Phytochemistry, 1972, Vol 11, pp 295 to 302 Pergamon Press Printed in England

MONOHEXOSIDE DERIVATIVES OF LONG-CHAIN POLYHYDROXY ALCOHOLS; A NOVEL CLASS OF GLYCOLIPID SPECIFIC TO HETEROCYSTOUS ALGAE

T A BRYCE, D WELTI, A E. WALSBY* and B. W. NICHOLS

Biosciences Group, Unilever Research Laboratories, Colworth House, Sharnbrook, Bedford

(*Received* 25 *March* 1971, *in revised form* 14 *May* 1971)

Abstract—Mono-glucoside and -galactoside derivatives of long-chain polyhydroxy alcohols were identified in an unsaponifiable fraction from *Anabaena cylindrica* The major alcohol moiety in these lipids was 1,3,25-trihydroxyhexacosane (I), which was accompanied by much smaller quantities of a tetrahydroxyalcohol, probably 1,3,25,27-tetrahydroxyoctacosane (III)

INTRODUCTION

CERTAIN filamentous blue–green algae possess unique non-saponifiable glycolipids which have not been detected in any other class of plant.[1] Their occurrence is restricted to algal species capable of fixing nitrogen and these lipids appear to be specific to heterocysts,[2,3] the specialized cells considered to be the site of nitrogen fixation [4]

Following the results of preliminary analyses already reported,[1] we now wish to present details of the structure of the heterocyst-specific glycolipid isolated from *Anabaena cylindrica*

RESULTS AND DISCUSSION

Eighty mg of glycosides were obtained from 1 8 g of lipids and pigments isolated from *A cylindrica*, i.e these compounds represented about 4 4% of the total lipid extract and about 0 5% of the total cell dry weight

The recrystallized material melted between 106 5 and 108° and had a specific rotation $[\alpha]_{Hg}^{546}$ of +58 6 in $CHCl_3$–MeOH (2 1, v/v) TLC on silicic acid indicated that the material was not homogeneous, a single major component being accompanied by a much smaller quantity of material giving similar reactions as the major component with specific and non-specific spray reagents, but having a slightly smaller R_f (0 45 as compared with 0 5) when $CHCl_3$–MeOH–HOAc–H_2O (85 15 10 4, v/v) was used for development.

Component Sugars

TLC of the water-soluble products of the glycolipid hydrolysate indicated that glucose and galactose were the only sugars present in significant proportions GLC of the reduced sugar acetates prepared from this fraction confirmed the identity of these components and showed that the mixture was comprised of 84% glucose and 16% galactose.

* Botany Dept , Westfield College, University of London.

[1] B W NICHOLS and B J B WOOD, *Nature, Lond* **217**, 767 (1968).
[2] A E WALSBY and B NICHOLS, *Nature, Lond* **221**, 673 (1969).
[3] C P WOLK and R D SIMON, *Planta* **86**, 92 (1969)
[4] P FAY, W D P STEWART, A E WALSBY and G E. FOGG, *Nature, Lond.* **220**, 810 (1968).

Determination of total hexose in the intact glycoside by the anthrone method indicated that hexose (as glucose) was present at a concentration of 308 μg/mg of lipid.

Aglycone Fraction

When the lipophilic products of hydrolysis were analysed by triple development TLC they were resolved into two alcoholic components. The relative concentration of these two alcohols was in the approximate ratio of 12 1, the major fraction being slightly more mobile (R_f 0 6) than the minor (R_f 0 5) when the mixture was analysed by single development TLC on silica gel employing $CHCl_3$–MeOH (2 1, v/v) as mobile phase

After recovery from the chromatograms the individual components were recrystallized from 80% MeOH, the major alcohol being obtained as an amorphous white powder m.p. 96 5–97 0°.

