert NEUBERGER
Department of Biochemistry, Charing Cross Hospital Medical School, Fulham Palace Road,
London W6 8RF, U.K.

and Malcolm A. O'NEILL and Robert R. SELVENDRAN
A.R.C. Food Research Institute, Colney Lane, Norwich NR4 7UA, U.K.

(Received 17 July 1981/Accepted 7 September 1981)

1. Methylation analysis of potato (*Solanum tuberosum*) lectin and thorn-apple (*Datura stramonium*) lectin confirmed previous conclusions that both glycoproteins contained high proportions of L-arabinofuranosides and lesser amounts of D-galactopyranosides. The arabinofuranosides are present in both lectins as short unbranched chains containing 1→2- and 1→3-linkages, which are known to be linked to hydroxyproline. Galactopyranosides are present as monosaccharides, which are known to be attached to serine, in potato lectin and as both the monosaccharide and the 1→3-linked disaccharide in *Datura* lectin. 2. Alkaline digestion of potato lectin and subsequent separation of the components by gel filtration led to the isolation of four fractions corresponding to the mono-, di-, tri- and tetra-arabinosides of hydroxyproline. The latter two fractions accounted for over 70% of the total hydroxyproline. 3. Methylation analysis was used to show that the triarabinoside contained only 1→2-linkages between sugars, but that the tetra-arabinoside contained both 1→2- and 1→3-linkages. Direct-insertion mass spectrometry of these compounds using electron impact and chemical ionization, in a comparison with other known structural patterns, was used to determine the sequences of the sugars, which were Ara f 1→2Ara f 1→2Ara f 1→Hyp and Ara f 1→3Ara f 1→2Ara f 1→2Ara f 1→Hyp. 4. On the basis of optical rotation it had previously been suggested [Allen, Desai, Neuberger & Creeth (1978) *Biochem. J.* **171**, 665–674] that all the arabinose of potato lectin was present as the β -L-furanoside. However, measurement of the optical rotations of the hydroxyprolyl arabinosides showed that whereas the diarabinoside had a molar rotation ($[\text{m}]$) value close to that predicted, the triarabinoside was more dextrorotatory and the tetra-arabinoside was less dextrorotatory than expected. Possible explanations for these findings are that, although the di- and tri-arabinosides contain exclusively β -arabinofuranosides, in the triarabinoside, interactions between pentose units lead to an enhanced positive rotation. The tetra-arabinoside, however, is proposed to contain a single α -arabinofuranoside residue, which is responsible for the lower than expected positive rotation. The observed rotation of the tetra-arabinoside was found to be close to the theoretical value predicted on that basis. Furthermore, the action of a specific α -arabinofuranosidase on the tetrasaccharide was to remove a single arabinose residue, presumably the terminal non-reducing sugar, and to produce a product that was indistinguishable on electrophoresis from the triarabinoside. Changes in rotation were compatible with this assumption. 5. It is concluded that the structures of the hydroxyprolyl tri- and tetra-arabinosides of potato lectin are: β Ara f 1→2 β Ara f 1→2 β Ara f 1→Hyp and α Ara f 1→3 β Ara f 1→2 β Ara f 1→2 β Ara f 1→Hyp. These are identical with compounds that have been isolated from the insoluble hydroxyproline-rich glycoproteins of plant cell walls.

Lectins with an unusual structure have been extracted from three species of the family Solanaceae: *Solanum tuberosum* (potato) (Allen & Neuberger, 1973; Allen *et al.*, 1978; Muray & Northcote, 1978); *Datura stramonium* (thorn-apple) (Hořejší & Kocourek, 1978; Desai *et al.*, 1981) and *Lycopersicon esculentum* (tomato) (Nachbar *et al.*, 1980). The most noteworthy features of these lectins are that they are glycoproteins that contain large amounts of hydroxyproline and arabinose. It has been shown that both potato lectin (Allen *et al.*, 1978) and *Datura* lectin (Desai *et al.*, 1981) have glycoprotein linkages between β -L-arabinofuranose and hydroxyproline as well as between α -D-galactopyranose and serine. These glycoproteins show a close resemblance in their glycosylated domains to hydroxyproline-containing fragments that have been isolated from insoluble plant-cell wall glycoproteins (Heath & Northcote, 1971; Lampert, 1977; O'Neill & Selvendran, 1980), and their similarities have been discussed by Ashford & Neuberger (1980). After detailed investigations (Akiyama & Katō, 1977; Akiyama *et al.*, 1980) of the hydroxyproline-containing glycopeptides from plant-cell-walls, we have further investigated the soluble lectins from potato and *Datura*, and have been able to show that the structures are similar to those of the plant-cell wall glycoproteins.

Experimental

Materials

Preparation of lectins. Potato lectin was prepared from potato tubers (*Solanum tuberosum* var. Ulster Sceptre) by the method of Desai & Allen (1979) and *Datura stramonium* (thorn-apple) lectin was isolated from seeds by the method of Desai *et al.* (1981). α -L-Arabinofuranoside from the fungus *Sclerotinia fructigena* (Laborda *et al.*, 1973) was a gift of Dr. A. H. Fielding, Long Ashton Research Centre, Bristol, U.K. One unit will hydrolyse 1.0 μ mol of 4-nitrophenyl α -L-arabinofuranoside/min at 30°C at pH 4.5.

Bio-Gel P-2 (–400 mesh) was from Bio-Rad Laboratories Ltd. (Watford, Herts., U.K.) and Sephadex G-10 was from Pharmacia (G.B.) Ltd. (Hounslow, Middx., U.K.). Both α - and β -methyl arabinofuranosides were prepared by the method of Augestad & Berner (1956) and were separated and purified on Dowex-1 resin (OH[–] form) by the procedure of Neuberger & Wilson (1971).

