Two New Acylated Flavonol Glycosides from the Roots of *Otostegia limbata*

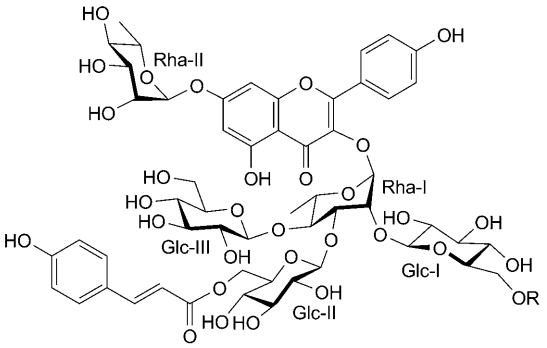
by **Afsar Khan**[a]), **Viqar U. Ahmad***[a]), and **Umar Farooq**[b])

[a]) H. E. J. Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Karachi-75270, Pakistan
(phone: +92-21-4819020; fax: +92-21-4819018-9; e-mail: vuahmad@yahoo.com)
[b]) Department of Environmental Sciences, COMSATS Institute of Information Technology, Abbottabad, Pakistan

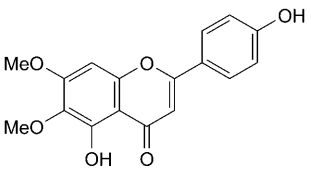
Two new acylated flavonol glycosides, 5-hydroxy-2-(4-hydroxyphenyl)-4-oxo-7-[(α-L-rhamnopyranosyl)oxy]-4*H*-chromen-3-yl β-D-glucopyranosyl-(1→2)-[β-D-glucopyranosyl-(1→4)]-[6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-D-glucopyranosyl-(1→3)]-α-L-rhamnopyranoside (**1**) and 5-hydroxy-2-(4-hydroxyphenyl)-4-oxo-7-[(α-L-rhamnopyranosyl)oxy]-4*H*-chromen-3-yl [6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-D-glucopyranosyl-(1→2)]-[β-D-glucopyranosyl-(1→4)]-[6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-D-glucopyranosyl-(1→3)]-α-L-rhamnopyranoside (**2**), were isolated from the roots of *Otostegia limbata*, along with two known flavones, cirsimaritin and 3′-*O*-methyleupatorin. Their structures were elucidated on the basis of spectroscopic evidences and chemical methods.

Introduction. – The genus *Otostegia* (Lamiaceae) comprises *ca.* 20 species with two disjunct areas of concentration, one in the Central Asian-Afghanistan-Pakistan region; the other in the mountains of East Africa [1]. In Pakistan, the genus *Otostegia* is represented by three species, namely *Otostegia persica*, *O. aucheri*, and *O. limbata*. The species *O. limbata* (BENTH.) BOISS. is endemic to Pakistan and is locally called '*spin aghzai*' or '*chiti booti*' [1]. It is used in the treatment of children gum diseases and for ophthalmia in man [2]. Moreover, the species of the genus *Otostegia* are widely used by the traditional practitioners against various diseases, and its constituents have been shown to possess antiulcer, antispasmodic, antidepressant, anxiolytic, sedative, and antioxidant activities [3]. Previously, we have isolated clerodane diterpenoids from *Otostegia limbata*, some of which were found active against acetyl cholinesterase and butyryl cholinesterase [4]. As part of our ongoing chemical study on *Otostegia limbata*, two new acylated flavonol pentaglycosides, **1** and **2**, were isolated from the roots along with two known flavones. The known compounds were identified as cirsimaritin (**3**) [5] and 3′-*O*-methyleupatorin (**4**) [6]. Compound **3** is reported for the first time from the title plant. We report herein the isolation and structural elucidation of these compounds.

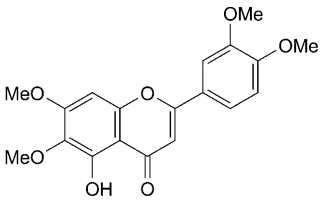
Results and Discussion. – Compound **1** was isolated as a pale yellow, gummy solid. The high-resolution electrospray-ionization quadrupole-time-of-flight (HR-ESI-Q-TOF)-MS (+ve) showed *quasi*-molecular-ion peaks at *m*/*z* 1233.3380 ($[M+Na]^+$) and 1211.3574 ($[M+H]^+$) indicating a molecular formula of $C_{54}H_{66}O_{31}$. The IR spectrum



