New Phenolic Constituents from the Fruit Juice of *Phyllanthus emblica*

Ying-Jun Zhang, Takashi Tanaka, Chong-Ren Yang, and Isao Kouno

Faculty of Pharmaceutical Sciences, Nagasaki University; Bunkyo Machi 1-14, Nagasaki 852-8521, Japan and Laboratory of Phytochemistry, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, P. R. China

Received October 25, 2000; accepted January 18, 2001

Six new phenolic constituents, L-malic acid 2-O- (1), mucic acid 2-O- (5), mucic acid 1,4-lactone 2-O- (6), 5-O- (8), 3-O- (10), and 3,5-di-O- (11) gallates, were isolated from the fruit juice of *Phyllanthus emblica* together with their methyl esters (2—4, 7, 9), and their structures were determined by spectral and chemical methods. Compounds 5, 6, and 8, the major phenolic constituents of the juice, were present as an equilibrium mixture in aqueous solution.

Key words *Phyllanthus emblica*; Euphorbiaceae; organic acid gallate; mucic acid

*Phyllanthus emblica* L. (Euphorbiaceae) is a shrub or tree growing in subtropical and tropical areas of the People’s Republic of China, India, Indonesia, and the Malay Peninsula. The whole plant, especially the fruit has been used as an anti-inflammatory and antipyretic drug in many local traditional medicines: Chinese herbal medicine, Tibetan medicine, and Ayurvedic medicine. Several reports about this plant have shown that the fruits are rich in vitamin C, mucic acid, and tannins. In our previous work, a novel highly oxygenated norbisabolane together with its methyl ester and three ester glycosides were isolated from the roots of the plant along with 15 tannins and related compounds. As part of our continuing studies on the plant, we report herein the isolation and structural elucidation of new phenolic constituents, L-malic acid 2-O- (1), mucic acid 2-O- (5), mucic acid 1,4-lactone 2-O- (6), 5-O- (8), 3-O- (10), and 3,5-di-O- (11) gallates, and their methyl esters (2—4, 7, 9) from the fruit juice of *P. emblica*.

Results and Discussion

A 60% aqueous acetone extract of the powdered form of fruit juice of *P. emblica* was subjected to MCI gel CHP 20P column chromatography to afford five fractions. Fraction 1 obtained by elution with H2O was further chromatographed successively over Sephadex LH-20, MCI gel CHP20P, Toyopearl HW-40F, and Cosmosil 75C18 OPN to afford compounds 1—11 and three known compounds. By comparison of the physical and spectral data with those of authentic samples, the known compounds were identified as gallic acid, chebulic acid, and 1-O-galloyl-β-D-glucose, respectively.

Compound 1 was obtained as a white amorphous powder and showed a dark blue coloration with ferric chloride reagent. Its molecular formula was assigned as C11H10O9 on the basis of the C13-NMR spectral data, the negative-ion FAB-MS (m/z 285, (M−H)−), and elemental analysis. The C13-NMR spectrum of 1 showed signals due to two carboxyl carbons at δ 170.9 and 170.6, an oxygen bearing methine at δ 69.4 and a methylene at δ 36.5, along with a set of carbons arising from a galloyl group. Hydrolysis of the galloyl ester with tannase yielded gallic acid and t-malic acid ([α]D −2.79°). In addition, the appearance of the methine proton signal of the malic acid moiety at low field (δ 5.60) indicated the galloyl group was attached to this position. Hence, the structure of 1 was determined to be t-malic acid 2-O-gallate.

The molecular formula of compound 2 was shown to be C15H18O12 based on the C13-NMR spectral data, the positive-ion FAB-MS (m/z 391, (M+H)+), and elemental analysis. The 1H- 13C-NMR, and heteronuclear single quantum coherence (HSQC) spectra of 2 indicated the presence of two carboxyl carbons (δ 170.6, 174.8), four oxygen bearing methines [δ 73.4, 71.9, 71.0, 70.6, corresponding to C-1, C-6, C-2, C-3, respectively], and two methoxyl groups [δ 52.7, 52.5, corresponding to δ 53.4, 53.2 (each 3H, s), and a galloyl group [δ 75.4 (2H, s)]. In the heteronuclear multiple bond connectivity (HMBC) spectrum, the two methoxyl groups (δ 37.3, 37.5) were correlated with the two carboxyl carbons at δ 170.6 (C-1) and δ 174.8 (C-6), respectively.