Methods

Analysis of amino acids and sugars. Amino acids were measured on a Locarte Mini amino acid analyser by reaction with ninhydrin. Samples were hydrolysed in 3M-toluene-4-sulphonic acid for 24 h at 110°C *in vacuo*. The hydrolysates were two-

thirds neutralized with NaOH and loaded directly on to the analyser. Acidic and neutral amino acids and hydroxyproline isomers were separated on a column (23 cm $\times$ 0.9 cm) by successive elution with 0.2 M-sodium citrate buffers: pH 3.15 for 35 min, pH 3.30 for 80 min and pH 4.25 for 90 min. The column was equilibrated at 38°C for the time of the first buffer and subsequently at 60°C.

Sugars were measured by g.l.c. after methanolysis and trimethylsilylation (Chambers & Clamp, 1971) with a Perkin–Elmer F.33 gas chromatograph fitted with dual columns and dual flame-ionization detectors. Separation was accomplished with 8.5% SE-30 on diatomite C as the stationary phase, with a linear temperature gradient from 120 to 210°C at 1.5°/min. Internal standards were used for both types of analysis to ensure comparability of the results.

Polarimetry. Optical rotation was measured on a Perkin–Elmer model 141 polarimeter in a 10 cm cell at 20°C. Molar rotation ($[\text{M}]$), and mean residue rotation of arabinose in oligoarabinosides ($[\text{M}]_{\text{Ara}}$), values were calculated from $[\alpha] \times \text{mol.wt.}/100$ for each wavelength.

Alkaline hydrolysis of potato lectin. Potato lectin (25 mg) was hydrolysed in 5 M-NaOH (2 ml) for 24 h at 100°C, in a sealed tube (with a polythene liner to prevent silicate formation).

Separation of hydroxyprolyl oligoglycosides. The alkaline hydrolysate was neutralized with HCl and applied to a column (90 cm $\times$ 2.5 cm) of Bio-Gel P-2 (–400 mesh) equilibrated and eluted with degassed, deionized water at 30 ml/h. The column was maintained at 65°C to increase the resolution. Column fractions were monitored for optical rotation at 365 nm, hydroxyproline content, by the method of Stegemann (1958), and carbohydrate content, by the method of Dubois *et al.* (1956), with L-arabinose as standard. Fractions comprising the major peaks were pooled, concentrated and desalted on a column (60 cm $\times$ 2.1 cm) of Sephadex G-10 equilibrated and eluted with water. The column effluent was monitored by optical rotation at 365 nm. The peak fractions were pooled and divided into portions for amino acid, sugar and methylation analysis, for optical-rotation measurements and for mass spectrometry.

α -L-Arabinofuranosidase treatment of hydroxyprolyl tetra-arabinoside. The conjugate, Hyp(Ara)₄ (0.5 μ mol) was incubated with 0.04 unit of α -L-arabinofuranosidase in 0.1 M-sodium acetate buffer, pH 4.5, in a final volume of 1.0 ml at 20°C. The optical rotation of the solution was monitored at 365 nm. After 24 h, samples of the reaction mixture were examined by g.l.c. without prior methanolysis and also by high-voltage electrophoresis.

High-voltage paper electrophoresis. A Shandon apparatus was used and separations were per-

formed on Whatman 3MM paper at pH 1.9 in an aqueous buffer [12.5 ml of 98% (v/v) formic acid and 43.5 ml of acetic acid/litre] for 90 min at 10 kV/m with constant cooling. The mobilities of components were expressed relative to free hydroxyproline (R_{Hyp}). Hydroxyproline-containing compounds were located by using a combination of isatin and dimethylaminobenzaldehyde reagents (Dawson *et al.*, 1969).

Methylation of native lectins. Purified lectins (3–5 mg) were vacuum dried over P_2O_5 for 24 h. The dried samples were dissolved in dry dimethyl sulphoxide (1 ml). Alkoxide formation and methylation was performed as described by Jansson *et al.* (1976). Partially methylated alditol acetates were prepared by the method of Ring & Selvendran (1978).

Methylation of hydroxyprolyl oligoarabinosides. Hydroxyprolyl-oligoarabinosides (0.5–1.0 mg) were vacuum-dried over P_2O_5 for 24 h. Dried samples were dissolved in dry dimethyl sulphoxide (500 μ l) and methylated. The methylated products were extracted into dichloromethane (Jansson *et al.*, 1976). The combined organic phases were concentrated by dryness under reduced pressure and dissolved in dichloromethane (100 μ l).

Mass spectrometry. (i) Partially methylated alditol acetates. Partially methylated alditol acetates were quantitatively analysed by g.l.c. and qualitatively identified by g.l.c.–mass spectrometry (Jansson *et al.*, 1976; Ring & Selvendran, 1978).

(ii) Permethylated hydroxyprolyl oligoarabinosides. A portion (5–10 μ l) of the permethylated hydroxyprolyl oligoarabinoside was evaporated to dryness in the heated direct-insertion probe tip before insertion. The electron impact (EI) spectra were recorded on an MS902 mass spectrometer (Kratos Ltd., Manchester, U.K.) connected to a DS50SM data system (Kratos) as follows: 70 eV ionization, 210°C source temperature and 27–847 m/e mass range scanned at 9 s/decade. Chemical ionization (CI) spectra using NH_3 gas were recorded on an MS3076 mass spectrometer (Kratos) connected to a DS50SM data system as follows: ionization energy 400 eV, 150°C source temperature and the probe heated at its maximum rate up to 250°C. Reagent gas was introduced into the source (source pressure between 6.7 mPa and 2.7 mPa). Mass range 50–850 m/e was scanned at 9 s/decade.