1 R = H
2 R = (*E*)-*p*-coumaroyl



3



4

indicated absorption bands typical of OH groups (3420 cm^{-1}), C–H stretching vibrations (2924 cm^{-1}), α,β-unsaturated CO (1694 and 1655 cm^{-1}), aromatic CO (1493 cm^{-1}), and of an *O*-glycosidic linkage (1108–1027 cm^{-1}). The UV-spectral data of **1**, recorded in MeOH, showed the characteristic maxima of kaempferol 3-*O*-glycosides or 3,7-*O*-glycosides at 268, 314, and 355 (sh) nm [7]. The maximum at 314 nm corresponded to the *p*-coumaroyl residue [8] and to the band I of flavonoid nucleus. Use of standard shift reagents showed that the OH groups at C(5) and C(4′) of kaempferol were free and that the OH groups at C(3) and C(7) were substituted [9]. The collision induced dissociation (CID)-MS/MS of compound **1** was found very useful as it provided information regarding the aglycone and glycone parts. Thus, MS/MS of the *quasi*-molecular-ion peak at *m*/*z* 1233.3 ($[M+Na]^+$) gave diagnostic fragment ion peaks at *m*/*z* 1087 ($[M-146+Na]^+$; loss of rhamnosyl), 801 ($[M-(286+146)+Na]^+$; loss of flavonol + rhamnosyl), and 147 ($[p$-coumaroyl$]^+$). The first fragmentation of the $[M+Na]^+$ ion was due to the preferential breakdown of the *O*-glycosidic bond at C(7) position [10]. The MS/MS of the *quasi*-molecular-ion peak at *m*/*z* 1211.3 ($[M+H]^+$) gave characteristic product-ion peaks at *m*/*z* 1049 ($[M-162+H]^+$; loss of glucosyl), 903 ($[M-(162+146)+H]^+$), 887 ($[M-(2\times162)+H]^+$; loss of two glucosyl units), 779 ($[M-(286+146)+H]^+$; loss of flavonol + rhamnosyl), 617 ($[M-(286+146+162)+H]^+$; loss of flavonol + rhamnosyl + glucosyl), 455 ([*p*-coumaroyl-glucosyl-rhamnosyl chain]$^+$), 433 ($[M-800+H]^+$; loss of acylated tetraglycosidic chain at C(3)), 309 ([*p*-coumaroyl-glucosyl chain]$^+$), and 147 ([*p*-coumaroyl]$^+$). The structure of **1** was deduced from the ^{1}H- and ^{13}C-NMR (*Table*),

^{1}H,^{1}H-COSY, 1D-TOCSY, HMBC, and ROESY data as 5-hydroxy-2-(4-hydroxyphenyl)-4-oxo-7-[(α-L-rhamnopyranosyl)oxy]-4*H*-chromen-3-yl β-D-glucopyranosyl-(1 → 2)-[β-D-glucopyranosyl-(1 → 4)]-[6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-D-glucopyranosyl-(1 → 3)]-α-L-rhamnopyranoside; which is unprecedented.

The ^{1}H- and ^{13}C-NMR spectra showed that compound **1** was a pentaglycoside with five anomeric H-atom resonances at δ(H) 5.73 (br. *s*), 5.55 (br. *d*, *J* = 2.3), 4.70 (*d*, *J* = 7.8), 4.67 (*d*, *J* = 8.0), 4.55 (*d*, *J* = 7.6), and the corresponding anomeric C-atom signals at δ(C) 101.2, 99.8, 105.3, 103.9, and 105.8, respectively. The chemical-shift and coupling-constant data for the aromatic H-atoms together with their corresponding ^{13}C-NMR chemical shifts obtained from HSQC and HMBC experiments confirmed the kaempferol aglycone (*Table*) [11]. In the ^{1}H-NMR spectrum, two H-atoms at δ(H) 6.37 (*d*, *J* = 1.9) and 6.52 (*d*, *J* = 1.7), respectively, correlated to δ(C) 100.5 and 95.3 *via* a HSQC experiment. These chemical shift values are characteristic for C(6) and C(8) in the ring *A* of flavonoids [12]. The coupling constants of the anomeric H-atoms and the corresponding C-atom resonances, together with two secondary Me signals at δ(H) 0.95 (*d*, *J* = 6.1) and 1.27 (*d*, *J* = 6.1) indicated that compound **1** contained three glucopyranosyl (Glc) and two rhamnopyranosyl (Rha) moieties. The anomeric configurations were assigned as β for the glucopyranosyl and α for the rhamnopyranosyl groups from their coupling constants. Despite the use of 600-MHz NMR spectrometer, the presence of five sugar units in **1** resulted in a considerable overlap of the other glycosidic H-atoms. The identification of the two rhamnopyranosyl and three glucopyranosyl units was realized from the combined analysis of COSY and 1D-TOCSY spectra, and using a 2D-*J* resolved H-atom experiment carried out to obtain unequivocal multiplicities. In the ^{13}C-NMR spectrum, the relative upfield shifts of the C(3) and C(7) signals to δ(C) 135.6 and 163.5 agreed with the glycosylation of kaempferol at the C(3) and C(7) position [13]. The sugar sequence was determined on the basis of both 1D- and 2D-NMR, particularly HMBC and ROESY experiments. The positions of attachment of the glycoside moieties with the aglycone were deduced by the ^{1}H,^{13}C long-range correlations observed for H−C(1‴) of Rha II (δ(H) 5.55)/C(7) of aglycone (δ(C) 163.5) and H−C(1″) of Rha I (δ(H) 5.73)/C(3) of aglycone (δ(C) 135.6). The point of attachment of Rha II was further confirmed through ROESY correlations observed for H−C(1‴) (δ(H) 5.55)/H−C(6/8) (δ(H) 6.37/6.52), indicating a 7-*O*-linkage in conjunction with COSY correlations. The chemical shift values for the C-atom resonances assigned to the Rha II unit were consistent for it being a terminal rhamnose attached directly to the aglycone [14]. On the other hand, the C(2″) (δ(C) 80.5), C(3″) (δ(C) 80.9), and C(4″) (δ(C) 78.3) resonances of the Rha I moiety were shifted downfield due to glycosylation and provided the sites of attachment of Rha I to Glc I, Glc II, and Glc III at these positions, respectively. These observations supported a tetraglycosidic oligosaccharide structure at C(3) of kaempferol having an inner rhamnose and three terminal glucose moieties. Additional long-range correlations were noted for H−C(1⁗) of Glc I (δ(H) 4.55)/C(2″) of Rha I (δ(C) 80.5), H−C(1⁗⁗) of Glc II (δ(H) 4.70)/C(3″) of Rha I (δ(C) 80.9), and H−C(1⁗⁗⁗) of Glc III (δ(H) 4.67)/C(4″) of Rha I (δ(C) 78.3). The inverse correlations were also observed for H−C(2″) of Rha I (δ(H) 4.43)/C(1⁗) of Glc I (δ(C) 105.8), H−C(3″) of Rha I (δ(H) 4.05)/C(1⁗⁗) of Glc II (δ(C) 105.3), and H−C(4″) of Rha I (δ(H) 3.72)/C(1⁗⁗⁗) of Glc III (δ(C) 103.9). The interglycosidic linkage of this tetraglycosyl chain was confirmed