Structural Analysis of the Major Alcohol Fraction

The mass spectrum obtained for the trimethylsilyl (TMSi) derivative of the major alcohol (Fig. 1) exhibited a molecular ion at *m/e* 630, while a more abundant [M–CH_3] ion,

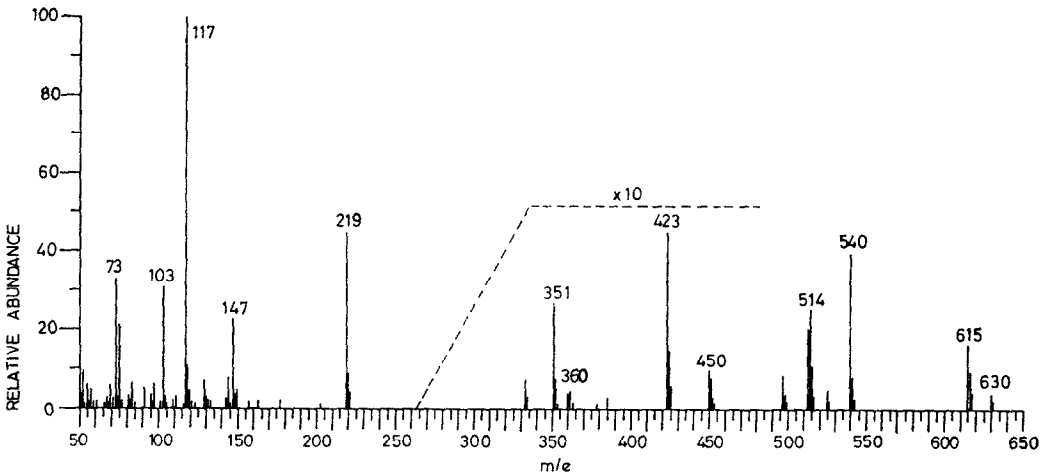

FIG 1. MASS SPECTRUM OF THE TMSi DERIVATIVE OF THE MAJOR ALCOHOL

characteristic of TMSi ethers, was observed at *m/e* 615 A series of peaks at *m/e* 540, 450 and 360, corresponding to the sequential loss of 1, 2 and 3 molecules of trimethylsilanol, indicated that three TMSi ether groups were present in the molecule, and because formation of a single TMSi ether group increases the molecular weight of an alcohol by 72 mass units the molecular weight of the initial alcohol was deduced to be 414 The ion of greatest mass (*m/e* 396) observed in the mass spectrum of the free alcohol (Fig 2) clearly corresponds to [M–H_2O]

These preliminary studies therefore clearly indicated that the major alcohol in the aglucone moiety of the glycolipid was a C_{26} trihydric alcohol, and this general formula was supported by chemical elemental analysis which gave C, 75 1%; H, 13 0%. Required values for $C_{26}H_{54}O_3$ are C, 75 3, H, 13 0%.

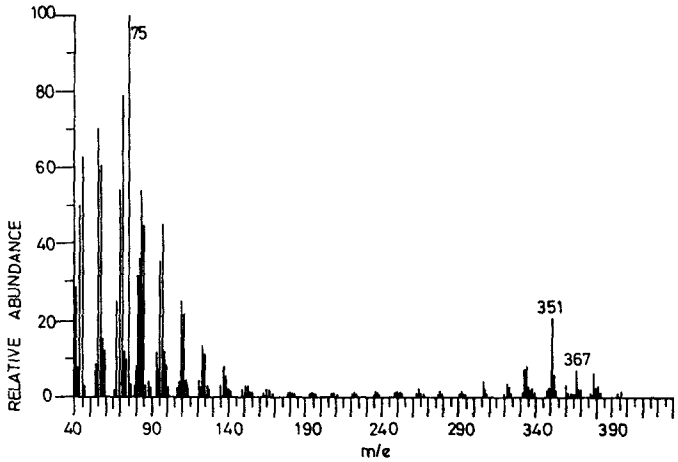

FIG. 2 MASS SPECTRUM OF THE MAJOR ALCOHOL

Position of the hydroxyl groups. The positions of the three hydroxyl groups were established by mass and NMR spectroscopy on the free alcohol and a variety of derivatives.

