

CHEMICAL STRUCTURE OF THE ETHER LIPIDS OF THERMOPHILIC ACIDOPHILIC BACTERIA OF THE *CALDARIELLA* GROUP

MARIO DE ROSA,* SALVATORE DE ROSA,* AGATA GAMBACORTA,* LUIGI MINALE* and JOHN D. BU'LOCK†

*Laboratorio per la Chimica M. I. B. del C. N. R. Via Toiano 2, Arco Felice, Napoli, Italy; †Department of Chemistry, University of Manchester, England

(Received 24 May 1977)

Key Word Index—*Caldariella*; thermophilic bacteria; ether lipids; biphytanyl ether lipids.

Abstract—The lipids of the *Caldariella* group of extremely thermophilic acidophilic bacteria are based on a 72-membered macrocyclic tetraether made up from two C_{40} diol units and either two glycerol units or one glycerol and one nonitol. The C_{40} components have the 16,16'-biphytanyl skeleton and the detailed structure of three of them is established.

INTRODUCTION

The prevalence of lipids based on isoprenoid glyceryl ethers in very thermophilic acidophilic bacteria such as *Sulfolobus*, *Thermoplasma*, and others, classified by us as the *Caldariella* group, [1-3] has been established as a characteristic of the whole group [4] and our preliminary characterisation of the major components as a mixture of cyclic ethers each combining a 2,3-glyceryl unit etherified with a C_{40} isoprenoid residue has been briefly described [4,5]. Since cleavage of the mixture with BCl_3 gave glycerol and the C_{40} dichlorides, $C_{40}H_{76-80}Cl_2$, the cyclic ethers were formulated as (1). Preliminary evidence was also presented to show that the C_{40} residues were formally derived by ω,ω' -linkage (i.e. head-to-head) of two O-phytanyl residues, with or without further cyclisations, giving an acyclic, a monocyclic, and a bicyclic C_{40} component; this conclusion was based on NMR and mass spectroscopy data obtained for the C_{40} diols and the corresponding hydrocarbons.

In this paper we report evidence that one of the two major series of ether lipids (obtained after hydrolysis of the total complex lipids) is to be formulated not as (1) but as (2) i.e. as a 72-membered macrocyclic tetraether combining two glycerols and two C_{40} residues, a conclusion which has also been reached independently by Langworthy [6]. We also report detailed structural assignments for the three types of C_{40} residue contained in the lipids of this series. These diglycerol di(biphytanyl) tetraethers are accompanied by a second lipid class, in which one of the two glycerols is replaced by a nonitol and the biphytanyl residues are more extensively cyclized; detailed description of this second series will be given in a later account.

RESULTS

The macrocyclic tetraether structure

Determinations of the molecular weight by vapour pressure osmometry of the mixed diglycerol tetraethers in $CHCl_3$ gave a value of $M = 1290$ [(1) requires 646-650; (2) requires 1292-1300], and for corresponding

diglycerol tetraether diacetate mixture $M = 1350$. In NMR spectra of the tetraether diacetate the signal due to the $-CH_2OAc$ protons (δ 4.14, multiplet) is coupled to the alkoxy methine proton signal (δ 3.6, complex and overlapping with alkoxymethylene signal) and this established structure (2) for the diglycerol tetraethers (without however distinguishing between the alternative attachments of the second carbinol group). Similar but more preliminary data on the glycerol nonitol tetraethers indicate an analogous structure, with the nonitol replacing one of the two glycerols in (2).

