Carbohydrate Research, 102 (1982) 111–119
Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

PREPARATION OF *C*-NUCLEOSIDES BY DEHYDRATION OF PENTAHYDROXYPENTYL-HETEROCYCLES: STERIC COURSE AND MECHANISM

JUAN A. GALBIS PÉREZ, EMILIO ROMÁN GALÁN, JOSÉ L. JIMÉNEZ REQUEJO, AND FLORENTINO POLO CORRALES

Departamento de Química Orgánica, Facultad de Ciencias, Universidad de Extremadura, Badajoz (Spain)

(Received June 8th, 1981; accepted for publication, September 28th, 1981)

ABSTRACT

The acid-catalysed dehydration of pentahydroxypentyl-heterocycles can take place between C-1′ and C-4′ to yield *C*-nucleosides having furanoid structures, or between C-1′ and C-5′ to yield pyranoid compounds. Each reaction can yield one of the two possible anomers or a mixture of both. We have dehydrated pentahydroxypentyl-heterocycles having D-*gluco*, D-*manno*, and D-*galacto* configurations. When the reactions were kinetically controlled, the 6,6-dimethyl-2-(pentitol-1-yl)-4,5,6,7-tetrahydroindol-4-ones having the D-*gluco* or D-*manno* configurations yielded 2-α-D-arabinofuranosyl-6,6-dimethyl-4,5,6,7-tetrahydroindol-4-one, but they yielded 2-α-D-arabinopyranosyl-6,6-dimethyl-4,5,6,7-tetrahydroindol-4-one under conditions of thermodynamic control. Dehydration of 3-acetyl-2-methyl-5-(D-*galacto*-pentitol-1-yl)-1-propylpyrrole gave a mixture of 3-acetyl-5-(α- and β-D-lyxopyranosyl)-2-methyl-1-propylpyrrole. These results are consistent with a mechanism involving an intermediate C-1′ carbocation.

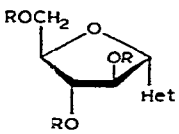
INTRODUCTION

One of the methods of preparation of *C*-nucleosides[1,2] involves dehydration of the polyhydroxyl chains joined to aromatic heterocycles[3–7]. Dehydration of pentahydroxypentyl-heterocycles can take place between C-1′ and C-4′, yielding furanoid compounds, or between C-1′ and C-5′, yielding pyranoid compounds. The configuration at C-1′ can be retained or inverted. In order to clarify this last point, Gómez Sánchez *et al.*[4] studied the dehydration of 2-(D-*arabino*-tetrahydroxybutyl)furans and proposed a mechanism involving an intermediate C-1′ carbocation. Likewise, El Khadem *et al.* studied the dehydration of sugar osazones and proposed an empirical rule[8] and a mechanism[5].

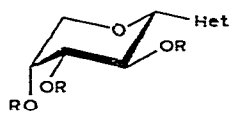
We now report on the dehydration of pentahydroxypentyl-heterocycles having D-*gluco*, D-*manno*, and D-*galacto* configurations, to obtain a more-detailed knowledge of the steric course and mechanism of this reaction.

RESULTS AND DISCUSSION

In studies of the preparation of new *C*-nucleoside analogues by intramolecular dehydration of pentahydroxypentyl-heterocycles, we have described[7] the synthesis of 2-α-D-arabinofuranosyl-6,6-dimethyl-4,5,6,7-tetrahydroindol-4-one (**1**) from the epimeric 6,6-dimethyl-2-(D-*gluco*- and D-*manno*-pentitol-1-yl)-4,5,6,7-tetrahydroindol-4-ones (**8** and **12**). Each reaction was carried out in concentrated, aqueous solution at room temperature, and **1** crystallised from the reaction mixture within a few minutes. We have now carried out the reactions using more-dilute solutions in order to avoid crystallisation of products. When such a solution was heated at 50°, **1** (R_F 0.66) was formed first and was gradually converted into a new product having R_F 0.42, which was isolated crystalline and identified as 2-α-D-arabinopyranosyl-6,6-dimethyl-4,5,6,7-tetrahydroindol-4-one (**3**). The conversion **1**→**3** also occurred on heating neutral and acidified, aqueous solutions. Compound **3** reduced 2 mol. equiv. of sodium metaperiodate, in accord with a pyranoid structure, and the p.m.r. spectrum [$(CD_3)_2SO$, Table I] contains three doublets for secondary hydroxyl groups. The $J_{1',2'}$ value (9.0 Hz) is consistent with a *trans*-diaxial arrangement for H-1',2', which is only possible for the α anomer in the $^1C_4$ conformation. A similar value (9.2 Hz) was also found for $J_{1',2'}$ of the triacetate **4**, in accord with the proposed C-1' configuration and $^1C_4$ conformation, and the $J_{4',5'}$ and $J_{4',5''}$ values (1.8 and 0.6 Hz, respectively) indicate H-4' to be equatorial.

