Carbohydrate Research, 95 (1981) 61–71
Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

PURIFICATION AND STRUCTURE DETERMINATION OF THREE α-D- GALACTOPYRANOSYLCYCLITOLS FROM SOYA BEAN

THOMAS F. SCHWEIZER* AND IAN HORMAN
Nestlé Research Department, CH-1814 La Tour-de-Peilz (Switzerland)
(Received December 8th, 1980; accepted for publication, February 5th, 1981)

ABSTRACT

Three α-D-galactopyranosylcyclitols previously isolated from soya bean are shown to be 1D-2-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol, 1D-5-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol, and 1D-2-*O*-(α-D-galactopyranosyl)-*chiro*-inositol. Assignments of the ^{13}C-n.m.r. spectra of these compounds and of 1L-1-*O*-(α-D-galactopyranosyl)-*myo*-inositol (galactinol) are presented. The mass spectra of derivatives prepared by permethylation or perdeuteriomethylation, followed by hydrolysis and acetylation or methylation, are discussed.

INTRODUCTION

We have described[1] the isolation of a mixture of three new α-D-galactopyranosylcyclitols from soya bean and their occurrence in other leguminous seeds. One of these compounds was tentatively identified as 1D-1-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol. The other two compounds were thought to be positional isomers and hence also α-D-galactopyranosylpinitols. Beveridge *et al.*[2] have reported the occurrence of two galactosylpinitols in pasture legumes and have identified the major isomer as 1D-2-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol (**1**).

We now describe the purification and structure determination of the three α-D-galactopyranosylcyclitols.

EXPERIMENTAL

Materials. — All reagents and solvents were of the purest grades commercially available. α-D-Galactosidase was obtained from Boehringer-Mannheim GmbH (Germany), and the reference cyclitols were kindly donated by Professors L. Anderson and S. J. Angyal.

General methods. — T.l.c., g.l.c., and the determinations of specific sugars were performed as described previously[1]. G.l.c.–m.s. was effected on a Pye 104 gas chro-

*Correspondence should be addressed to this author at Nestlé Products Technical Assistance Co. Ltd., CH-1814 La Tour-de-Peilz, Switzerland.

matograph, using a 50′ × 0.02″ OV-225 SCOT column, with helium as carrier gas at 5 ml/min and a temperature programme of 5°/min from 150 → 230°. The column effluent was introduced into an AEI MS-30 mass spectrometer through an all-glass jet separator maintained at 250°. E.i.-mass spectra were obtained at 70 eV and 300 μA with an ion-source temperature of 230°, and were recorded at a resolving power of 1500 on a Kratos DS-50 data system with perfluorokerosine as the external standard. Trimethylsilyl (TMS) derivatives were prepared by the method of Sweeley *et al.*[3]. Procedures for methylation and trideuteriomethylation, followed by hydrolysis with 0.5M sulphuric acid, neutralisation with barium carbonate, reduction with sodium borohydride, and acetylation, were as described by Hakomori[4] and Björndal *et al.*[5]. ^{13}C-N.m.r. spectra were recorded at 30° with a Varian FT-80A spectrometer operating at 20 MHz. Compounds were dissolved in 300 μl of D_2O in a 5-mm tube to give ~150mM solutions. Proton-decoupled spectra were processed and recorded after 50k–120k transients. Line positions were defined by using 1,4-dioxane as the internal standard, and are reported as chemical shifts downfield from the signal for Me_4Si relative to which the 1,4-dioxane signal appears at δ 67.4.