Results

Separation of hydroxyprolyl oligoglycosides

The elution profile of hydroxyprolyl oligoglycosides on Bio-Gel P-2 (–400 mesh) is shown in Fig. 1. The four peaks that contain hydroxyproline also contain carbohydrate (Fig. 1a). They also display the positive optical rotation (Fig. 1b) seen in the

Fig. 1. Elution profile on Bio-Gel P-2 (–400 mesh) of the hydroxyprolyl arabinosides derived from potato lectin (a) ●, Carbohydrate content (μ g/ml) measured at 480 nm after reaction by the method of Dubois *et al.* (1956). ○, Hydroxyproline content (μ g/ml) measured at 550 nm after reaction by the method of Stegemann (1958). (b) ■, Optical rotation of the eluate as measured at 365 nm in a 10-cm cell.

34 000-mol.wt. glycopeptide of potato lectin (Allen *et al.*, 1978). Peaks II–IV were desalted on Sephadex G-10 and all gave single symmetrical peaks of optically active material. Peak I had insufficient material to be handled in this way and was concentrated and freeze-dried. Amino acid and sugar analyses of the individual peaks are given in Table 1. Alkaline hydrolysis of L-hydroxyproline residues has been shown to produce an equimolar mixture of L-hydroxyproline and D-allo-hydroxyproline (Robinson & Greenstein, 1952). These isomers of hydroxyproline accounted for over 80% of the amino acids in each fraction and, in the case of peaks IV and III, 94.8 and 90.8% respectively. The contaminating amino acids were predominantly serine, glutamic acid, threonine, aspartic acid, glycine and alanine. These amino acids were also found to be present in a previous unseparated alkaline hydrolysate of potato lectin (Allen *et al.*, 1978).

Table 1. *Analyses of hydroxyprolyl oligoglycosides from potato lectin*
For experimental details, see the text.

Amino acid*	Peak ...	Content (relative to Hyp + <i>allo</i> -Hyp = 1.0)			
		IV	III	II	I
Hyp + <i>allo</i> -Hyp		1.000	1.000	1.000	1.000
Asp		0.005	0.010	0.023	0.016
Thr		0.010	0.010	0.021	0.016
Ser		0.011	0.025	0.050	0.034
Glu		0.004	0.010	0.029	0.031
Pro		0.000	0.000	0.000	0.033
Gly		0.016	0.031	0.069	0.047
Ala		0.004	0.011	0.017	0.015
Ile		0.002	0.002	0.006	0.006
Leu		0.002	0.002	0.012	0.011
Tyr		0.001	0.000	0.000	0.005
Sugar					
Ara		3.778	3.215	2.074	1.416
Gal		0.000	0.000	0.000	0.042
Identification		Hyp(Ara) ₄	Hyp(Ara) ₃	Hyp(Ara) ₂	HypAra

* Basic amino acids and tryptophan were not determined, and cystine, valine, methionine and phenylalanine were not detectable.

Sugar analysis revealed only arabinose in peaks II–IV and arabinose and a small amount of galactose in peak I. Tentative identifications of the peaks were made on the basis of the analytical data. Peak IV has an arabinose/hydroxyproline ratio of 3.78:1 and was considered to be the tetra-arabino-side of hydroxyproline, Hyp(Ara)₄. By the same reasoning, peaks III, II and I, with ratios of 3.22:1, 2.07:1 and 1.42:1 respectively, were designated as the tri-, di- and (mainly) mono-arabinosides, Hyp(Ara)₃, Hyp(Ara)₂ and HypAra, and will be referred to as such from now on. Hyp(Ara)₄ and Hyp(Ara)₃ are present in approximately equimolar proportions and together comprise over 70% of the glycosylated hydroxyproline (37.3 and 34.5% respectively), whereas Hyp(Ara)₂ and HypAra comprise only 12.5 and 15.5% respectively. Therefore 72% of the hydroxyproline of potato lectin carries 84% of the arabinose. No free hydroxyproline was detected in the Bio-Gel P-2 column effluent.

Optical-rotation measurements

Optical rotations of Hyp(Ara)₄ and Hyp(Ara)₃ were measured at 589 and 365 nm, and specific rotations ($[\alpha]$) and molar rotations were calculated at these wavelengths for the whole molecule ($[m]$) and for the arabinose units after correcting for the rotation due to the mixture of hydroxyproline isomers, ($[m]_{\text{Ara}}$). Insufficient Hyp(Ara)₂ was available for these measurements, but a single value was calculated at 365 nm from previous data. These values are listed in Table 2. It can be seen that the $[\alpha]_{\text{Ara}}$ and $[m]_{\text{Ara}}$ for Hyp(Ara)₄ are less than expected by use of Hudson's rules, as applied by

Ferrier (1980) to oligosaccharides for a molecule containing only β Ara_f. The converse is true for Hyp(Ara)₃, with $[\alpha]_{\text{Ara}}$ and $[m]_{\text{Ara}}$ values that are greater than expected. Hyp(Ara)₂, however, has $[\alpha]_{\text{Ara}}$ and $[m]_{\text{Ara}}$ values close to the theoretical value. The possibility that the lower-than-expected rotation values obtained for Hyp(Ara)₄ could be explained by the presence of α -linked arabinose, as found recently by Akiyama *et al.* (1980) for the tetra-arabinoside of hydroxyproline from tobacco (*Nicotiana tabacum*) cell-wall glycoproteins, was tested. The theoretical molar rotations of molecules with either three β -linked arabinofuranosides and one α -arabinofuranoside or four β -linked arabinofuranosides linked to hydroxyproline were calculated and compared. The values were calculated from the $[m]_{\lambda}$ values of the respective methyl arabinofuranosides and an equimolar mixture of L-hydroxyproline and D-*allo*-hydroxyproline. These values are given in Table 3, together with the theoretical values for β -linked Hyp(Ara)₃ and Hyp(Ara)₂. As can be seen, the observed values for Hyp(Ara)₄ agree best with the theoretical values for a structure containing a single α -arabinofuranose and three β -arabinofuranosides. Hyp(Ara)₃, as was seen in Table 2, has a much larger $[m]$ value than the theoretical value. Hyp(Ara)₂, again as seen previously, gives a close fit with the calculated data.