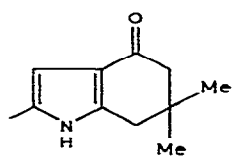


1 R = H
2 R = Ac



3 R = H
4 R = Ac

Het =



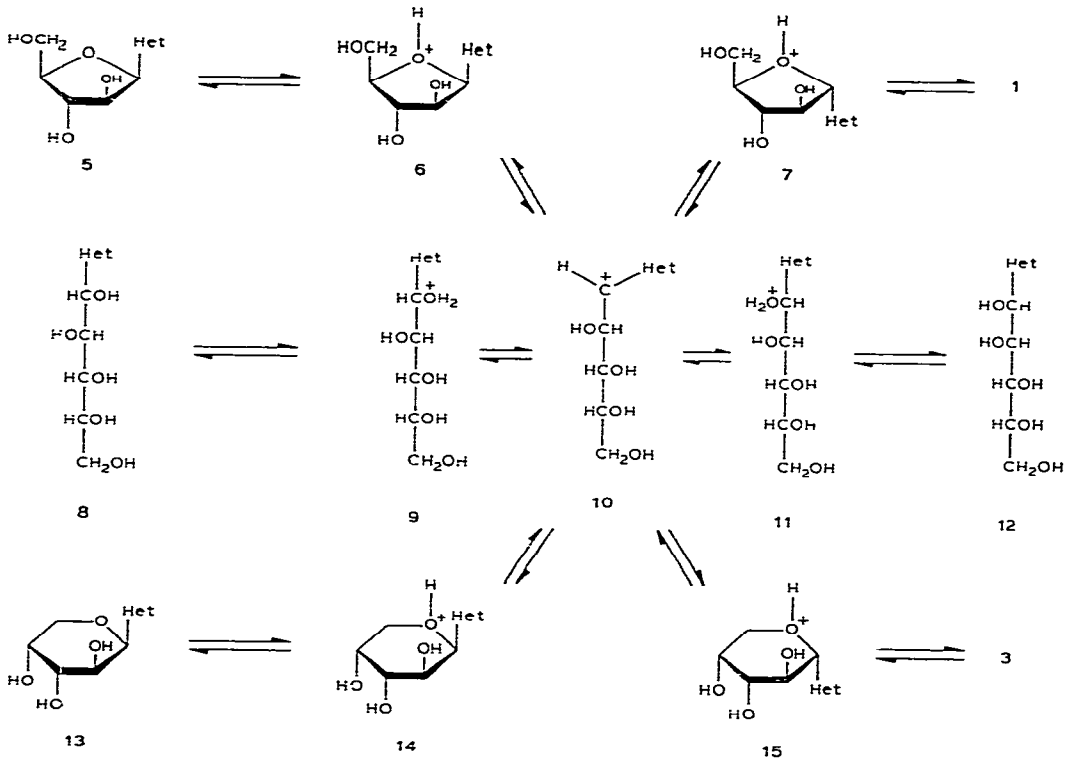
As previously indicated[7], the isolation of **1** from the pentahydroxypentyl-4,5,6,7-tetrahydroindol-4-ones having D-*gluco* (**8**) or D-*manno* (**12**) configurations supports the proposed mechanism for the dehydration of polyhydroxyalkyl-heterocycles[4] through an intermediate C-1' carbocation (Scheme 1). Compound **1**, formed first, must be the kinetically controlled product, whereas the more-stable **3** must be the thermodynamically controlled product. Under kinetic control, **1** is formed because