Purification of the galactosylcyclitols. — Mixtures of the galactosylcyclitols **1**–**3** were isolated from soya as already described[1]. Aliquots (~750 mg) of the mixture (**1**:**2**:**3** ~10:8:1) were eluted from a column (2.5 × 45 cm) of charcoal–Celite (1:1) with 2% aqueous ethanol (500 ml) followed by 5% aqueous ethanol. Fractions (50 ml) were taken, and the sugars therein were analysed by g.l.c. of the TMS derivatives. Fig. 1 shows the fractionation thus obtained. Melibiose was sometimes present in small amounts owing to incomplete elimination as the phenylosazone[1]. Fractions were combined as indicated by A–C in Fig. 1. The ratios of **1**–**3** in A–C were as follows: A (30 mg) 0:1:6; B (310 mg) 25:50:1; and C (260 mg) 6:1:0. The total recovery was 85–90%. Rechromatography of A–C, as described above, gave pure **1** and **3**. Compound **2** could not be purified by this procedure, and a solution of a fraction B containing 60 mg of **1** and 140 mg of **2** in distilled water was therefore incubated with 50 U of α-D-galactosidase (1 ml of 3.2M ammonium sulphate suspension) at 37° for 15 h. The mixture was concentrated *in vacuo* and then freeze-dried, and the residue was subjected to chromatography on charcoal–Celite. Galactose, pinitol, and the salt were eluted with distilled water (1 litre). Elution with 5% aqueous ethanol then gave 45 mg of **2** together with a mixture containing 25 mg of **2** and 6 mg of **1**. Compounds **1**–**3** were obtained, after freeze-drying, as white powders, and were homogeneous by t.l.c.[1] with R_G values of 0.72, 0.58, and 0.16, and $[\alpha]_D^{23}$ values (*c* 0.2, water) of +181, +159, and +170°, respectively.

RESULTS AND DISCUSSION

The galactosylcyclitols **1** and **3** could be purified by chromatography on charcoal–Celite (Fig. 1). Compound **2** was isolated after treatment of a mixture of **1** and **2** with α-D-galactosidase, which hydrolyses **1** much faster than **2** (Fig. 2). Treatment of **1** and **2** with α-D-galactosidase gave equal amounts of D-galactose and pinitol (**5**),

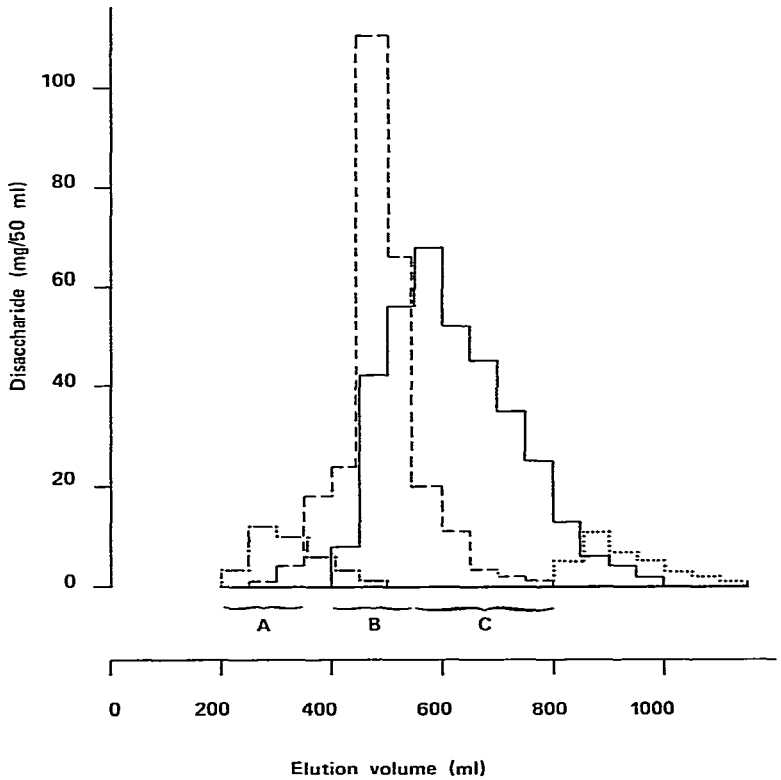
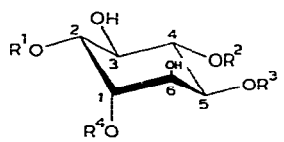
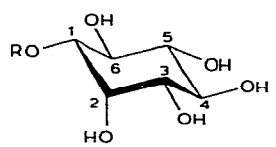


Fig. 1. Fractionation of α-D-galactosylcyclitols and melibiose on charcoal–Celite (1 : 1). Order of elution: **3** (—·—·—·), **2** (———), **1** (———), and melibiose (·····). A, B, and C show the fractions that were combined.