Mass spectral studies. The mass spectrum of the TMSi derivative (Fig. 1) showed abundant ions at *m/e* 103, 117 and 219. From the patterns of fragmentation generally established for TMSi ethers of aliphatic alcohols[5–8] the first two ions could be considered as corresponding to $CH_2{=}O^+Si(CH_3)_3$ and $CH_3CH{=}O^+Si(CH_3)_3$ respectively, while that of *m/e* 219 could be attributed to either of two possible ions, namely

$$CH_3{-}CH(OSi(CH_3)_3){-}CH{=}O^+Si(CH_3)_3 \quad \text{(a)} \qquad \text{or} \qquad (CH_3)_3SiO{-}CH_2{-}CH_2{-}CH{=}O^+Si(CH_3)_3. \quad \text{(b)}$$

The presence of peaks at *m/e* 513, 423 and 333 are in agreement with these deductions. Loss of 117 mass units from the molecular ion gives rise to *m/e* 513, and the further consecutive loss of two molecules of trimethylsilanol accounts for the formation of the peaks at *m/e* 423 and 333. An indication of which of the two structures for *m/e* 219 is the correct one was obtained by a comparison of the mass spectra of the alcohol and its TMSi derivative with those of 1,3-dihydroxyhexadecane and its TMSi ether (Figs. 3 and 4). As expected, the mass spectrum of the TMSi ether of the diol contained peaks at *m/e* 103 (corresponding to $CH_2{=}O^+Si(CH_3)_3$) and *m/e* 219, corresponding to structure (b). More significantly, the base peak in the spectrum of both the free diol and triol (Fig. 2) was *m/e* 75 of composition

[5] G EGLINTON, D H HUNNEMAN and A MCCORMICK, *Org Mass Spectros* **1**, 593 (1968).
[6] J DIEKMAN, J B THOMSON and C DJERASSI, *J Org Chem* **33**, 2271 (1968)
[7] W J. ESSELMAN and C O CLAGETT, *J Lipid Res* **10**, 234 (1969).
[8] J. KOSSANYI, J. P. MONZUR, B FIRTH, J WIEMANN, A M DUFFIELD and C D JERASSI, *Org. Mass Spectros* **1**, 777 (1968).

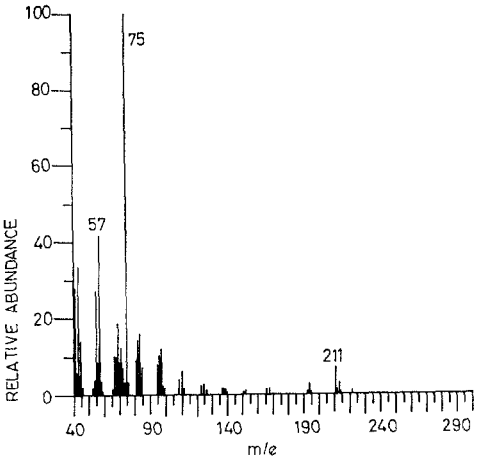

FIG 3 MASS SPECTRUM OF 1,3-DIHYDROXYHEXADECANE

$C_3H_7O_2$[9] corresponding to the structure HO—CH_2—CH_2—CH=O^+H The alternative fragment of the same composition which could be derived from a 2,3-dihydroxy compound, i e HO^+=CH—CHOH—CH_3, is less likely to account for the base peak in the triol spectrum since this type of fragmentation is normally of minor importance in the spectra of *vic*-diols [7 10]

Evidence indicating the position of the third hydroxyl group was provided by analytical studies on the acetonide derivative of the triol, in which the alcohol groups at the 1- and 3-positions were condensed with acetone leaving the third hydroxyl free Preparation of the TMSi derivative of this isolated hydroxyl group in the acetonide was followed by mass spectral analysis of the final TMSi-acetonide which showed no molecular ion, although the

