Carbohydrate Research, 154 (1986) 289–295
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

Note

Glycosylated hydroxyproline derivatives from *Acacia erioloba* exudate

DAVID W. GAMMON AND ALISTAIR M. STEPHEN
Department of Organic Chemistry, University of Cape Town, Rondebosch 7700 (South Africa)
(Received December 28th, 1985; accepted for publication, March 31st, 1986)

The general structural features of glycoproteins found in the stem exudate of *Acacia erioloba* have recently been described[1]. The nature of the evidence for covalent attachment of carbohydrate to hydroxyproline (Hyp) in the protein portion of the exudate required clarification, and this has now been sought by isolating and characterising some of the hydroxyproline glycosides released on alkaline hydrolysis.

A complex mixture of oligomers containing both carbohydrate and Hyp and totalling 36% by weight of starting material was formed on hydrolysis of *A. erioloba* gum sample *B* with saturated aqueous barium hydroxide. During this treatment, the Hyp residues were epimerised at C-2 to give a mixture of the naturally occurring isomer, *cis*-L-Hyp, and the C-2 epimer, *allo*-D-Hyp. Glycosides of Hyp are therefore present as diastereomeric mixtures, contributing further to the complexity of the mixture of products and accounting for their behaviour in t.l.c. Each of the glycosides G1–G11 gave, after p.c. or t.l.c., colour responses similar to that of Hyp in forming a red chromophore which changed to yellow–orange on further heating or storage for ~2 h. The positive $[\alpha]_D$ values recorded for these glycosides contrast with the negative value for the gum sample as a whole, reflecting the laevorotatory contribution of the protein. Some free Hyp was detected at V_T of the column (Fig. 1), but this was present in low proportion relative to glycosylated Hyp. G1 and G4, the two major components (33 and 22%) of the mixture of glycosides, represent neutral and acidic oligosaccharides, respectively, and are of relatively low molecular weight and structural complexity (Table I).

β-L-Ara*f*-(1→4)-DL-Hyp **(1)**

The two diastereomers **(1)** comprising G1 were partially separated by p.c. (solvent *B*), with the component of lower mobility in t.l.c. yielding Ara and *allo*-D-Hyp on hydrolysis. The latter had a chromatographic mobility that was the same as that of the diastereomer formed on heating L-Hyp in acetic acid–acetic anhydride[2]. Methylation analysis showed Ara to be present solely as terminal, non-reducing Ara*f*. The ^{1}H-n.m.r. spectrum of G1 showed, in addition to complex signals centred