High-voltage electrophoresis and the effect of α -L-arabinofuranosidase

High-voltage electrophoresis of Hyp(Ara)₃ and Hyp(Ara)₄ resolved each hydroxyprolyl arabinoside into double bands with R_{Hyp} values of 0.53, 0.59 and

Table 2. Optical rotations of hydroxyprolyl arabinosides derived from potato lectin and of methyl arabinofuranosides

Sample	λ (nm)	$\alpha(^{\circ})$	$[\alpha](^{\circ})$	$[\alpha]_{\text{Ara}}(^{\circ})$	$[m](^{\circ})$	$[m]_{\text{Ara}}(^{\circ})$
Hyp(Ara) ₄	589	+0.020	+60.7	+64.0	+400.0	+105.6
	365	+0.061	+185.1	+193.3	+1220.0	+318.9
Hyp(Ara) ₃	589	+0.046	+174.6	+178.7	+920.0	+314.5
	365	+0.115	+436.4	+446.6	+2300.0	+786.0
Hyp(Ara) ₂	365	+0.016	+242.5	+256.1	+957.9	+507.5
α -Me-Araf	589	—	-127.0	—	-208.3*	—
	365	—	-360.5	—	-591.2	—
β -Me-Araf	589	—	+117.0	—	+191.9*	—
	365	—	+338.1	—	+554.5	—

* Literature values of $[m]_D$ from Augestad & Berner (1956): α -Me-Araf = -209.0°; β -Me-Araf = +194.0°.

Table 3. A comparison of theoretical and observed $[m]_D$ values for hydroxyprolyl oligoarabinosides

Sample	λ (nm)	Theoretical values (°)		Observed values (°)
		All β	$\beta_3\alpha_1$	
Hyp(Ara) ₄	589	+746	+346	+400
	365	+2164	+1019	+1220
Hyp(Ara) ₃	589	+554	—	+920
	365	+1610	—	+2300
Hyp(Ara) ₂	365	+1056	—	+958

lysis demonstrated that free arabinose was present. Electrophoresis of the α -L-arabinofuranosidase-treated Hyp(Ara)₄ showed that the product had R_{Hyp} values of 0.53 and 0.59, which were identical with those for Hyp(Ara)₃. No other compounds were observed.

Mass spectrometry of native lectins

Both potato and *Datura* lectins were subjected to standard methylation analysis and the results are shown in Table 5 (columns a and b). Recoveries of permethylated alditol acetates were found to be in good agreement with sugar values obtained by direct analysis (Allen *et al.*, 1978). Methylation of the glycoproteins was found to be complete by examination of the i.r. spectrum of the permethylated products (Ring & Selvendran, 1978).

Potato lectin was found to yield four permethylated alditol acetates corresponding to terminal, 1→2 and 1→3 arabinose units and terminal galactose units. The occurrence of 2,3,5-tri-*O*-methylarabinitol confirms the data from polarimetry and partial acid hydrolysis (Allen *et al.*, 1978) that arabinose occurs in the furanose form. The three types of Araf units (terminal, 1→2 and 1→3) correspond to those found in hydroxyproline-rich glycoproteins from higher-plant cell walls (Akiyama & Katō, 1976, 1977; O'Neill & Selvendran, 1980). However, their molar ratio 1.0:1.0:0.6 does not correspond to that expected from a glycoprotein containing approximately equal amounts of hydroxyprolyl tetra- and tri-arabinosides. The reasons for this are not clear, but may be due to an underestimation of 1→2Araf. However, it must be emphasized that standard methylation analysis of 'complete' glycoproteins such as potato lectin can only give qualitative information on the nature of the major *O*-glycosidic linkages. Galactose was found to occur solely as terminal galactopyranose residues. Galactose has previously been shown to be α -*O*-glycosidically linked to serine, probably as single units (Allen *et al.*, 1978). Our data confirm this, but unambiguous

Table 4. Effect of α -L-arabinofuranosidase on Hyp(Ara)₄ as measured by polarimetry at 365 nm

t (h)	Observed rotation (°)	$[m](^{\circ})$
0	+0.060	1200
2	+0.082	1640
5	+0.101	2020
20	+0.110	2200
Predicted for an equimolar mixture of Hyp(Ara) ₃ + Ara ...	+0.137	2750

0.48, 0.54 respectively. The doublets arise from the presence of two isomers of hydroxyproline after alkali treatment. The R_{Hyp} values are in good agreement with those reported by Lamport (1967) and Muray & Northcote (1978) for Hyp(Ara)₃ and Hyp(Ara)₄ from respectively potato callus cell wall and potato lectin.

Treatment of Hyp(Ara)₄ with α -L-arabinofuranosidase was monitored by polarimetry. There was a linear increase in rotation over a period of 2 h, and after 20 h there was no further change in rotation (Table 4). The final rotation was tending towards that calculated for an equimolar mixture of Hyp(Ara)₃ and free arabinose. Analysis of this final reaction mixture by g.l.c. without prior methano-

Table 5. *Partially methylated alditol acetates obtained from potato lectin, Datura lectin and from the hydroxypropyl arabinosides derived from potato lectin*

The results are expressed as molar percentages of the total alditol acetates that were found.