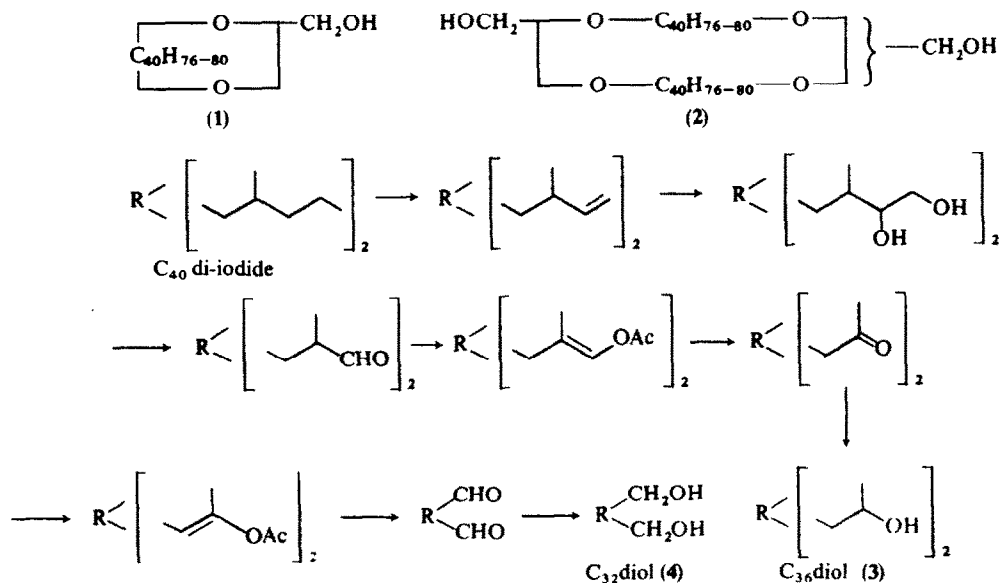
The C_{40} ω,ω' -biphytanyl components

Information on the isoprenoid structures was obtained mainly by detailed GC-MS studies on the derived hydrocarbons and by 1H NMR; using the europium shift reagent on the separate C_{40} diols and on C_{36} and C_{32} diols derived by chain-shortening reactions.

