Journal of Natural Products
Vol. 51, No. 1, pp. 103-109, Jan-Feb 1988

STRUCTURES OF THE PREGNENEDIOL[1] TRI- AND DI-GLUCOSIDES FROM EGGS OF THE TOBACCO HORNWORM, *MANDUCA SEXTA*[2]

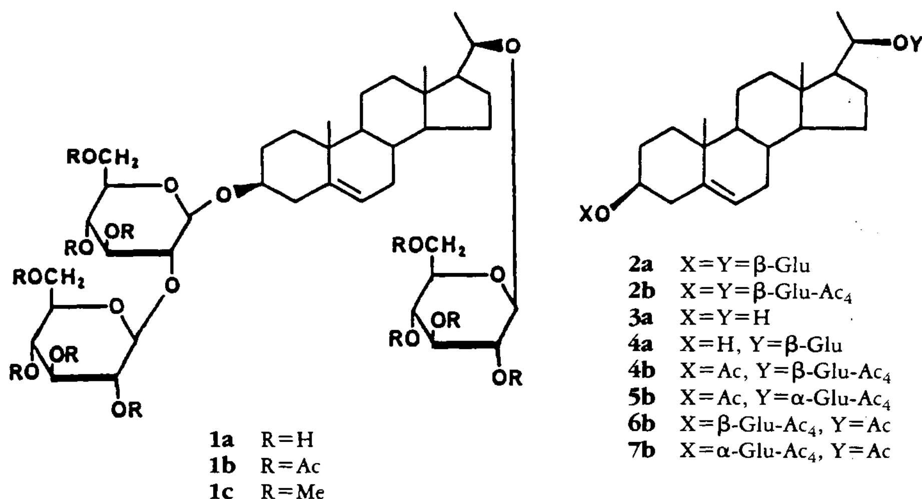
JAMES E. OLIVER,* WILLIAM R. LUSBY, ROLLAND M. WATERS,[3] and MALCOLM J. THOMPSON

USDA, ARS, Insect and Nematode Hormone Laboratory, Building 467, BARC-East, Beltsville, Maryland 20705

ABSTRACT.—A metabolite of ^{14}C-cholesterol isolated from eggs and ovaries of the tobacco hornworm, *Manduca sexta*, has been identified as 3β,20(*R*)-pregn-5-ene-3-[*O*-β-D-glucopyranosyl (1→2)-β-D-glucopyranoside]-20-[*O*-β-D-glucopyranoside] [**1a**]. Smaller amounts of the related 3β,20(*R*)-pregn-5-ene-3-[*O*-β-D-glucopyranoside]-20-[*O*-β-D-glucopyranoside] [**2a**] were also isolated.

We recently reported (1) that a radiolabeled triglucoside of pregn-5-ene-3β,20(*R*)-diol had been isolated from eggs and ovaries of tobacco hornworms [*Manduca sexta* (L.) Lepidoptera: Sphingidae] that had been treated as pupae with ^{14}C-cholesterol. We proposed on the basis of the ^{1}H-nmr spectrum that at least one of the glucose units was associated with the 20-OH of **3a**, but the locations and configurations of the sugars were otherwise undefined. We have since found that **1a** is accompanied by a smaller amount of a related diglucoside **2a** and here report the complete structures of these unusual cholesterol metabolites as deduced through ^{1}H-nmr spectroscopic considerations along with partial enzymatic and chemical degradations to independently synthesized (2) standards.

Our initial work determined that **1a** could be isolated from ovaries and eggs of the tobacco hornworm, but that it was not found in newly emerged larvae (1). It was subsequently found that **1a** remained with the egg shells after hatch (R. Lozano, unpublished observations) and, finally, that **1a** was in fact on the surface of unhatched eggs and could be removed by simply rinsing the eggs with water or MeOH. (These treatments, incidentally, did not affect either egg hatch or subsequent development of lar-



[1]All pregnenediol derivatives discussed in this paper have the Δ^5-3β,20(*R*)-configuration. To avoid possible confusion, this stereochemistry is to be understood, and all subsequent use of the α- and β-descriptors will be reserved for discussion of stereochemistry at the anomeric carbons of glucose units.

