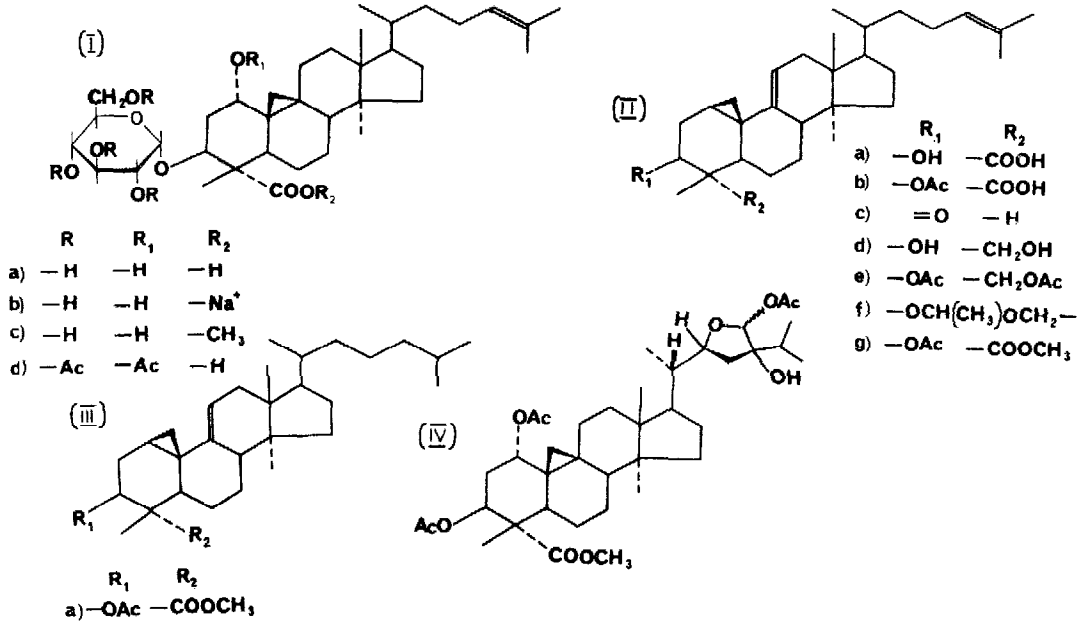
Tetrahedron Letters No. 47, pp 4299 - 4302, 1976. Pergamon Press. Printed in Great Britain.

MOLLIC ACID 3-β-D-GLUCOSIDE, A NOVEL 1α-HYDROXYCYCLOARTANE SAPONIN FROM *COMBRETUM MOLLE* (COMBRETACEAE).

Karl H. Pegel and Colin B. Rogers*
Chemistry Department, University of Natal, Durban, Republic of South Africa.

(Received in UK 20 August 1976; accepted for publication 2 October 1976)

The major constituent of the acetone extract of *Combretum molle* leaves is a colourless, sparingly soluble, crystalline triterpene acid saponin which we have named mollic acid glucoside (Ia), m.p. 248-250° dec., $[\alpha]_D$ + 38° (pyridine), $C_{36}H_{58}O_9$. Titration of this glucoside acid with NaOH shows that it has only one carboxyl function forming a highly water soluble sodium salt (Ib), m.p. 305-307° dec., $C_{36}H_{57}O_9Na$. Mollic acid glucoside readily forms a monomethyl ester (Ic), m.p. 225-226°, $[\alpha]_D$ - 89° (EtOH), $C_{37}H_{60}O_9$, M^+ 648, and a pentaacetate (Id), m.p. 198-200°, $[\alpha]_D$ +22° ($CHCl_3$), $C_{46}H_{68}O_{14}$, M^+ 844, with no hydroxyl bands in the i.r. spectrum.



The ^1H n.m.r. ($CHCl_3$) spectrum of (Id) presents acetoxymethyne signals between δ 5.20 and 4.00, characteristic of the acetylated glucoside moiety. In particular a doublet at δ 4.53 (J8Hz) is assignable to the axially oriented C-1 acetal proton of the glucoside moiety establishing the β-nature of the glucoside link. In addition the mass spectrum of (Id) shows prominent peaks at

m/e 331, 169, 109 and 43 which are characteristic fragments given by glucoside acetates.[1]

