Carbohydrate Research, 206 (1990) 13–19
Elsevier Science Publishers B.V., Amsterdam

Elucidation of the structure of an unusual cyclic glycolipid from *Torulopsis apicola*

Lutz Weber, Joachim Stach, Günter Haufe,
Department of Chemistry, Karl-Marx-University, Talstr. 35, DDR-7010 Leipzig (G.D.R.)

Rolf Hommel, and H.-P. Kleber
Department of Biosciences–Biochemistry, Karl-Marx-University, Talstr. 33, DDR-7010 Leipzig (G.D.R.)

(Received October 2nd, 1989; accepted for publication, January 11th, 1990)

ABSTRACT

The yeast *Torulopsis apicola* produces a microcrystalline mixture of glycolipids, the major component of which has been identified as the novel 16-(2′-*O*-β-D-glucopyranosyl-β-D-glucopyranosyloxy)hexadecanoic acid 1′,4″-lactone 6′,6″-diacetate (**2**) by spectroscopic methods.

INTRODUCTION

Cultivation of the yeast *Torulopsis apicola* on D-glucose and/or octadecane yields a mixture of extracellular glycolipids[1], which consists of ω-1 hydroxylated fatty acid glycosides of partially acylated sophorose (2-*O*-β-D-glucopyranosyl-D-glucopyranose).

The structures of several, mainly unsaturated glycolipids, derived after cultivation of the yeast on octadecane or oleic acid, have been determined[2–4] by ^{1}H-n.m.r. spectroscopy, chemical degradation, synthesis, and derivatization. Hitherto, no sophorolipids, the lipophilic moiety of which consists of ω-hydroxylated fatty acids, have been found[5].

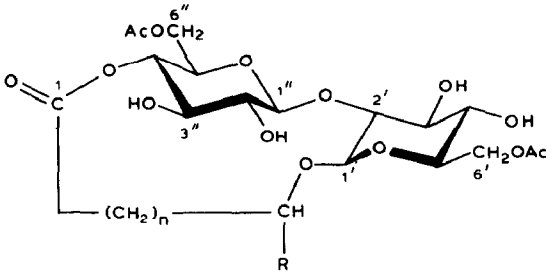
Fermentation of *T. apicola* on hexadecane under modified conditions[6,7] yielded a crude, crystalline ~ 1:1 mixture of the sophorosides of 15-hydroxyhexadecanoic acid (**1**) and 16-hydroxyhexadecanoic acid (**2**). By varying the conditions of fermentation[8], a 3:7 microcrystalline mixture of **1** and **2** was obtained. Attempts to resolve **1** and **2** by t.l.c. or h.p.l.c. failed. However, double-quantum (DQ) and double-quantum-filtered (DQF) RELAY spectroscopy in combination with quantitative ^{13}C-n.m.r. spectroscopy allowed the structure of the novel lipid **2** to be determined without derivatization or chemical degradation experiments, and the results are now reported.

RESULTS AND DISCUSSION

The molecular weight of 662 of both **1** and **2** was verified by f.a.b.-mass spectrometry. On addition of KCl/NaCl to the glycerol matrix, spectra were obtained that contained $[M + K]^+$ and $[M + Na]^+$ ions (*m/z* 701 and 685, respectively). Loss of monoacetylated glucopyranose residues was shown by the peaks at *m/z* 441 for $[221]^+$ and 459 for $[M - 203]^+$, whereas the hexadecanoyl residue appeared at *m/z* 237.

0008-6215/90/$03.50

The ^{13}C- and ^{1}H-n.m.r. spectra of **1** and **2** were similar (Tables I–II) but could be identified from the spectra of mixtures with various compositions. The numbers of C, CH, CH_2, and CH_3 atoms/ groups were determined using the standard DEPT pulse sequence and quantitative ^{13}C-n.m.r. spectroscopy[9]. From these data (Tables I and II), the main features of **1** and **2** were identified by the signals of two pairs of anomeric carbon (> 100 p.p.m.) and hydrogen atoms (< 4.5 p.p.m.) and $J_{H\text{-}1,H\text{-}2}$ values (7.8 Hz) typical of β-linked glucose residues[10]. The two disaccharide units of **1** and **2** are both 6′,6″-diacetylated, and sixteen, further carbon signals corresponding to the hexadecanoyl residue were found for each lipid.