![Chart 1](image-url)
respectively. In addition, correlations of a doublet methine proton at $\delta$ 5.55 (H-2) with the C-1 and the carboxyl carbon ($\delta$ 166.7) of galloyl group indicated that 2 was a 2-O-gallate of 2,3,4,5-tetrahydroxyladipic acid (mucic acid) dimethyl ester. The dimethyl ester was obtained by hydrolysis of 2 with tannase and identified as mucic acid dimethyl ester by direct comparison with an authentic sample. Accordingly, the structure of 2 was established as mucic acid dimethyl ester 2-O-gallate.

Compounds 3 and 4 showed the same (M+H)$^+$ ion peak at $m/z$ 377 in the positive-ion FAB-MS, which was 14 mass units less than that of 2 and suggested the lack of a methoxyl group. The $^1$H- and $^{13}$C-NMR spectral data of 3 and 4 closely resembled those of 2, except for the absence of one of the two methoxyl signals. The presence of a mucic acid core in their molecules was confirmed by tannase hydrolysis followed by methylation affording mucic acid dimethyl ester. In the HMBC spectra of 3 and 4, H-2 ($\delta$ 5.58 for 3, $\delta$ 5.55 for 4) was correlated with C-1 ($\delta$ 170.6 for 3, $\delta$ 170.2 for 4) and the carboxyl carbon of galloyl group ($\delta$ 166.2 for 3, $\delta$ 166.3 for 4). However, the methoxyl proton ($\delta$ 3.71) of 3 was correlated with the C-6 ($\delta$ 174.6), while the methoxyl proton ($\delta$ 3.73) of 4 was coupled with the C-1 ($\delta$ 170.2). This observation indicated that these two compounds were positional isomers differing in the location of their methoxyl group. Hence, the structures of 3 and 4 were established as mucic acid 6-methyl ester 2-O-gallate (3) and mucic acid 1-methyl ester 2-O-gallate (4), respectively.

The molecular formula of compound 5 was assigned as $C_{13}H_{12}O_{12}$ based on the $^{13}$C-NMR spectral data, the negative-ion FAB-MS [$m/z$ 361, (M−H$^-$)], and elemental analysis. Chemical shifts and coupling patterns of the $^1$H-NMR spectrum of 5 as well as its $^{13}$C-NMR spectral data were very similar to those of 2−4, except for the absence of methyl signals, indicating that this compound is the free acid form of 2−4. Thus, the structure of 5 was determined as mucic acid 2-O-gallate.

Compound 6 was obtained as a white amorphous powder. Its $^1$H-, $^{13}$C-NMR, and HSQC spectra revealed the presence of two carboxyl groups ($\delta$ 171.0, 173.0), four oxygen-bearing methines ($\delta_c$ 82.0, 75.5, 72.4, 68.2, corresponding to $\delta_{\text{H}}$ 4.76 (dd, $J = 1.8, 8.4$ Hz), 5.97 (d, $J = 8.7$ Hz), 4.87 (dd, $J = 8.7, 8.4$ Hz), 4.49 (d, $J = 1.8$ Hz), respectively), and a galloyl group [$\delta_{\text{H}}$ 7.18 (2H, s)]. The composition of 6 was similar to that of 5; however, their chemical shifts and coupling constants were significantly different. The core alcohol of 6 was shown to be mucic acid because it became an equilibrium mixture with 5 in aqueous solution and even in the NMR sample tube (acetone-$d_6$, $D_2$O). The positive-ion FAB-MS of 6 showed the (M+H)$^+$ ion peak at $m/z$ 345, which was 18 mass units less than that of 5, indicating that 6 is a lactone form of 5. Among the above-mentioned four methine protons, two doublet signals at $\delta$ 5.97 and 4.49 were assignable to H-2 or H-5 of the mucic acid core. In the HMBC spectrum, one of the doublet signals at $\delta$ 5.97 was correlated with two carboxyl signals at $\delta$ 171.0 and 166.5, the latter being assignable to the carboxyl carbon of the galloyl group, indicating that the galloyl group was located at this position. Furthermore, the chemical shift of the other doublet signal indicated that this position was not acylated. Since the lactone formation usually occurs at $\gamma$ (1,4-lactone) or $\delta$ (1,5-lactone) position, this observation suggested that the mucic acid moiety in 6 existed as a 1,4-lactone form. Although locations of the two esters could not be determined by the HMBC experiment, this was achieved by comparison of the $^1$H-NMR spectrum with that of $d$-saccharic acid 1,4-lactone (12), which is a C-4 epimer of mucic acid with the same relative configurations at C-2 and C-3 positions. The H-2 signal of 12 appeared as a doublet signal with large coupling constant ($J_{5,6}=8.7$ Hz, $J_{3,4}=7.8$ Hz, $J_{5,3}=2.8$ Hz), and the $J$ value coincided with that observed for the methine signal at $\delta$ 5.97 of 6 where the galloyl group was located. This indicated that the galloyl group of 6 was attached to the hydroxyl group adjacent to the lactone carbonyl carbon. Based on the above evidence, the structure of 6 was concluded to be mucic acid 1,4-lactone 2-O-gallate.