Partially methylated alditol acetate	Content (mol%)				Deduced linkage
	Potato lectin (a)	Datura lectin (b)	Hyp(Ara) ₄ (c)	Hyp(Ara) ₃ (d)	
2,3,5-Tri- <i>O</i> -methylarabinitol	34.5	31.3	27.2	35.3	Araf1→
3,5-Di- <i>O</i> -methylarabinitol	34.4	35.5	46.6	64.7	→2Araf1→
2,5-Di- <i>O</i> -methylarabinitol	24.4	19.8	26.2	0	→3Araf1→
2,3,4,6-Tetra- <i>O</i> -methylgalactitol	6.7	8.5	0	0	Galp1→
2,4,6-Tri- <i>O</i> -methylgalactitol	0	4.9	0	0	→3Galp1→

Table 6. *Diagnostic ions obtained from the direct insertion mass spectra of permethylated hydroxypropyl tri- and tetra-arabinosides derived from potato lectin*

See Fig. 2 and the text for further details; the symbols used correspond to those of Kochetkov & Chizhov (1966).

Ion (<i>m/e</i>)	Relative abundance				Symbol
	Electron impact		Chemical ionization		
	Hyp(Ara) ₄	Hyp(Ara) ₃	Hyp(Ara) ₄	Hyp(Ara) ₃	
829	—	—	30.1	—	
828	<0.1	—	74.9	—	(<i>M</i> ₁ + CH ₃) ⁺
813	0.4	—	0.7	—	<i>M</i> ₁ ⁺
812	1.3	—	2.2	—	(<i>M</i> - 1) ⁺
769	36.0	—	9.8	—	(<i>M</i> ₁ + CH ₃) - 59
768	100.0	—	25.1	—	<i>M</i> ₁ - 45
754	0.5	—	—	—	<i>M</i> ₁ - 59
690	—	—	12.1	—	
688	—	—	8.6	—	
669	—	—	—	22.9	
668	—	<0.1	—	72.8	(<i>M</i> ₁ + CH ₃) ⁺
655	—	—	—	—	dcba A ₁
653	—	2.9	—	—	<i>M</i> ₁ ⁺
652	0.1	9.0	—	2.0	dcba A ₂
623	0.5	—	0.2	—	
622	1.9	1.1	1.1	—	
609	—	20.1	—	2.7	(<i>M</i> ₁ + CH ₃) - 59
608	—	62.2	—	11.3	<i>M</i> ₁ - 45
594	0.2	1.3	—	—	<i>M</i> ₁ - 59
530	—	1.4	0.8	12.9	
528	—	1.3	0.7	15.0	
495	0.2	3.1	1.1	—	cba A ₁
463	0.4	4.1	0.9	—	cba A ₂
462	1.1	4.2	0.8	—	
431	0.8	5.8	3.3	2.3	cba A ₃
335	1.6	7.2	10.6	6.1	ba A ₁
303	1.5	8.0	6.1	11.8	ba A ₂
302	1.1	0.7	1.4	—	
271	0.8	6.3	1.1	5.2	ba A ₃
189	10.1	22.2	13.8	20.9	
175	11.3	35.7	16.9	42.0	a A ₁
173	1.1	6.1	1.3	7.9	
170	5.9	6.0	10.9	7.5	
161	2.3	1.9	6.4	62.6	
157	3.6	—	4.5	11.8	c A ₁
156	8.9	—	17.1	20.0	
143	52.2	10.3	100.0	67.3	a A ₂
142	2.7	33.8	4.9	13.8	
101	37.8	100.0	57.5	44.0	F ₁
96	75.2	—	81.7	92.3	
45	35.3	—	—	—	

identification will require the isolation of seryl galactosides. It is noteworthy that recoveries of galactose from methylation and direct analysis are similar, suggesting little loss of galactose attached to serine by alkali-induced β -elimination during derivatization. Similar observations have been made from methylation studies on hydroxyproline-rich glycoproteins from plant cell walls (O'Neill & Selvendran, 1980). These data add further support to the previous suggestion that ionization of Ara/hydroxy groups may significantly decrease β -elimination (Allen *et al.*, 1978) of adjacent glycosylated serine residues.

Methylation analysis of *Datura* lectin yielded essentially similar results for Ara f residues (Table 5, column b), but approx. 37% of the galactose in *Datura* was found to occur as non-terminal 1 $\rightarrow$ 3Gal p units. Also, in contrast with potato lectin with a ratio of glycosylated serine to galactose of 1:1, the ratio of glycosylated serine to galactose in *Datura* lectin is 1:1.5. This suggests that some serine residues are monosubstituted, but others are *O*-glycosylated with oligosaccharides of structure Gal p 1 $\rightarrow$ 3Gal p 1 $\rightarrow$ Ser. This is in agreement with the observation (Desai *et al.*, 1981) that the ratio of the loss of galactose to the destruction of serine on β -elimination of *Datura* lectin glycopeptide was 1.8:1. However, confirmation of this again requires the isolation of seryl galactosides.

Partially methylated alditol acetates from Hyp(Ara) $_4$ and Hyp(Ara) $_3$ derived from potato lectin

Alkaline hydrolysis of potato lectin followed by gel filtration allowed hydroxyprolyl tetra- and tri-arabinosides to be isolated. Insufficient hydroxyprolyl di-arabinoside was available for complete analysis.