[2]Mention of a proprietary product does not necessarily constitute an endorsement by the USDA.

[3]USDA, ARS, Insect Chemical Ecology Laboratory, Building 007, BARC-West, Beltsville, Maryland 20705.

vae.) This latter finding greatly facilitated isolation of **1a** inasmuch as aqueous rinses provided considerably cleaner samples than were obtained by homogenization of eggs or ovaries. Thus, by thoroughly rinsing 250 g of eggs with H_2O, then partitioning and chromatographing, we were able to obtain 170 mg of a fairly pure **1a**, which could be crystallized from iPrOH plus MeOH but which was too hygroscopic to handle in a humid environment. These egg washes also contained much smaller amounts of a diglucoside eventually determined to be the 3β,20β-isomer **2a**; no isomeric diglucosides nor any monoglucosides could be detected by hplc or hptlc.

In the initial publication (1) we reported that **1a** was fairly resistant to several commercially available β-glucosidases, but that it could be cleaved to pregnenediol by certain β-glucuronidases. In the current work, we found that a commercially available sulfatase (also possessing β-glucuronidase activity) from *Helix pomatia* cleaves **1a** at a convenient rate and in a stepwise manner as seen in Figure 1. Hplc-monitoring showed that disappearance of **1a** was accompanied by the appearance of one, and, subsequently, a second, new component. As the reaction progressed, the first new component gave way to the second, and, ultimately (overnight), the second was converted to pregnenediol [**3a**] (not seen under the hplc conditions illustrated). By comparison to standards (2), it

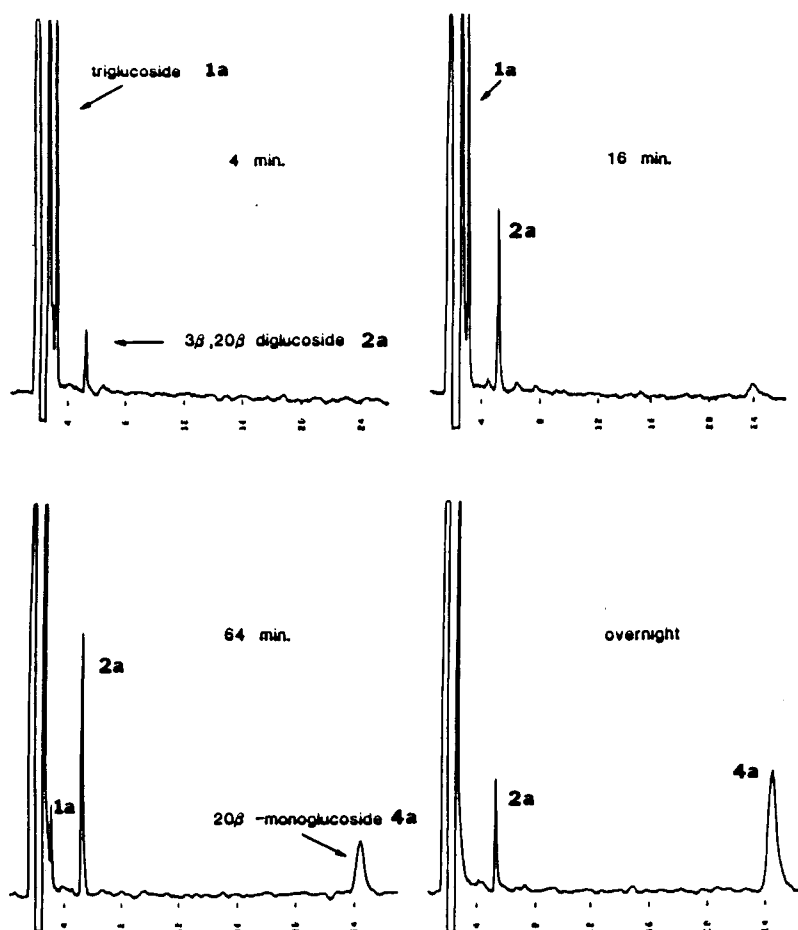


FIGURE 1. Stepwise hydrolysis of triglucoside **1a** with *Helix pomatia* enzyme. A 25 cm ODS column was eluted with 30% MeCN at 1.5ml/min and monitored at 214 nm.

was ascertained that the first product was the 3β,20β-diglucoside **2a** and that the second was the 20β-monoglucoside **4a**. Thus, we knew that each of the pregnenediol oxygens was glucosylated and that each of these anomeric linkages was β, i.e., **1a** was a derivative of **2a** with the position of one final glucose yet to be determined.