Table. *^{1}H- and ^{13}C-NMR Data* (600 and 150 MHz, resp., CD_3OD) *of* **1** *and* **2**. δ in ppm, J in Hz.

	1		**2**	
	δ(H)	δ(C)	δ(H)	δ(C)
Aglycone				
C(2)		158.6		158.1
C(3)		135.6		136.2
C(4)		179.3		179.3
C(5)		162.6		162.7
H–C(6)	6.37 (*d*, *J* = 1.9)	100.5	6.31 (*d*, *J* = 1.8)	100.5
C(7)		163.5		163.3
H–C(8)	6.52 (*d*, *J* = 1.7)	95.3	6.36 (br. *s*)	95.4
C(9)		157.7		157.5
C(10)		107.5		107.6
C(1′)		122.3		122.3
H–C(2′/6′)	7.76 (*d*, *J* = 8.6, 2 H)	131.9	7.66 (*d*, *J* = 8.6, 2 H)	132.0
H–C(3′/5′)	6.94 (*d*, *J* = 8.6, 2 H)	116.6	6.90 (*d*, *J* = 8.6, 2 H)	116.5
C(4′)		161.8		161.7
Rhamnose-I				
H–C(1″)	5.73 (br. *s*)	101.2	5.68 (br. *s*)	101.6
H–C(2″)	4.43 (br. *d*, *J* = 3.6)	80.5	4.47 (br. *d*, *J* = 3.7)	81.8
H–C(3″)	4.05 (*dd*, *J* = 3.0, 9.9)	80.9	4.11 (*dd*, *J* = 3.7, 9.7)	80.9
H–C(4″)	3.72 (*dd*, *J* = 8.92, 9.17)	78.3	3.75 (*t*, *J* = 9.7)	78.37
H–C(5″)	3.25 – 3.21 (*m*)	70.87	3.49 – 3.45 (*pseudo dq*/*m*, *J* = 6.7, 9.9)	70.82
Me(6″)	0.95 (*d*, *J* = 6.1)	17.95	1.03 (*d*, *J* = 5.99)	17.99
Rhamnose-II				
H–C(1‴)	5.55 (br. *d*, *J* = 2.3)	99.8	5.50 (br. *d*, *J* = 2.2)	99.9
H–C(2‴)	4.03 (*dd*, *J* = 1.6, 3.3)	71.7	4.03 (br. *d*, *J* = 4.0)	71.7
H–C(3‴)	3.83 (*dd*, *J* = 3.4, 9.4)	72.0	3.83 (*dd*, *J* = 3.8, 9.3)	72.1
H–C(4‴)	3.48 (*t*, *J* = 9.5)	73.6	3.48 (*t*, *J* = 9.4)	73.7
H–C(5‴)	3.66 – 3.63 (*pseudo dq*/*m*, *J* = 6.1, 9.4)	71.1	3.70 – 3.64 (*pseudo dq*/*m*, *J* = 6.0, 9.9)	71.1
Me(6‴)	1.27 (*d*, *J* = 6.1)	18.1	1.28 (*d*, *J* = 6.0)	18.1
Glucose-I				
H–C(1⁗)	4.55 (*d*, *J* = 7.6)	105.8	4.51 (*d*, *J* = 8.1)	106.1
H–C(2⁗)	3.20 (*dd*, *J* = 7.97, 9.12)	75.30	3.25 (*t*, *J* = 8.5)	75.3
H–C(3⁗)	3.37 (*t*, *J* = 8.8)	77.85	3.40 (*t*, *J* = 8.8)	78.30
H–C(4⁗)	3.29 (*dd*, *J* = 7.78, 9.89)	71.36	3.31 (*t*, *J* = 9.4)	71.9
H–C(5⁗)	3.28 – 3.26 (*m*)	77.87	3.46 (br. *dd*, *J* = 7.3, 10.4)	75.2
CH_2(6⁗)	3.68 (*dd*, *J* = 4.5, 12.1), 3.78 (br. *d*, *J* = 11.7)	62.8	4.17 (*dd*, *J* = 5.8, 11.7), 4.49 (*dd*, *J* = 2.7, 11.2)	64.5
Glucose-II				
H–C(1⁗′)	4.70 (*d*, *J* = 7.8)	105.3	4.73 (*d*, *J* = 7.8)	105.4
H–C(2⁗′)	3.31 (*dd*, *J* = 8.1, 9.2)	75.33	3.32 (*dd*, *J* = 7.8, 9.4)	75.4
H–C(3⁗′)	3.42 (*t*, *J* = 9.0)	78.1	3.42 (*t*, *J* = 8.9)	77.8
H–C(4⁗′)	3.27 (*dd*, *J* = 8.67, 9.7)	72.6	3.28 (*dd*, *J* = 8.6, 9.8)	72.5
H–C(5⁗′)	3.73 (*ddd*, *J* = 2.2, 9.2, 9.8)	76.0	3.71 (br. *dd*, *J* = 6.8, 9.7)	76.0
CH_2(6⁗′)	4.38 (br. *d*, *J* = 10.77), 4.60 (*dd*, *J* = 8.3, 11.55)	65.2	4.41 (br. *d*, *J* = 10.5), 4.55 (*dd*, *J* = 6.4, 11.2)	65.3

Table (cont.)