A single proton dd signal in the ^{1}H n.m.r. spectrum of (Id) at δ 4.70 (J_1J_2 3Hz) is assigned to an equatorial acetoxymethyne proton at C-1 of the triterpene moiety. We believe that the shielding of this proton by the 9,19-cyclopropane ring accounts for its relatively high field position. In the same spectrum the five acetoxy methyl group signals appear as sharp singlets between δ 2.09 and 1.96; of the six aliphatic methyl groups two signals at δ 1.70 and 1.63 are typical of the side chain isopropylidene methyl groups, a three proton singlet at δ 1.11 is typical of an axial methyl group α to a carboxyl function, while a broad nine proton singlet at δ 0.93 accounts for the three remaining methyl groups. A pair of one proton doublets at δ 0.73 and 0.50 (J4Hz), the former partly obscured by a methyl group signal, indicates the presence of the isolated cyclopropyl methylene group in (Id). The proximity of the C-1 hydroxyl group explains the uncharacteristic downfield shift of these protons.[2] By using C_6D_6 as solvent these two doublets shift upfield to δ 0.54 and 0.17 (J4Hz) at virtually identical positions to those found in the ^{1}H n.m.r. (C_6D_6) spectrum of triacetyl passifloric acid methyl ester (IV).[3] This similarity and the fact that almost all cyclopropane triterpenes are based on a cycloartane skeleton, suggests that the cyclopropane ring is indeed at the C-9,19 position. The acid induced rearrangement described below confirms this placement.

Acid hydrolysis of (Ia) in THF gives a mixture of aglycones plus glucose, which was identified by co-chromatography and enzymatic indicator paper. Acetylation of the aglycone mixture, followed by chromatographic separation yields as the major product the artifact dehydromollic acid acetate (IIb), m.p. 220-222°, $[\alpha]_D$ + 132°, $C_{32}H_{48}O_4$, M^+ 496. The ^{1}H n.m.r. spectrum of (IIb), shows a three proton multiplet at δ 5.15 due to one acetoxymethyne proton, the single olefinic proton at C-24 of the side chain and a second single olefinic proton introduced as a result of the dehydration reaction. The trisubstituted nature of this introduced double bond was confirmed by ^{13}C n.m.r. spectroscopy. Since the dd signal at δ 4.70 due to the equatorial C-1 proton in the ^{1}H n.m.r. spectrum of (Id) is absent in the spectrum of (IIb), the axial 1α-hydroxy function is, by implication, involved in the acid catylised dehydration reaction as observed and reported for (IV).[3] The remaining acetoxy group in (IIb) and hence the hydroxy group in (IIa) is placed at C-3 by virtue of the cumulative evidence presented here and since the presence of an oxygen function at that position is ubiquitous in all triterpenes. By using C_6D_6 as a solvent for (IIb) the ^{1}H n.m.r. multiplet at δ 5.15 is resolved into a dd at δ 5.71 ($J_1$12, $J_2$6Hz) due to the proton at C-3 and two broad doublets at δ 5.34 (J4Hz), and δ 5.07 (J3Hz) due to the two olefinic protons. The large J_1 value obtained for the C-3 proton indicates that it must be axial and therefore the C-3 oxygen function has to be equatorial. The introduced endocyclic double bond in (IIb) has a marked shielding effect on two of the methyl group ^{1}H n.m.r. signals; these occur at δ 0.93 for (Id) but move to δ 0.83 and 0.57 for (IIb) thereby exposing a three proton d (J5Hz) at δ 0.93. This is characteristic of the C-21 side chain methyl group. The characteristic signals due to the cyclopropane methylene protons also undergo a pronounced environmental change with only a one proton signal visible as a dd ($J_1$8Hz, $J_2$4Hz) at δ 0.08.

The double bond introduction on acid hydrolysis of (Ia) and the resultant upfield shift of the two methyl- and the cyclopropyl proton signals in (IIb) can be rationalised on the basis

that a concerted elimination-rearrangement sequence arises from the protonation and elimination of an axial C-1 hydroxyl group. This elimination is followed by β-migration of the cyclopropane group from C-9,19 to C-1,19 with sequential elimination of the axial 11α-hydrogen. This then explains the high-field shift of the two methyl group signals, since a double bond at 9(11) will shield the axial C-18 and C-28 methyl groups and these, incidentally, are the only methyl group signals that had not yet been assigned. In addition, the splitting of the 1,19-cyclopropane, 19-methylene group signal is due to the 19α-proton coupling with the 19β- and 1β-protons, while the shielding effect of the 9(11)-double bond accounts for the upfield shift of the 19α-proton as has been reported by Bombardelli and co-workers.[3]