The partially methylated alditol acetates from Hyp(Ara) $_4$ and Hyp(Ara) $_3$ were qualitatively identified by g.l.c.-mass spectrometry and quantified by g.l.c. and the results are shown in Table 5 (columns c and d). Since only methylated arabinitol residues were detected, the previous conclusions that only oligosaccharides containing arabinose are attached to hydroxyproline have been confirmed. This is similar to hydroxyprolyl oligoarabinosides from higher-plant cell walls (Akiyama & Katō, 1976, 1977; O'Neill & Selvendran, 1980), although certain algal-cell-wall glycoproteins may contain other sugars attached to hydroxyproline (Miller *et al.*, 1972; O'Neill & Roberts, 1981).

The ratio terminal/1 $\rightarrow$ 2/1 $\rightarrow$ 3 residues for Hyp(Ara) $_4$ is 1.0:1.7:1.0 and for Hyp(Ara) $_3$ is 1.0:1.8:0. This suggests that analogous structures to those proposed by Akiyama & Katō (1976, 1977) for hydroxyprolyl oligoarabinosides from tobacco cell walls may be present.

Mass-spectral analysis of permethylated hydroxyprolyl oligoarabinosides

Analysis of partially methylated alditol acetates from Hyp(Ara) $_4$ revealed two types of interchain linkages, namely 1 $\rightarrow$ 2 and 1 $\rightarrow$ 3. However, these data do not allow the order of the *O*-glycosyl units to be determined. Recent work on permethylated oligosaccharide alditols has shown that sequences may be determined by the type(s) of fragmentation in the mass spectrum (Kovacik *et al.*, 1968; Kärkkäinen, 1969, 1971; Moor & Waight, 1975; Fournet *et al.*, 1980; Strecker *et al.*, 1981; Ring & Selvendran, 1980, 1981).

Permethylated derivatives of hydroxyprolyl tetra- and tri-arabinosides were analysed by direct-insertion mass spectrometry. Diagnostic ions obtained from both electron-impact (EI) and chemical-ionization (CI) spectra are shown in Table 6 and the origins of the ions for the tetrasaccharide and by analogy for the trisaccharide are shown in Fig. 2. The symbols used correspond to those described by Kochetkov & Chizhov (1966).

On methylating the hydroxyprolyl oligoarabinosides, the imino nitrogen of the hydroxyprolyl moieties would be mono- and di-*N*-methylated:

Under the high-temperature conditions of the mass spectrometer the quaternary nitrogen derivative may undergo pyrolysis causing ring cleavage (Finar, 1960):

This could explain the occurrence of fragments in the mass spectrum (containing the derivatized imino acid) with m/e one mass unit less than expected (e.g. m/e 157 and 156). The EI spectra contained only small amounts of ions at m/e 828 (from hydroxyprolyl tetra-arabinoside) and 668 (from hydroxyprolyl triarabinoside) corresponding to $(M + CH_3)^+$ ions. However, the $(M + CH_3)^+$ ions, which probably arise from the quaternary nitrogen salt, were predominant in the CI spectra. Assuming the molecular weights of the respective permethylated hydroxyprolyl oligoarabinosides to be 828 and 668, the corresponding underivatized imino acid gly-

cosides would have molecular weights of 659 and 527. Thus a single hydroxyproline residue is *O*-glycosylated with oligosaccharides containing four or three arabinose residues.

Mass-spectral analysis of permethylated derivatives of hydroxyprolyl diarabinoside gave poor spectra, owing to high background contamination. However, it could be inferred that the derivative of the conjugate had a molecular weight of 508 and therefore 395 underivatized.

Evidence from analysis of permethylated alditol acetates from permethylated hydroxyprolyl tetra-

arabinoside has shown the tetrasaccharide to contain a terminal arabinofuranoside unit *O*-glycosidically linked to either C2 or C3 of the penultimate arabinofuranoside unit. Kärkkäinen (1971) and Moor & Waight (1975) have shown that 1→2- and 1→3-linked glycosyl residues show differences in the A-series of fragmentation. The A-series is characterized by loss of substituent at C-1 with subsequent elimination of other substituents such as CH_3OH m/e 32 (Kochetkov & Chizhov, 1966). The diagnostic ions of the A-series of fragmentation are observed with both EI and CI spectra (Table 6). Chemical ionization gives more intense ions and thus when considering relative abundance is probably more reliable.

The EI spectrum of permethylated hydroxyprolyl tetra-arabinoside gave a small, but significant, amount of the molecular ion at m/e 813 and a relatively large amount of the ion at m/e 768 ($M^+ - \text{CH}_2\text{OCH}_3$), corresponding to the loss of 45 mass units from M^+ . The ions at 768 and 754 probably arise from loss of 59 mass units (CO_2CH_3) from ($M^+ + \text{CH}_3$)⁺ and M^+ ions respectively. The ions at m/e 157, 302, 462 and 622 arise from the permethylated tetra-arabinoside, as shown in Fig. 2.

An indication of the sequence in which the Araf residues a, b, c and d are linked together can be obtained by considering the ions at m/e 175, 335, 495 and 623 (655–32). A relatively high percentage of ion m/e 175 shows the presence of the non-reducing terminal pentose residue. The ion at m/e 335 shows the occurrence of the non-reducing pentose-pentose. From the work of Kärkkäinen (1971), Moor & Waight (1975) and Ring & Selvendran (1981) with various permethylated pentose-containing oligosaccharides, it has been shown that $\text{ba A}_1 > \text{ba A}_2 > \text{ba A}_3$ (nomenclature of Kochetkov & Chizhov, 1966) if the terminal non-reducing pentose is *O*-glycosidically linked to C3 of the penultimate pentose. The permethylated

Fig. 2. Mass-spectral fragmentation pattern of the methylated hydroxyprolyl tetra-arabinoside and the origins of some pertinent ions

See the text and Table 6 for details of the relative abundance of the ions.