	1		**2**	
	δ(H)	δ(C)	δ(H)	δ(C)
Glucose-III				
H−C(1‴‴‴)	4.67 (*d*, *J* = 8.0)	103.9	4.70 (*d*, *J* = 8.0)	104.0
H−C(2‴‴‴)	3.04 (*dd*, *J* = 8.3, 8.9)	75.7	3.07 (*dd*, *J* = 8.49, 8.66)	75.8
H−C(3‴‴‴)	3.27 (*t*, *J* = 9.0)	78.7	3.30 (*dd*, *J* = 8.97, 9.23)	78.7
H−C(4‴‴‴)	3.23 (*dd*, *J* = 7.79, 9.5)	71.8	3.25 (*t*, *J* = 8.5)	71.8
H−C(5‴‴‴)	3.20–3.18 (*m*)	77.57	3.22 (*ddd*, *J* = 3.6, 5.7, 9.3)	77.7
CH_2(6‴‴‴)	3.60 (*dd*, *J* = 6.3, 11.4), 3.79 (br. *d*, *J* = 10.9)	62.8	3.62 (*dd*, *J* = 6.0, 11.4), 3.80 (*dd*, *J* = 2.4, 11.3)	62.9
Glc II-(*E*)-*p*-coumaroyl				
C(1⁗⁗⁗)		168.9		168.9
H−C(2⁗⁗⁗)	6.18 (*d*, *J* = 15.9)	114.9	6.14 (*d*, *J* = 15.88)	114.8
H−C(3⁗⁗⁗)	7.47 (*d*, *J* = 15.9)	146.5	7.45 (*d*, *J* = 15.87)	146.5
C(4⁗⁗⁗)		126.6		126.6
H−C(5⁗⁗⁗/9⁗⁗⁗)	7.00 (*d*, *J* = 8.5, 2 H)	130.3	6.95 (*d*, *J* = 8.45, 2 H)	130.4
H−C(6⁗⁗⁗/8⁗⁗⁗)	6.31 (*d*, *J* = 8.5, 2 H)	116.3	6.28 (*d*, *J* = 8.41, 2 H)	116.3
C(7⁗⁗⁗)		160.7		160.7
Glc I-(*E*)-*p*-coumaroyl				
C(1⁗⁗⁗)				169.0
H−C(2⁗⁗⁗)			6.12 (*d*, *J* = 15.88)	115.0
H−C(3⁗⁗⁗)			7.43 (*d*, *J* = 15.85)	146.7
C(4⁗⁗⁗)				127.0
H−C(5⁗⁗⁗/9⁗⁗⁗)			7.25 (*d*, *J* = 8.5, 2 H)	131.1
H−C(6⁗⁗⁗/8⁗⁗⁗)			6.65 (*d*, *J* = 8.5, 2 H)	116.6
C(7⁗⁗⁗)				161.1

in the ROESY plot by the observation of ROEs between the anomeric H-atoms and the glycosidic H-atoms of glycosylated positions. Thus, ROEs were observed for H−C(1‴) of Glc I/H−C(2″) of Rha I, H−C(1⁗) of Glc II/H−C(3″) of Rha I, and H−C(1‴‴) of Glc III/H−C(4″) of Rha I. The ^{1}H-NMR spectrum of **1** also showed signals at δ(H) 6.18 (*d*, *J* = 15.9, 1 H), 7.47 (*d*, *J* = 15.9, 1 H), 7.00 (*d*, *J* = 8.5, 2 H), and 6.31 (*d*, *J* = 8.5, 2 H). The ^{13}C-NMR spectrum contained the corresponding C-atom resonances at δ(C) 114.9, 146.5, 130.3, and 116.3, plus a CO signal at δ(C) 168.9 and two quaternary C-atom signals at δ(C) 126.6 and 160.7. These observations indicated the presence of a (*E*)-*p*-coumaroyl moiety in compound **1** [15]. As the H−C(6⁗) signals of Glc II were deshielded at δ(H) 4.38 (br. *d*, *J* = 10.77, 1 H) and 4.60 (*dd*, *J* = 8.3, 11.55, 1 H), C(6⁗) had to be the point of attachment of the (*E*)-*p*-coumaroyl moiety. Accordingly, the C(6⁗) resonance was shifted downfield at δ(C) 65.2 while the C(5⁗) resonance was shifted upfield at δ(C) 76.0 due to the β-effect [16]. The HMBC experiment confirmed the site of attachment of (*E*)-*p*-coumaroyl moiety, showing a ^{1}H,^{13}C long-range correlation between the C(1⁗⁗⁗) (δ(C) 168.9) of (*E*)-*p*-coumaroyl moiety and the H−C(6⁗) (δ(H) 4.38 and 4.60) of Glc II unit. Consequently, the sequence of the acylated tetraglycosidic carbohydrate portion was established as β-D-glucopyranosyl-(1 → 2)-[β-D-glucopyranosyl-(1 → 4)]-{6-*O*-[(2*E*)-3-(4-hydroxyphe-

nyl)prop-2-enoyl]-β-D-glucopyranosyl-(1 → 3)}-α-L-rhamnopyranoside. The ^{1}H- and ^{13}C-NMR spectra of **1** were in full agreement with the proposed structure.