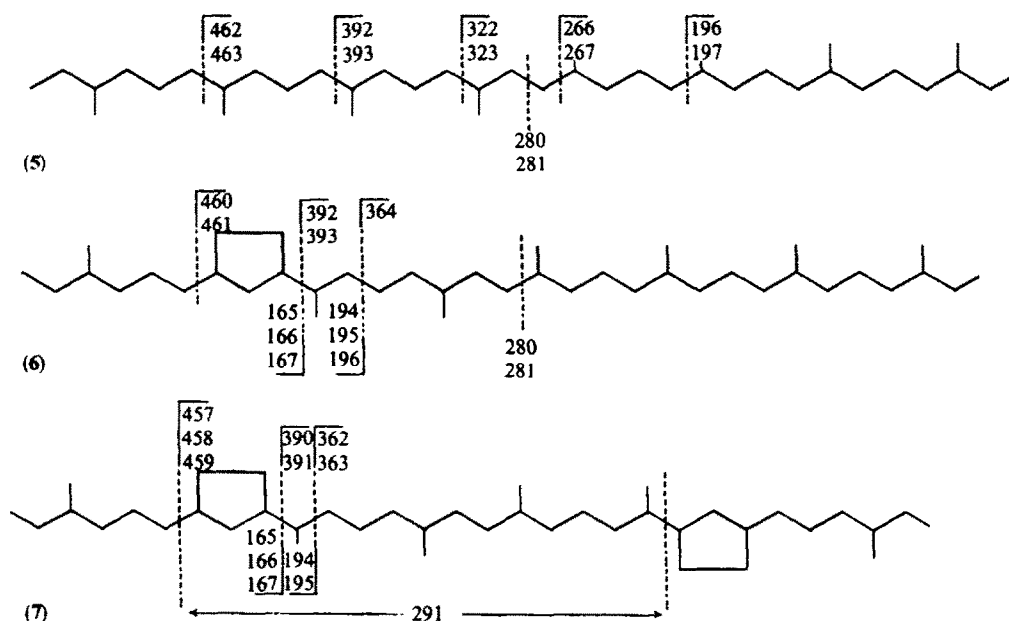
The C_{40} hydrocarbons

Conversions of the tetraethers into the C_{40} diiodides with HI and of the diiodides into the C_{40} hydrocarbons by $LiAlH_4$, have already been described as has the GLC separation of the hydrocarbon mixture. Hydrocarbons with $-C^1H_2^2H$ terminal groups were also prepared (by using $LiAl^2H_4$) and these 2H_2 -hydrocarbons were useful in assigning certain ions in the mass spectra. Whereas the diglycerol tetraethers afforded the hydrocarbons $C_{40}H_{78}$, $C_{40}H_{80}$, and $C_{40}H_{82}$, discussed in detail below, the corresponding hydrocarbons from the glycerol nonitol tetraethers covered the range $C_{40}H_{74}$, $C_{40}H_{76}$, and $C_{40}H_{78}$, the last being identical with the bicyclic hydrocarbon from the diglycerol tetraether mixture.

The MS of the saturated acyclic hydrocarbon $C_{40}H_{82}$ ($M^+ m/e$ 562; 2H_2 analogue, 564) can be rationalized as in Scheme 2 in terms of (a) a series of cleavages, both with and without H transfer to the uncharged fragment, α to $>CHMe$ groups and removing successively 7,5,5,4 and 5C atoms, finally generating the base peaks at m/e 196/197 and (b) central cleavage



Scheme 1.



Scheme 2.

(again with or without H transfer), β to two $>\text{CHMe}$ groups, to fragments at m/e 280,281. The two types of fragmentation are independent since in chemical ionization spectra (kindly provided by C. J. W. Brooks) only the molecular ion and the base peak from series *a* appear. All the fragmentation peaks were shifted by 1 mass unit in the spectra of the $^2\text{H}_2$ -hydrocarbon. With the NMR evidence that the corresponding functionalized derivatives [e.g. the diol $\text{C}_{40}\text{H}_{80}(\text{OH})_2$] contain 8 Me groups (see below), and the biogenetic evidence for an isoprenoid skeleton [5,7], the data require structure (5) for the hydrocarbon $\text{C}_{40}\text{H}_{82}$, as in Scheme 2. In particular they provide strong

evidence for the central head-to-head linkage in the ω,ω' -biphytanyl skeleton.

The MS of the monocyclic hydrocarbon $\text{C}_{40}\text{H}_{80}$ (m/e 560 and m/e 562 in the $^2\text{H}_2$ -material) shows significant differences. While the molecular ion and the first major fragments at m/e 460,461 have one formal unsaturation the next fragments, at m/e 392 and 393, do not. Cleavage at the same bond also generates a group of strong peaks at m/e 165–167 (base peak, 166). There are corresponding cleavages apparently at an adjacent bond, differing by 28 mass units; the 'central' cleavage gives only fragments of m/e 280 and 281, presumably because the other moiety, for which peaks at m/e 278

and 279 would be expected, generates instead the smaller fragments at m/e ca 166 and ca 194. As before all the fragmentation peaks were shifted by 1 mass unit in the spectrum of $^2\text{H}_2$ -material. Since the NMR evidence (below) shows that in the corresponding monocyclic diol there are just 7 CHMe groups we may conclude that the eighth Me group of the parent skeleton is used in forming the ring; this leads to structure (6) and the rationalization of the MS data shown in Scheme 2.

Finally for the bicyclic hydrocarbon $\text{C}_{40}\text{H}_{78}$ the molecular ion m/e 558 (560 in the $^2\text{H}_2$ -material) and the fragments at m/e 457–459, 390–391, and 362–363 (all shifted by 1 mass unit in the spectra of the $^2\text{H}_2$ -material) show a further deficit of 2 mass units when compared with the data for (6), the important counterparts at m/e 165–166 and 194–195 appearing as before. The next major fragment in the spectrum has m/e 291 corresponding to a C_{21} fragment with two formal unsaturations. This peak is not shifted in the spectrum of the $^2\text{H}_2$ -hydrocarbon; it is therefore formed by cleavage at each end of the molecule and moreover it retains only one of the two rings originally present. Again, the NMR evidence (below) shows that in the corresponding diol etc. there are only six methyl groups, all secondary, which in the europium-expanded spectrum give three separated doublets, implying that they are symmetrically disposed in the molecule. This leads to structure (7) for the bicyclic hydrocarbon and the rationalization shown in Scheme 2 for its MS fragmentation.