0008-6215/86/$ 03.50

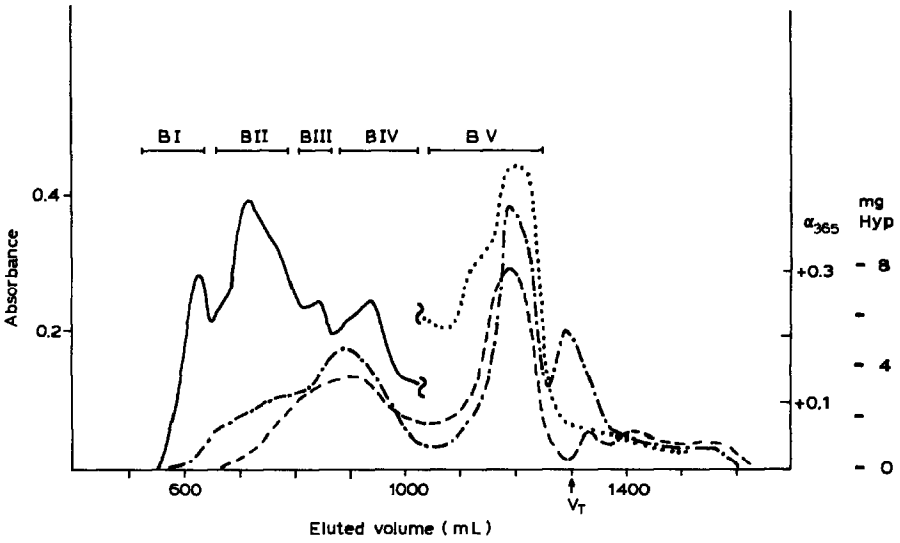

Fig. 1. Preparative chromatography (see Experimental) of products from *B* on Bio-Gel P-2 (V_o, $\overline{M}_w$ ~1800 on basis of dextran calibration, at 600 mL); absorbance at 490 nm following addition of phenol and H_2SO_4 to 20-μL (——) or 100-μL (·····) aliquots withdrawn from collected fractions; optical rotation at 365 nm (-----); Hyp content of collected fractions of 18 mL each (-·-·-·-).

at δ 4.54, 3.45, and ~2.3 due to Hyp proton resonances[3], two doublets ($J_{1,2}$ ~4.1 Hz) at δ 5.06 and 5.09. The spectrum of the component in G1 containing D-Hyp had one doublet at δ 5.07 (J 4.2 Hz), indicating that the two doublets in the spectrum of G1 arise from differences in the chemical shift of the anomeric proton of Ara*f* induced by configurational and conformational changes in the aglycon, an effect analogous to that observed for sugar residues adjacent to reducing residues in oligosaccharides[4]. The ^{1}H chemical shifts and coupling constants for the doublets

TABLE I

GLYCOSIDES FROM PRODUCTS OF ALKALINE HYDROLYSIS OF *A. erioloba* EXUDATE

	G1[a]	*G4*
Wt. % of *B*	12	8
R_{Hyp}[b]	0.81	0.52
R_{Hyp}[c]	0.93	1.06
	0.78	0.84
R_{Hyp}[d]	0.87	0.47
	0.75	0.39
$[\alpha]_D$ (degrees)	+82	+30
Neutral sugar (mol %)		
Arabinose	100	23
Galactose	—	77

[a]Found: N, 5.1. Calc. for $C_{10}H_{17}NO_7$: N, 5.3%. [b]P.c. (solvent *A*). [c]T.l.c. (solvent *C*). [d]T.l.c. (solvent *A*).

are assigned to glycosidically linked β-L-Ara*f*[5,6], and the ratio of integrated intensities of these signals to those of the Hyp protons indicated a 1:1 ratio of Ara and Hyp. Confirmation of the structure proposed for G1 was obtained by ^{13}C-n.m.r. spectroscopy (Table II), the relatively high-field position of the anomeric carbon resonances indicating β-L-Ara*f*-(1→. The anomeric carbon and certain Hyp carbons gave double resonances, attributable to the presence of diastereomers in G1 and consistent with similar observations in ^{13}C-n.m.r. studies of hydroxyproline arabinosides derived by alkaline hydrolysis from suspension-cultured tobacco *(Nicotiana tabacum)* cells[8]. Since G1 has a high positive $[\alpha]_D$ value (+82°), the absolute configuration L-Ara*f* is established.

G4, containing Gal:Ara ~3:1, has been shown to include **2** as a structural feature, although Ara is also present, in an as yet undetermined structural role. By following its methanolysis, it was shown that 4MeGlcA and Gal were present in the ratio ~1:1, and limited hydrolysis yielded **3**. Hydrolysis of G4 under conditions known to split Ara*f*→Hyp, but not Gal*p*→Hyp, linkages[9] yielded Ara*f* but not free Hyp, as well as a component with a mobility the same as that of G4, indicating the presence of a Gal*p*→Hyp linkage in G4. Methylation analysis confirmed the proportions and modes of linkages of 4MeGlcA and Gal. Since (*a*) Ara*f* was shown to be present predominantly as terminal non-reducing residues, (*b*) no branched-sugar residues were detected, and (*c*) no free oligosaccharide was found in G4, Ara*f* is probably also attached to Hyp, within a dipeptide.