Compound **2** was isolated as a pale yellow, gummy solid. The HR-ESI-Q-TOF-MS (+ve) showed *quasi*-molecular-ion peaks at *m*/*z* 1379.3810 ($[M+Na]^+$) and 1357.3999 ($[M+H]^+$), indicating a molecular formula of $C_{63}H_{72}O_{33}$. The MS/MS of the *quasi*-molecular peak at *m*/*z* 1379.3 ($[M+Na]^+$) gave diagnostic product ions at *m*/*z* 1233 ($[M-146+Na]^+$; loss of rhamnosyl), 947 ($[M-(286+146)+Na]^+$; loss of flavonol + rhamnosyl), 455 ([*p*-coumaroyl-glucosyl-rhamnosyl chain]$^+$), 309 ([*p*-coumaroyl-glucosyl chain]$^+$), and 147 ([*p*-coumaroyl]$^+$). The first fragmentation of the ($[M+Na]^+$) ion was due to the preferential breakdown of the *O*-glycosidic bond at the C(7) position. The MS/MS of the *quasi*-molecular peak at *m*/*z* 1357.3 ($[M+H]^+$) gave product ions at *m*/*z* 1195 ($[M-162+H]^+$; loss of glucosyl), 1049 ($[M-(162+146)+H]^+$; loss of glucosyl + rhamnosyl), 925 ($[M-(286+146)+H]^+$; loss of flavonol + rhamnosyl), 763 ($[M-(286+146+162)+H]^+$; loss of flavonol + rhamnosyl + glucosyl), 455 ([*p*-coumaroyl-glucosyl-rhamnosyl chain]$^+$), 309 ([*p*-coumaroyl-glucosyl chain]$^+$), and 147 ([*p*-coumaroyl]$^+$). The IR spectrum of compound **2** showed strong absorption bands typical of OH groups (3386 cm^{-1}), C–H stretching vibrations (2925 cm^{-1}), α,β-unsaturated CO groups (1654 cm^{-1}), aromatic moieties (1602 cm^{-1}), and O-glycosidic linkage (1074 – 1027 cm^{-1}). The UV spectral data of **2**, recorded in MeOH, showed the characteristic maxima at 271, 312, and 359 (sh) nm of acylated kaempferol 3-*O*-glycosides or 3,7-*O*-glycosides. The maximum at 312 nm corresponded both to the *p*-coumaroyl residue and the band I of a flavonoid nucleus. Use of standard shift reagents showed that the OH groups at C(5) and C(4′) of kaempferol were free and that the OH groups at C(3) and C(7) were substituted. The structure of **2** was deduced from the ^{1}H- and ^{13}C-NMR (*Table*), ^{1}H,^{1}H-COSY, 1D-TOCSY, HMBC, and ROESY data as 5-hydroxy-2-(4-hydroxyphenyl)-4-oxo-7-[(α-L-rhamnopyranosyl)oxy]-4*H*-chromen-3-yl [6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-D-glucopyranosyl-(1 → 2)]-[β-D-glucopyranosyl-(1 → 4)]-{6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-D-glucopyranosyl-(1 → 3)}-α-L-rhamnopyranoside; which is unprecedented.

Compound **2** was also a pentaglycoside as shown by five anomeric H-atom signals at δ(H) 5.68 (br. *s*), 5.50 (br. *d*, *J* = 2.2), 4.73 (*d*, *J* = 7.8), 4.70 (*d*, *J* = 8.0), and 4.51 (*d*, *J* = 8.1) with the corresponding anomeric C-atom resonances at δ(C) 101.6, 99.9, 105.4, 104.0, and 106.1, respectively. The ^{1}H- and ^{13}C-NMR spectra of **2** were very similar to those of **1** (*Table*), except for additional signals due to a second (*E*)-*p*-coumaroyl moiety at δ(H) 6.12 (*d*, *J* = 15.8, 1 H), 7.43 (*d*, *J* = 15.8, 1 H), 7.25 (*d*, *J* = 8.5, 2 H), and 6.65 (*d*, *J* = 8.5, 2 H). The ^{13}C-NMR spectrum contained the corresponding C-atom resonances at δ(C) 115.0, 146.7, 131.1, and 116.6, plus a CO signal at δ(C) 169.0 and two quaternary signals at δ(C) 127.0 and 161.1. These observations plus molecular mass difference of 146 u between the two compounds, clearly suggested that **2** was a (*E*)-*p*-coumaric ester of **1**. As the H–C(6‴) signals of Glc I were deshielded to δ(H) 4.17 (*dd*, *J* = 11.7, 5.8, 1 H) and 4.49 (*dd*, *J* = 11.2, 2.7, 1 H), C(6‴) had to be the point of attachment of the second (*E*)-*p*-coumaroyl moiety. Accordingly, the C(6‴) resonance was shifted downfield at δ(C) 64.5. The site of attachment of the second (*E*)-*p*-coumaroyl moiety was also confirmed by a HMBC long-range correlation between the C(1⁗) (δ(C)169.0) of (*E*)-*p*-coumaroyl and the H–C(6‴) (δ(H) 4.17 and 4.49) of

the Glc I unit (*Figure*). The ^{1}H- and ^{13}C-NMR spectra of **2** were in full agreement with the proposed structure.