Mass spectra of **1a** were recorded under various conditions: the ci spectrum with NH_3 as reagent gas contained peaks at *m/z* $[M+NH_4]^+$ 822 (33%) and $[M+NH_4-H_2O]^+$ 804 (24%), a strong peak at *m/z* 660 (100%) with *m/z* 642 (20%), and another strong peak at *m/z* 498 (also 100%) with *m/z* 480 (28%). Relatively weak (<20%) peaks were apparent at *m/z* 318, 300, and 284, along with *m/z* 180 (30%). The CH_4ci spectrum of **1a** was dominated by three major ions (*m/z* 463, 301, 283), the relative intensities resembling those of the 20α- and 20β-monoglucosides (2). Additional ions in the spectrum of **1a** were observed at *m/z* 325 (15%), 317 (6%), 625 (5%), and 643 (8%). The $[M+H]^+$ peak was detectable (ca. 1%) at *m/z* 805.

We pointed out elsewhere (2) that the CH_4ci spectra of pregnenediol-20-monoglucosides differ significantly from those of the 3-monoglucosides in that prominent high mass ions of the 20-glucoside (*m/z* $[M+H-H_2O]^+$ 463) are barely detectable in the 3-glucoside where *m/z* $[M+H-\text{glucose}]^+$ 301 represented the highest mass ions of significance. These differences can be explained by the preferential elimination of the C-3 substituent (H_2O from 20-monoglucosides; glucose from 3-monoglucosides) due to the participation of the Δ^5-double bond in the ion stability of the resulting fragment ion. The CH_4ci spectrum of **1a** resembles those of the 20-monoglucosides in that the highest mass ion of significance is observed at *m/z* 463; this would be most easily explained by a single glucose (glu) at position 20 and ionization of a glu-glu unit from position 3; had the glu-glu unit been located at position 20, with a single glu at position 3, $[M+H-\text{glucose}]^+$ (*m/z* 625) might have been expected to be a more predominant ion.

Triglucoside **1a** gave an ei spectrum containing major ions at *m/z* 123 (100%), 283 (65%), 300 (28%), and 301 (20%) but little that seemed of diagnostic importance.

Methylation hydrolysis, reduction, and acetylation by standard techniques (3,4) gave partially methylated alditol acetates **8** and **9** via **1c** (Figure 2) that were easily

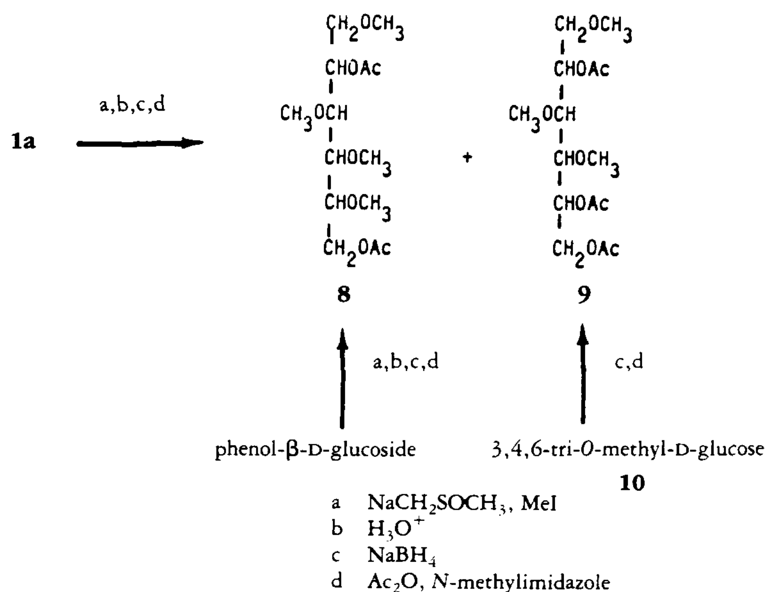


FIGURE 2. Methylated glucitol acetates from triglucoside **1a**.