1 R^1 = α-D-Gal, R^2 = Me, R^3 = R^4 = H
2 R^1 = R^4 = H, R^2 = Me, R^3 = α-D-Gal
3 R^1 = α-D-Gal, R^2 = R^3 = R^4 = H
4 R^1 = R^3 = H, R^2 = R^4 = Me
5 R^1 = R^3 = R^4 = H, R^2 = Me
6 R^1 = Me R^2 = R^3 = R^4 = H
7 R^1 = R^2 = R^3 = R^4 = H



8 R = α-D-Gal
9 R = Me
10 R = H

whereas **3** gave D-galactose and D-*chiro*-inositol (**7**). The D configuration of galactose was proved by oxidation to galactonic acid with NAD in the presence of β-D-galactose-dehydrogenase. The identities of **5** and **7** were shown by g.l.c.–m.s. of their TMS derivatives. Since α- but not β-D-galactosidase hydrolysed **1–3**, α linkages were present. Methylation analysis[5] of **1–3** gave 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methylgalactitol, indicating the galactosyl group to be pyranoid.

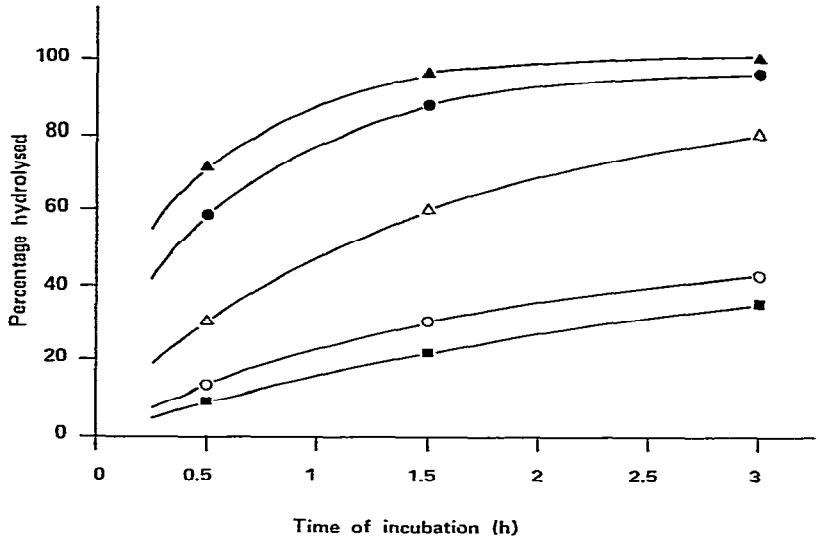


Fig. 2. Enzymic hydrolysis of α-D-galactosylcyclitols and raffinose (1 ml of 5mM solution) with 2U of α-D-galactosidase. Raffinose, ▲; **1**, ●; **2**, ■; **3**, ○; and **8**, △.

Table I contains the ^{13}C-n.m.r. data for **1**–**3** and for some reference compounds. Signals from the galactose moieties of **1**–**3** and **8** were readily identifiable, because substitution at C-1′ of α-D-galactose has little influence on the chemical shifts of other carbons. The absolute assignments were based on recent work[6,7].

The signals of the cyclitol moieties were compared with those of partially methylated inositols[8]. Thus, **3** and **8** show signals at δ 68.5 and 68.8, respectively, typical of a carbon atom bearing an axial hydroxyl group and vicinal to an equatorial methoxyl substituent[8,9]. Accordingly, the α-D-galactosyl group in **3** must be attached to C-2 or C-5 (identical positions) of D-*chiro*-inositol. Based on the C-6 resonance[9] of pinpollitol (**4**) at δ 69.1, the galactosyl group could have been located at C-1 or C-6 (identical positions) in D-*chiro*-inositol. However, this possibility can be eliminated, because the resonances of C-3 and C-4 along with those of C-2 and C-5 would then have appeared in the same place as for **7**. This is not the case, so that **3** must be 1D-2-*O*-(α-D-galactopyranosyl)-*chiro*-inositol. The usual downfield shift (8–10 p.p.m.) in the resonance of an α-carbon induced by methylation of its hydroxyl group is diminished by 4.7 and 4.1 p.p.m., respectively, in **8** and in **3**, showing that the galactosyl group has a smaller inductive effect.

