Acyclic Diterpene Glycosides, Capsianosides VIII, IX, X, XIII, XV and XVI from the Fruits of Paprika *Capsicum annuum* L. var. *grossum* Bailey and Jalapeño *Capsicum annuum* L. var. *annuum*

Jong-Hyun Lee, Naoko Kiyota, Tsuyoshi Ikeda, and Toshihiro Nohara*

Faculty of Medical and Pharmaceutical Sciences, Kumamoto University; 5–1 Oe-Honmachi, Kumamoto 862–0973, Japan.
Received March 13, 2006; accepted June 16, 2006

Paprika and Jalapeño are used as vegetables and spices. We have obtained six new acyclic diterpene glycosides, called capsianosides XIII (2), XV (3), IX (4), XVI (5), X (6) and VIII (7) together with known capsianoside II (1) from the fruits of the Paprika and Jalapeño. The structures of these compounds have been elucidated by the ^{1}H- and ^{13}C-NMR spectra and two-dimensional NMR methods.

Key words Paprika; Jalapeño; capsianoside; acyclic diterpene glycoside; Solanaceae; *Capsicum annuum*

Paprika is known for its high vitamin C content and has been isolated from Hungarian paprika in large amounts.[1] Jalapeño is a stimulating spice which contains capsaicin and related compounds in its fruits and veins. Carotenoids,[2,3] lipids and capsaicin[4] have been well studied in the *Capsicum* species. Now our study has focused on the water-soluble constituents of Paprika and Jalapeño fruits. Previous studies in this laboratory led to the isolation of novel acyclic diterpene glycosides, capsianoside A—H and I—VI from *Capsicum* plants.[5] In this study, we isolated six new acyclic diterpene glycosides termed capsianosides XIII (**2**), XV (**3**), IX (**4**), XVI (**5**), X (**6**) and VIII (**7**) along with one known compound.[5] Since the known compound, capsianoside II[5] had an error in sugar sequence, we revised it. Distribution of these new compounds is summarized in Table 1.