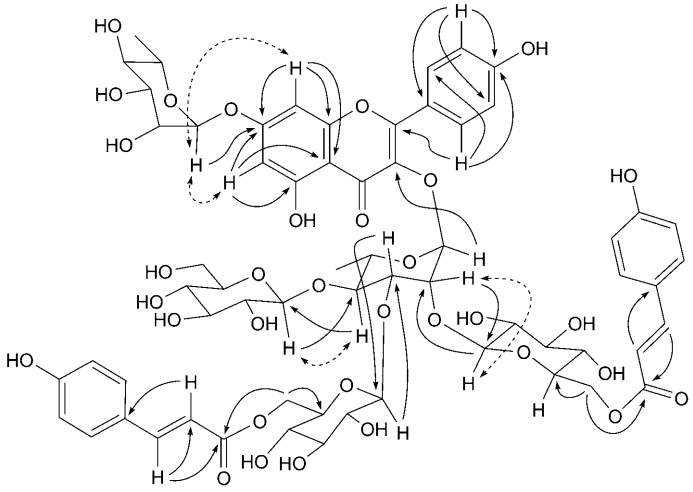


Figure. *Important HMBC* (→) *and ROESY* (←- - -→) *interactions of* **2**

The absolute configuration of the sugar units was assigned after hydrolysis of the crude fraction (*Fr. 14*) with HCl. The hydrolyzed sugars were converted into acetylated thiazolidine derivatives, and the GC retention time of each sugar was compared with those of authentic sugar samples prepared in the same manner.

Flavonoid *p*-coumaroyl glucosides are commonly found in some genera of the Lamiaceae, and they are generally considered as valuable chemotaxonomic markers in this family [17]. The isolation of compounds **1** and **2** is of chemotaxonomic interest as acylated flavonoid glycosides are here reported for the first time from the genus *Otostegia*. These compounds may be useful for the authentication and standardization of extracts of medicinal interest.

We thank Dr. *David Smith*, Emeritus Professor of Chemistry (University of Nebraska, Lincoln, NE 68006 USA) for helpful discussion on ESI spectra. Financial support by the *Higher Education Commission* (HEC) *of Pakistan* is also gratefully acknowledged.

Experimental Part

General. Column chromatography (CC): silica gel *60* (SiO_2; 230–400 mesh, *Macherey–Nagel*), *Sephadex LH-20* (25–100 μm; *Sigma-Aldrich*). Gas Chromatography: *GC-17A* (*Shimadzu*) instrument; detector: FID; column: *Optima-5-Accent* (0.25 μm, ϕ 0.25 mm × 30 m; *Macherey–Nagel*); carrier gas: N_2; column initial temp.: 50°; temp. increased: 5°/min; column final temp.: 240°; inj. port temp.: 240°. MPLC: *Eyela VSP-3050* instrument attached with a column (ϕ 25 mm × 200 mm; *Eyela*) filled with *Polygroprep C-18* (25–40 μm, 10 nm; *Macherey–Nagel*). Recycling prep. HPLC: *LC-908W* instrument (*Japan Analytical Industries (JAI) Co. Ltd.*) equipped with *J'sphere ODS-M80* column (S-4 μm, 8 nm, ϕ 20 mm × 250 mm; *YMC*) or *Hibar LiChrosorb RP-18* column (7 μm, ϕ 25 mm × 250 mm; *Merck*) using a *JAI RI-5* refractive index detector and a *JAI UV-310* detector (254 nm). TLC: *RP-18* F_{254s} plates (*Merck*); spots observed first under UV (254/366 nm), and then stained with cerium(IV)sulphate spray

reagent in 10% H_2SO_4 soln. and heated until coloration developed. Optical rotations: *JASCO DIP-360* digital polarimeter or *Polartronic D* polarimeter. UV Spectra: *Hitachi U-3200* spectrophotometer; λ_{max} (log ε) in nm. IR Spectra: *Bruker Vector 22 FT-IR* spectrometer; KBr pellets; in cm^{-1}. NMR Spectra: *Bruker AV* (*Avance*) instrument; at 600 (^{1}H) or 150 MHz (^{13}C) resp.; δ in ppm rel. to the solvent peaks (δ(H) 3.31 and δ(C) 49.0 from CD_3OD), *J* in Hz. ESI-MS (pos.): *QStar XL Hybrid LC/MS/MS* (*Applied Biosystems*) spectrometer; in *m/z*. CID-MS/MS: 25, 30 or 55 eV; N_2: collision gas. EI-MS: *MAT312*.

Plant Material. The root parts of *Otostegia limbata* (Lamiaceae) were collected in July 2002 from Abbottabad, Pakistan, and identified by Dr. *Manzoor Ahmad* of the Department of Botany, Post-Graduate College, Abbottabad, Pakistan. A voucher specimen has been deposited with the herbarium of the same department (No. 6872).

