Constitution of Lucumin and Its Related Glycosides from *Calocarpum sapota* MERRILL

Tadahiro TAKEDA,*,[a] Ryoko GONDA,[a] and Keiichiro HATANO[b]

Kyoritsu College of Pharmacy,[a] Shibakoen 1–5–30, Minato-ku, Tokyo 105, Japan and Faculty of Pharmaceutical Sciences, Nagoya City University,[b] Tanabe-dori, Mizuho-ku, Nagoya 467, Japan.

Received October 18, 1996; accepted December 2, 1996

Three aromatic glycosides, lucumin, lucuminamide and lucuminic acid, were isolated from the seeds of *Calocarpum sapota* MERRILL (Sapotaceae). Their structures were established by spectroscopic methods, chemical evidence, and X-ray crystallography.

Key words *Calocarpum sapota*; Sapotaceae; lucumin; lucuminamide; lucuminic acid; crystal structure

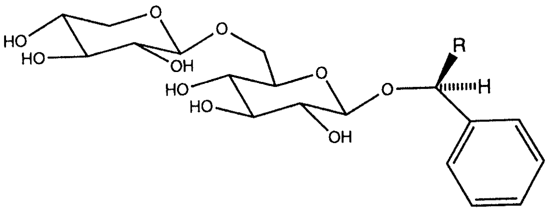
Aqueous extracts of the seeds of *Calocarpum sapota* MERRILL (Sapotaceae) are used as a hairtonic in Mexico.[1] In 1948 Bachstez *et al.*[2] isolated a crystalline nitrogenous glycoside which they termed lucumin from the seeds of *Lucuma mammosa* GAERTN of the same family. At that time they postulated lucumin to be the monohydrate of diarabinoside of benzaldehyde cyanohydrin. In 1971, Eyjolfsson's reinvestigation[3] of this compound resulted in the liberation of D-xylose, D-glucose as the component sugars by complete acid hydrolysis. On the basis of partial acid hydrolysis, and Klyne's rule he postulated lucumin to be 2(*R*)-[6-*O*-(β-D-xylopyranosyl)-β-D-glucopyranosyloxy]-2-phenylacetonitrile. This result is not complete because of poor yield of the partial acid hydrolysis product and vagueness of the absolute configuration.

We studied the seeds of *Calocarpum sapota* with the aim of isolating the biologically active constituents. In this paper we wish to report the structure elucidation of three aromatic glycosides, lucumin, lucuminamide and lucuminic acid. The crushed seeds were extracted with hexane, chloroform, acetone and methanol successively. The methanol extract was partitioned between *n*-butanol and water. The water soluble fraction of the methanol extract was partitioned with chloroform–methanol–water (4 : 4 : 3) and the upper layer was evaporated. The residue was dissolved in water and then was chromatographed on a Diaion HP-20 column using water–methanol (0—100%, linear gradient) as eluate to give six fractions, frs. 1—6. Compound **1** was obtained from fr. 4 by recrystallization from water after repeated porous polymer gel column chromatography. Fraction 3 was further chromatographed on semi-preparative HPLC on an Asahipak ODP-50 column and gave compound **2**. Compound **3** was obtained from fraction 2 by silica gel column chromatography.