1 n = 13, R = CH_3

2 n = 14, R = H

TABLE I

^{13}C-N.m.r. chemical shift data (p.p.m.) for **1** and **2**

Atom	*$CDCl_3$*		*$(CD_3)_2SO$*	
	1	**2**	**1**	**2**
C-1	173.3	173.3	172.1	172.1
C=O	171.4	171.5	170.1	170.2
	170.5	170.6	169.9	169.8
C-1″	104.0	104.2	104.2	104.9
C-1′	102.5	101.8	102.2	101.5
C-2′	82.2	82.0	82.7	83.5
C-4′	75.7	75.6	75.6	75.6
C-5′,5″	75.4	74.7	75.6	75.4
C-5′,5″	74.0	73.8	73.2	73.4
C-3′,3″	73.2	73.4	72.8	73.0
C-3′,3″	72.6	72.7	71.5	71.5
C-2″	70.4	70.3	70.3	69.9
C-4″	69.7	69.7	70.3	69.2
C-6′	63.6	63.5	63.8	63.8
C-6″	62.1	62.1	62.3	62.3
C-2–13	24.7–30.0		24.4–30.2	
CH_3C=O	20.9	20.9	21.0	21.0
	20.9	20.9	20.8	20.8
C-16	20.9	70.5	21.1	70.4
C-15	79.4	34.3	77.8	33.8
C-14	37.6		37.3	

TABLE II

^{1}H-N.m.r. data (δ in p.p.m., J in Hz) for **2** in $(CD_3)_2SO$

Atom	δ	J	*Atom*	δ	J
H-1′	4.33	7.8	H-1″	4.42	7.8
H-2′	3.21	7.8, 9.2	H-2″	3.12	7.8, 9.5, 5.7
H-3′	3.40	9.2, 9.0, 3.4	H-3″	3.45	9.5, 9.5, 5.5
H-4′	3.15	9.0, 10.0, 4.4	H-4″	4.72	9.5, 10.0
H-5′	3.38	10.0, 6.3, 2.7	H-5″	3.59	10.0, 4.8
H-6a′	4.03	12.3, 6.3	H-6″	3.97	4.8
H-6b′	4.24	12.3, 2.7			
HO-3′	5.64	3.4	HO-2″	5.33	5.7
HO-4′	5.44	3.4	HO-3″	5.29	5.5
H-16a	3.72	8.9, 6.3			
H-16b	3.39	8.9, 6.3			
H-2	2.31	6.3, 6.3			
Ac	2.02 and 2.01				
Chain CH_2	1.65–1.15				

TABLE III

Characteristic ^{1}H-n.m.r. chemical shift data (p.p.m.) for **1** and **2** in $CDCl_3$

Atom	1	2
H-4″	4.95	4.95
H-1″	4.53	4.52
H-1′	4.41	4.39
H-6′	4.32	4.33
H-6″	4.16	4.16
H-2′	3.53	3.50
H-15	3.75	2.43
H-16	1.22	3.93 and 3.40

The spectral parameters of **1** are almost identical with those observed[2–4] for the homologue that contains the 17-hydroxyoctadecanoic acid residue. On the basis of the ^{1}H–^{1}H COSY and ^{13}C–^{1}H HETCOR correlated 2D spectra and considerations analogous to those outlined below, **1** was identified as 15-(2′-*O*-β-D-glucopyranosyl-β-D-glucopyranosyloxy)hexadecanoic acid 1′,4″-lactone 6′,6″-diacetate.

The ω-1 hydroxylated hexadecanoyl residue was identified by the ω-methyl group doublet in the ^{1}H-n.m.r. spectrum and the signals of the CH_3–CHR–O– moiety in the ^{13}C-n.m.r. spectrum. The ^{13}C signals of the methylene chain in the macrocyclic lactones **1** and **2** were well resolved between 23.9 and 29.9 p.p.m., in contrast to the overlapped signals at 29.9 p.p.m. of the acyclic ω-1 15-hydroxyhexadecanoic acid.