Capsianoside II (**1**), an amorphous powder, $[\alpha]_D$ −31.2°, showed peaks due to $[M+Na]^+$ at *m/z* 1108, [*m/z* 1108-deoxyhexose]$^+$ at 962 and [*m/z* 962-hexose-deoxyhexose]$^+$ at 654 in positive FAB-MS and gave a molecular formula $C_{50}H_{84}O_{25}Na$ at *m/z* 1107.5374 by HR-FAB-MS. Moreover, the ^{1}H-NMR spectrum of **1** disclosed the presence of four methyl groups at δ 1.39 (3H, s), 1.61 (6H, s) and 1.77 (3H, s), three olefinic protons at δ 5.22 (1H, br d, J=11.0 Hz), 5.23 (1H, br d, J=18.3 Hz) and 6.12 (1H, dd, J=11.0, 18.3 Hz) arising from a mono-substituted double bond, three olefinic protons at δ 5.12 (2H, m) and 5.40 (1H, t) adjacent to methylene group and seven methylene groups at δ 1.61 (2H, m), 1.95—2.20 (10H, m), 4.13 (1H, d, J=11.6 Hz) and 4.32 (1H, d, J=11.6 Hz). The ^{13}C-NMR spectrum of **1** as listed in Table 2 showed total twenty carbon signals composed of three tri-substituted double bonds at δ 125.8, 125.9, 131.2, 132.4, 135.5 and 136.0, a mono-substituted double bond at δ 116.0 and 144.4, three methyl groups at δ 16.3 (2×C), 23.4 and seven methylene groups. Since all the signals due to the aglycone moiety of capsianoside II (**1**) in the ^{1}H- and ^{13}C-NMR spectra were identical with those of capsianoside A reported previously,[5] the configuration at C-3 was characterized as *S*. The above NMR signals were identical with those of the aglycone moiety, 17-hydroxy-6*E*,10*E*,14*Z*-(3*S*)-geranyl-linalool, of capsianoside II previously reported.[5] Other sapogenols of capsianosides XIII (**2**), IX (**4**), XVI (**5**), X (**6**) and VIII (**7**) were also coincident with that of **1**. The ^{1}H-NMR spectrum due to the sugar part of **1** showed signals ascribable to 3 mol of hexosyl and 2 mol of deoxyhexosyl moieties from the evidence of signals ascribable to five anomeric protons at δ 4.21 (1H, d, J=7.3 Hz), 4.47 (1H, d, J=7.3 Hz), 4.56 (1H, d, J=7.3 Hz), 4.71 (1H, s) and 4.83 (1H, s). The ^{13}C-NMR spectra also exhibited five anomeric carbon signals at δ 105.9, 102.7, 102.2, 101.6 and 98.3. The following HMBC correlations shown in Fig. 1 revealed the sugar connections respectively: correlations between the signal at δ_H 4.47 (1H, d, J=7.3 Hz) and the signal at δ_C 82.1 (aglycone C-3) indicating the Glc I H-1 to be linked to the C-3 of the aglycone; the signal at δ_H 4.56 (1H, d, J=7.3 Hz) and the signal at δ_C 83.3 (Glc I C-2) showing the Glc II H-1 to be linked to the C-2 of the Glc I; the signal at δ_H 4.21 (1H, d, J=7.3 Hz) and the signal at δ_C 67.8 (aglycone C-17) exhibiting that the Glc IV H-1 was linked to the C-17 of the aglycone; the signal at δ_H 4.83 (1H, s) and the signal at δ_C 79.3 (Glc IV C-4) indicating that the Rha I H-1 was bound to the Glc IV C-4; the signal at δ_H 4.71 (1H, s) and the signal at δ_C 66.9 (Glc IV C-6) showing the Rha II H-1 to be linked to the Glc IV C-6. The sugar structure of **1** previously reported should be revised. The structure of **1** was determined to be 3-*O*-β-D-glucopyranosyl-(1→2)-β-D-glucopyranosyl 17-hydroxy-6*E*,10*E*,14*Z*-(3*S*)-geranyllinalool 17-*O*-α-L-rhamnopyranosyl-(1→4)-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside.

Capsianoside XIII (**2**), an amorphous powder, $[\alpha]_D$ −38.9°, showed a peak due to $[M+Na]^+$ at *m/z* 1108 in positive FAB-MS and gave a molecular formula $C_{50}H_{84}O_{25}Na$ at *m/z* 1107.5350 by HR-FAB-MS. Compound **2** was constituted of 3 mol of hexosyl and 2 mol of deoxyhexosyl moieties based on the signals of anomeric protons at δ 4.20 (1H,

Table 1. Distribution of Capsianosides in Paprika and Jalapeño

Capsianoside	II (**1**)	VIII (**7**)	IX (**4**)	X (**6**)	XIII (**2**)	XV (**3**)	XVI (**5**)
C. annuum L. var. *grossum* (Paprika)	○		○	○	○	○	○
C. annuum L. var. *annuum* (Jalapeño)	○	○	○	○	○		

* To whom correspondence should be addressed. e-mail: none@gpo.kumamoto-u.ac.jp

d, J=7.3 Hz), 4.39 (1H, d, J=7.3 Hz), 4.41 (1H, d, J=7.3 Hz), 4.71 (1H, s) and 4.82 (1H, s). Moreover, the ^{13}C-NMR spectra exhibited five anomeric carbon signals at δ 104.6, 102.7, 102.2, 101.6 and 99.4. Compared with the NMR signals of **1**, the signal due to the C-4 of Glc I appeared at δ_C 81.0 (+9.6 ppm shift), while the signals due to C-3 and C-5 shifted by −1.5 and −1.6 ppm, respectively. No glycosylation shift at C-2 in Glc I was observed, however. The signals due to the sugar moiety bound to C-17-OH were superimposable on those of compound **1**. The HMBC correlations between the signal at δ_H 4.39 (1H, d, J=7.3 Hz) and the signal at δ_C 81.5 (aglycone C-3), and between the signal at δ_H 4.41 (1H, d, J=7.3 Hz) and that at δ_C 81.0 (Glc I C-4) clarified its sugar connections to be clear. Thus, it was suggested that the sugar part is constituted of 3-*O*-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl residues.