Compound **1** was obtained as colorless plates. Compound **1** revealed a quasimolecular peak at *m*/*z* 428 $[M+H]^+$ in the fast atom bombardment mass spectrum (FAB-MS). The molecular formula was determined to be $C_{19}H_{25}NO_{10}$ from high-resolution MS measurement. Complete acid hydrolysis of **1** resulted in the liberation of hydrogen cyanide (color test),[3] benzaldehyde (TLC), and D-glucose and D-xylose as the component sugars (TLC, GLC). A mild acid hydrolysis of **1** with 1 N-HCl containing benzene gave mandelonitrile as an aglycone. The ^{1}H-NMR spectrum of compound **1** exhibited a characteristic methine hydrogen singlet signal of H-7 at δ 6.53, signals due to the anomeric protons of the carbohydrate moieties at δ 5.16 (d, $J=7.9$ Hz, Xyl H-1) and δ 4.90 (d, $J=6.7$ Hz, Glc H-1). Furthermore, signals of five protons appeared in the aromatic region of about 7—8 ppm. In the ^{13}C-NMR spectrum, the carbon signals at δ 119.43 and 67.80 ppm further corroborated the presence of a phenylacetonitrile group. In the heteronuclear multiple bond connectivity (HMBC) spectrum, the anomeric proton of β-D-xylopyranose at δ 5.16 was correlated with the signal at δ 69.95, which was assigned to C-6 of the β-D-glucopyranose. Therefore, it was deduced that the β-D-xylopyranosyl moiety with a signal of an anomeric proton at δ 5.16 ppm was substituted at C-6 of the β-D-glucopyranosyl moiety. From the above data, we concluded that compound **1** was lucumin, which was identical to Eyjolfsson's lucumin (mp, ^{1}H-NMR) by comparison with the literature data. X-ray analysis determined the absolute stereostructure of **1** to be 2(*R*)-[6-*O*-(β-D-xylopyranosyl)-β-D-glucopyranosyloxy]-2-phenylacetonitrile, as shown in Figs. 1 and 2 proving the *R* configuration of the chiral center in the aglycone of lucumin.

Amygdalin[4] is a recognized cyanogenetic glycoside that occurs naturally in the seeds of members of the Rosaceae family such as almonds and cherries. Amygdalin can undergo degradation by hydrolysis of the nitrile group to yield amygdalinic acid *via* amygdalinamide.

Comparison of the ^{1}H-NMR and the ^{13}C-NMR data of compounds **2** and **3** with those of compound **1** (lucumin) showed them to be similar except for the signals of the aglycone moiety. Comparing the ^{1}H-NMR spectra of compounds **1**, **2** and **3**, the aglycone methine hydrogen of compounds **2** and **3** (**2**; δ 5.38 ppm, **3**; δ 5.23 ppm) appeared



1: R = CN
2: R = $CONH_2$
3: R = COOH

Fig. 1. Structure of Compounds **1**—**3**

---

* To whom correspondence should be addressed.