TABLE I

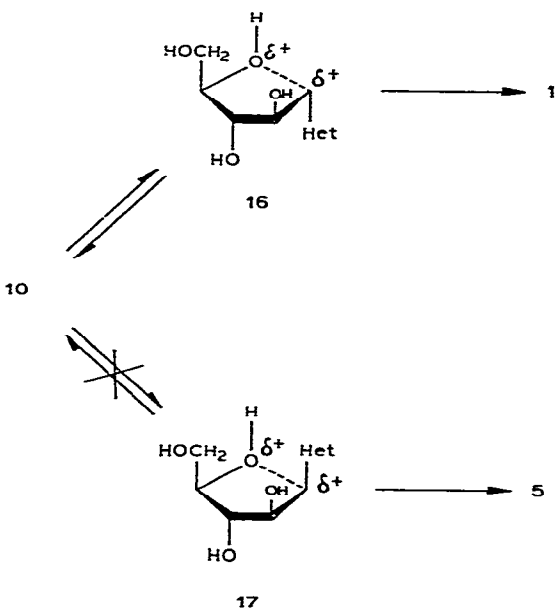
N.M.R. DATA[a] (90 MHz) FOR 2–4

Compound	Carbohydrate moiety								Aglycon				
	H-1'	*H-2'*	*H-3'*	*H-4'*	*H-5'*	*H-5''*	*OH*	*OAc*	*H-1*	*H-3*	*H-5,5*	*H-7,7*	*Me-6,6*
2[b]	5.05d $J_{1',2'}$ 2.8	← 5.55–5.15m →			← 3.85m →			2.17s (3H) 2.07s (6H)	9.35[c]	6.43d $J_{1,3}$ 2.3 $J_{1',3} \sim 1$	2.68s	2.36s	1.12s
3[d]	3.92d $J_{1',2'}$ 9.0	← 3.80–3.10m →					4.80d 4.60d 4.45d			6.17d $J_{1,3}$ 2.0 $J_{1',3} \sim 1$	2.62s	2.17s	1.03s
4[b]	4.36d $J_{1',2'}$ 9.2	← 5.50–5.00m →			4.09dd $J_{4',5'}$ 1.8 $J_{5',5''}$ −12.5	3.80dd $J_{4',5''}$ 0.6		2.18s (3H) 2.01s (3H) 1.91s (3H)	8.77[c]	6.42d $J_{1,3}$ 2.5 $J_{1',3} \sim 1$	2.65s	2.33s	1.09s

[a]Recorded at 35.5°, δ scale (internal Me_4Si), J in Hz. [b]In $CDCl_3$, [c]Broadening due to ^{14}N-quadrupole relaxation. [d]In $(CD_3)_2SO$.



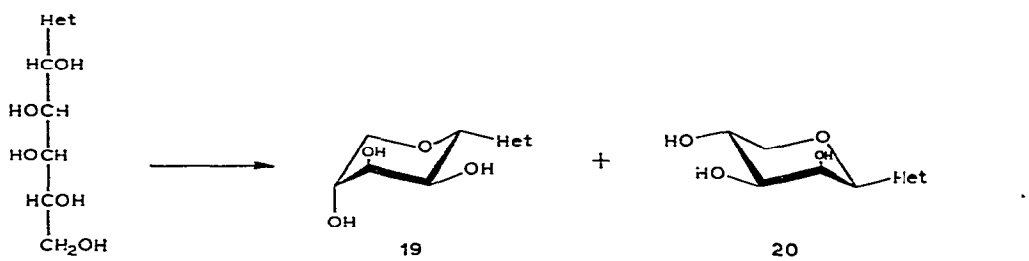
Scheme 1

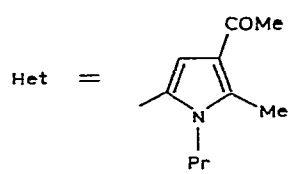


the subsequent transition-state (**16**), leading from the carbocation **10** to the α-D-arabinofuranoside (**1**), should be more stable than that (**17**) leading to the β anomer (**5**). The higher energy of **17** arises from the steric repulsions between HO-2′ and the C-1′ substituent, which are *cis*.