The $^1$H-NMR spectrum of compound 7 was almost superimposable on that of 6, except for the appearance of an additional methyl signal at $\delta$ 3.81(s). The positive-ion FAB-MS of 7 showed the (M+H)$^+$ ion peak at $m/z$ 359, which was 14 mass units more than that of 6, indicating the occurrence of a methyl group in 7. Accordingly, 7 was determined as mucic acid 1,4-lactone 6-methyl ester 2-O-gallate.

Compound 8 was obtained as a white amorphous powder. Its molecular composition $C_{13}H_{12}O_{13}$ was identical to that of 6. Coupling patterns of four methine signals observed in the $^1$H-NMR spectrum were similar to those of 6; however, the H-5 of 8 resonated at lower field [$\delta$ 5.57 (d, $J = 2.1$ Hz)] compared with that of 6 [$\delta_{\text{H}}$ 4.49 (d, $J = 1.8$ Hz)], and instead, the H-2 was shifted to higher field [$\delta$ 4.71 (d, $J = 8.7$ Hz), $\delta$ 5.97 (d, $J = 8.7$ Hz) for 8] for 8. This observation clearly indicated that the galloyl group was attached to C-5 in 8, which was also supported by the lower field shift of C-5 ($\delta$ 74.4) and the upper field shift of C-2 ($\delta$ 74.0) compared to those of 6 [$\delta$ 68.2 (C-5), $\delta$ 75.5 (C-2)]. According to the above evidence, the structure of 8 was established as mucic acid 1,4-lactone 5-O-gallate.

Compound 9 exhibited the (M+H)$^+$ ion peak at $m/z$ 359, and $^1$H-NMR spectrum was very similar to that of 8 except for occurrence of an additional methyl signal at $\delta$ 3.80, suggesting that 9 was the methyl ester of 8. Therefore, 9 was determined to be mucic acid 1,4-lactone 6-methyl ester 5-O-gallate.

Compound 10 was found to be an isomer of 6 and 8 by analysis of the $^{13}$C-NMR spectral data and the result of positive-ion FAB-MS [$m/z$ 343, (M−H$^-$)]. The $^1$H-NMR spectrum showed a large downfield shift of H-3 [$\delta$ 5.73 (d, $J = 7.2, 6.6$ Hz)] instead of H-2 of 6, indicating that the galloyl group was attached to C-3 position in 10. Thus, the structure of 10 was established as mucic acid 1,4-lactone 3-O-gallate.