Fig. 3. Proposed structures for the hydroxyprolyl arabinosides of potato lectin (a) Hydroxyprolyl triarabinoside; (b) hydroxyprolyl tetra-arabinoside.

hydroxypropyl tetra-arabinoside (Table 6 and Fig. 2) shows a relative abundance of the ions at m/e 335 (ba A₁), 303 (ba A₂) and 271 (ba A₃), which correspond to 1.6:1.5:0.8 (EI) and 10.6:6.1:1.1 (CI). This suggests that the terminal arabinofuranoside is linked C1→C3 to the penultimate arabinofuranoside.

The above studies showed that if, on the other hand, the terminal pentose sugar were *O*-glycosidically linked to C2 of the penultimate sugar ba A₂ > ba A₁ > ba A₃. It can be seen from Table 6 that, for the derivative of hydroxypropyl triarabinoside, the relative abundances of ions at m/e 303 (ba A₂), 335 (ba A₁) and 271 (ba A₃) are 8.0:7.2:6.3 (EI) and 11.8:6.1:5.2 (CI), which indicates that the terminal linkage of the trisaccharide is C1→C2. From the results of the analyses of the partially methylated alditol acetates the remaining arabinofuranosyl linkages can be deduced to be linked C1→C2 and the sequences of the tri- and tetra-arabinosides of hydroxyproline are therefore as shown in Fig. 3.

Discussion

It is apparent from the results of methylation and mass spectrometry presented here, and from previous work (Allen *et al.*, 1978; Desai *et al.*, 1981), that both potato and *Datura* lectins contain arabinose exclusively as the furanoside and galactose as the pyranoside. In both lectins the arabinofuranosides are present as unbranched chains containing 1→2- and 1→3-linkages and these chains are linked to the polypeptide chain through hydroxyproline. The galactopyranosides are linked to serine residues of the polypeptide chain; in potato lectin there is a single α -linked substituent, but in *Datura* lectin, although some of the serine residues have a single α -linked substituent, at least half appear to be substituted with a disaccharide of composition Galp1→3Galp-Ser.

Fractionation of the hydroxypropyl oligoglycosides released by alkaline digestion of potato lectin indicated that mono-, di-, tri- and tetra-saccharides were all present, with the tri- and tetra-saccharides predominating. This is in contrast with the results of Muray & Northcote (1978), who only detected the tri- and tetra-saccharides. Although no unglycosylated hydroxyproline was detected in the present work, with a more sensitive assay Allen *et al.* (1978) detected 6% unglycosylated hydroxyproline in potato lectin. Mass spectrometry of the tri- and tetra-saccharides showed that they had the structures Araf1→2Araf1→2Araf→Hyp and Araf1→3Araf1→2Araf1→2Araf→Hyp respectively (see Fig. 3).

The anomeric nature of these linkages was examined by polarimetry, and the $[m]_{\text{Ara}}$ values for the tri- and tetra-saccharides were found to differ

markedly from the value for the model compound. This finding is at variance with that reported previously (Allen *et al.*, 1978) that the arabinofuranosides of potato lectin were all of the β -configuration. The previous conclusion was based on the very good agreement between the mean residue rotation ($[m]_{589}$) for the arabinoside in the glycoprotein of +198° and the value of +194° for the $[m]_{589}$ of methyl β -L-arabinofuranoside, assuming complete additivity of $[m]$ for the component sugars. It is now apparent that whereas the molar rotations of the disaccharides [Hyp(Ara)₂] show good agreement with the theoretical values (Table 3) the corresponding values for the trisaccharides [Hyp(Ara)₃] show greater dextrorotation, and those of the tetrasaccharides [Hyp(Ara)₄] show less dextrorotation than expected. The combined effect of these deviations from additivity fortuitously gave average molar rotations which came close to the theoretical value. An explanation of the data for the trisaccharide is that it contains β -arabinofuranosides that interact by the formation of hydrogen bonds to give an enhanced positive rotation. In the case of the tetrasaccharide, which has a lower molar rotation than the trisaccharide, we considered the possibility that the terminal arabinofuranoside is present in the α -configuration ($[m]_{589} = -209^\circ$) and a theoretical calculation on that basis agrees well with the observed value (see Table 3). The polarimetric observations were confirmed by the action of a specific α -L-arabinofuranosidase (Laborda *et al.*, 1973) on the tetra-arabinoside; this enzyme released arabinose from the compound to yield the tri-arabinoside, but did not cause any further degradation that could be detected by electrophoresis. Together these observations indicate that the tetrasaccharide of potato lectin has a non-reducing terminal α -arabinofuranoside. In this respect it is relevant that glycopeptides from tobacco cell wall that were isolated by Akiyama & Katō (1977) showed a similar pattern in that the trisaccharide was much more dextrorotatory than the tetrasaccharide. In a later paper Akiyama *et al.* (1980) showed by ¹³C n.m.r. that the terminal arabinofuranoside of the tetrasaccharide was of the α -configuration, whereas the other arabinofuranoside linkages were of the β -configuration. It therefore seems that the structures of the potato lectin hydroxypropyl tri- and tetra-arabinoside are as shown in Fig. 3 and that they are identical with those found in the cell-wall preparations of tobacco cells by Akiyama *et al.* (1980) and presumably with those of hydroxyproline-rich glycopeptides of other higher plants. This means that potato lectin is likely to be a useful soluble model for studies of the biosynthesis of insoluble hydroxyproline-rich glycoprotein in cell walls, as has already been indicated by the work of Owens & Northcote (1981).