Capsianoside XV (**3**), an amorphous powder, $[\alpha]_D$ −36.9°, showed peaks due to $[M+Na]^+$ at m/z 1123, [m/z 1123-hexose]$^+$ at 961, [m/z 961-deoxyhexose]$^+$ at 815, [m/z 815-deoxyhexose]$^+$ at 669 and [m/z 669-hexose]$^+$ at 507 in positive FAB-MS and gave a molecular formula $C_{50}H_{84}O_{26}Na$ at m/z 1123.5298 by HR-FAB-MS. Compound **3** displayed the signals from three hexosyl and two deoxyhexosyl anomeric protons [δ 4.21 (1H, d, J=7.3 Hz), 4.39 (1H, d, J=7.9 Hz), 4.41 (1H, d, J=7.9 Hz), 4.71 (1H, s) and 4.82 (1H, s)] in the ^{1}H-NMR spectrum. The ^{13}C-NMR spectrum also exhibited five anomeric carbon signals at δ 104.4, 102.5, 102.0, 101.4 and 99.2. In a comparative study of the NMR signals with those of capsianoside II,[5] the signal due to Glc I C-4 to appear at δ 80.8 (+9.5 ppm shift), while the signals due to C-3 and C-5 shifted −0.2 and −1.4 ppm, respectively. But no glycosylation shift at C-2 was observed, suggesting that the sugar part is constituted with 3-*O*-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl residue. The signals ascribable to the sugar moiety were superimposable on those of **2**, however, the signals due to the aglycone moiety were different. In **3**, the signal due to one hydroxymethyl group appeared at δ 4.09 (2H, s) instead of one methyl group in **2**. The signals due to C-6, C-7 and C-8 in **3** appeared at δ 129.0, 139.2 and 36.0, respectively, showing hydroxylation shifts[5] by +3.2, +3.7 and −4.7 ppm, respectively. From the above evidence, a methyl group at C-19 in **2** changed to a hydroxymethyl group in **3**.