separated by gc. The earlier-eluting **8** matched (by gc retention time and ms) the diacetate of 2,3,4,6-tetra-*O*-methylglucitol (prepared independently from phenol-β-D-glucoside). The ms of the remaining glucitol was consistent only with that of 3,4,6-tri-*O*-methylglucitol triacetate (3); this assignment was confirmed by independently preparing the latter compound from 3,4,6-tri-*O*-methylglucose **10** (5). Thus, one of the oxygens of **1a** is substituted with a 1↦2 diglucose unit and the other with a single glucose.

^{1}H-nmr spectra were recorded on peracetylated derivatives. Peracetylation (6–8) simplified the spectra by avoiding the *CHOH* coupling on the glucose moieties, removing O-*H* signals, and improving solubility in $CDCl_3$ and C_6D_6. Spectra were recorded in both these solvents; $CDCl_3$ was generally more useful for literature comparisons, but, in several cases, superior resolution was realized with C_6D_6. Systematic decouplings of signals of the glucosyl hydrogens were conducted to ensure unambiguous assignments.

Having determined that triglucoside **1a** was a derivative of the 3,20-diglucoside **2a**, and that the third glucose was attached to a 2-position of a glucose of **2a**, the final questions to be addressed were (a) what was the anomeric configuration of the 3rd glucose, and (b) to which of the glucoses of **2a** was it attached? An additional objective was to confirm the β,β-anomeric configuration of **2a**, since that assignment had been initially made on the basis of hplc retention times (2). Spectra of the synthetic monoglucosides **4b**–**7b** were in full accord with earlier reports (9–11) that H-1s of α-glucopyranosides are deshielded relative to H-1s of corresponding β-anomers (e.g., compare **5b** with **4b**, **7b** with **6b**, Table 1). This relationship held in both solvents. Smaller coupling constants (3.5–3.8 Hz vs. 7.5–8.1 Hz) for H-1s of α-glucosides than for β-glucosides were also consistent with expectations. Thus, both chemical shifts and coupling constants confirm the 3β,20β-configuration for diglucoside **2b**.

The glucose H-1 signals in C_6H_6 were also useful in location of the third glucoside moiety in **1b**. Comparison of the three H-1 absorptions of **1b** (δ = 4.35, 4.25, 4.70) to those of **4b** (4.33), **6b** (4.46), and **2b** (4.34, 4.45) shows that the H-1 signal associated with the glucoside at the 20-position of pregnenediol (signal D-1′, Table 1) remains essentially uninfluenced by subsequent substitutions, whereas the H-1 signal of the glucose on the A ring (signal A-1′, Table 1) is significantly shifted when the third glucose is added (i.e., as **2b**↦**1b**). This strongly indicates that the glucose on the A ring is the one further substituted. Coupling constants of these three signals were all in the 7.5–9.0 Hz range, consistent only with β-glucosides; thus, the configuration of the third glucose is also β.

The H-3 chemical shifts of the peracetylated glucose units supported anomeric assignments. The olefinic H (position 6) of the pregnenediol nucleus provided a convenient reference marker for these signals; β-glucosides (**4b**, **6b**, **2b**) produced H-3 signals in $CDCl_3$ upfield (δ = 5.21) from the 6-H signal, whereas the α-glucosides (**5b**, **7b**) displayed H-3 signals deshielded (δ = 5.50 and 5.49, respectively) relative to the marker. Peracetylated triglucoside **1b** produced no signals at lower field than the 6-H absorption, further indicating that all of the glucosidic linkages were β.

Irradiation of the 3.6–3.7 ppm region of the spectrum of **1b** (C_6D_6) resulted in collapse of the doublet at δ = 4.25 (H on anomeric C). Consequently, one of the H-2s absorbed at 3.6–3.7 ppm, too far upfield for *H*-C-OAc but consistent with *H*-C-O-C-O, i.e., the position of attachment of the third glucose unit. This is in full accord with the conclusions drawn from the methylation/degradation experiments discussed earlier.