NMR data: C_{40} diols

High-resolution NMR measurements (300 MHz; CDCl_3) of a tetraether sample in which the bicyclic C_{40} component predominated clearly indicated that in this component there are six $>\text{CHMe}$ groups in the molecule. However, in order to obtain definitive data it was clearly essential to investigate the separate C_{40} components; preparative GLC of the C_{40} diiodides was impossible because of their thermal instability but the di-acetates were successfully separated by this technique, and converted into the separate C_{40} diols: $\text{C}_{40}\text{H}_{80}(\text{OH})_2$ (acyclic), $\text{C}_{40}\text{H}_{78}(\text{OH})_2$ (monocyclic), and $\text{C}_{40}\text{H}_{76}(\text{OH})_2$ (bicyclic). The relevant data obtained by NMR measurements using the europium shift reagent are summarized in Table 1.

With additions of up to 1:1 mole ratio of $\text{Eu}(\text{fod})_3$ the signals for protons up to the fourth carbon from

Table 1. Values of δ (from Me_4Si) for CDCl_3 and for 1:1 mole ratio $\text{Eu}(\text{fod-d}_9)_3$ in CDCl_3 for the C_{40} diols (9), (10), and (11)

Assignment*	δ (CDCl_3)	δ (Eu shifted)	Present in (integral)
(protons on C no.)			
1,1'	3.67	8.87	9,10,11 (4H)
2,2'	1.30	4.70, 4.30†	9,10,11 (4H)
3,3'	1.55	3.65	9,10,11 (2H)
4,4'	1.28	2.58, 2.18†	9,10,11 (4H)
17,17'	0.90	2.15	9,10,11 (6H)
18,18'	0.85	0.97	9 (6H), 10 (3H)

*Numbering as in formulae (9–11); assignments supported by integration, multiplicity, and decoupling.

†Methylene protons non-identical with respect to adjacent chiral $>\text{CHMe}$.

the alcohol oxygen were fully resolved, and showed that in all three C_{40} diols there are two identical end-groups $-\text{CH}_2\text{CHMeCH}_2\text{CH}_2\text{OH}$. It is convenient to use the convention accepted for polyprenols and designate these C_5 units as $\alpha\alpha'$, and similarly, inwards from the end-groups the remainder as $\beta\beta'$, $\gamma\gamma'$, and $\delta\delta'$. The spectra obtained for the acyclic diol also showed a resolved signal corresponding to two further CHMe groups (in the $\beta\beta'$ units) at 0.85 shifting to 0.97 at 1:1 mole ratio; the same signal, but of half the intensity (i.e. for only one CHMe) appeared in spectra of the monocyclic diol but was completely absent in the spectra obtained with the bicyclic diol. Thus the cyclizations involve the methyl carbons of the β units; the β and β' units are identical in the acyclic diol, different in the monocyclic, and again identical in the bicyclic diol. Similarly in the NMR spectra of the C_{40} diacetates measured in C_6D_6 the acyclic diacetate shows three separate doublets, at δ 0.83 (6H), 0.92 (6H) and 0.95 (12H) assigned to the $\alpha\alpha'$, $\beta\beta'$ and ($\gamma\gamma'$ + $\delta\delta'$) units respectively; the bicyclic diacetate shows δ 0.83 (6H), no signal at 0.92, and a pair of doublets at 0.94 and 0.96—i.e. the CHMe of the β and β' units are missing and the $\gamma\gamma'$ and $\delta\delta'$ units are now somewhat differentiated.