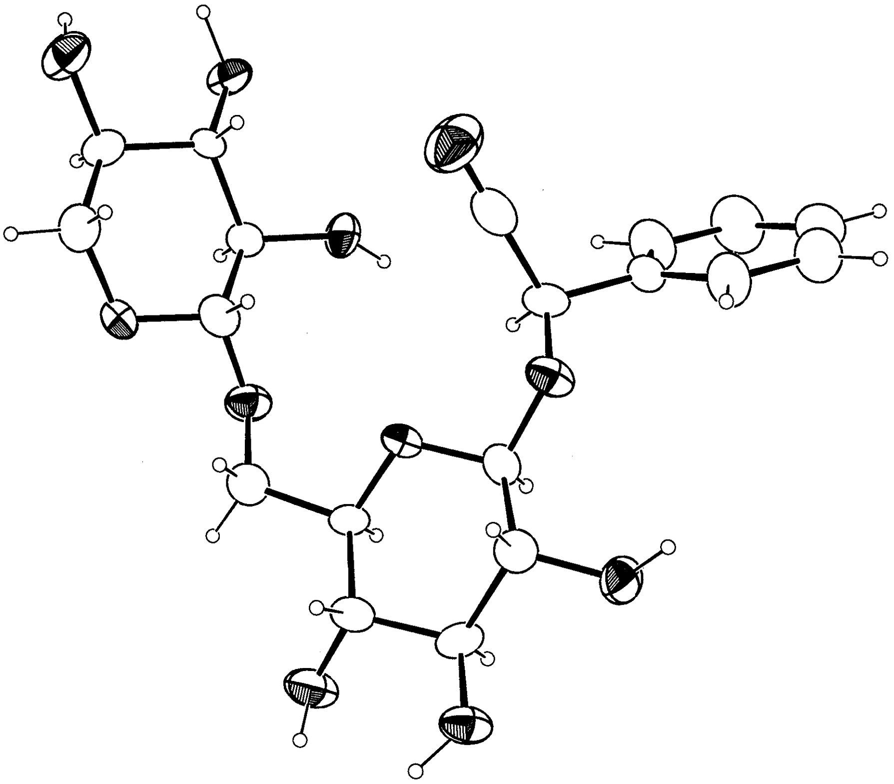


Fig. 2. Crystal Structure of **1**

Octant shaded ellipsoids indicate oxygen and nitrogen atoms. Hydrogens are shown by arbitrary circles.

Table 1. ^{13}C-NMR Spectral Data of Compounds **1**—**3**[a]

Carbon	**1**	**2**	**3**
Aglycon			
1	134.53	137.81	139.46
2	128.22	130.90	130.87
3	129.36	131.74	131.53
4	129.96	137.81	131.53
5	129.36	131.74	131.53
6	128.22	130.90	130.87
7	67.80	81.30	82.44
8	119.43	178.35	180.26
Sugar			
Glc			
1	102.57	101.69	101.16
2	74.71	75.49	75.54
3	78.29	75.49	77.49
4	71.53	72.07	72.16
5	78.29	78.39	78.34
6	69.95	71.17	71.14
Xyl			
1	106.21	106.32	106.27
2	75.20	75.71	75.71
3	78.00	77.65	77.95
4	71.22	71.91	71.92
5	67.14	67.83	67.83

[a] The spectra of compound **1** was measured in C_5D_5N and that of compounds **2** and **3** in D_2O.

at higher field than **1** (**1**; δ 6.53 ppm).

FAB-MS spectrum of compound **2** indicated a quasimolecular peak at 468 $[M+Na]^+$. The molecular formula of **2** was determined to be $C_{19}H_{27}NO_{11}$ from high-resolution MS measurement. The IR spectrum indicated two absorption bands of carbonylamide groups at 1676 cm^{-1} in addition to a broad absorption band at 3383 cm^{-1}.

Compound **3** revealed a quasimolecular peak at m/z 469 $[M+Na]^+$ in the FAB-MS, and the molecular formula was determined to be $C_{19}H_{26}O_{12}$ from high-resolution MS measurement. The IR spectrum indicated an absorption band characteristic of carboxylic acid at 1739 cm^{-1} instead of compound **1** which had a very weak absorption band at 2262 cm^{-1} characteristic of the nitrile group.

Thus, the structure of **2** was established as shown in Fig. 1, and it was called lucuminamide; similarly, the structure of **3** called lucuminic acid was estimated as indicated in the figure. Compounds **2** and **3** would be derived from compound **1**. The identity of both derived lucuminamide and lucuminic acid was confirmed by ^{1}H- and ^{13}C-NMR spectra and HPLC.

This is therefore the first report of the isolation of **1** from *Calocarpum sapota*. Compounds **2** and **3** are the first cyanogenic glycoside related compounds.[5] There has been no such precedent before. The possibilities of artifacts for these compounds **2** and **3** formed during storage or in the extraction or isolation procedure may not be entirely negligible, because we cannot examine whether these compounds exist in the fresh seed owing to the difficulty in the collection of this plant. HPLC and TLC analyses showed the presence of compounds **2** and **3** in the MeOH extract of the dried seed prepared by extraction after only 3 h at room temperature. This fact at least ruled out the possibility of formation of these compounds during extraction or isolation. Judging from these results, it is probable that compounds **2** and **3** are the genuine constituents of *Calocarpum sapota*.

Experimental

General Procedures The melting points were determined on a Yanagimoto microapparatus and are uncorrected. ^{1}H- and ^{13}C-NMR spectra, with tetramethylsilane (TMS) and/or sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as internal standard were recorded on a JEOL FT-NMR A500. FAB-MS was measured on a JEOL JMS-SX 102. GLC was run on a Shimadzu GC-14B gas chromatograph. IR spectra (on KBr pellets) were measured on a JASCO FT/IR-5300 spectrometer. TLC was conducted on precoated silica gel plates (Merck $60F_{254}$ and RP-18). Column chromatography was carried out on silica gel (Merck Kiesel gel 60) and a porous polymer gel (Mitsubishi Diaion HP-20). HPLC was carried out on a column of Asahipak ODP-50.