Extraction and Isolation. The air-dried roots of *Otostegia limbata* (30 kg) were extracted repeatedly with MeOH (3 × 50 l) at r.t., every time for 10 d. The solvent was evaporated under reduced pressure to give a dark residue (450 g), which was partitioned between equal volumes (4 l) of each solvent; hexane (50 g), $CHCl_3$ (80 g), AcOEt (120 g), BuOH (170 g), and H_2O (30 g). The BuOH extract was subjected to *LH-20* CC and eluted with H_2O (2.0 l), followed by increasing concentrations of MeOH in H_2O (0 → 100% MeOH, in steps of 25% of MeOH, each 2.0 l) as eluent to yield five fractions. The fraction eluted with MeOH/H_2O (1:1) was then applied to repeated (three times) CC on *Sephadex LH-20* with increasing concentrations of MeOH in H_2O (0 → 100% MeOH) to remove the saponin part. The flavonoid part was then subjected to MPLC (*Polygroprep C-18* column) and eluted with H_2O followed by increasing concentrations of MeOH in H_2O (0 → 100% MeOH) to give 14 fractions. *Fr. 14* upon further purification on prep. recycling HPLC (*J'sphere ODS-M80* column, MeOH/H_2O (1:1); flow rate 3 ml/min) resulted in compound **1** (193 mg; t_R 45 min) and compound **2** (66 mg; t_R 58 min). The elution of these compounds was monitored through UV and RI detectors, simultaneously. The $CHCl_3$ extract was subjected to CC (SiO_2 *60*; hexane/$CHCl_3$ 1:0 → 0:1) to afford eleven fractions. *Fr. 6* was loaded on an SiO_2 *60* column and eluted with AcOEt/hexane (2:8) to give two known compounds cirsimaritin (**3**; 5.0 mg) and 3′-*O*-methyl eupatorin (**4**; 7.3 mg).

*5-Hydroxy-2-(4-hydroxyphenyl)-4-oxo-7-[(α-*L*-rhamnopyranosyl)oxy]-4*H*-chromen-3-yl β-*D*-Glucopyranosyl-(1 → 2)-[β-*D*-glucopyranosyl-(1 → 4)]-{6-*O*-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-*D*-glucopyranosyl-(1 → 3)}-α-*L*-rhamnopyranoside* (**1**). Pale yellow, gummy solid. $[\alpha]_D^{29} = -56.4$ ($c = 0.242$, MeOH). UV (MeOH): 224.6 (4.46), 268 (4.56), 314 (4.61), 355 (sh). UV (MeOH – $AlCl_3$): 233, 277, 304, 389 (sh). UV (MeOH – $AlCl_3$ – HCl): 234, 277, 320, 389. UV (MeOH – AcONa): 268, 314, 340 (sh), 370 (sh). IR (KBr): 3420, 2924, 1694, 1655, 1605, 1515, 1493, 1452, 1352, 1264, 1207, 1172, 1108 – 1027, 965, 830. ^{1}H- and ^{13}C-NMR: *Table*. HR-ESI-Q-TOF-MS: 1233.3380 ($[M+Na]^+$, $C_{54}H_{66}NaO_{31}^+$; calc. 1233.3486), 1211.3574 ($[M+H]^+$, $C_{54}H_{67}O_{31}^+$; calc. 1211.3666). ESI-Q-TOF-MS/MS (pos.; 1233.3): 1233 (100, $[M+Na]^+$), 1087 (10, $[M-146+Na]^+$), 801 (29, $[M-(286+146)+Na]^+$), 625 (6), 475 (13), 203 (7), 147 (56, [*p*-coumaroyl]$^+$). ESI-Q-TOF-MS/MS (pos.; 1211.3): 1211 (46, $[M+H]^+$), 1049 (41, $[M-162+H]^+$), 903 (9.8, $[M-(162+146)+H]^+$), 887 (9.8, $[M-(2\times162)+H]^+$), 779 (31.8, $[M-(286+146)+H]^+$), 617 (63.8, $[M-(286+146+162)+H]^+$), 455 (51.8, [*p*-coumaroyl-glucosyl-rhamnosyl]$^+$), 433 (86, $[M-800+H]^+$), 309 (100, [*p*-coumaroyl-glucosyl]$^+$), 147 (47, [*p*-coumaroyl]$^+$).

*5-Hydroxy-2-(4-hydroxyphenyl)-4-oxo-7-[(α-*L*-rhamnopyranosyl)oxy]-4*H*-chromen-3-yl {6-O-[(2*E*)-3-(4-Hydroxyphenyl)prop-2-enoyl]-β-*D*-glucopyranosyl-(1 → 2)}-[β-*D*-glucopyranosyl-(1 → 4)]-{6-O-[(2*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]-β-*D*-glucopyranosyl-(1 → 3)}-α-*L*-rhamnopyranoside* (**2**). Pale yellow, gummy solid. $[\alpha]_D^{29} = -117.6$ ($c = 0.034$, MeOH). UV (MeOH): 212 (4.45), 229.4 (4.51), 271 (4.56), 312 (4.76), 359 (sh). UV (MeOH – $AlCl_3$): 210, 230.8, 279.8, 314, 389 (sh). UV (MeOH – $AlCl_3$ – HCl): 210.8, 232.8, 281, 317.6, 389 (sh). UV (MeOH – AcONa): 271, 312, 377 (sh). IR (KBr): 3386, 2925, 1694, 1654, 1602, 1512, 1447, 1351, 1266, 1209, 1172, 1074 – 1027, 827. ^{1}H- and ^{13}C-NMR: *Table*. HR-ESI-Q-TOF-MS: 1379.3810 ($[M+Na]^+$, $C_{63}H_{72}NaO_{33}^+$; calc. 1379.3854), 1357.3999 ($[M+H]^+$, $C_{63}H_{73}O_{33}^+$; calc. 1357.4034). ESI-Q-TOF-MS/MS (pos.; 1379.3): 1379 (100, $[M+Na]^+$), 1233 (0.16, $[M-146+Na]^+$), 947 (0.35, $[M-(286+146)+Na]^+$), 455 (0.16, [*p*-coumaroyl-glucosyl-rhamnosyl]$^+$), 309 (7.9, [*p*-coumaroyl-glucosyl]$^+$), 147 (24.8, [*p*-coumaroyl]$^+$). ESI-Q-TOF-MS/MS (pos.; 1357.3): 1357 (40, $[M+H]^+$), 1195 (8, $[M-162+H]^+$), 1049 (6, $[M-(162+146)+H]^+$), 925 (25.7, $[M-(286+146)+H]^+$), 763 (52, $[M-(286+146+162)+H]^+$), 455 (84, [*p*-coumaroyl-glucosyl-rhamnosyl]$^+$), 309 (100, [*p*-coumaroyl-glucosyl]$^+$), 147 (48, [*p*-coumaroyl]$^+$).