NMR data: C_{36} and C_{32} bicyclic diols

Further data were obtained from NMR shift studies on the C_{36} and C_{32} bicyclic diols obtained from the C_{40} bicyclic diacetate by the route shown in Scheme 1 (see Experimental), which incidentally provides direct chemical evidence for the identical α and α' units. The NMR results are summarized in Tables 2 and 3.

Table 2. Values of δ (from Me_4Si) for CDCl_3 and for 1:1 mole ratio $\text{Eu}(\text{fod-d}_9)_3$ in CDCl_3 for the bicyclic C_{36} diol (3)

Assignment*	δ (CDCl_3)	δ (Eu shifted)	Integral
(protons on C no.)			
3,3'	3.00	14.00	2H
17,17'	1.16	5.96	6H
4,4'	1.40	7.40, 6.20†	4H
5,5'	1.28	5.18	4H
6,6'	1.28	4.03	4H
7,7'	1.70	3.20	2H

*Numbering as in formulae 9–11; assignments supported by integration, multiplicity, and decoupling.

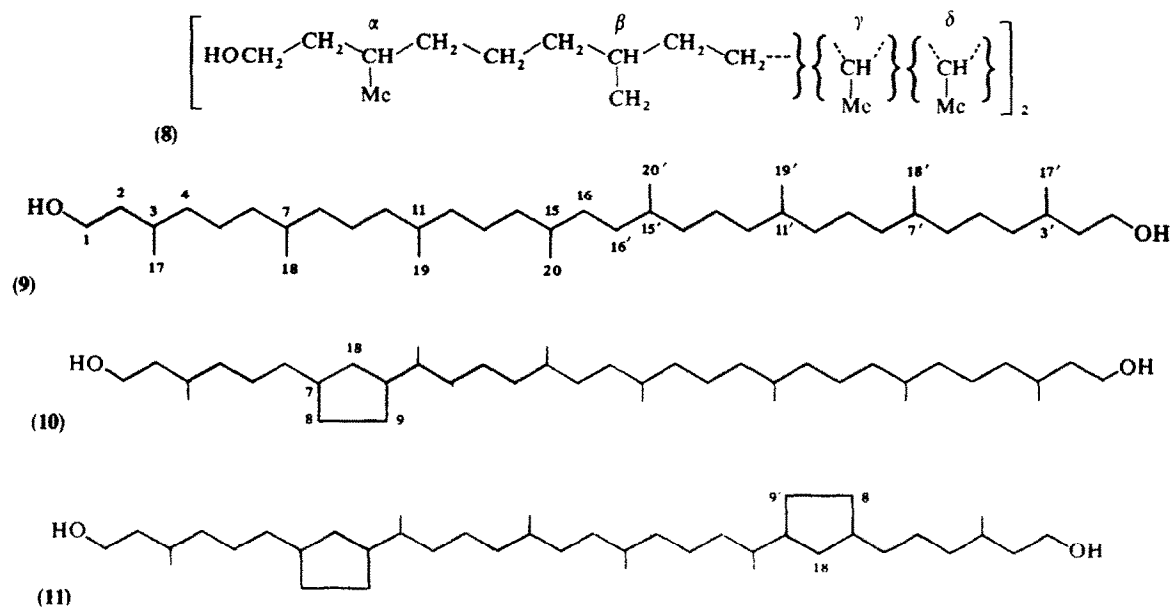
†Methylene protons non-identical with respect to adjacent chiral $>\text{CHMe}$

Table 3. Values of δ (from Me_4Si) for CDCl_3 and for 1:1 mole ratio $\text{Eu}(\text{fod-d}_9)_3$ in CDCl_3 for the bicyclic C_{32} diol (4)

Assignment*	δ (CDCl_3)	δ (Eu shifted)	Integral
(protons on C no.)			
4,4'	3.63	9.7	4H
5,5'	1.38	5.2	4H
6,6'	1.28	4.1	4H
7,7'	1.70	3.3	2H
19,19'	0.85	0.95	6H
20,20'	0.85	0.85	6H

*Numbering as in formulae 9–11; assignments supported by integration, multiplicity, and decoupling.