The α-pyranoid compound **3** in the $^1C_4$ conformation must be the more stable structure since it has fewer, bulky, axial groups than the alternative structures. These results accord with the general picture of cyclisation reactions[9] in which products containing five-membered rings are generally the kinetically controlled products, whereas those containing six-membered rings are the thermodynamically controlled products. This cyclisation mechanism *via* an S_N1 process involving a C-1′ carbocation contrasts with the cyclisation of 5-(pentahydroxypentyl)uracils[3], which may involve an S_N2 type of displacement of protonated HO-1′ by HO-4′. The difference could arise from the greater ability of our heterocycles to stabilise a "benzyl-type" C-1′ carbocation.

On the other hand, trifluoroacetic acid-catalysed dehydration of 3-acetyl-2-methyl-5-(D-*galacto*-pentitol-1-yl)-1-propylpyrrole[10] (**18**) yields an αβ-mixture of 3-acetyl-5-D-lyxopyranosyl-2-methyl-1-propylpyrrole (**19** and **20**), which were separated by fractional crystallisation.

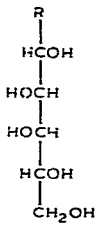




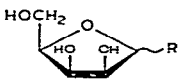
Compounds **19** and **20** consumed 2 mol. equiv. of sodium metaperiodate, and showed three doublets for the hydroxyl protons in their p.m.r. spectra [$(CD_3)_2SO$], in accord with the proposed pyranoid structures. The $J_{1',2'}$ values of 9.0 and ~1 Hz for **19** and **20**, respectively, indicated α and β configurations and $^1C_4$ and $^4C_1$ conformations, and are of the same magnitude as those for related compounds[7,11,12].

These results are analogous to those obtained by dehydration of other D-*galacto*-pentahydroxypentyl-heterocycles[7,12] that always yield anhydro derivatives having pyranoid structures, but apparently do not accord with those of Sallam[13], who described the dehydration of D-*galacto*-heptulose phenylosazone and proposed α-

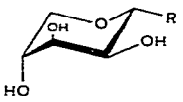
and β-furanoid structures (**22**) for the resulting anhydro derivatives. However, we believe this assignment to be in error, since the p.m.r. and c.d. data[13] are consistent with pyranoid structures. When we dehydrated D-*galacto*-heptulose phenylosazone (**21**) under similar conditions, a product was isolated which appeared to be identical with that described by Sallam as β-furanoid, but which consumed 2 mol. equiv. of sodium metaperiodate and had three doublets for secondary hydroxyl groups in its p.m.r. spectrum (see Table II), in accord with the pyranoid structure **23**. The α-anomeric configuration and the $^1C_4$ conformation are indicated by the $J_{1',2'}$ value (10.0 Hz).



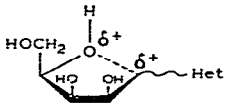
21



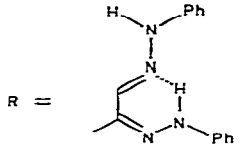
22



23



24



The mechanisms of dehydration of pentahydroxypentyl-heterocycles having the D-*galacto*, D-*gluco*, or D-*manno* configurations are probably similar[7]. Furanoid products are not formed from D-*galacto*-pentahydroxypentyl-heterocycles, because the associated transition-state **24** would be destabilised by steric repulsions[14] between the hydroxymethyl group, HO-2′, and HO-3′, which are all *cis*.