The ^{13}C-n.m.r. spectrum of **1** is closely similar to that of **3**. The signal at δ 68.2 suggests substitution at C-2 or C-5. However, unlike **3**, these two positions in **1** are no longer equivalent. Substitution at C-5 would be expected to leave the resonances of C-2 and C-3 unchanged relative to those of pinitol (**5**). With the galactosyl residue at C-2, the signal for C-5 should remain unchanged, and that for C-3 might be expected to move ~0.8 p.p.m. upfield, as typically observed in substituted *chiro*-inositols for C-OH*eq* vicinal to C-OMe. Table I shows that the latter interpretation fits with the observed line positions. Further evidence for structure **1** lies in the near-

TABLE I

^{13}C-N.M.R. CHEMICAL SHIFTS[a] (p.p.m.) OF GALACTOSYLCYCLITOLS AND RELATED COMPOUNDS IN D_2O

Compounds	*C-1′*	*C-2′*	*C-3′*	*C-4′*	*C-5′*	*C-6′*	*C-1*	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*	*OCH_3*
α-D-Galactose (**11**)	93.15	69.25	70.05	70.15	71.25	62.0							
Methyl α-D-galactopyranoside (**12**)	100.15	69.0	70.3	70.05	71.4	62.0							
Galactinol (**8**)	96.0	68.9§	70.0*	69.8*	71.7†	61.6	76.25	68.8§	72.95	71.55†	74.85	71.55†	
Bornesitol (**9**)							80.85	68.3	72.6	71.45	74.75	72.0	57.2
myo-Inositol (**10**)							72.9	72.65	72.9	71.6	74.8	71.6	
chiro-Inositol (**7**)							72.2	71.0	73.25	73.25	71.0	72.2	
Quebrachitol (**6**)							67.6	80.55	72.35	73.25	70.8	71.75	57.3
3	96.5	69.1	70.15*	69.95*	71.65§	61.7	68.5	76.4	72.05†	73.5	71.2§	71.85†	
1	96.4	68.95	70.0*	69.85*	71.35§	61.6	68.2	76.4	71.75§	83.3	70.35	71.5§	60.3
Pinitol (**5**)							71.95†	71.05	72.6	83.2	70.3	72.2†	60.2
2	95.5	68.6	70.05*	69.55*	71.3	61.3	71.3	70.9	72.8	81.8	74.75	67.65	60.9

[a]Downfield from tetramethylsilane and rounded to the nearest 0.05 p.p.m. 1,4-Dioxane (δ 67.4) was used as internal standard. The assignments of the resonances marked with *, §, or †, respectively, may be interchanged. The numbering throughout is according to the formula diagrams. Thus, for example, **5** is 4-*O*-methyl-*chiro*-inositol instead of the usual 1D-3-*O*-methyl-*chiro*-inositol.

identical resonances for C-1,2,3,6 compared to those in **3**, and for C-4 in **1** and **5**. Thus, **1** is 1D-2-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol.

Since **2** also contains pinitol, the signal at δ 67.65 suggested substitution at C-5, the methoxyl-group peak being shifted 0.7 p.p.m. downfield relative to that in pinitol (**5**). As expected, the C-2 and C-3 resonances are unchanged. Thus, **2** appeared to be 1D-5-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol. The alternative possibility, namely substitution at C-1, was considered by comparison of the spectra of **2** and pinpollitol (**4**). However, there is no obvious reason for the C-4 signal to be shifted upfield by 1.4 p.p.m. Proof that **2** is 1D-5-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol was obtained by mass spectrometry.

The mass spectra[1] of the TMS derivatives of **1**–**3** allowed only a tentative identification of structure. The mass spectra of the trideuteriomethylated **1**–**3** and **8**