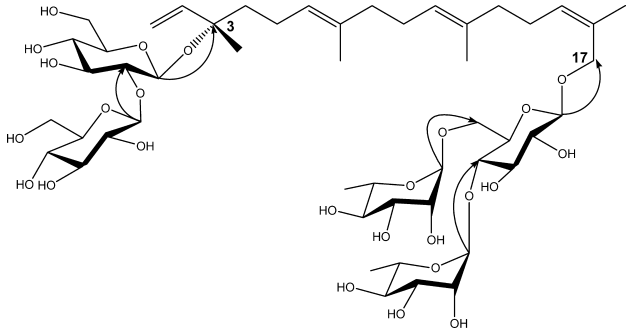


Fig. 1. Key HMBC of **1**

Capsianoside IX (**4**), an amorphous powder, $[\alpha]_D$ −32.7°, showed a peak due to $[M+Na]^+$ at m/z 1269 in positive FAB-MS and gave a molecular formula $C_{56}H_{94}O_{30}Na$ at m/z 1269.5649 by HR-FAB-MS. Compound **4** was shown to be composed of 4 mol of hexosyl and 2 mol of deoxyhexosyl moieties from evidence due to anomeric proton signals at δ 4.20 (1H, d, J=7.9 Hz), 4.41 (1H, d, J=7.9 Hz), 4.50 (1H, d, J=7.9 Hz), 4.58 (1H, d, J=7.9 Hz), 4.71 (1H, s) and 4.83 (1H, s). Moreover, the ^{13}C-NMR spectrum exhibited six anomeric carbon signals at δ 106.0, 104.6, 102.7, 102.2, 101.6 and 98.3. By comparing the ^{13}C-NMR signals with those of **1**, it was recognized that the signal of Glc I C-4 shifted to δ_C 80.8 (+9.4 ppm shift), while the signals due to C-3 and C-5 shifted by −0.2 and −1.0 ppm, respectively. The following HMBC showed the connections to the respective sugar: correlations between the signal at δ_H 4.50 (1H, d, J=7.9 Hz) and the signal at δ_C 82.1 (aglycone C-3) showing the Glc I H-1 to be linked to the C-3 of the aglycone; the signal at δ_H 4.58 (1H, d, J=7.9 Hz) and that at δ_C 82.8 (Glc I C-2) indicating the Glc II H-1 to be linked to the C-2 of the Glc I; the signal at δ_H 4.41 (1H, d, J=7.9 Hz) and that at δ_C 80.8 (Glc I C-4) showing the Glc III H-1 to be linked to C-4 of Glc I. Therefore, the 3-*O*-sugar part was constituted with β-D-glucopyranosyl-(1→4)-[β-D-glucopyranosyl-(1→2)]-β-D-glucopyranosyl residue.

Capsianoside XVI (**5**), an amorphous powder, $[\alpha]_D$ −35.5°, showed a peak due to $[M+Na]^+$ at m/z 1431.6107 (Calcd for $C_{62}H_{104}O_{35}Na$, 1431.6256) in HR-ESI-MS. Compound **5** showed five hexosyl and two deoxyhexosyl anomeric proton signals at δ 4.20 (1H, d, J=7.3 Hz), 4.41 (1H, d, J=7.9 Hz), 4.51 (1H, d, J=7.9 Hz), 4.58 (1H, d, J=7.9 Hz), 4.58 (1H, d, J=7.9 Hz), 4.72 (1H, s) and 4.84 (1H, s). The ^{13}C-NMR spectrum also exhibited seven anomeric carbon signals at δ 106.0, 105.6, 104.6, 102.4, 102.2, 101.6 and 98.3. The HMBC spectrum revealed the respective connectivities: the Glc I H-1 at δ_H 4.51 (1H, d, J=7.9 Hz) and the C-3 at δ_C 82.1 of the aglycone; the Glc II H-1 at δ_H 4.58 (1H, d, J=7.9 Hz) and the Glc I C-2 δ_C 82.9;

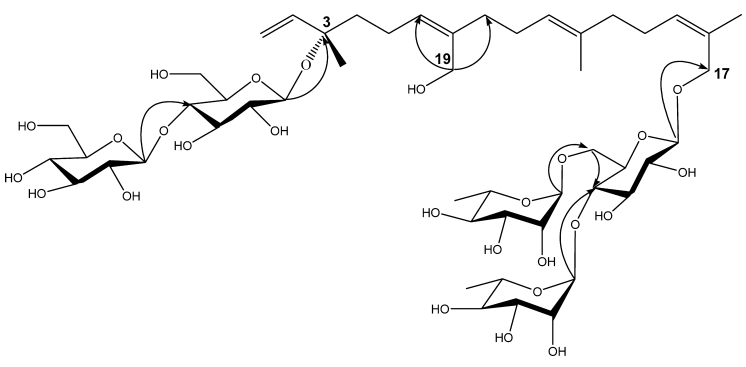


Fig. 2. Key HMBC of **3**

Table 2. ^{13}C-NMR Data of Compounds **1**, **2**, **3**[a], **4** and **5** in CD_3OD (125 MHz)