Thus, the structure of triglucoside **1a** is 3β,20(*R*)-pregn-5-ene-3-[*O*-β-D-glucopyranosyl (1↦2)-β-D-glucopyranoside]-20-[*O*-β-D-glucopyranoside], the (1↦2)-β

TABLE 1. Selected ^{1}H-nmr Data for Pregnenediol Glucosides[a]

Glucoside (anomeric configuration)	H-6		D[b]-1′(J)		A[c]-1′(J)		G[d]-1′(J)		D[b]-3′		A[c]-3′	
	C_6D_6	$CHCl_3$	C_6D_6	$CHCl_3$	C_6D_6	$CHCl_3$	C_6D_6	$CHCl_3$	C_6D_6	$CHCl_3$	C_6D_6	$CHCl_3$
4b (β)	5.36	5.36	4.33 (8.1)	4.55 (7.8)					5.43	5.21		
5b (α)	5.35	5.38	4.97 (3.64)	5.10 (3.6)					5.86	5.50		
6b (β)	—[e]	5.36			4.46 (7.8)	4.59 (8.0)					5.26–5.50[e]	5.21
7b (α)	5.40	5.35			5.34 (3.48)	5.23 (3.8)					5.99	5.49
2b (β, β)	5.50	5.35	4.34 (8.1)	4.55 (8.1)	4.45 (7.8)	4.59 (7.8)			5.15–5.3[e]	5.21	5.1–5.5[e]	5.1–5.3[e]
1b (β, β, β)	5.52	5.35	4.35 (9.0)	4.55 (7.5)	4.25 (7.5)	4.55 (7.5)	4.70 (7.5)	4.81 (7.8)	5.15–5.3[e]	5.21	5.1–5.5[e]	5.1–5.3[e]

[a]A more comprehensive compilation of ^{1}H-nmr data for these and related compounds will be published elsewhere (J.D. Warthen, Jr., R.M. Waters, and J.E. Oliver, *Org. Magn. Reson.*, submitted).
[b]D designates the glucose attached to the D ring of pregnendiol.
[c]A designates the glucose attached to the A ring of pregnenediol.
[d]G designates the glucose attached via other glucose molecule to the pregnenediol.
[e]Incomplete resolution.

linkage making it a derivative of the diglucoside sophorose. The function of this unusual cholesterol metabolite is still under study.

EXPERIMENTAL

GENERAL.—Gc was performed on a Shimadzu model GC-9A instrument equipped with a flame ionization detector, a split-splitless injector operated as a ca. 50:1 split ratio, and a 15-m DB-1 fused silica column with He as a carrier. Hplc was carried out on a Spectra Physics model 8700 system with a variable wavelength uv detector operated at 210 nm, fitted with a 250 × 4.6-mm Whatman Partisil 5 ODS-3 column or with a 250 × 4.6-mm Dupont Zorbax-CN column. In both cases, H_2O containing 0.1% HOAc and MeCN were the primary and secondary solvents, respectively.

Mass spectra were obtained from a Finnigan model 4510 gas chromatograph-mass spectrometer. Samples were introduced either via a direct exposure probe or through a 30-m × 0.32-mm i.d. DB-1 fused silica column. Ei spectra were collected at 70 eV and a source block temperature of 150°. Data were analyzed via an Incos data system. The ^1H-nmr spectra were obtained from 5 mg/0.3 ml $CDCl_3$ or C_6D_6 + TMS (internal standard) samples using a General Electric QE-300 nmr spectrometer. Data acquisition parameters were: 30 degree flip angle, 16K data points, 1.34–3.35 sec acquisition time, pulse repetition = acquisition time + 1 sec, and 12 bit digital resolution with spectral width and number of transients adjusted to the individual sample. Decoupling experiments used single-frequency, low-power, on-resonance conditions to remove the target absorption. ^1H chemical shift assignments were made by decoupling experiments. Pregnenediol and the *Helix pomatia* sulfatase were purchased from Sigma.