Cleavage of the tetraethers. The diglycerol tetraether mixture (100 mg) was treated with BCl_3 (5 ml in 5 ml CHCl_3 , 18° , 12 hr); the reaction mixture was evapd under N_2 and chromatographed; CHCl_3 eluted the C_{40} dichlorides (93 mg) already described [4].



C₄₀ diols. The mixed di-iodides (2 g) were treated with AgOAc (4 g) (in AcOH, 300 ml; 24 hr reflux). Water was added and the mixture extracted with Et₂O; the Et₂O layer was washed successively with H₂O, aq. NaHCO₃, aq. Na₂S₂O₃, and H₂O, dried and evapd. The residue was chromatographed in petrol. Et₂O, and with 20% Et₂O the C₄₀ diacetates (1.4 g) were eluted. The mixed diacetates were separated by prep GLC

(recovery ca 60%) to give the acyclic (24 mg), monocyclic (112 mg), and bicyclic (630 mg) C_{40} diacetates (ν_m 1745, 1235, 1030 cm^{-1} for all three). Acyclic C_{40} diacetate [for numbering, cf. (9)]: $[\alpha]_D^{20}$ 0.00° (c, 1.0); m/e 678 (M^+ , < 1%), 616 (M^+ -AcOH, 2), 558 (M^+ -2AcOH, 2), 459 (M^+ -2AcOH- C_7H_{15} , 2), 389(15), 319(12), 279(25), 195(100); δ (C_6D_6) 3.97 [4H, t (J = 7 Hz), CH_2-CH_2-OAc], 1.92 (6H, s, CH_3-CO), 1.55 (ca 8H, b, CHMe), 0.95 [12H, d (J = 5.5 Hz), $CH-CH_3$ (19,19', 20,20')], 0.92 [6H, d (J = 5.5 Hz), $CH-CH_3$ (18,18')], 0.83 [6H, d (J = 5.5 Hz), $CH-CH_3$ (17,17')]. Monocyclic C_{40} diacetate [for numbering, cf. (10)]: $[\alpha]_D^{20}$ -13.3° (c, 1.0); m/e 676 (M^+ , 2%), 616 (M^+ -2AcOH, 2), 556 (M^+ -2AcOH, 2), 517 (M^+ -AcOH- C_7H_{15} , 2), 457 (M^+ -2AcOH- C_7H_{15} , 2), 389(6), 223(35), 165(100); δ (C_6D_6) 3.97 [4H, t (J = 7 Hz), CH_2-CH_2-OAc], 1.92 (6H, s, CH_3-CO), 1.80 (ca 2H, b, ring CH), 1.55 (ca 7H, b, CHMe), 0.95 [12H, d (J = 5.5 Hz), $CH-CH_3$ (19,19', 20,20')], 0.92 [3H, d (J = 5.5 Hz), $CH-CH_3$ (18')], 0.83 [6H, d (J = 5.5 Hz), $CH-CH_3$ (17,17')]. Bicyclic C_{40} diacetate [for numbering, cf. (11)]: $[\alpha]_D^{20}$ +5.2° (c, 2.0); m/e 674 (M^+ , 2%), 614 (M^+ -AcOH, 2), 554 (M^+ -2AcOH, 2), 515 (M^+ -AcOH- C_7H_{15} , 2), 455 (M^+ -2AcOH- C_7H_{15} , 2), 291(5), 223(40), 165(100); δ (C_6D_6) 3.97 [4H, t (J = 7 Hz), CH_2-CH_2-OAc], 1.92 (6H, s, Me-CO), 1.80 (ca 4H, b, ring CH), 1.55 (ca 6H, b, CHMe), 0.96 [6H, d (J = 5.5 Hz), $CH-CH_3$ (20,20')], 0.94 [6H, d (J = 5.5 Hz), $CH-CH_3$ (19,19')], 0.83 [6H, d (J = 5.5 Hz), $CH-CH_3$ (17,17')]. The separate diacetates were saponified (10% aq. KOH, 6 hr reflux); the hydrolyzate was diluted with H_2O and extracted several times with Et_2O to afford, separately, the C_{40} diols (9), (10), and (11); for 1H NMR and Eu shift data cf. Table 1.