Position	1	2	3	4	5	Position	1	2	3	4	5	Position	1	2	3	4	5
1	116.0	115.8	115.6	116.0	116.0	Glc I-1	98.3	99.4	99.2	98.3	98.3	Glc V-1					105.6
2	144.4	144.4	144.1	144.4	144.4	2	83.3	74.9	74.8	82.8	82.9	2					74.9
3	82.1	81.5	81.3	82.1	82.1	3	77.5	76.6	76.4	76.5	76.8	3					78.2
4	43.0	42.6	42.8	43.0	43.1	4	71.4	81.0	80.8	80.8	80.9	4					71.4
5	23.6	23.5	23.6	23.6	23.6	5	77.7	76.1	76.1	76.7	76.7	5					77.6
6	125.9	125.8	129.0	125.8	125.8	6	62.7	62.2	62.0	62.7	62.7	6					62.8
7	135.5	135.5	139.2	135.5	135.5	Glc II-1	105.9			106.0	106.0	Rha I-1	102.7	102.7	102.5	102.7	102.4
8	40.8	40.7	36.0	40.8	40.8	2	76.5			76.0	76.6	2	72.3	71.4	71.3	72.2	72.2
9	27.7	27.6	27.9	27.7	27.7	3	78.2			78.0	78.3	3	72.5	72.4	72.3	72.4	71.6
10	125.8	125.9	125.7	125.9	125.9	4	71.7			72.3	72.3	4	73.8	74.0	73.7	73.8	83.3
11	136.0	135.9	135.3	136.0	136.0	5	78.1			77.5	78.0	5	69.8	69.8	69.7	69.8	69.3
12	40.8	40.9	40.9	40.8	40.9	6	62.7			62.4	62.4	6	17.9	17.9	17.9	17.9	18.0
13	27.3	27.3	27.4	27.3	27.3	Glc III-1		104.6	104.4	104.6	104.6	Rha II-2	101.6	101.6	101.4	101.6	101.6
14	131.2	131.2	131.1	131.2	131.3	2		75.2	75.1	74.9	75.3	2	72.2	72.2	72.2	72.3	72.4
15	132.4	132.5	132.2	132.4	132.4	3		77.9	77.8	77.9	78.3	3	72.4	72.5	72.2	72.5	72.4
16	21.9	21.9	22.0	21.9	21.9	4		72.3	72.4	71.4	72.3	4	74.0	73.8	74.0	74.0	74.0
17	67.8	67.8	67.6	67.8	67.8	5		78.1	78.0	77.5	77.9	5	70.6	70.6	70.5	70.6	69.8
18	16.3	16.3	16.4	16.3	16.3	6		62.5	62.4	62.1	62.2	6	18.2	18.2	18.3	18.2	18.2
19	16.3	16.1	59.9	16.3	16.3	Glc IV-1	102.2	102.2	102.0	102.2	102.2						
20	23.4	23.3	23.2	23.5	23.5	2	75.2	74.9	74.8	75.2	75.5						
						3	76.7	76.8	76.7	76.5	76.8						
						4	79.3	79.4	79.2	79.3	79.1						
						5	75.4	75.4	75.3	75.4	76.1						
						6	66.9	66.9	66.8	66.9	66.9						

a) Spectra recorded in CD_3OD at 100.4 MHz.

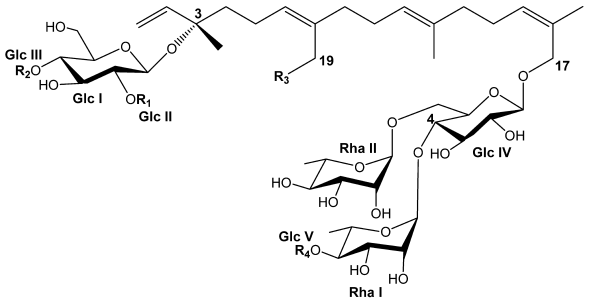
the Glc III H-1 at δ_H 4.41 (1H, d, J=7.9 Hz) and the Glc I C-4 at δ_C 80.9; the Glc IV H-1 at δ_H 4.20 (1H, d, J=7.3 Hz) and the C-17 at δ_C 67.8 of the aglycone; the Rha I H-1 at δ_H 4.84 (1H, s) and the Glc IV C-4 δ_C 79.1; the Rha II H-1 at δ_H 4.72 (1H, s) and the Glc IV C-6 δ_C 66.9; the Glc V H-1 at δ_H 4.58 (1H, d, J=7.9 Hz) and the Rha I C-4 δ_C 83.3. This suggested that the sugar part was constituted with 3-*O*-β-D-glucopyranosyl-(1→4)-[β-D-glucopyranosyl-(1→2)]-β-D-glucopyranosyl residue and another branched sugar part was identical to 17-*O*-β-D-glucopyranosyl-(1→4)-α-L-rhamnopyranosyl-(1→4)-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranosyl residue.