Structural analysis of lipid **2** was started with 2D J-resolved ^{1}H-n.m.r. spectroscopy which gave the data in Table II. The large $^3J_{H,H}$ values (>7.5 Hz) and the ^{13}C shift data allowed the disaccharide unit to be assigned as β-glucopyranosyl-β-glucopyranose.

The advantage and limitations of ^1H–^1H-correlated 2D-n.m.r. methods in oligosaccharide analysis have been presented[11].

COSY-90, COSY-45, DQF-COSY, DQF-RELAY, and DQ 2D methods were applied to determine the proton connectivities in **2**. The COSY experiments were of limited use because several cross-peaks near the diagonal associated with vicinal protons having similar shift values were obscured by strong peaks on the diagonal. Also, there was poor resolution (natural line-width of ~ 3 Hz) caused by the large molecular rotational correlation time (τ_c) that resulted from aggregation of molecules which exhibit surfactant properties (the critical micellar concentration of the glycolipid mixture in water was 45 mg/L).

DQ 2D experiments[12] were performed with selection times according to coupling constants of 11 and 4 Hz. Fig. 1 shows 9 cross-peaks attributed to 6a′–6b′ (4.03 and 4.24 p.p.m.), 6a′–5′ (3.38 p.p.m.), 5′–4′ (3.15 p.p.m.), 4′–3′ (3.40 p.p.m.), 2′–1′ (4.33 p.p.m.), 5″–4″ (3.59 and 4.72 p.p.m.), 4″–3″ (3.45 p.p.m.), 2″–1″ (3.12 and 4.42 p.p.m.), and

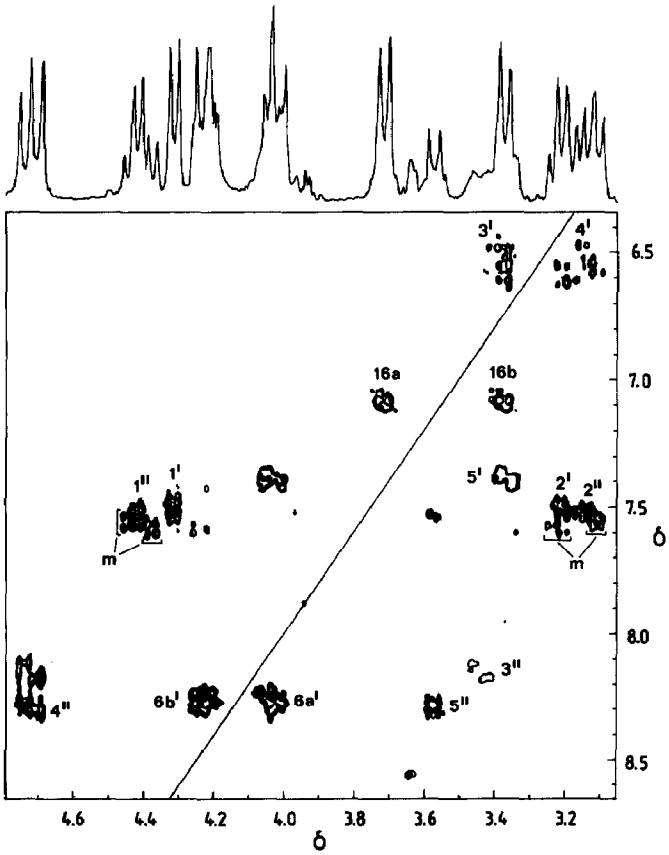


Fig. 1. 2D 300-MHz DQ spectrum of a 3:7 mixture of **1** and **2** in $(CD_3)_2SO$ with a selection time according to 11 Hz. The DQ diagonal $F_1 = 2F_2$ is shown, and contributions from **1** are marked m. The projection above is along the F_2 dimension.