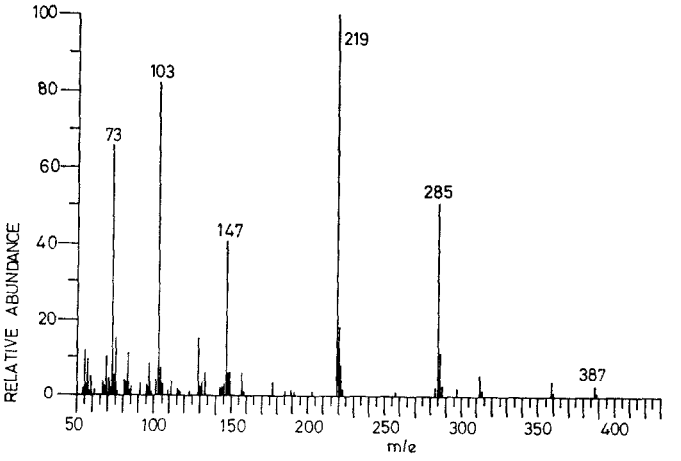

FIG 4 MASS SPECTRUM OF THE TMSi DERIVATIVE OF 1,3-DIHYDROXYHEXADECANE

[9] Confirmed by high resolution mass spectroscopy
[10] E O A HAAHTI and H M FALES, *J Lipid Res* **8**, 131 (1967)

abundant [M–CH_3] ion had the expected composition of $C_{31}H_{63}O_3Si$[9] thus confirming the molecular formula of the parent trihydric alcohol as $C_{26}H_{54}O_3$ Moreover, the base peak *m/e* 117 had the composition $C_5H_{13}SiO$,[9] corresponding to the structure CH_3—CH=$O^+Si(CH_3)_3$, which indicated that the isolated hydroxy group in the triol must have been in the ω-1 position. The position of the acetonide moiety was confirmed by recognizing the peaks derived from the cleavages alpha to the 1,3-dioxane ring.[11] The α-cleavage peaks are *m/e* 115, which had the expected composition of $C_6H_{11}O_2$,[9] and (M–CH_3)

On the basis of the evidence described above, the structure proposed for the triol was 1,3,25-trihydroxyhexacosane (I) and further support for this structure was obtained from NMR studies on the alcohol and various derivatives

NMR spectral studies The chemical shift values (δ ppm) of the protons in the triol, its acetate, the reference compound 1,3,-dihydroxyhexadecane, and its acetate, are shown in Table 1

$HOCH_2$—CH_2—CHOH—$(CH_2)_{21}$—CHOH—CH_3 (I)

$HOCH_2$—CH_2—$(CH_2)_{20}$—CHOH—CH_2—CHOH—CH_3 (II)

$HOCH_2$—CH_2—CHOH—$(CH_2)_{21}$—CHOH—CH_2—CHOH—CH_3 (III)

TABLE 1 CHEMICAL SHIFTS (δ-ppm) OF THE PROTONS IN 1,3,25-TRIHYDROXYHEXACOSANE, 1,3-DIHYDROXYHEXADECANE, AND THE CORRESPONDING ACETATES

Type of proton	a	b	c	d	e	f	g	Acetate CH_3
Triol (in CD_3OD at 55 60°)	1 14	1 27	1 43	1 66	3 69	(3 69)	(3 69)	—
Diol (in $CDCl_3$ at Ambient temp and ~35°	—	1 23	1 46	1 70	3 86	—	(3 86)	—
Triacetate (in CCl_4 at Ambient temp and 50°)	1 165	1 25	1 51	1 78	4 00	4 79	4 88	1 939 1·968 1 975
Diacetate (in CCl_4 at Ambient temp)	—	1 24	1 52	1 80	4 00	—	4 86	1 955
	Relative number of Protons from Spectral Integration							
Triol	3	—†	4	2	e + f + g = 4			—
Diol	—*	—†	2	2	e + g = 3			—
Triacetate	3	—†	4	2	2	f + g = 2		9
Diacetate	—*	—†	2	2	2	—	1	6

e d g c b c f a
CH_2—CH_2—CH—CH_2—$(CH_2)_{19}$—CH_2—CH—CH_3
| | | |
OH OH OH

* There is no CH_3 CHOH group in the diol

† The integrations of the $(CH_2)_n$ groups were not very accurate compared with those of the groups containing 1, 2 or 3 protons

[11] J A McCLOSKEY and M J McCLELLAND, *J Am Chem Soc* **87**, 5090 (1969), R E WOLFF, G WOLFF and J A McCLOSKEY, *Tetrahedron* **22**, 3093 (1966)