Bicyclic C_{36} diol (cf. Scheme 1). The bicyclic C_{40} diacetate (0.60 g) was treated with HI (50 ml of 57% HI; 24 hr reflux); the product taken up in petrol, was washed successively with H_2O , satd aq. K_2CO_3 , aq. $Na_2S_2O_3$, H_2O , and H_2O -MeOH (1:9) and finally chromatographed in *n*-hexane. The bicyclic C_{40} di-iodide thus prepared (0.61 g) was treated with M KOBut in dry ButOH (8.5 ml) at 85°; after 12 hr the reaction mixture was diluted with aq. MeOH and extracted with petrol; prep TLC afforded the bicyclic C_{40} diene $C_{40}H_{74}$ (0.25 g), R_f 0.8 (in Et_2O -petrol (1:9) on $AgNO_3$ -treated Si gel), R_i 20.3 min (on 1% OV-1 at 235°); $[\alpha]_D^{20}$ -4.4° (c, 3.0); m/e 554 (M^+ , 9%), 483(2), 470(3), 455(3), 329(3), 291(3), 205(7), 192(18), 191(18), 165(100); ν_m 1654, 1000 and 915 ($-CH=CH_2$) cm^{-1} ; δ (CCl_4) 5.66-5.5 (2H, seven lines, $CHMe-CH=CH_2$), 4.94-4.84 (4H, 15 lines, $CH=CH_2$), 1.73 (7-8H, m), 1.26 (b), 0.97 [6H, d (J = 7.5 Hz), $MeCH-CH=CH_2$], 0.85 [15H, two d (J = 5.5, 6 Hz), $CH-CH_3$]. The diene (0.25 g) was stirred with OsO_4 (0.24 g in 10 ml dry C_5H_5N) at 18° for 5 hr, then aq. $NaHSO_3$ (3 g in 20 ml) and C_5H_5N (3 ml) were added and stirred for 30 min. Extraction with CH_2Cl_2 afforded the bicyclic C_{40} tetrol (0.32 g), ν_m 3350 and 1050 cm^{-1} , δ 3.5 (complex, $CHOH$ and CH_2OH), which without purification was treated with $Pb(OAc)_4$ (0.43 g) in AcOH (8 ml) and MeOH (940 ml) at 18° for 4 hr. Normal work-up gave the C_{38} dialdehyde (0.32 g), ν_m 1710 cm^{-1} , (CCl_4) 9.48 (d (J = 1.5 Hz) $CH-CHO$). The dialdehyde was refluxed for 3 hr in isopropenyl acetate (8 ml) containing conc H_2SO_4 (1 drop); the excess isopropenyl acetate was evapd under red. press. H_2O was added, and the whole extracted with Et_2O . The Et_2O extract, washed with H_2O , aq. Na_2CO_3 ,