Compound 11 was obtained as a white amorphous powder, and its $^1$H- and $^{13}$C-NMR spectra revealed the presence of two galloyl groups ($\delta$ 7.23, 7.17), and a mucic acid core. This was supported by the positive-ion FAB-MS showing the (M−H$^-$) peak at $m/z$ 495. In the $^1$H-NMR spectrum, the signals due to H-3 and H-5 were observed at the lower field [$\delta$ 5.62 (dd, $J = 7.5, 6.9$ Hz) and $\delta$ 5.65 (d, $J = 1.8$ Hz), respectively], while H-2 and H-4 signals remained at the upper field [$\delta$ 5.10 (d, $J = 7.5$ Hz), $\delta$ 5.14 (dd, $J = 1.8, 6.9$ Hz)]. Hence, the structure of 11 was determined to be mucic acid 1,4-lactone 3,5-di-O-gallate.
These mucic acid gallates were unstable in the aqueous solution and became an equilibrium mixture. Actually, treatment of 2 with H2O-MeOH (4:1) at room temperature for 36 h afforded compounds 2—9, and 5 also gave a mixture of compounds 2—9 under similar conditions, confirming the equilibrium between these compounds (Fig. 1). Taking the equilibrium into account, the fact that these gallates are optically active implied the occurrence of an enantiomeric speciation at C-2 position of the optically inactive mucic acid in *P. emblica*. Interestingly, the mucic acid core of 8 corresponds to the enantiomer of 6.

Most of the methyl esters were probably artifacts generated during the separation; however, direct analysis of the aqueous solution was acidified to pH 2. Repeated chromatography of the mixture over Sephadex LH-20 and Cosmosil 75C18-OPN column gave compounds 1 (230 mg), 5 (400 mg), 4 (114 mg), 6 (122 mg), 7 (10 mg), 8 (141 mg), 9 (44 mg), 10 (15 mg), and chebulic acid (35 mg).

**1-Malic Acid 2-O-Gallate (1)**
White amorphous powder, [α]D **22** +51.0° (c = 0.41, MeOH). 1H-NMR (acetone-**d**6, 500 MHz): δ 3.75 (3H, s, 6-OCH3), 3.79 (3H, s, 1-OCH3), 4.40 (1H, dd, J = 2.0, 10.0 Hz, H-3), 4.62 (1H, dd, J = 1.5, 15.5 Hz, H-5), 5.55 (1H, d, J = 2.0 Hz, H-2), 7.24 (2H, s, galloyl H-2', 6'). 13C-NMR (acetone-**d**6, 125 MHz): δ 43.9 (C-3), 68.7 (C-2), 172.9, 174.0 (COOH). FAB-MS m/z: 285 (M+Na) **1**.

**Mucic Acid Dimethyl Ester 2-O-Gallate (2)**
White amorphous powder, [α]D **22** +50.0° (c = 0.41, MeOH). 1H-NMR (acetone-**d**6, 300 MHz): δ 3.75 (3H, s, OCH3), 3.79 (3H, s, OCH3), 4.10 (1H, dd, J = 1.5, 10.0 Hz, H-4), 4.40 (1H, dd, J = 2.0, 10.0 Hz, H-3), 4.62 (1H, dd, J = 1.5, 15.5 Hz, H-5), 5.55 (1H, d, J = 2.0 Hz, H-2), 7.24 (2H, s, galloyl H-2', 6'). 13C-NMR (acetone-**d**6, 125 MHz): δ 37.5 (C-3), 68.7 (C-2), 172.9, 174.0 (COOH). FAB-MS m/z: 291 (M+H) **2**.

**Mucic Acid 1-Methyl Ester 2-O-Gallate (3)**
White amorphous powder, [α]D **22** +43.9° (c = 0.08, MeOH). 1H-NMR (acetone-**d**6, 500 MHz): δ 3.71 (3H, s, OCH3), 4.08 (1H, dd, J = 1.5, 10.0 Hz, H-4), 4.60 (1H, dd, J = 1.5, 15.5 Hz, H-5), 4.44 (1H, dd, J = 2.0, 10.0 Hz, H-2), 7.24 (2H, s, galloyl H-2', 6'). 13C-NMR (acetone-**d**6, 125 MHz): δ 37.5 (C-3), 68.7 (C-2), 172.9, 174.0 (COOH). FAB-MS m/z: 343 (M+Na) **3**.

**Mucic Acid 6-Methyl Ester 2-O-Gallate (4)**
White amorphous powder, [α]D **22** +38.1° (c = 0.22, MeOH). 1H-NMR (acetone-**d**6, 500 MHz): δ 3.73 (3H, s, OCH3), 4.12 (1H, brd, J = 10.0 Hz, H-4), 4.36 (1H, dd, J = 2.0, 10.0 Hz, H-3), 4.57 (1H, d, J = 1.0 Hz, H-5), 5.55 (1H, d, J = 2.0 Hz, H-2), 7.23 (2H, s, galloyl H-2', 6'). 13C-NMR (acetone-**d**6, 125 MHz): δ 52.3 (OCH3), 71.3 (C-5), 71.4 (C-3), 72.5 (C-4), 73.1 (C-2), 110.3 (galloyl C-2', 6'), 121.4 (C-1'), 138.9 (C-4'), 145.9 (C-3', 5'), 166.2 (C-7'), 170.6 (C-1'), 174.6 (C-6). FAB-MS m/z: 391 (M+Na) **4**.