TABLE II

N.M.R. DATA[a] (90 MHz) FOR **19**, **20**, AND **23**

Compound	*Carbohydrate moiety*			*Aglycon*					
	H-1′	*H-2′,3′,4′,5′*	*OH*	*H-4*	*Me*	*COMe*	*H-C=N*	*Ph*	*NH*
19[b]	4.45d $J_{1',2'}$ 9.0	4.20–3.30m	4.90d 4.85d 4.42d	6.52s[c] $J_{1',4} \sim 1$	2.50s	2.30s			
20[b]	4.33m[d] $J_{1',2'} \sim 1$	4.00–3.10m	4.80d 4.75d 4.60d	6.75s[c] $J_{1',4} \sim 1$	2.48s	2.27s			
23[b]	4.21d $J_{1',2'}$ 10.0	4.10–3.30m	5.00d 4.88d 4.55d				7.85s	7.60–6.70m	12.35s 10.76s

[a]Recorded at 35.5°, δ scale (internal Me_4Si), *J* in Hz. [b]In $(CD_3)_2SO$. [c]Slightly broadened by $J_{1',4}$. [d]Narrow multiplet.

EXPERIMENTAL

General methods. — Solutions were concentrated *in vacuo* at temperatures <40°. Melting points were determined with a Gallenkamp apparatus, and are uncorrected. Optical rotations were measured at 20 ±2° with a Perkin–Elmer 141 polarimeter (10-cm cell). I.r. spectra (KBr discs) were recorded with Beckman IR-33 and Perkin–Elmer 399 spectrophotometers, and u.v. spectra with a Beckman 25 instrument, P.m.r. spectra (90 MHz, internal Me_4Si) were recorded with a Perkin–Elmer R-32 spectrometer, and coupling constants were measured directly from spectra recorded at 300-Hz sweep-width. Assignments were confirmed by double-resonance experiments. T.l.c. was performed on silica gel GF_{254} (Merck) with ethyl acetate–ethanol (3:1) and detection with u.v. light and iodine vapour. Consumption of periodate was determined as previously described[15].

6,6-Dimethyl-2-(2,3,5-tri-O-acetyl-α-D-arabinofuranosyl)-4,5,6,7-tetrahydroindol-4-one (**2**). — Conventional treatment of 2-α-D-arabinofuranosyl-6,6-dimethyl-4,5,6,7-tetrahydroindol-4-one[7] (**1**) with pyridine (1 mL) and acetic anhydride (0.6 mL) at 0° for 36 h gave **2** (0.37 g, 52%), m.p. 116–118°, $[\alpha]_{578}$ −33°, $[\alpha]_{546}$ −38°, $[\alpha]_{436}$ −60°, $[\alpha]_{365}$ −87° (*c* 0.5, chloroform); λ_{max}^{EtOH} 212, 243, and 280 nm (ε 16,700, 7,500, and 6,900); ν_{max} 3320 (NH), 1750 (C=O ester), 1640 (C=O ketone), 1580 and 1480 (C=C pyrrole) cm^{-1}. P.m.r. data are given in Table I.

Anal. Calc. for $C_{21}H_{27}NO_8 \cdot H_2O$: C, 57.39; H, 6.65; N, 3.19. Found: C, 57.36; H, 6.48; N, 3.55.

2-α-D-Arabinopyranosyl-6,6-dimethyl-4,5,6,7-tetrahydroindol-4-one (**3**). — (*a*) A solution of 6,6-dimethyl-2-(D-*gluco*-pentitol-1-yl)-4,5,6,7-tetrahydroindol-4-one[16] (**8**; 1.1 g, 3.51 mmol) in water (44 mL) was treated with trifluoroacetic acid (0.25 mL). After 1 h at 50°, t.l.c. revealed the absence of **8**, and one product (R_F 0.66) corresponding to **1**. Heating was continued for 50 h, and t.l.c. then showed that **1** had almost disappeared; a product, R_F 0.42, was present. The mixture was neutralised with Amberlite IR-45 (HO^-) resin and concentrated, and acetone was evaporated from the syrupy residue. The resulting, white solid was **3** (R_F 0.42) contaminated with **1**. Crystallisation from acetone gave **3** (0.34 g, 33%), m.p. 205–207° (from butanone), $[\alpha]_D$ −39°, $[\alpha]_{578}$ −40°, $[\alpha]_{546}$ −46°, $[\alpha]_{436}$ −75.5° (*c* 0.5, water); λ_{max}^{EtOH} 212, 244, and 281 nm (ε 16,400, 7,400, and 6,300); ν_{max} 3430–3300 (NH, OH), 1610 (C=O), 1575 and 1475 (C=C pyrrole) cm^{-1}. P.m.r. data are given in Table I.