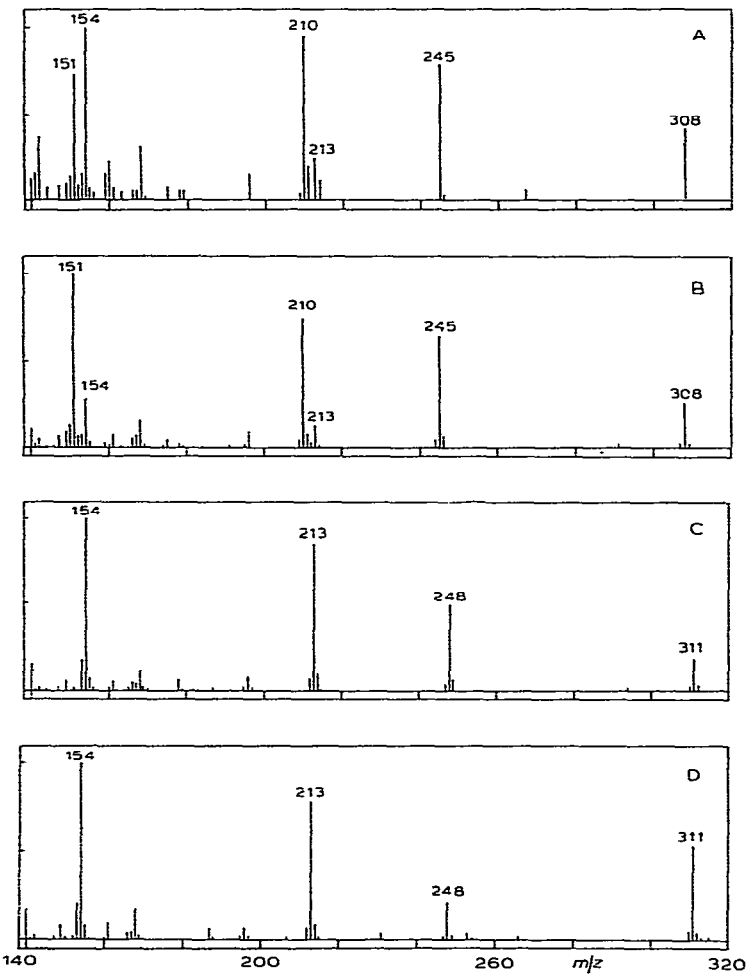


Fig. 3. Mass spectra (e.i., 70 eV) of pertrideuteriomethylated α-D-galactosylcyclitols **1** (A), **2** (B), **3** (C), and **8** (D).