and H_2O , dried and evapd, was chromatographed in petrol with up to 10% Et_2O , which eluted the C_{38} di-(enol acetate) (0.13 mg), ν_m 1750 and 1220 cm^{-1} . This was ozonised in $EtOAc$ at -20°; evaporation of the solvent gave the crude C_{36} diketone (0.09 g), ν_m 1715 cm^{-1} ; δ 2.30 [4H, t (J = 7 Hz), CH_2-CH_2-CO], 2.02 (6H, s, CH_3-CO), 0.88 [14-15H, d (J = 5 Hz), $CH-CH_3$]. Part of this diketone (20 mg) with $LiAlH_4$ (in dry Et_2O , 12 hr), worked up in the normal way, and chromatographed in petrol- Et_2O (2:3) gave the bicyclic C_{36} diol (3) (16 mg). Part of this was used for 1H NMR and Eu shift measurements (Table 2) and part was acetylated ($Ac_2O-C_5H_5N$) to give the bicyclic C_{36} diacetate: m/e 618 (M^+ , 2%), 558 (M^+ -AcOH, 2), 498 (M^+ -2AcOH, 427(3), 137(100)); δ (C_6D_6) 4.78 (2H, m $CHOAc$), 1.91 (6H, s, CH_3COO), 1.15 [6H, d (J = 6 Hz), CH_3-CH-O], 0.96 [6H, d (J = 5.5 Hz), $CH-CH_3$ (20,20')], 0.94 [6H, d (J = 5.5 Hz), $CH-CH_3$ (19,19')].

Bicyclic C_{32} diol (4). Using the methods described above, the remainder of the C_{36} diketone (70 mg) treated with isopropenyl acetate gave the C_{36} di-(enol acetate) (20 mg) ν_m 1255 and 1215 cm^{-1} , ozonolysis of which gave the C_{32} dialdehyde (14 mg), ν_m 1710 cm^{-1} , δ 9.45 [t (J = 6 Hz), CH_2-CHO], reduced with $LiAlH_4$ to the bicyclic C_{32} diol (4) (6 mg). Part of this was used for 1H NMR and Eu shift measurements (Table 3) and the remainder acetylated to give the C_{32} diacetate: m/e 562 (M^+ , 2%), 502 (M^+ -AcOH, 2), 442 (M^+ -2AcOH, 2), 399(2), 109(100); δ (C_6D_6) 3.97 [4H, t (J = 7 Hz), CH_2-CH_2-OAc], 1.92 (6H, s CH_3-CO), 0.96 [6H, d (J = 5.5 Hz), $CH-CH_3$ (20,20')], 0.94 [6H, d (J = 5.5 Hz), $CH-CH_3$ (19,19')].

Acknowledgements—The authors thank Messrs. Enrico Esposito and Salvatore Sodano for technical assistance, Mr. Alfredo Milone for GC-MS and Mr. Corrado Di Pinto for NMR measurements at Arco Felice. We also thank Dr. C. J. W. Brooks for chemical ionization MS, and Dr. R. Warrent and Mrs. L. Clough for high-resolution NMR measurements at Manchester.

REFERENCES

1. de Rosa, M., Gambacorta, A., Millonig, G. and Bu'Lock, J. D. (1974) *Experientia* **30**, 860.
2. de Rosa, M., Gambacorta, A. and Bu'Lock, J. D. (1975) *J. Gen. Microbiol.* **86**, 156.
3. Millonig, G., de Rosa, M., Gambacorta, A. and Bu'Lock, J. D. (1975) *J. Gen. Microbiol.* **86**, 165.
4. de Rosa, M., Gambacorta, A. and Bu'Lock, J. D. (1976) *Phytochemistry* **15**, 143.
5. de Rosa, M., Gambacorta, A., Minale, L. and Bu'Lock, J. D. (1974) *Chem. Commun.* 543.
6. Langworthy, T. A. and Mayberry, W. R. (1976) *Soc. Gen. Microbiol. Proc.* **3**, 165.
7. de Rosa, M., de Rosa, S. and Gambacorta, A. (1977) *Phytochemistry* **16**, 1909.
8. Darland, G., Brook, T. D., Samsonoff, W. and Conti, S. F. (1970) *Science* **170**, 1416.
9. Kates, M., Yengoyan, L. S. and Sastry, P. S. (1965) *Biochim. Biophys. Acta* **98**, 252.