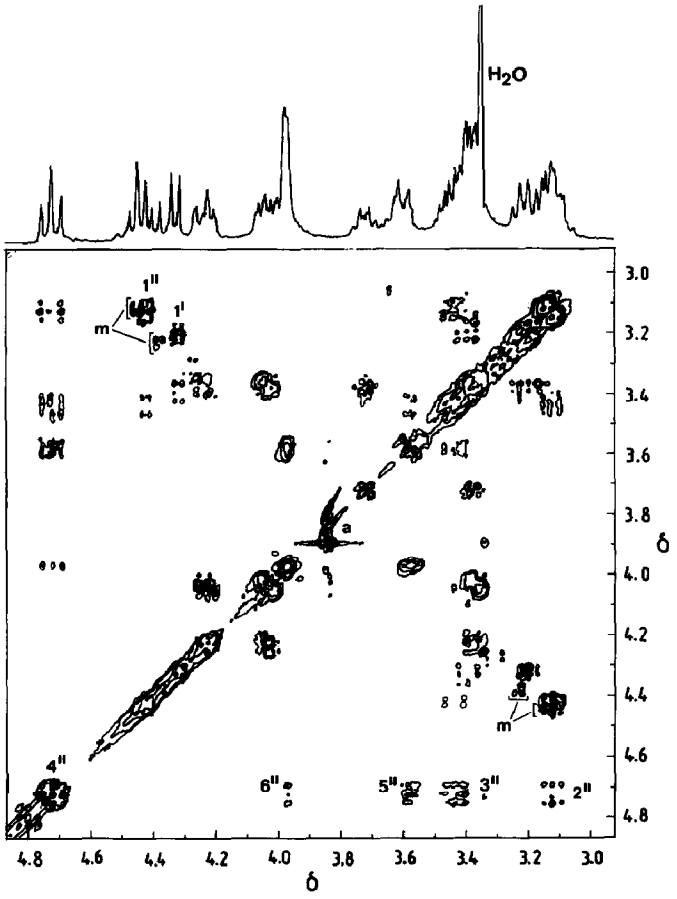


Fig. 2. Unsymmetrized 300-MHz RELAY spectrum of a 3:7 mixture of **1** and **2** in $(CD_3)_2SO$. Contributions from **1** are marked m and artefacts at the carrier frequency with a. The projection above is along the F_2 dimension.

16a–16b (3.72 and 3.39 p.p.m.). In the 4-Hz DQ ^1H–^1H 2D spectrum, the correlations from the anomeric protons were significantly smaller, but 6″–5″ and 6b′–5′ cross-peaks were also observed; the 3″–2″ correlation was not observed.

The dependence of signal intensities on the mixing time and time-consuming acquisition are disadvantages of the DQ method. Most effective was the DQF RELAY method[13] (Fig. 2). Remote connectivities (*i.e.*, relay correlations: A→M→X) were detected as weak cross-peaks at σ_A/σ_X frequency in addition to strong vicinal correlations (A→M). All the proton shifts (and carbon shifts *via* the HETCOR experiment) could be assigned, *e.g.*, for H-4″ which shows vicinal correlations with H-5″ and H-3″ as well as relay correlations with H-6″ and H-2″, respectively.

Determination of inter-residue linkage. — The DQF COSY spectrum of **2** in $(CD_3)_2SO$ revealed the location of the hydroxyl groups *via* the vicinal couplings of hydroxyl protons at 5.29–5.64 p.p.m. to protons at positions 3′,4′,2″, and 3″; therefore, the linkage positions were established at positions 1′,2′,1″, and 4″.

The 1D selective INEPT method, which can be useful for detection of long-range $^{1}H-^{13}C$ couplings and, thus, in the determination of inter-residue linkages[14], showed only a polarization transfer from H-4″ (4.72 p.p.m.) to the hexadecanoyl carbonyl atom C-1 (172.1 p.p.m.). No transfer was observed from the anomeric protons or H-2′ due to fast relaxation ($T_1 < 100$ ms).

The 1D n.O.e. difference spectra (NOEDSY) of **2** showed several diagnostic negative signals. Through-space interactions were observed upon irradiation of H-1′ with H-3′, H-5′, and H-16b of the hexadecanoyl residue and, similarly, upon irradiation of H-1″ with H-3″, H-5″, and H-2′. The n.O.e. data reflected the axial *cis* relationship of H-1,3,5 of the glucose moieties as well as the connections Glc-(1→2)-Glc and Glc-(1→16)-hexadecanoic acid, thereby completing determination of the structure.