The free triol was sparingly soluble in the common NMR solvents, requiring temperatures of 55–60° before it would dissolve in CD_3OD The 60 MHz spectrum was dominated by the absorptions due to H_2O/OH, the CH_3 impurity from CD_3OD, and the long chain CH_2 protons. The weaker but more important absorptions due to only one or two protons were more easily observed in the 220 MHz spectrum which showed a doublet ($J = \sim 6$ Hz) centered at δ1·14 ppm which was attributed to the protons of the CH_3CHOH group After integration, an unresolved multiplet at δ1 43 ppm was attributed to two CH_2 groups each of which was β to a single OH group Another multiplet at δ1 66 ppm was attributed to a CH_2 group which was β to two OH groups. A triplet at δ3 69 ppm integrated to the equivalent of four protons, and was associated with the protons attached to the same carbon atoms as the OH groups Because of some doubts in the interpretation of this spectrum, including the possibility that additional CHOH and CH_3 proton absorptions might be concealed beneath the water and $(CH_2)_n$ absorptions respectively, and also because of the solubility problems, the acetate derivative was prepared and used for subsequent NMR studies.

Triol triacetate. The 60-MHz spectrum of the acetate was not well enough defined between δ1 0 and δ2 0 ppm for a full interpretation to be made, and it was again necessary to obtain a 220-MHz spectrum

A distinctive doublet due to the CH_3 protons in the $CH_3CH(OAc)$-group appeared at δ1 165 ppm. A multiplet at δ4 00 ppm which integrated to two protons was attributed to the protons of the primary —$\underline{CH_2}$OAc group, while a broad multiplet in the range δ4 7–4 95 ppm, which also integrated to two protons, was interpreted as being a composite band with two centres at δ4 79 ppm and δ4 88 ppm caused by the absorption of the two secondary —$\underline{CH}$OAc protons This was partially confirmed when the —$\underline{CH}$OAc proton in 1,3-diacetoxyhexadecane was correlated with an absorption centred at 4 86 ppm

This evidence reinforced the probability that the molecule had chain endings of —CH_2OAc and $CH_3CH(OAc)$— respectively, and that it contained another secondary —CHOAc group within the chain The absence of a CH_3 absorption with a chemical shift lower than δ1 0 ppm suggested that there were no side chains· any further methyl substitution on a —C(OH) group would have produced a methyl singlet absorption. The mass spectra had already shown that the molecule was acyclic and fully saturated, and this was partly confirmed by the absence of any olefinic proton absorption in the NMR spectra

Of the remaining absorptions in the spectrum, an unresolved multiplet at δ1 51 ppm, which integrated to an equivalent of four protons was attributed to two pairs of —CH_2—$\underline{CH_2}$—CH(OAc)— protons The difference in shift between a multiplet at δ1 78 ppm and the $(CH_2)_n$ absorption at δ1 25 ppm was approximately double that between the δ1 51 ppm and δ1 25 ppm absorptions It was reasonable therefore to attribute the δ1 78 ppm multiplet to the protons of the CH_2 group lying between two —CH(OAc)— groups

On this basis alone the triacetate therefore had either of the two possible structures derived from I and II However, in the 220-MHz spectrum both the δ4 00 ppm $\underline{CH_2}$OAc and δ1 78 ppm —CH(OAc)—$\underline{CH_2}$—CHOAc absorptions were multiplets These multiplets must have been caused by the non-equivalence of both pairs of protons, which can only occur in the case of structure (I)

The structure of the triacetate was confirmed by a double decoupling experiment performed on a 100-MHz spectrometer Irradiation of the CH_2 multiplet at δ1 78 ppm reduced the δ4 00 ppm 'triplet' to a singlet and simplified the absorption around δ4 8–4 9 ppm Irradiation of the CH_2 'triplet' at δ4 00 ppm reduced the CH_2 multiplet at δ1 78 ppm to a

doublet. Irradiation of the CH_3 doublet at δ1 165 ppm also simplified the absorption around δ4 8–4·9 ppm, and irradiation of the latter multiplet reduced the CH_3 doublet to a singlet.