ISOLATION OF **1a** AND **2a**.—Eggs of the tobacco hornworm, *M. sexta*, that had been laid on tobacco leaves over a 3-day period, removed, and stored at −20° (250 g) were suspended in cheesecloth and agitated several minutes in 400 ml H_2O. The H_2O was decanted, and the process was repeated with 4 × 250-ml portions of H_2O. The combined rinses were concentrated in vacuo and the residue partitioned between *n*-BuOH and H_2O in 15-ml centrifuge tubes (5 portions of *n*-BuOH vs. 3 portions of H_2O). Concentration of the *n*-BuOH phase gave 0.38 g of residue that was evaporated onto Celite®, which in turn was added to the top of a 15-g Florisil® column. The column was eluted with $CHCl_3$ followed by increasing portions of 95% EtOH in $CHCl_3$; triglucoside **1a** (170 mg) was eluted with 45–70% EtOH. This material was crystallized from iPrOH + MeOH and collected on a sintered glass funnel; in humid weather it immediately became syrupy but upon drying overnight in vacuo, ca. 100 mg of solid **1a** was obtained that was virtually pure by tlc and hplc. Triglucoside **1a** does not have a well-defined mp; darkening and shrinking occurred above ca. 175° at a rate that depended upon initial bath temperature and rate of heating. A typical determination in a Kofler block (uncorrected) showed softening at 180–185°, resolidification, melting at 205–209° with resolidification again at 220–225°, and final melting with decomposition at 265–279°. The chromatographic fractions eluting immediately prior to those containing **1a** could be combined and rechromatographed (similar conditions except smaller increments of 95% EtOH in $CHCl_3$) to provide ca. 5–6 mg of the 3β,20β-diglucoside **2a**.

ENZYMATIC HYDROLYSIS OF **1a**.—*Product evaluation*.—To 4.4 mg **1a** in 0.42 ml acetate buffer (16.4 g NaOAc and 15 ml 0.4 M HOAc added to 500 ml H_2O, pH ca. 6.2) was added 20 μl of a commercial enzyme preparation, sulfatase type H-2 from *Helix pomatia*. At intervals of 1, 2, 4, 8, 16, 32, and 64 min, 50-μl aliquots were withdrawn and added to 100 μl cold MeOH; these were subsequently diluted with H_2O and passed through a C_{18} extraction cartridge. After rinsing with H_2O, products were eluted with MeOH, concentrated, and made up to 200 μl with MeOH for hplc analysis. The remainder of the sample was similarly worked up after standing overnight at room temperature. Results are displayed in Figure 1.

Preparative run: Preparation of **4b** *and* **2b**.—To a solution of ca. 77 mg **1a** in 7.3 ml acetate buffer was added 340 μl of the enzyme solution. After 1 h at room temperature, the solution was passed twice through a C18 extraction cartridge which was then rinsed with a little H_2O and eluted with MeOH to give, upon evaporation, 31.5 mg of an oil that was chromatographed on 10 g of Florisil®. Elution with $CHCl_3$ containing increasing portions of 95% EtOH provided 5.2 mg of **4a** (16–24% EtOH) followed by 6.8 mg **2a** (32–45% EtOH). Acetylation of these products with Ac_2O/pyridine gave **4b** and **2b**, respectively. Sample **4b** was collected by preparative hplc (C18, 90% MeCN, 1.5 ml/min) for nmr studies; **2b** was obtained pure by elution from a silica extraction cartridge with Et_2O (after first rinsing with CH_2Cl_2).

METHYLATION OF **1a**.—Sodium hydride (0.9 g of a 50% mineral oil suspension) was transferred to a sintered glass funnel, washed with hexane, then stirred in dry DMSO (9 ml) under N_2 at 65–70° until no further change or gas evolution could be detected. A solution of 15 mg **1a** in 0.7 ml DMSO was treated with 1 ml of this dimsyl sodium solution at room temperature followed after 1.5 h with 1 ml MeI (exothermic reaction). After 2 days, the mixture was added to acetate buffer and passed through a C_{18} cartridge.