4-*O*-Me-α-D-Glc*p*A-(1→4)-β-D-Gal-(1→4)-Hyp
2

4-*O*-Me-α-D-GlcA-(1→4)-D-Gal
3

Signals in the 500-MHz ^{1}H-n.m.r. spectrum of G4 were assigned on the basis of those observed in spectra of aldobiouronic acids I and II (≡ **3** and **4**) obtained in earlier work[1] and from the Hyp proton resonances for *cis*-L-Hyp and *allo*-D-Hyp[3]. Since all of the uronic acid residues in the glycoproteins have been shown[1] to

TABLE II

^{13}C-N.M.R. DATA[a] FOR G1 (**1**)

		Chemical shifts δ (p.p.m.)				
		C-1	*C-2*	*C-3*	*C-4*	*C-5*
Residue in G1	Hyp	?	60.8	37.4	77.2	52.1
	Ara	101.0 100.8	77.0	75.2	82.8	63.9

[a]Assignments made by reference to Allerhand *et al.*[7].

be in the form of 4MeGlcA, the integral ratio of the lowest-field doublet due to H-1 of α-4MeGlcA and the singlet at δ 3.51 indicates that all of this uronic acid is α-linked. The presence of diastereomers of Hyp is illustrated most clearly by the two sets of overlapping doublets centred (at 30°) at δ 4.37 and 4.32, respectively, with a 3:2 ratio of integrated intensities, which are due to H-2 of L-Hyp and *allo*-D-Hyp, respectively. The same integral ratio was observed for doublets centred at 4.58 and 4.50, the region of β-Gal-(1→ anomeric proton resonances. The fact that the ratios of integrated intensities of the anomeric proton doublet for α-4MeGlcA, the combination of doublets between δ 4.5 and 4.6, and those between δ 4.3 and 4.4 are ~1:1:1 is further evidence for the presence of **2**. The absence of a signal above δ 5.0, and the presence of a broad singlet at δ 4.79 (at 80°), could be indications of the presence of α-Ara*f*-(1→.

Further hydrolysis of G4 in $Ba(OH)_2$, followed by p.c., showed that, while a component having the same mobility as G4 remained, three new components (R_{Hyp} 0.57, 0.95, and 1.08) were also detected, suggesting that G4 is a mixture of Hyp glycosides and small glycopeptides.

Comparative structural investigation of oligomers G5–G11. — The $[\alpha]_D$ values were all positive, ranging from +3° (G9, G10) to +21° (G6), except for G11 which was not measurable. The molar proportions of Ara varied from 28 to 44%, Gal constituting the only other neutral sugar residue. The overall ratio of Ara to Gal within the products G1–G11 was 0.9, similar to the value for the intact glycoprotein.

Methylation studies, including analyses of the products of reduction of the permethylated oligomers with lithium aluminium deuteride, showed that 4MeGlcA is present in each of G5–G11 solely as non-reducing terminal residues, whereas Ara*f* residues are preponderantly in this form but some are (1→2)-linked. Terminal Gal*p* residues were detected only in G10 and G11, although a significant proportion was detected also in G3, a minor, uronic acid-free component which, from the end-groups present, is probably a mixture of Hyp glycosides of Ara and Gal. The high proportions of end-groups found, the ubiquity of →4)-Gal and →3,6)-Gal, and the occurrence in the more complex products (G9–G11) of →6)-Gal are consistent with units **1, 2,** and variants involving some branching of Gal residues being present. As previously reported[1], the number of end-groups in the larger molecules is more than required for their apparent molecular weights (see Fig. 1), this suggesting that two or more contiguous glycosylated Hyp residues remain within these oligomeric products. Microheterogeneity of the glycoproteins would be expected to give rise to mixtures of closely related and hence chromatographically indistinguishable components.