The spectra of the reference compounds hexadecane-1,3-diol and hexadecane-1,3-diacetate were very similar to the equivalent triol and triacetate spectra (see Table 1) and confirmed their interpretation

Structure of the Minor Alcohol Fraction

The mass spectrum of the minor aglucone alcohol showed a very small (<0 5%) ion at *m/e* 440, while the molecular weight of its TMSi derivative (Fig. 5) was 746. These data

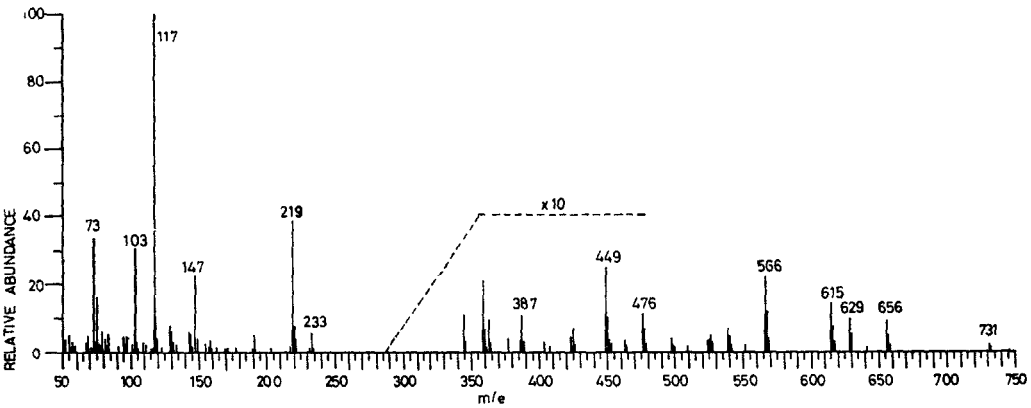

FIG 5 MASS SPECTRUM OF THE TMSi DERIVATIVE OF THE MINOR ALCOHOL

indicated that the alcohol contained four hydroxyl groups, had a molecular weight of 458 (746 = 458 + 4 × 72) and the molecular formula $C_{28}H_{58}O_4$. In the mass spectrum of the TMSi derivative (Fig. 5) the ions *m/e* 656, 566, 476 and 386 indicated consecutive losses of trimethylsilanol and confirmed that there were four TMSi ether groups in the molecule. The abundant ions *m/e* 103, 117 and 219 also present in this spectrum indicated that the tetrol has like the triol, a 1,3,ω-1 substitution pattern. The position of the fourth hydroxyl group is not certain, although the mass spectrum of the free alcohol contains the ions *m/e* 75 and 89 which could clearly correspond to the groups. $HOCH_2—CH_2—CH=O^+H$ and $HO^+=CH—CH_2—CHOH—CH_3$ respectively The presence of the latter fragment suggests that the fourth hydroxyl is in the ω-3 position and that the tetrol has the structure (III). The mass spectrum of the TMSi ether was also in agreement with this structure, exhibiting ions of *m/e* 219 and 233 (Fig. 5) which correspond to the TMSi derivatives of the *m/e* 75 and 89 ions observed in the free alcohol spectrum

Structure of the Glycolipid

The sugar content of the glycolipid, as determined by the anthrone method, was 308 μg hexose (calculated as glucose) per mg of lipid. This agrees very closely with the theoretical value for a glycoside in which a single hexose unit is bound to a C_{26} trihydric alcohol, although we have so far been unable to determine the specific hydroxyl group involved in the glycosidic linkage.

The unsaponifiable glycolipid fraction present in the heterocysts of *A. cylindrica* is

therefore mainly composed of monosaccharide derivatives of 1,3,25-trihydroxyhexacosane (I) and 1,3,25,27-tetrahydroxyoctacosane (III) in which the monosaccharide moiety is either glucose or galactose

To our knowledge, polyhydric alcohols such as we describe here have not previously been detected in natural tissues Their role in the metabolism of heterocysts is not obvious, although it is possible that they fulfil some function analogous to that of bactoprenol and related substances in microbial metabolism, namely that of acting as carriers and intermediates in cell wall biosynthesis [12–14]