Anal. Calc. for $C_{15}H_{21}NO_5$: C, 61.01; H, 7.11; N, 4.74. Found: C, 60.80; H, 7.30; N, 4.69. Periodate consumption: 2.03 mol.

(*b*) Compound **3** was also prepared from a solution of **1** (0.1 g) in water (25 mL), under the following conditions: the solution was treated with 6 drops of trifluoroacetic acid, kept at 50° for 7 h, and then neutralised with Amberlite IR-45 (HO^-) resin; the solution was treated with 2 drops of trifluoroacetic acid and warmed at 50° for 36 h; the neutral solution was heated at 70° for 10 days.

Each reaction was monitored by t.l.c. which revealed the conversion of **1** (R_F 0.66) into **3** (R_F 0.42), which could be isolated as described in (*a*).

6,6-Dimethyl-2-(2,3,4-tri-O-acetyl-α-D-arabinopyranosyl)-4,5,6,7-tetrahydroindol-4-one (**4**). — Conventional treatment of **3** (0.15 g, 0.51 mmol) with acetic anhydride (0.18 mL) and pyridine (0.20 mL) gave **4** (0.10 g, 46%), m.p. 172–174°, $[\alpha]_D$ −32°, $[\alpha]_{578}$ −33°, $[\alpha]_{546}$ −37.5°, $[\alpha]_{436}$ −60°, $[\alpha]_{365}$ −89° (*c* 0.5, chloroform); λ_{max}^{EtOH} 213, 242, and 278 nm (ε 19,700, 7,700, and 7,600); ν_{max} 3240 (NH), 1735 (C=O ester), 1640 (C=O ketone), 1575 and 1480 (C=C pyrrole) cm^{-1}. P.m.r. data are given in Table I.

Anal. Calc. for $C_{21}H_{27}NO_8$: C, 59.85; H, 6.46; N, 3.32. Found: C, 59.54; H, 6.35; N, 3.41.

3-Acetyl-5-α-D-lyxopyranosyl-2-methyl-1-propylpyrrole (**19**) *and 3-acetyl-5-β-D-lyxopyranosyl-2-methyl-1-propylpyrrole* (**20**). — A suspension of 3-acetyl-2-methyl-5-(D-*galacto*-pentitol-1-yl)-1-propylpyrrole[10] (0.40 g, 1.27 mmol) in water (2 mL) was treated with 3 drops of trifluoroacetic acid and then kept at room temperature with occasional shaking until the solution was clear. After 15 min, a white product crystallised (0.20 g, 51%) which was shown by t.l.c. to contain two compounds (R_F 0.65 and 0.59). Several recrystallisations from ethanol gave **20** (0.12 g, 32%), R_F 0.59, m.p. 214–215°, $[\alpha]_D$ −47°, $[\alpha]_{578}$ −49°, $[\alpha]_{546}$ −56°, $[\alpha]_{436}$ −86°, $[\alpha]_{365}$ −124° (*c* 0.4, pyridine); λ_{max}^{EtOH} 216, 251, and 282 nm (ε 13,400, 5,500, and 3,800); ν_{max} 3395 (OH), 1633 (C=O), 1550 and 1500 (C=C pyrrole) cm^{-1}. P.m.r. data are given in Table II.

Anal. Calc. for $C_{15}H_{23}NO_5$: C, 60.59; H, 7.80; N, 4.71. Found: C, 60.58; H, 7.85; N, 4.84. Periodate consumption: 2.08 mol.

Concentration of the mother liquors gave **19** as a minor product (0.07 g, 18%), R_F 0.65, m.p. 170–171° (from ethanol), $[\alpha]_D$ +19°, $[\alpha]_{578}$ +20°, $[\alpha]_{546}$ +24°, $[\alpha]_{436}$ +53°, $[\alpha]_{365}$ +121° (*c* 0.4, pyridine); λ_{max}^{EtOH} 225, 254, and 285 nm (ε 6,400, 3,300, and 2,300); ν_{max} 3390 (OH), 1625 (C=O), 1560 and 1510 (C=C pyrrole) cm^{-1}. P.m.r. data are given in Table II.