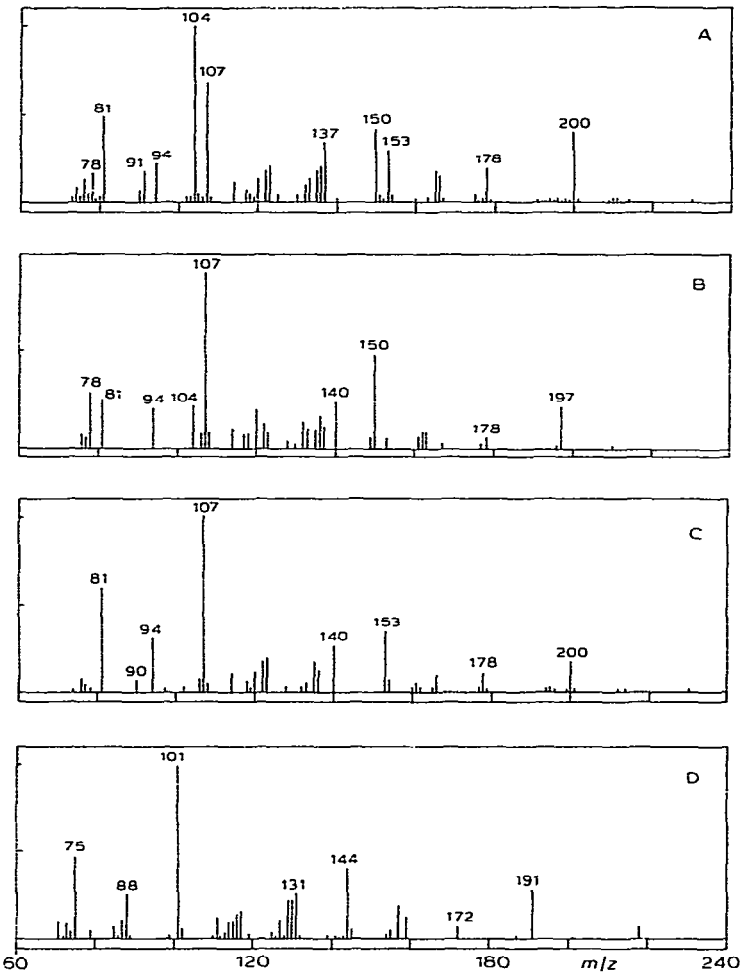


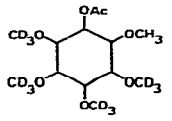
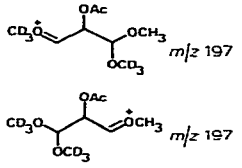
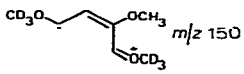
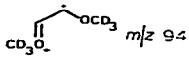
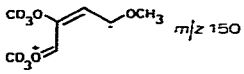
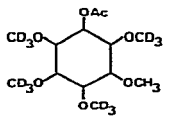
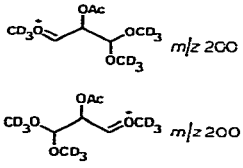
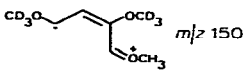
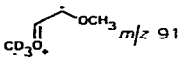
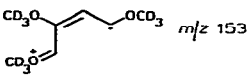
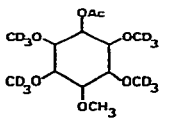
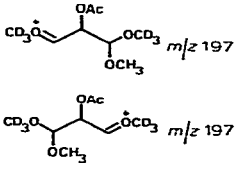
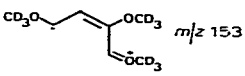
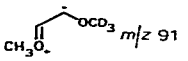
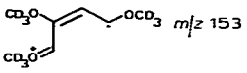
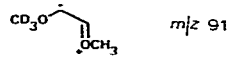
Fig. 4. Mass spectra (e.i., 70 eV) of the cyclitol acetates prepared from **1** (A), **2** (B), and **3** (C) by pertrideuteriomethylation, hydrolysis, and acetylation, and from **3** (D) by permethylation, hydrolysis, and acetylation. Peak intensities are ×5 from *m/z* 110 and ×10 from *m/z* 160.

(Fig. 3) confirm that the cyclitol moiety of **3** is not methylated, whereas those of **1** and **2** both carry a methyl group. Peaks at *m/z* 245 and 308 in Figs. 3A and 3B move to *m/z* 248 and 311 in Figs. 3C and 3D. The peak intensities for specific ions in the high-mass region of the spectra of trideuteriomethylated **1–3**, where the cyclitol is D-*chiro*-inositol, are very similar, but different to those of the corresponding derivative of galactinol (**8**), where the cyclitol is *myo*-inositol. In particular, the *m/z* 248 peak, which corresponds to the ionised cyclityl residue, is much less intense.

Hydrolysis and acetylation of trideuteriomethylated **1–3** gave cyclitols, bearing trideuteriomethoxyl groups in the positions of the original hydroxyl groups, and an acetoxyl group in the position originally occupied by the galactosyl residue. The mass spectra of these derivatives are given in Fig. 4. Three possible configurations (*IIa–IVa*

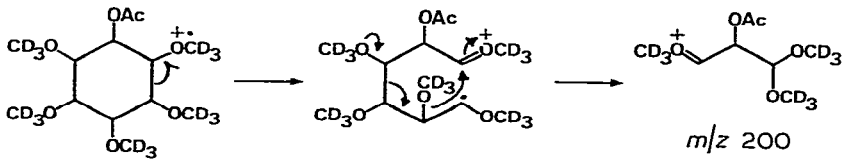
TABLE II

STRUCTURALLY SIGNIFICANT IONS FROM THE CYCLITOL DERIVATIVES *IIa–IVa*

Compound	*Structure*	*Ions formed according to fragmentation scheme 1*	*Ions formed according to fragmentation scheme 2*	*Ions formed according to fragmentation scheme 3*
IIc		 *m/z* 197	 *m/z* 150	 *m/z* 94
			 *m/z* 150	 *m/z* 94
IIIa		 *m/z* 200	 *m/z* 150	 *m/z* 91
			 *m/z* 153	 *m/z* 94
IVa		 *m/z* 197	 *m/z* 153	 *m/z* 91
			 *m/z* 153	 *m/z* 91

in Table II) had to be considered for **1** and **2**, whereas 3 could only have given rise to *Ia*. The formation of the structurally significant ions at *m/z* 200, 153, and 94 (Fig. 4C) from *Ia* is shown in Schemes 1–3, respectively.*

Transannular hydroxyl-group shifts analogous to that shown in Scheme 1



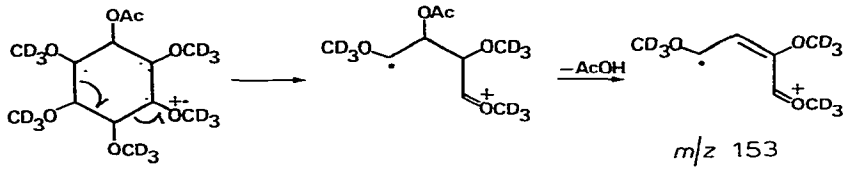
Ia

Scheme 1

have been cited in the formation of rearrangement ions for cyclohexanediols[10].