Acid Hydrolysis. A soln. of semi-purified *Fr. 14* (20 mg) was dissolved in 10% aq. HCl and refluxed for 3 h. On cooling, the reaction mixture was extracted with AcOEt. After separating the org. layer, the aq. phase was neutralized with Na_2CO_3 and concentrated. The sugars were identified as glucose and rhamnose by co-TLC with authentic samples by using the standard method [18].

Determination of Absolute Configuration. The concentrated residue of the hydrolyzed sugars in pyridine (0.04M) and L-cysteine ethyl ester hydrochloride (0.06M) were mixed, and the soln. was warmed at 60° for 1 h. Ac_2O (150 µl) was then added, and the mixture was warmed at 90° for another 1 h. After evaporation of pyridine and Ac_2O *in vacuo*, each residue was dissolved in acetone (350 µl), and the soln. (1 µl) was subjected to GC [19]. The peaks for peracetylated thiazolidine derivatives with t_R of 31.68 and 44.57 min were observed for the samples which were, resp., identical to the derivatives of authentic D-glucose and L-rhamnose prepared in the same manner.

REFERENCES

[1] I. C. Hedge, in 'Flora of Pakistan', Eds. S. I. Ali, Y. J. Nasir, BCC & T Press, University of Karachi, Karachi, 1990, No. 192, p. 177.
[2] R. N. Chopra, S. L. Nayar, I. C. Chopra, 'Glossary of Indian Medicinal Plants', NISC & IR (CSIR) Press, New Delhi, 1956, p. 183.
[3] K. Vural, N. Ezer, K. Erol, F. P. Sahin, *J. Fac. Pharm. Gazi.* **1996**, *13*, 29; F. Shrififar, N. Yassa, A. Shafiee, *Iranian J. Pharm. Res.* **2003**, *2*, 235.
[4] V. U. Ahmad, A. Khan, U. Farooq, F. Kousar, S. S. Khan, S. A. Nawaz, M. A. Abbasi, M. I. Choudhary, *Chem. Pharm. Bull.* **2005**, *53*, 378; V. U. Ahmad, A. Khan, U. Farooq, F. Kousar, S. S. Khan, J. Hussain, *J. Asian Nat. Prod. Res.* **2007**, *9*, 91.
[5] D. Youssef, A. W. Frahm, *Planta Med.* **1995**, *61*, 570.
[6] M. Ikram, S. Jehangir, S. Razaq, M. Kawano, *Planta Med.* **1979**, *37*, 189.
[7] K. R. Markham, T. J. Mabry, in 'The Flavonoids', Eds. J. B. Harborne, T. J. Mabry, H. Mabry, Academic Press, New York, 1975, p. 45.
[8] F. Tomás, J. L. Nieto, F. A. T. Barberán, F. Ferreres, *Phytochemistry* **1986**, *25*, 1253.
[9] K. R. Markham, in 'Methods in Plant Biochemistry', Eds. P. M. Dey, J. B. Harborne, Academic Press, New York, 1989, Vol. 1, p. 197; T. J. Mabry, K. R. Markham, M. B. Thomas, 'The Systematic Identification of Flavonoids', Springer-Verlag, Berlin, 1970.
[10] F. Vallejo, F. A. Tomas-Barberán, F. Ferreres, *J. Chromatogr., A* **2004**, *1054*, 181.
[11] M.-L. Crublet, C. Long, T. Sévenet, H. A. Hadi, C. Lavaud, *Phytochemistry* **2003**, *64*, 589.
[12] T. Fossen, O. M. Andersen, in 'Flavonoids: Chemistry, Biochemistry and Applications', Eds. O. M. Andersen, K. R. Markham, CRC Press/Taylor & Francis, Boca Raton, 2006, p. 52.
[13] N. Mulinacci, F. F. Vincieri, A. Baldi, M. Bambagiotti-Alberti, A. Sendl, H. Wagner, *Phytochemistry* **1995**, *38*, 531.
[14] M. Veit, G. F. Pauli, *J. Nat. Prod.* **1999**, *62*, 1301.
[15] C. Fiorini, B. David, I. Fourasté, J. Vercauteren, *Phytochemistry* **1998**, *47*, 821.
[16] P. K. Agrawal, M. C. Bansal, in 'Carbon-13 NMR of Flavonoids', Ed. P. K. Agrawal, Elsevier, Amsterdam, 1989, p. 283.
[17] F. A. Tomas-Barberán, M. I. Gil, F. Ferreres, F. Tomás-Lorente, *Phytochemistry* **1992**, *31*, 3097.
[18] E. Bedir, I. A. Khan, *J. Nat. Prod.* **2000**, *63*, 1699.
[19] Q. Ye, W. Zhao, *Planta Med*. **2002**, *68*, 723.

Received September 16, 2008