Capsianoside X (**6**), an amorphous powder, $[\alpha]_D$ −16.7°, showed a peak due to $[M+Na]^+$ at *m/z* 1286 in positive FAB-MS and gave a molecular formula $C_{56}H_{94}O_{31}Na$ at *m/z* 1285.5880 by HR-FAB-MS. Compound **6** displayed five hexosyl and one deoxyhexosyl anomeric proton signals at δ 4.36 (1H, d, J=7.9 Hz), 4.42 (1H, d, J=7.9 Hz), 4.50 (1H, d, J=7.3 Hz), 4.58 (1H, d, J=7.3 Hz), 4.62 (1H, d, J=7.9 Hz) and 4.75 (1H, s). The ^{13}C-NMR spectrum also exhibited six anomeric carbon signals at δ 106.1, 104.8, 104.6, 102.1, 101.1 and 98.3. The following HMBC correlations were observed between the signal at δ_H 4.50 (1H, d, J=7.3 Hz) and the signal at δ_C 82.1 (aglycone C-3): the Glc I H-1 linked to the C-3 of the aglycone; the signal at δ_H 4.58 (1H, d, J=7.3 Hz) and the signal at δ_C 82.9 (Glc I C-2): the Glc II H-1 linked to the C-2 of the Glc I; the signal at δ_H 4.42 (1H, d, J=7.9 Hz) and the signal at δ_C 80.8 (Glc I C-4): the Glc III H-1 linked to the C-4 of the Glc I; the signal at δ_H 4.36 (1H, d, J=7.9 Hz) and the signal at δ_C 68.3 (aglycone C-17): the Glc IV H-1 linked to the C-17 of the aglycone; the signal at δ_H 4.62 (1H, d, J=7.9 Hz) and the signal at δ_C 82.1 (Glc IV C-2): (1→2) linkage between the two glucoses. The linkage of rhamnose to the Glc IV C-6 was also determined by the HMBC correlation between the signal at δ_H 4.75 (1H, s) and the signal at δ_C 67.7 (Glc IV C-6). From the above evidence, **6** was determined as 3-*O*-β-D-glucopyranosyl-(1→4)-[β-D-glucopyranosyl-(1→2)]-β-D-glucopyranosyl 17-hydroxy-6*E*,10*E*,14*Z*-(3*S*)-geranyllinalool 17-*O*-β-D-glucopyranosyl-(1→2)-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside.