The 500-MHz ^{1}H-n.m.r. spectra of the oligomers G7, G8, and G9 provided further evidence for the structural features described. The anomeric doublets for α-4MeGlcA-(1→ are clearly distinguishable in the region δ 4.93–4.97 of spectra run at 80°, while the ratios of their integrated intensities to those of the methoxyl singlets around δ 3.50 (1:6 in G7, 1:3 in G8, and 1:6 in G9) provide evidence for

the presence (in G7 and G9) or absence (in G8) of β-4MeGlcA-(1→. This agrees with the detection, by p.c., of aldobiouronic acids **3** and **4** in products of limited hydrolysis of G7 and G9. As in the spectra (at 30° and 80°) of G4, the proton signals for Hyp H-2 (δ 4.27–4.37), H-5 (3.28–3.36), and H-3 (2.1–2.66) were clearly resolved from the carbohydrate proton resonances. The ratios of the integrated intensities of these signals and that of H-1 of α-4MeGlcA are indications of the multiplicity of Hyp residues present. For G7 and G8, the α-4MeGlcA:Hyp ratio is ~1:2, whereas it is ~1:1 for G9.

4-*O*-Me-β-D-GlcA-(1→6)-D-Gal
4

However, further alkaline hydrolysis of G7–G11 released more than two Hyp glycosides from each, suggesting that they contain mixtures of related glycopeptides incorporating contiguous Hyp residues. The hydrolysis conditions [$Ba(OH)_2$, 8 h, 100°] were clearly insufficient to bring about degradation to single glycosylated Hyp residues, which are the fundamental structural subunits. This finding is consistent with the earlier observation[1] of enrichment of a protease-resistant portion of *A. erioloba* exudate in carbohydrate and Hyp, suggesting a concentration of glycosylated Hyp residues in domains within the glycoproteins.

Exudates from *A. erioloba* thus contain glycoproteins, with a low overall content of carbohydrate, which show unique structural features. Both Ara→Hyp and Gal→Hyp glycosidic linkages occur in the same preparations, the Ara*f* being attached to Hyp as single residues and not in chains[10]. 4MeGlcA has not been previously shown to be a constituent of oligosaccharide substituents in glycoproteins[10], although some high-protein exudates have been reported to be relatively rich in this component[11,12]. Terminal non-reducing Ara*f* and aldobiouronic acid units **3** and **4**, here found to be attached to Hyp residues, are commonly encountered as peripheral features of *Acacia* gum polysaccharides; these glycosylated hydroxyproline segments may either contribute to the outer chains of complex acidic arabinogalactans, through a transglycosylation mechanism, or serve as an alternative repository for the sugar moieties.

EXPERIMENTAL

General methods. — These were as described[1]. P.c. and t.l.c. were performed with *A*, 2:1:1 1-butanol–acetic acid–water; *B*, 18:3:1:4 ethyl acetate–acetic acid–formic acid–water; *C*, 20:20:7 chloroform–methanol–water. Preparative p.c. was performed on Whatman 3MM sheets. N.m.r. spectra were recorded for deuterium-exchanged samples with a Bruker WH-90 spectrometer (22.6 MHz, ^{13}C) or a Bruker WM-500 spectrometer (500 MHz, ^{1}H). The ^{1}H chemical shifts were measured with reference to internal acetone (δ 2.230) and the ^{13}C chemical shifts

at 22.6 and 125.7 MHz with reference to internal 1,4-dioxane (δ 67.4) and TSP, respectively.

The origin and isolation of gum sample *B* from *Acacia erioloba* has been described[1].

Alkaline hydrolysis. — Gum nodules from sample *B* (10 g) were stirred at room temperature in 0.22M $Ba(OH)_2$ (250 mL) for 2 h, and the mixture was then heated at 100° for 8 h; the gum nodules dissolved completely after ~3 h. Products of hydrolysis were recovered from the yellow solution[1], yielding alkali-degraded material (7 g; $[\alpha]_D$ positive). P.c. and t.l.c. revealed several amino acids, and components containing carbohydrate which gave a response with ninhydrin similar to that of Hyp. Portions (2 g) of this mixture were fractionated on a preparative column (153 × 4.5 cm) of Bio-Gel P-2, eluted with water at ~100 mL/h (Fig. 1), to give fractions *B*I–*B*V. Hyp was measured on aliquots, diluted ten-fold, by Leach's method[1,13]. These fractions were separated further, in 150-mg portions, by preparative p.c. (solvent *A*), yielding oligomers G1–G11, in proportions by weight, relative to *B*, as follows: 12, 3, 2, 8, <1, 2, 3, 1, 2, 1, and 1%. The yield of oligomers (36%) was high in view of the carbohydrate content of *B* being ~47%.