After rinsing with H_2O, crude **1c** was eluted with MeOH, concentrated, taken up in EtOAc, and washed with H_2O and then with brine. The solution was dried and concentrated, and the residue was chromatographed on 2 g Unisil* silicic acid. Elution with hexane, 25, 50, 75, and 100% $CHCl_3$ followed by 10, 20, 30, and 40% 95% EtOH in $CHCl_3$ gave 10.7 mg **1c** in the 10% and 20% EtOH fractions. That methylation was complete was demonstrated by a strong ion at *m/z* 976 (ammonium adduct) in the NH_3ci mass spectrum.

Hydrolysis of **1c** was achieved by boiling the above sample 0.5 h in 1.5 ml HCO_2H containing 100 μl H_2O; then 2 ml 0.25 M H_2SO_4 was added, heating was continued another 0.75 h, and another 2 ml portion of 0.25 M H_2SO_4 was added. After an additional 2.5 h, the solution was divided, and one portion was cooled and neutralized with concentrated NH_4OH, then was treated with 10 ml of a solution of $NaBH_4$ in DMSO (1 g/50 ml) (40°, 0.75 h). With cooling, 1.5 ml HOAc was slowly added, followed by *N*-methylimidazole (2 ml), and, finally, Ac_2O (20 ml). After 0.25 h at 40°, the mixture was cooled, carefully diluted with H_2O, then thoroughly extracted with Et_2O-hexane (1:1). The extract was washed well with H_2O, then with 1 N $NaHCO_3$, and then dried to give 2.6 mg of crude residue that was added in C_6H_6/hexane to a silica cartridge. After rinsing with hexane then C_6H_6, EtOAc eluted the methylated glucitol acetates **8** and **9**, shown by gc-ms to consist of 2,3,4,6-tetra-*O*-methylglucitol-1,6-diacetate **8** (identical by gc-ms to a sample prepared from phenol-β-D-glucoside by the procedure just outlined) and 3,4,6-tri-*O*-methylglucitol-1,2,5-triacetate **9** which was initially identified by comparison of its mass spectrum to published values (3) and subsequently confirmed by comparison to a sample independently prepared from 3,4,6-tri-*O*-methylglucose **10** (5) via the last two steps (reduction and acetylation) of the sequence just outlined.

ACKNOWLEDGMENTS

We appreciate the assistance of J. Quigley, K. Wilzer, and D. Harrison, and thank Drs. B. Li and H. Sinclair for helpful discussions.

LITERATURE CITED

1. M.J. Thompson, J.A. Svoboda, W.R. Lusby, H.H. Rees, J.E. Oliver, G.F. Weirich, and K.R. Wilzer, *J. Biol. Chem.*, **260**, 15410 (1985).
2. J.E. Oliver and W.R. Lusby, submitted for publication in *Steroids*.
3. H. Bjorndal, C. Hellerqvist, B. Lindberg, and S. Svensson, *Agnew. Chem. Int. Ed. Eng.*, **9**, 610 (1970).
4. B. Lindberg, in: "Methods in Enzymology." Ed. by V. Ginsburg, Academic, New York, 1972, Vol. XXVIII, pp. 178–195.
5. A.K. Mitra, D.H. Ball, and L. Long, Jr., *J. Org. Chem.*, **27**, 160 (1962).
6. M. Rychener, P. Bigler, and H. Pfander, *Helv. Chim. Acta*, **67**, 378 (1984).
7. C. Nishino, S. Manabe, N. Enoki, T. Nagata, T. Tsushida, and E. Hamaya, *J. Chem. Soc., Chem. Commun.*, 720 (1986).
8. Y. Fujimoto, N. Ikekawa, M. Ogiso, and E. Ohnishi, *Experientia*, **42**, 567 (1986).
9. K. Bock and H. Pedersen, *J. Carbohydr. Chem.*, **3**, 581 (1984).
10. R. Faghih, C. Fontaine, I. Horibe, P.M. Imamure, G. Lukars, A. Olesker, and S. Seo, *J. Org. Chem.*, **50**, 4918 (1985).
11. A. De Bruyn, M. Anteunis, and G. Verhegge, *Bull. Soc. Chim. Belg.*, **84**, 721 (1975).

Received 29 July 1987