Materials Seeds of *Calocarpum sapota* were collected in Morelos, Mexico in 1987 and voucher specimens were deposited in the Jardin de Etno-botanico, Instituto Nacional de Antropologia e Historia. The material was identified as *C. sapota* Merrill by Dr. G. S. Ortega (Botanical Garden Director at Jardin de Etnobotanica, Morelos, Mexico).

Extraction and Isolation Crushed seeds except the shell (inteeguments) of *Calocarpum sapota* (500 g) were extracted successively with hexane, chloroform, acetone and methanol (1 l × 2, 6 h in each) under reflux. The methanol extract was concentrated under reduced pressure and the residue (52.6 g) was suspended in water (500 ml). The suspension was extracted with *n*-butanol (500 ml) and then the water soluble fraction was concentrated *in vacuo* to give a residue (36.7 g). This residue was partitioned with chloroform–methanol–water (4 : 4 : 3) (550 ml) and the upper layer was concentrated *in vacuo* to give a residue (32.3 g). The residue was dissolved in water and then was chromatographed on a porous polymer gel Mitsubishi Diaion HP-20 column. After washing the column with water, the absorbed materials were eluted with water–methanol (0→100% linear gradient) to give six fractions (frs. 1—6). Fraction 4 was further chromatographed on porous polymer gel HP-20 to give a compound **1** rich fraction, which was crystallized from water and gave compound **1** (400 mg). Fraction 3 was chromatographed on semi-preparative HPLC (Asahipak ODP-50, 10 × 250 mm: 10% acetonitrile) and gave compound **2** (630 mg). Fraction 2 was repeatedly chromatographed over silica gel ($CHCl_3$–MeOH–H_2O, 9 : 1 : 0.1) to afford compound **3** (670 mg). HPLC: t_R value; Asahipak ODP-50 (10% CH_3CN) **1**: 8.69, **2**: 3.54, **3**: 2.53; TLC: *Rf* value; RP-18 (CH_3CN : H_2O = 1 : 2) **1**: 0.49, **2**: 0.71, **3**: 0.80. $60F_{254}$ (AcOEt : MeOH : H_2O : AcOH = 11 : 3 : 1 : 1) **1**: 0.49, **2**: 0.29, **3**: 0.11.

Acid Hydrolysis of Compound 1, 2 and 3 a) Determination of Sugar Species: Each sample (1 mg) was heated at 100 °C with 2 N H_2SO_4 (0.2 ml) for 6 h. The reaction mixture was neutralized with Dowex 2 (OH^-) (Dow Chemical Co.) and filtrated. The residue was dissolved in small volumes of water and the solution was subjected to cellulose TLC analysis (developing solvent: *n*-butanol–pyridine–water, 6 : 4 : 3), together with authentic samples (glucose and xylose). Spots were visualized with 3% *p*-anisidine reagent. Glucose and xylose were detected. *Rf*: glucose 0.37, xylose 0.50. Determination of their absolute configuration was made according to the method reported by Hara *et al.*[6] using GLC. GLC conditions: column; 3% ECNSS-M (0.3 mm × 2 m); column temperature 190 °C; injection temperature 210 °C; retention times (min): xylose 19.3, glucose 49.2. The remaining residue was dissolved in water (1 ml) and then reduced with $NaBH_4$ (*ca.* 5 mg) for 2 h at room temperature. After neutralization with Dowex 50W-X8 (H^+) (Dow Chemical Co.), filtrated and concentrated to dryness. Boric acid was removed by repeated addition and evaporation of methanol. The product was acetylated with acetic anhydride and pyridine (1 ml each) at 100 °C for 1 h. The reagents were evaporated and the residual alditol acetates were analyzed by GLC. From **1**, **2** and **3**, xylitol acetate and glucitol acetate were detected in the ratio of 1 : 1. GLC conditions: column; Supelco SP-2380 fused silica capillary column (0.53 mm × 15 m) (Spelco Co.) ; column temperature 150→220 °C, 4 °C/min; injection temperature 250 °C; retention times (min): glucitol acetate 9.93, xylitol acetate 6.65.