**Hydrolysis of 4 with Tannase**
A solution of 4 (25 mg) in water (1 ml) was incubated with tannase (2 mg) at room temperature overnight. The reaction mixture was subjected to Chromatorex ODS column chromatography with H2O-MeOH (1 : 0 — 0 : 1) to give compounds 5 (230 mg), 6 (400 mg), 7 (114 mg), 8 (122 mg), 9 (44 mg), 10 (15 mg), and chebulic acid (35 mg).
−25.3° (c = 0.28, H2O). 1H-NMR (D2O, 300 MHz) δ: 4.08 (1H, br d, J = 9.3 Hz, H-4), 4.29 (1H, br d, J = 9.3 Hz, H-3), 4.52 (1H, br s, H-5), 5.51 (1H, s, H-2), 7.31 (2H, s, galloyl H-2', 6'). 13C-NMR (D2O, 75 MHz) δ: 70.4, 71.6 (C-3, 4, 5), 74.3 (C-2), 110.4 (galloyl C-2', 6'), 120.2 (C-1'), 139.9 (C-4'), 145.0 (C-3', 5'), 167.6 (C-7'), 172.2, 174.2 (C-1, 6). FAB-MS m/z: 361 (M−H)

Anal. Caled for C13H12O11 · 3/2H2O: C, 42.06; H, 4.07. Found: C, 42.30; H, 4.28.

Mucic Acid 1.4-Lactone 3,5-Di-O-gallate (11) White amorphous powder, [α]D25 25 96.5° (c = 0.20, MeOH). 1H-NMR (acetone-d6, 300 MHz) δ: 5.10 (1H, d, J = 7.5 Hz, H-2), 5.14 (1H, dd, J = 1.8, 6.9 Hz, H-4), 5.62 (1H, dd, J = 7.5, 6.9 Hz, H-3), 5.65 (1H, d, J = 1.8 Hz, H-5), 7.17, 7.23 (each 2H, s, galloyl H-2', 2', 6', 6'). 13C-NMR (acetone-d6, 75 MHz) δ: 71.1 (C-5), 72.4 (C-2), 76.3 (C-3), 79.1 (C-4), 110.1 (galloyl C-2', 2', 6', 6'), 120.2, 119.8 (C-1', 1'), 138.3 (C-3', 4', 5'), 146.0 (C-3', 3', 5', 5'), 166.6, 165.6 (C-7', 7'), 168.2 (C-6), 172.9 (C-1). FAB-MS m/z: 495 (M−H)

Anal. Caled for C20H16O15 · 5/2H2O: C, 44.37; H, 3.91. Found: C, 44.23; H, 4.22.

Equilibrium of Compounds 2—9 A solution of 2 (150 mg) in H2O/MeOH (4:1, 10 ml) was left at room temperature for 36 h. Repeated chromatography of the mixture over Sephadex LH-20 and Cosmosil 75C18 OPN column afforded compounds 3 (20 mg), 4 (17 mg), 5 (30 mg), 6 (10 mg), 7 (3 mg), 8 (30 mg), 9 (7 mg), and a recovery of 2 (14 mg). These compounds were identified by comparison of the 1H-NMR spectral data and the [α]D25 values with those of the authentic samples. A similar experiment was also carried out for compound 5: a solution of 5 (1 mg) in H2O/MeOH (4:1, 10 ml) was left at room temperature for 48 h and it became a mixture of compounds 2—9, which were identified directly by TLC comparison with the authentic samples (benzene–ethyl formate–formic acid, 1:7:1).

Acknowledgements The authors are grateful to Mr. K. Inada and Mr. N. Yamaguchi (Nagasaki University) for NMR and MS measurements. One of the authors (Y. J. Zhang) is grateful to the Japanese Government for the award of a scholarship.

References