EXPERIMENTAL

Fractionation of lipids Cultures of *A cylindrica* were centrifuged and the culture medium decanted from the packed cells which were then shaken with 100 vol of $CHCl_3$–MeOH (2 1, v/v) After standing at room temp for about 1 hr, the extraction mixtures were filtered and the residues re-extracted with $CHCl_3$–MeOH The extracts were combined, concentrated *in vacuo* and the residues redissolved in a small volume of $CHCl_3$–MeOH (2 1, v/v) which was then shaken with 1/5 vol of 0 7% saline to remove water-soluble material

Isolation and purification of glycolipid 1 8 g of the lipid mixture was dissolved in pure $CHCl_3$ and applied to a 50 × 3 5 cm column of DEAE cellulose (acetate form) [15] Pigment and neutral lipid were eluted from the column with 2L of $CHCl_3$–MeOH (97 3, v/v), and a crude glycolipid fraction was then displaced with 1 5 l of $CHCl_3$–MeOH (95 5 v/v)

The non-saponifiable glucoside was purified by saponification of this fraction with 2 N methanolic KOH at room temp for 3 hr Following dilution of the reaction mixture with 100 vol water the mixture was brought to pH 2 with acid and was then extracted with $CHCl_3$ (×3) This extract, which contained the glycolipid, pigment and free fatty acid, was applied to a column of silic acid which was then eluted with (a) pure $CHCl_3$ and (b) $CHCl_3$–MeOH (2 1, v/v) The fraction eluted by the latter solvent was recrystallized from 80% MeOH (×3) to yield 80 mg of a white, amorphous solid

Hydrolysis of the glycolipid The glycolipid was hydrolysed by refluxing for 28 hr with 2 N H_2SO_4 in MeOH–H_2O (9 1, v/v) The reaction mixture was then diluted with water, extracted with $CHCl_3$ and the $CHCl_3$ and aqueous layers separated and analysed for lipophilic compounds and sugars respectively

Analysis of the aglucone fraction A solution of the crude aglucone was streaked onto thin layer chromatograms of silica gel and the applied material was then resolved by triple development in $CHCl_3$–MeOH (95·5, v/v) The individual components were detected by spraying the chromatograms with water which revealed the separated substances as opaque bands on a translucent background The relevant areas were scraped from the chromatograms and the adsorbed compounds eluted with $CHCl_3$–MeOH (2 1, v/v) Trimethylsilyl ethers,[5] acetonide[10] and acetate derivatives were prepared by standard methods

Mass spectroscopic analyses were performed on an AEI MS-902 apparatus using the direct inlet system The ionizing current was 100 μA and the electron energy was 70 eV NMR spectra were obtained with Perkin–Elmer R10 60 MHz and Varian HR 220 MHz spectrometers, a variable temperature probe being used with the 220 MHz apparatus De-coupling experiments were performed on a Varian HA 100 spectrometer

Sugar analyses Portions of the water-soluble hydrolysis products were fractionated by TLC on air-dried cellulose powder by triple development with acetone-*n*-BuOH–H_2O (10 8 3, v/v), using aniline phthalate for the location of sugars Preparation and GLC of the reduced sugar acetates was according to the method of Sawardeker *et al* [16] Quantitative determination of sugar in the intact glycoside was obtained by the anthrone method

Acknowledgements—The authors thank Dr P E J Verwiel of the Central Laboratory, T N O, Delft, Holland for the 220-MHz spectra, Mr I A Fowlis for the 60-MHz and Mr D J Frost of U R L Vlaardingen for the de-coupling data

[12] K J I THORNE and E KODICEK, *Biochem J* **99**, 123 (1966)
[13] G SIEWERT and J L STROMINGER, *Proc Nat Acad Sci U S A* **57**, 767 (1967)
[14] J G CHRISTENSON, S K GROSS and P W ROBBINS, *J Biol Chem* **244**, 5436 (1969).
[15] B W NICHOLS and A T JAMES, *Fette, Seifen Anstrichmittel* **66**, 1003 (1964)
[16] J S SARWARDEKER, J H SLONEKER and A JEANES, *Analyt Chem* **37**, 1602 (1965)

Key Word Index—*Anabaena cylindrica*, Cyanophyta, glycolipids, polyhydroxy alcohols, galacto lipids.