EXPERIMENTAL

F.a.b.-mass spectra were recorded in a glycerol matrix with a VG-ZAB-HSQ spectrometer. N.m.r. spectra were obtained for 0.25M solutions in $(CD_3)_2SO$ or $CDCl_3$ (internal Me_4Si) with a Bruker MSL-300 spectrometer (^{1}H, 300.13 MHz; ^{13}C, 75.479 MHz). The quantitative ^{13}C-n.m.r. spectra were recorded after addition of 5% of chromium(III) acetonyl acetonate to the sample (16k data points, zero-filled to 64k and weighted with a 2-Hz line-broadening prior to transformation and spectra subtraction).

2D Heteronuclear-correlated n.m.r. spectra. — Coherence delays were used according to 145 Hz for $^{1}J_{C,H}$. A total of 256 experiments were acquired, each with 32 transients and a relaxation delay of 1.5 s. The resulting matrix of 1024 × 256 data points was zero-filled in the F_1 dimension to 512 points and multiplied by a sine-bell function ($\pi/3$). The F_2 domain was multiplied by a line-broadening function (3 Hz) prior to processing.

2D J-Resolved spectra. — Spectra were recorded with spectral widths of 1700 Hz and 8k data points in the F_2 dimension, and 26 Hz and 64 data points in the F_1 dimension. The data matrix was zero-filled in F_1 to 128 data points and multiplied with sine-bell functions in F_1 ($\pi/4$) and F_2 ($\pi/8$).

2D Homonuclear-correlated n.m.r. spectra. — COSY-90, COSY-45, DQF COSY, DQF RELAY, and DQ 2D n.m.r. spectra were used with a 512 × 1024 data matrix, which was zero-filled and sine-bell shifted by $\pi/8$ in each dimension. According to phase-cycling in the pulse programs, 16, 16, 32, 64, and 64 scans were accumulated for each experiment with a relaxation delay of 2 s. The total mixing time was set to 30 ms for the DQF RELAY. Quadrature detection was employed in both the F_1 and F_2 dimensions, completing the original DQF RELAY[13] to a 64-step phase cycle. The phase cycling for quadrature detection in DQ 2D[12] n.m.r. spectra was supplied by G. E. Martin, which is gratefully acknowledged. The sweep widht in the DQ F_1 dimension was chosen twice as large as in F_2.

REFERENCES

1 A. P. Tulloch, in L. A. Wittig (Ed.), *Glycolipid Methodology*, Am. Oil Chem. Soc., Champaign, Illinois, 1976, pp. 326–344.
2 A. P. Tulloch, *Can. J. Chem.*, 46 (1968) 3727–3730.
3 A. P. Tulloch, A. Hill, and J. F. T. Spencer, *Can. J. Chem.*, 46 (1968) 3337–3351.
4 H.-J. Asmer, S. Lang, F. Wagner, and V. Wray, *J. Am. Oil Chem. Soc.*, 65 (1988) 1460–1466.
5 D. Haferburg, R. Hommel, R. Claus, and H.-P. Kleber, *Adv. Biochem. Eng./Biotechnol.*, 33 (1987) 53–93.
6 R. Hommel, O. Stüwer, W. Stuber, D. Haferburg, and H.-P. Kleber, *Appl. Microbiol. Biotechnol.*, 26 (1987) 199–205.
7 O. Stüwer, R. Hommel, D. Haferburg, and H.-P. Kleber, *J. Biotechnol.*, 6 (1987) 259–269.
8 R. Hommel, unpublished data.
9 L. Weber and G. Haufe, *J. Prakt. Chem.*, 330 (1988) 319–322.
10 J. Dabrowski, P. Hanfland, and H. Egge, *Biochemistry*, 19 (1980) 5652–5658.
11 J. Dabrowski, in W. R. Croasmun and R. M. K. Carlson (Eds.), *Methods in Stereochemical Analysis*, Vol. 9, VCH Publishers, Deerfield Beach, 1987, pp. 349–386.
12 D. A. Craig and G. E. Martin, *J. Nat. Prod.*, 49 (1986) 456–465.
13 P. L. Weber and L. Mueller, *J. Magn. Reson.*, 73 (1987) 184–190.
14 A. Bax, *J. Magn. Reson.*, 57 (1984) 314–318.