Only a mechanism involving loss of acetic acid (Scheme 2) can explain the specific retention of the pinitol methyl group to give an ion at *m/z* 150 in the spectrum

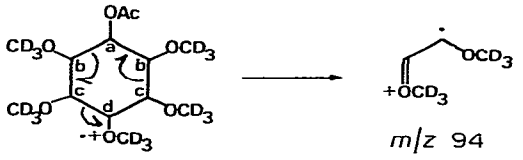
*The base-peak ion at *m/z* 107 apparently does not arise from a unique fragmentation pathway and is not considered as a characteristic fragment for structure determination.



Ia

Scheme 2

of Fig. 4B. The ion *m/z* 94 apparently consists of two adjacent ring-carbons and their attached CD_3O groups. Specific non-retention of the pinitol methyl group in the spectrum of Fig. 4B and partial retention in the spectrum of Fig. 4A indicate that the ring carbons come almost exclusively from positions *c* and *d*, as shown in Scheme 3.



Ia

Scheme 3

For **3**, after permethylation, hydrolysis, and acetylation, the fragments analogous to *m/z* 200, 153, and 94 in Schemes 1–3 are then found at *m/z* 191, 144, and 88, respectively (Fig. 4D).

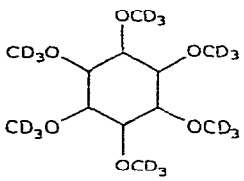
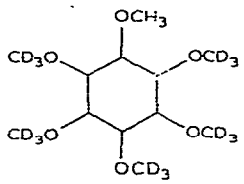
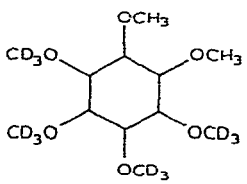
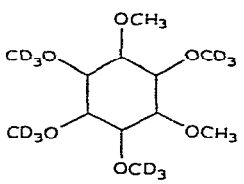
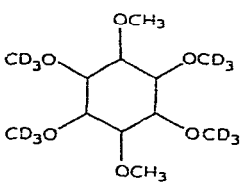
Table II shows the fragments which would be formed according to Schemes 1–3 in the mass spectra of structures *IIa-IVa*. Since the cyclitol derivatives formed after trideuteriomethylation, hydrolysis, and acetylation of **1** and **2** will contain the original pinitol methyl group in addition to four CD_3-groups, retention of the former will give fragments that are 3 mass-units smaller than those in the spectrum of *Ia* (Fig. 4C). Similarly, loss of the methyl group will give ions of the same mass, and partial retention will give pairs of ions at masses separated by 3 mass-units. Thus, comparison of the data in Table II with those in Figs. 4A and 4B shows that the derivatives of **1** and **2** are compatible with structures *IIIa* and *IIa*, respectively, but not with *IVa*.

This finding supports the ^{13}C-n.m.r. identifications described above, but is in conflict with the structure previously proposed[1] for **2**. A second series of derivatives was therefore prepared from **1** and **2** by trideuteriomethylation, hydrolysis, and methylation, to give substituted inositols carrying two methoxyl groups. One of these comes from pinitol and the other is in the position of the original link to galactose. In the mass spectra of these compounds, the formation of ions analogous to the *m/z* 94 ion in Scheme 3 is no longer directed by the presence of an acetyl group. Such ions may therefore be expected to contain any two adjacent ring-carbons and their substituents with approximately equal probability. The relative intensities of ions at *m/z* 88, 91, and 94 should thus reflect directly the relative positions of the methoxyl groups on the inositol ring.