b) Determination of Aglycon: Compound **1** (100.7 mg) was dissolved in 20 ml of 1 N HCl and 10 ml of benzene. The solution was refluxed for 6 h, then organic layer was evaporated and gave mandelonitrile (21.1 mg). ^{1}H-NMR ($CDCl_3$) δ: 7.52—7.50 (2H, m, H-3,5), 7.44—7.42 (3H, m, H-2,4,6), 3.23 (1H, s, OH), 5.51 (1H, s, H-7). ^{13}C-NMR ($CDCl_3$) δ: 135.23 (C-1), 129.88 (C-4), 129.22 (C-3,5), 126.67 (C-2,6), 118.81 (C-8), 63.63 (C-7).

Lucumin (1): [6-*O*-(β-D-Xylopyranosyl)-β-D-glucopyranosyloxy]-2-phenylacetonitrile Colorless plates, mp 186 °C. High-resolution positive-mode FAB-MS: Calcd for $C_{19}H_{26}NO_{10}$ $(M+H)^+$: 428.4151; Found: 428.4174. IR $\nu_{\max}^{\text{KBr}}$ cm^{-1}: 3443, 2262. ^{1}H-NMR (pyridine-d_5) δ: 7.85—7.83 (2H, m, H-3,5), 7.33—7.32 (3H, m, H-2,4,6), 6.53 (1H, s, H-7), 5.16 (1H, d, J=7.9 Hz, Xyl H-1), 4.90 (1H, d, J=6.7 Hz, Glc H-1), 4.84 (1H, d, J=10.4 Hz, Glc H-6a), 4.36—4.30 (2H, m, Xyl H-5a, Glc H-6b), 4.26—4.24 (2H, m, Xyl H-2,4), 4.08—4.05 (5H, m, Glc H-2,3,4,5, Xyl H-3), 3.78—3.74 (1H, m, Xyl H-5b). ^{13}C-NMR (pyridine-d_5): Table 1.

X-Ray Analysis of **1**: Crystal Data; $C_{19}H_{25}NO_{10}\cdot H_2O$, $M_r = 445.4$, orthorhombic, $P2_12_12_1$, $a = 5.073(1)$ Å, $b = 17.155(4)$ Å, $c = 24.328(5)$ Å, $V = 2117.1(13)$ A^3, $Z = 4$, $D_c = 1.397$ gcm^{-3}, $\mu = 1.08$ cm^{-1}, $F(000) = 944$. A colorless plate crystal with the dimensions 0.06 × 0.15 × 0.36 mm, which was obtained from water solution, was used for X-ray measurements at 293 K on an Enraf-Nonius CAD4 diffractometer equipped with graphite-monochromated Mo*Kα* radiation ($\lambda = 0.71073$ A). Cell constants were determined from 25 reflections in the range of $13° < 2\theta < 24°$. The total number of independent reflections measured was 2549 of which 1476 were believed to be observed $[F_0 > 3.0\sigma(F_0)]$. No absorption correction was applied. The structure was solved by the direct method and refined by full-matrix least-squares, with anisotropic temperature factors for non-H atoms and anisotropic equivalent thermal factors for H atoms. The refinement converged to $R = 0.057$, $R_w = 0.058$.