In our opinion, a minimum of three enzymes would be involved in the transfer of arabinose from UDP-arabinose to form the complete tetrasaccharide. The first responsible for the formation of a β -linkage with hydroxyproline, the second forming a β 1 $\rightarrow$ 2-linkage between arabinofuranose residues and the third to add an α -arabinofuranoside residue to the 3-position of the penultimate residue. Furthermore, since both anomers are present in the structure, if a single nucleotide-sugar is the precursor, it follows that one of the enzymic transfers would require a double Walden inversion and the possible involvement of an intermediate acceptor (see Ashford & Neuberger, 1980).

We would like to thank Mr. R. Self, Mr. J. Eagles and Mrs. K. Barnes of the Food Research Institute for performing the mass-spectral analysis. The work done at Charing Cross Hospital Medical School was supported by the Agricultural Research Council, The Nuffield Foundation and the Central Research Fund of London University.

References

- Akiyama, Y. & Katō, K. (1976) *Agric. Biol. Chem.* **40**, 2343–2348
- Akiyama, Y. & Katō, K. (1977) *Agric. Biol. Chem.* **41**, 79–81
- Akiyama, Y., Mori, M. & Katō, K. (1980) *Agric. Biol. Chem.* **44**, 2487–2489
- Allen, A. K. & Neuberger, A. (1973) *Biochem. J.* **135**, 307–314
- Allen, A. K., Desai, N. N., Neuberger, A. & Creeth, J. M. (1978) *Biochem. J.* **171**, 665–674
- Ashford, D. & Neuberger, A. (1980) *Trends Biochem. Sci.* **5**, 245–248
- Augustad, I. & Berner, E. (1956) *Acta Chem. Scand.* **10**, 911–916
- Chambers, R. E. & Clamp, J. R. (1971) *Biochem. J.* **125**, 1009–1018
- Dawson, R. M. C., Elliott, D. C., Elliot, W. H. & Jones, K. M. (1969) *Data for Biochemistry*, 2nd edn., p. 532, Oxford University Press, London
- Desai, N. N. & Allen, A. K. (1979) *Anal. Biochem.* **93**, 88–90
- Desai, N. N., Allen, A. K. & Neuberger, A. (1981) *Biochem. J.* **197**, 345–353
- Dubois, M., Gillies, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) *Anal. Chem.* **28**, 350–356
- Ferrier, P. J. (1980) in *The Carbohydrates, Chemistry and Biochemistry* (Pigman, W. & Horton, D., eds.), pp. 1354–1375, Academic Press, New York
- Finar, I. L. (1960) *Organic Chemistry*, vol. 2, 2nd edn., pp. 584–586, Longmans, London
- Fournet, B., Dhalluin, J.-M., Strecker, G. P., Monteil, J., Bosso, C. & Defaye, J. (1980) *Anal. Biochem.* **108**, 35–56
- Heath, M. F. & Northcote, D. H. (1971) *Biochem. J.* **125**, 953–961
- Horejši, V. & Kocourek, J. (1978) *Biochim. Biophys. Acta* **532**, 92–97
- Jansson, P. E., Kene, L., Liedgren, H., Lindberg, B. & Lönngren, J. (1976) *Chem. Commun. Univ. Stockholm* **8**, 1–75
- Kärkkäinen, J. (1969) *Carbohydr. Res.* **11**, 247–256
- Kärkkäinen, J. (1971) *Carbohydr. Res.* **17**, 1–10
- Kochetkov, N. K. & Chizhov, O. S. (1966) *Adv. Carbohydr. Chem. Biochem.* **21**, 39–93
- Kovacik, V., Bauer, S., Rozik, J. & Kovac, P. (1968) *Carbohydr. Res.* **8**, 282–290
- Laborda, F., Fielding, A. H. & Byrde, R. J. W. (1973) *J. Gen. Microbiol.* **79**, 321–329
- Lamport, D. T. A. (1967) *Nature (London)* **216**, 1322–1324
- Lamport, D. T. A. (1977) *Recent Adv. Phytochem.* **11**, 79–115
- Miller, D. H., Lamport, D. T. A. & Miller, M. (1972) *Science* **176**, 918–920
- Moor, J. & Waight, E. S. (1975) *Biomed. Mass Spectrom.* **2**, 36–45
- Murray, R. H. A. & Northcote, D. H. (1978) *Phytochemistry* **17**, 623–629
- Nachbar, M. S., Oppenheim, J. D. & Thomas, J. O. (1980) *J. Biol. Chem.* **255**, 2056–2061
- Neuberger, A. & Wilson, B. M. (1971) *Carbohydr. Res.* **17**, 89–95
- O'Neill, M. A. & Roberts, K. (1981) *Phytochemistry* **20**, 25–28
- O'Neill, M. A. & Selvendran, R. R. (1980) *Biochem. J.* **187**, 53–63
- Owens, R. J. & Northcote, D. H. (1981) *Biochem. J.* **195**, 661–667
- Ring, S. G. & Selvendran, R. R. (1978) *Phytochemistry* **17**, 745–752
- Ring, S. G. & Selvendran, R. R. (1980) *Phytochemistry* **19**, 1723–1730
- Ring, S. G. & Selvendran, R. R. (1981) *Phytochemistry* **20**, 2511–2519
- Robinson, D. S. & Greenstein, J. P. (1952) *J. Biol. Chem.* **195**, 383–388
- Stegemann, H. (1958) *Z. Physiol. Chem.* **311**, 41–45
- Strecker, G., Pierce-Cretal, A., Fournet, B., Spik, G. & Montreuil, J. (1981) *Anal. Biochem.* **111**, 17–26