Table III shows the theoretical and observed relative intensities of these three

TABLE III

RELATIVE INTENSITIES OF IONS AT *m/z* 88, 91, AND 94 IN THE 70-eV MASS SPECTRA OF CYCLITOL DERIVATIVES PREPARED BY PERTRIDEUTERIOMETHYLATION OF REFERENCE CYCLITOLS AND BY PERTRIDEUTERIOMETHYLATION, HYDROLYSIS, AND METHYLATION OF **1** AND **2**

Structure	*Origin*	*Relative peak intensities*					
		Theoretical[a]			*Observed*		
		m/z *88*	m/z *91*	m/z *94*	m/z *88*	m/z *91*	m/z *94*
	*chiro*-inositol	0	0	6	0.30	0.48	6
	bornesitol	0	2	4	0.38	2.20	4
	quebrachitol	0	2	4	0.34	2.22	4
	sequoyitol	0	2	4	0.30	1.84	4
	pinitol	0	2	4	0.30	1.82	4
	**2** (theoretical structure *IIb*)	1	2	3	0.91	1.33	3
	**1** (theoretical structure *IIIb*)	0	4	2	0.38	3.00	2
	(theoretical structure *IVb*)	0	4	2	—	—	—

[a]Calculated by assuming equal probabilities for the formation of ions, analogous to the *m/z* 94 of scheme 3, incorporating any two adjacent carbons of the cyclitol ring. Total theoretical possibilities add up to six in all cases.

fragments, as compared with structures *IIb-IVb* and several, related inositol derivatives. Good agreement was obtained. The minor discrepancies probably reflect steric

differences. The *m/z* 91 peak from the pinitol derivative is slightly lower than its theoretical value relative to *m/z* 94. This is also seen in the spectra of the corresponding compounds from the pinitol derivatives **1** and **2**. The relative intensities observed for the derivative from **1** are compatible with both *IIIb* and *IVb*, but the latter can be eliminated for reasons based on ^{13}C-n.m.r. data and on the mass spectrum of Fig. 4A, as already described. The relative intensities observed for the derivatives from **2** are much closer to those predicted for *IIb*, confirming that, in **2**, the galactosyl group is attached to the pinitol ring at a position adjacent to the methoxyl group.

Thus, it is concluded that the α-D-galactopyranosylcyclitols isolated from soya bean are 1D-2-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol (**1**), 1D-5-*O*-(α-D-galactopyranosyl)-4-*O*-methyl-*chiro*-inositol (**2**), and 1D-2-*O*-(α-D-galactopyranosyl)-*chiro*-inositol (**3**). Compound **1** is identical with the major galactosylpinitol of *Trifolium subterraneum*, first described by Beveridge *et al.*[2]. Considering optical rotation values and t.l.c. data, it is likely that the minor galactosylpinitol mentioned by the same authors is identical to **2**.

ACKNOWLEDGMENTS

We thank Professors L. Anderson and S. J. Angyal for generous gifts of *chiro*-inositol, quebrachitol, sequoyitol, bornesitol, *scyllo*-inositol, and *muco*-inositol, Mrs. S. Reimann and Mr. J. Llewellyn for assistance with the chromatography, and Mr. B. Dreux and Mr. E. Brambilla for spectroscopic measurements.

REFERENCES

1 T. F. Schweizer, I. Horman, and P. Würsch, *J. Sci. Food Agric.*, 29 (1978) 148–154.
2 R. J. Beveridge, C. W. Ford, and G. N. Richards, *Aust. J. Chem.*, 30 (1977) 1583–1590.
3 C. C. Sweeley, R. Bentley, M. Makita, and W. W. Wells, *J. Am. Chem. Soc.*, 85 (1963) 2497–2507.
4 S. Hakomori, *J. Biochem. (Tokyo)*, 55 (1964) 205–208.
5 H. Björndal, C. G. Hellerqvist, B. Lindberg, and S. Svensson, *Angew. Chem.*, 82 (1970) 643–674.
6 T. E. Walker, R. E. London, T. W. Whaley, R. Barker, and N. A. Matwiyoff, *J. Am. Chem. Soc.*, 98 (1976) 5807–5813.
7 P. E. Pfeffer, K. M. Valentine, and F. W. Parrish, *J. Am. Chem. Soc.*, 101 (1979) 1265–1274.
8 D. E. Dorman, S. J. Angyal, and J. D. Roberts, *J. Am. Chem. Soc.*, 92 (1970) 1351–1354.
9 J. W. Blunt, M. H. G. Munro, and A. J. Paterson, *Aust. J. Chem.*, 29 (1976) 1115–1118.
10 F. W. McLafferty, *Anal. Chem.*, 28 (1956) 306–316.