Lucuminamide (2): [6-*O*-(β-D-Xylopyranosyl)-β-D-glucopyranosyloxy]-2-phenylacetamide A white powder. High-resolution positive-mode FAB-MS: Calcd for $C_{19}H_{27}NNaO_{11}$ $(M+Na)^+$: 468.4122; Found: 468.4197. IR $\nu_{\max}^{\text{KBr}}$ cm^{-1}: 3383, 1676. ^{1}H-NMR (D_2O) δ: 7.48 (5H, br s, arom.), 5.38 (1H, s, H-7), 4.46 (1H, d, J=7.9 Hz, Xyl H-1), 4.27 (1H, d, J=7.9 Hz, Glc H-1), 4.14 (1H, d, J=11.6 Hz, Glc H-6a), 3.96 (1H, dd, J=5.5, 5.5 Hz, Xyl H-5a), 3.85 (1H, dd, J=2.5, 3.7 Hz, Glc H-6b), 3.67—3.62 (1H, m, Xyl H-4), 3.47—3.43 (4H, m, Glc H-4,5, Xyl H-3), 3.35—3.31 (3H, m, Glc H-3, Xyl H-2,5b). ^{13}C-NMR (D_2O): Table 1.

Lucuminic Acid (3): [6-*O*-(β-D-Xylopyranosyl)-β-D-glucopyranosyloxy]-2-phenylcarboxylic Acid A white powder. High-resolution positive-mode FAB-MS: Calcd for $C_{19}H_{26}NaO_{12}$ $(M+Na)^+$: 469.1397; Found: 469.1339. IR $\nu_{\max}^{\text{KBr}}$ cm^{-1}: 3385, 1739. ^{1}H-NMR (D_2O) δ: 7.48—7.42 (5H, m, arom.), 5.23 (1H, s, H-7), 4.46 (1H, d, J=7.3 Hz, Xyl H-1), 4.16 (1H, d, J=8.0 Hz, Glc H-1), 4.12 (1H, d, J=11.6 Hz, Glc H-6a), 3.96 (1H, dd, J=11.6, 5.5 Hz, Xyl H-5a), 3.83 (1H, dd, J=11.6, 5.5 Hz, Glc H-6b), 3.67—3.62 (1H, m, Xyl H-4), 3.48—3.44 (4H, m, Glc H-2,3,4,5), 3.41—3.29 (3H, m, Xyl H-2,3,5b). ^{13}C-NMR (D_2O): Table 1.

Conversion of Compound 1 into Compound 2 and 3 Dissolve 25 mg of compound **1** in 2 ml concentrated $NH_3\cdot H_2O$ contained in a 5 ml screw-cap vial equipped with a stopper. Seal the vial and suspend in a boiling water bath for 60 min. Cool the vessel, unscrew the cap, raise the stopper, and evaporate to give compound **2** (20 mg).

Compound **3** (18 mg) was prepared according to the procedure described by Viehoever and Mack[7] using saturated barium hydroxide solution (compound **1**: 25 mg in 2 ml). The solution was boiled gently for 60 min until ammonia no longer evolved.

Acknowledgements The authors thank Dr. K. Kojima for measurements of FAB-MS at Nagoya City University, Nagoya, Japan. Financial support of this work by Mitsui Pharmaceuticals Inc. is gratefully acknowledged. The work was also supported by a Grant-in-Aid for Encouragement of Young Scientists, No. 08772033 from the Ministry of Education, Science, Sports and Culture of Japan.

References

1) Diaz J. L., "Index y Sinonoma de las Plantas Medicinales de Mexico," Instituto Mexicano para el Estudio de las Plantas Medicinales, Mexico City, 1976, p.16.
2) Bachstez M., Santisteban Prieto E., Canales Gaja A. M., *Ciencia* (Mexico City), **9**, 200—203 (1948).
3) Eyjolfsson R., *Acta Chem. Scand.*, **25**, 1898—1900 (1971).
4) Turczan J. W., Medwick T., *J. Assoc. Off. Anal. Chem.*, **62**, 190—196 (1979).
5) Seigler D. S., *Phytochemistry*, **14**, 9—29 (1975).
6) Hara S., Okabe H., Mihashi K., *Chem. Pharm. Bull.*, **35**, 501—506 (1987).
7) Viehoever A., Mack H., *Am. J. Pharm.*, **107**, 397—450 (1935).