Anal. Found: C, 60.23; H, 8.11; N, 5.02. Periodate consumption: 2.03 mol.

3,7-Anhydro-D-talo-heptulose phenylosazone (**23**). — A solution of D-*galacto*-heptulose phenylosazone[17] (2.50 g, 6.44 mmol) in methanolic 0.05% sulfuric acid (300 mL) was boiled under reflux for 6 h and then poured into hot water (300 mL), and the methanol was evaporated under diminished pressure. The resulting precipitate was collected, washed with water, and dried (1.35 g). Recrystallisation from methanol, with decolorisation with charcoal, gave **23** (0.29 g) as yellow needles, m.p. 237–239°, $[\alpha]_D$ +58° (initial; *c* 0.5, pyridine). Sallam[13] gave m.p. 234–235°, $[\alpha]_D$ +52° (initial; *c* 0.93, pyridine). P.m.r. data are given in Table II. Periodate consumption: 2.04 mol.

ACKNOWLEDGMENTS

We thank Mr. A. Ali for revision of the manuscript, and one of us (F.P.C.) thanks the Ministry of Education and Universities of Spain for the award of a scholarship.

REFERENCES

1 S. HANESSIAN AND A. G. PERNET, *Adv. Carbohydr. Chem. Biochem.*, 33 (1976) 111–188.
2 J. M. J. TRONCHET, *Colloq. INSERM*, 81 (1978) 117–144.
3 U. LERCH, M. G. BURDON, AND J. G. MOFFATT, *J. Org. Chem.*, 36 (1971) 1507–1513.
4 A. GÓMEZ SÁNCHEZ AND M. A. RODRÍGUEZ ROLDÁN, *Carbohydr. Res.*, 22 (1972) 50–62.
5 H. S. EL KHADEM, *Carbohydr. Res.*, 23 (1972) 311–315.
6 J. FERNÁNDEZ-BOLAÑOS, J. FUENTES MOTA, M. A. PRADERA DE FUENTES, AND F. RENDÓN SAINZ, *An. Quim.*, 73 (1977) 1184–1187.
7 F. GARCÍA GONZÁLEZ, M. GÓMEZ GUILLÉN, J. A. GÁLBIS PÉREZ, AND E. ROMÁN GALÁN, *Carbohydr. Res.*, 80 (1980) 37–43.
8 L. MESTER, H. S. EL KHADEM, AND G. VASS, *Tetrahedron Lett.*, (1969) 4135–4138.
9 E. L. ELIEL, *Stereochemistry of Carbon Compounds*, McGraw-Hill, New York, 1962, p. 198.
10 J. A. GÁLBIS PÉREZ AND J. L. JIMÉNEZ REQUEJO, unpublished results.
11 L. D. HALL, L. HOUGH, K. A. MCLAUCHLAN, AND K. PACHLER, *Chem. Ind. (London)*, (1962) 1465–1466.
12 M. GÓMEZ GUILLÉN, J. A. GÁLBIS PÉREZ, E. ROMÁN GALÁN, AND J. ESPINOSA GARCÍA, *An. Quim.*, 75 (1979) 745–748.
13 M. A. E. SALLAM, *Carbohydr. Res.*, 85 (1980) 93–105.
14 B. G. HUDSON AND R. BARKER, *J. Org. Chem.*, 32 (1967) 3650–3658.
15 F. GARCÍA GONZÁLEZ, J. FERNÁNDEZ-BOLAÑOS, AND M. A. PRADERA DE FUENTES, *An. Quim.*, 70 (1974) 57–59.
16 F. GARCÍA GONZÁLEZ, M. GÓMEZ GUILLÉN, J. A. GÁLBIS PÉREZ, AND E. ROMÁN GALÁN, *Carbohydr. Res.*, 78 (1980) 17–23.
17 H. S. EL KHADEM, M. M. ABDEL RAHMAN, AND M. A. E. SALLAM, *J. Chem. Soc., C*, (1968) 2411–2414.