*B*I gave G7 and G9–G11, *B*II gave G4–G9, *B*III gave G4, *B*IV gave G3 and G4, and *B*V gave G1 and G2. Each fraction migrated as one zone in p.c. and (except for the two slowest moving) as two components in t.l.c. (solvent *A*).

Graded acid hydrolysis of oligomers obtained after alkaline hydrolysis. — In a typical experiment, a portion (2–3 mg) of the oligomer was dissolved in 0.1M trifluoroacetic acid (final concentration, 5 mg/mL) in a small, screw-capped vial, and the solution was heated at 100° for 1 h and then cooled. Two aliquots (80 μL) were removed, freeze-dried, and examined by p.c. The remaining solution was adjusted to 2M with respect to trifluoroacetic acid, heated at 100° for 3 h, and cooled; two further aliquots were then removed and freeze-dried for p.c. analysis. Heating at 100° was continued for 15 h, aliquots were taken for p.c. as before, the remainder was freeze-dried, and the residue was converted into alditol acetates for g.l.c. analysis. The products of each stage of hydrolysis were examined by p.c. in duplicate on the same paper sheet, one sample of each being sprayed and heated with *p*-anisidine hydrochloride reagent and the other with ninhydrin.

ACKNOWLEDGMENTS

We thank the C.S.I.R. and the Council of the University of Cape Town for financial support, Miss Bridget Williamson (Mass Spectrometry Unit) for technical assistance, Dr. P. L. Wessels (National Chemical Research Laboratory) and Dr. K. Koch (Analytical Science) for the high-resolution n.m.r. measurements, and Dr. S. C. Churms for helpful discussions.

REFERENCES

1 D. W. GAMMON, A. M. STEPHEN, AND S. C. CHURMS, *Carbohydr. Res.*, in press.
2 D. S. ROBINSON AND J. P. GREENSTEIN, *J. Biol. Chem.*, 195 (1952) 383–388.
3 R. J. ABRAHAM AND K. A. MCLAUGHLAN, *Mol. Phys.*, 5 (1962) 195–203.
4 J. F. G. VLIEGENTHART, H. VAN HALBEEK, AND L. DORLAND, *Pure Appl. Chem.*, 53 (1981) 45–77.
5 Y. AKIYAMA AND K. KATÔ, *Agric. Biol. Chem.*, 41 (1977) 79–81.
6 S. J. ANGYAL, *Carbohydr. Res.*, 77 (1979) 37–50.
7 A. ALLERHAND, K. DILL, AND E. BERMAN, *Carbohydr. Res.*, 97 (1981) 331–336.
8 Y. AKIYAMA, M. MORI, AND K. KATÔ, *Agric. Biol. Chem.*, 44 (1980) 2487–2489.
9 D. T. A. LAMPORT, *Recent Adv. Phytochem.*, 11 (1977) 79–115.
10 R. R. SELVENDRAN AND M. A. O'NEILL, in F. A. LOEWUS AND W. TANNER (Eds.), *Encyclopaedia of Plant Physiology*, New Series, Vol 13A, Springer-Verlag, 1982, pp. 515–583.
11 D. M. W. ANDERSON, J. G. K. FARQUHAR, AND C. G. A. MCNAB, *Phytochemistry*, 22 (1983) 2481–2484.
12 D. M. W. ANDERSON AND A. HENDRIE, *Carbohydr. Res.*, 20 (1971) 259–268.
13 A. A. LEACH, *Biochem. J.*, 74 (1960) 70–71.