Confirmation of the position and orientation of the carboxyl function was obtained. The β-hydroxyacid (IIa), m.p. 250-253°, $C_{30}H_{46}O_3$, M^+ 454, prepared by base hydrolysis of (IIb) is oxidised by Jones' reagent to the corresponding β-keto acid, which readily decarboxylates to the nor-ketone (IIc), m.p. 196-203°, $C_{29}H_{44}O$, M^+ 408. Since the hydroxyl-, and keto functions in (IIa) and (IIc) respectively must be at C-3, the eliminated carboxylic acid group has to be at C-4. The ^1H n.m.r. spectrum of the nor-ketone (IIc) no longer has a methyl group signal at δ 1.11. The configuration of the carboxylic acid function was established by reducing the methyl ester of (IIb) with LAH to the diol (IId), m.p. 160-165°, $C_{30}H_{48}O_2$ and acetylating it to the diacetate (IIe), m.p. 167-170°, $C_{34}H_{52}O_4$. The ^1H n.m.r. spectrum of (IIa) shows a two proton singlet at δ 3.83 which is characteristic of a primary equatorial acetoxymethyl group.[4] In addition the diol (IId) forms an ethylidene derivative (IIf), m.p. 160-165°, $C_{32}H_{50}O_2$, which further confirms the 1,3 relationship of the oxygen functions at C-3 and C-4.

The side chain is typically that of cycloartane according to the ^1H n.m.r. evidence outlined so far. This is supported by mass spectral data which shows that all the mollic acid derivatives have fragments at (M^+ - 111) and (M^+ - 111 - 42) due to the loss of the unsaturated C_8H_{15} side chain and side chain plus ring D respectively.[5] On hydrogenation of the side chain double bond, these fragments appear at (M^+ - 113) and (M^+ - 113 - 42). This also provides evidence that no oxygen functions exist on the side chain or on the D ring. Finally, ^{13}C n.m.r. ($CDCl_3$) spectroscopy confirmed the similarity of the mollic acid side chain to that of lanosterol and cycloartenol as presented in the TABLE. A detailed report on the structure determination of mollic acid and related compounds will be presented elsewhere.

TABLE. Carbon-13 Chemical Shifts.[a]

	C-17	C-20	C-21	C-22	C-23	C-24	C-25	C-26	C-27
Cycloartanol[6]	52.5	36.0	18.3	36.4	24.0	39.4	28.0	22.5	22.7
Lanost-8-ene-3β-ol[7]	50.7	36.5	18.8	36.5	24.2	39.6	28.1	22.6	22.8
(IIIa)	52.1	36.6	18.8	36.5	24.1	39.6	28.1	22.6	22.8
Cycloartenol[8]	52.4	36.0	18.0	36.4	25.0	125.3	130.8	17.6	25.7
Lanost-8,24-diene-3β-ol[7]	50.7	36.6	18.8	36.3	25.0	125.3	130.8	17.6	25.7
(IIc)	52.1	36.7	18.8	36.5	25.0	125.3	130.8	17.6	25.7

[a] In parts per million relative to internal TMS.

Acknowledgements.

We thank the C.S.I.R. of the Republic of South Africa and the University of Natal Research Fund for the financial support of this work.

References.

1. I.A. Pearl and S.F. Darling, Phytochemistry, 1968, 7, 831-837.

2. R.H. Takata and P.J. Scheuer, Tetrahedron, 1976, 32, 1077-1080.

3. E. Bombardelli, A. Bonati, B. Gabetta, E.J. Martinelli, G. Mustich and B. Danieli, Phytochemistry, 1975, 14, 2661-65.

4. A. Gaudamer, J. Polonsky and F. Wenkert, Bull. Soc. Chim., 1964, 407.

5. H. Budzikiewicz, C. Djerassi and D.H. Williams, Structural Elucidation of Natural Products by Mass Spectrometry, Volume II, Holden-Day, Inc., San Francisco, California, 1964, 94.

6. F. Khuong-Huu, M. Sangare, V.M. Chari, A. Bekaert, M. Devys, M. Barbier and G. Lukacs, Tetrahedron Letters, 1975, 22 + 23, 1787-90.

7. S.A. Knight, Organic Magnetic Resonance, 1974, 6, 603-611.

8. L. Radics, M. Kajtar-Peredy, S. Corsano and L. Standoli, Tetrahedron Letters, 1975, 48, 4287-4290.