Capsianoside VIII (**7**), an amorphous powder, $[\alpha]_D$ −20.4°, showed a peak due to $[M+Na]^+$ at *m/z* 1124 in positive FAB-MS and gave a molecular formula $C_{50}H_{84}O_{26}Na$ at *m/z* 1123.5348 by HR-FAB-MS. Compound **7** showed 4 mol of hexosyl and one deoxyhexosyl anomeric proton signals at δ 4.35 (1H, d, J=7.9 Hz), 4.39 (1H, d, J=7.9 Hz), 4.41 (1H, d, J=7.9 Hz), 4.63 (1H, d, J=7.9 Hz) and 4.75 (1H, s). The ^{13}C-NMR spectrum exhibited five anomeric carbon signals at δ 104.8, 104.6, 102.1, 101.1 and 99.3. The sequence of the sugar part was determined by the HMBC spectrum to provide the following connections: correlations between the signal at δ_H 4.39 (1H, d, J=7.9 Hz) and the signal at δ_C 81.5 (aglycone C-3) indicating the Glc I H-1 to be linked to the C-3 of the aglycone; the signal at δ_H 4.41 (1H, d, J=7.9 Hz) and the signal at δ_C 81.0 (Glc I C-4) indicating the Glc III H-1 to be linked to the C-4 of the Glc I; the signal at δ_H 4.35 (1H, d, J=7.9 Hz) and the signal at δ_C 68.3 (aglycone C-17) indicating the Glc IV H-1 to be linked to the C-17 of the aglycone; the signal at δ_H 4.63 (1H, d, J=7.9 Hz) and the signal at δ_C 82.0 (Glc IV C-2) suggesting the two glucosyl moieties have a (1→2) linkage; the rhamnosyl H-1 to the C-6 of Glc IV was also determined by the HMBC correlation between the signal at δ_H 4.75 (1H, s) and the signal at δ_C 67.6 (Glc IV C-6). From the above evidence, **7** was determined as 3-*O*-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl 17-hydroxy-6*E*,10*E*,14*Z*-(3*S*)-geranyllinalool 17-*O*-β-D-glucopyranosyl-(1→2)-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside.

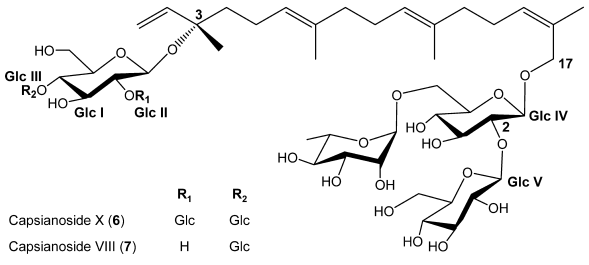
Table 3. ^{13}C-NMR Data of Compounds **6** and **7** in CD_3OD (125 MHz)

Position	6	7	Position	6	7	Position	6	7
1	116.0	115.8	Glc I-1	98.3	99.3	Glc IV-1	101.1	101.1
2	144.4	144.4	2	82.9	74.9	2	82.1	82.0
3	82.1	81.5	3	76.1	76.6	3	77.5	77.7
4	43.0	42.6	4	80.8	81.0	4	71.4	71.5
5	23.6	23.5	5	76.7	76.1	5	76.5	76.6
6	126.0	125.8	6	62.8	62.4	6	67.7	67.6
7	135.4	135.4	Glc II-1	106.1		Glc V-1	104.8	104.8
8	40.8	40.7	2	76.6		2	76.0	75.9
9	27.7	27.6	3	77.8		3	78.0	78.0
10	125.8	125.9	4	71.4		4	71.4	71.4
11	136.0	135.9	5	77.8		5	77.8	77.8
12	40.9	40.8	6	62.1		6	62.8	62.8
13	27.3	27.2	Glc III-1	104.6	104.6	Rha 1	102.1	102.1
14	131.2	131.2	2	74.9	74.9	2	72.2	72.1
15	132.4	132.4	3	78.0	78.0	3	72.4	72.4
16	21.9	22.0	4	71.5	71.4	4	74.0	74.0
17	68.3	68.3	5	78.4	78.2	5	69.7	69.7
18	16.3	16.2	6	62.4	62.2	6	18.1	18.1
19	16.3	16.1						
20	23.4	23.3						



	R_1	R_2	R_3	R_4
Capsianoside II* (**1**)	Glc	H	H	H
Capsianoside XIII (**2**)	H	Glc	H	H
Capsianoside XV (**3**)	H	Glc	**OH**	H
Capsianoside IX (**4**)	Glc	Glc	H	H
Capsianoside XVI (**5**)	Glc	Glc	H	**Glc**

* : Revised structure



	R_1	R_2
Capsianoside X (**6**)	Glc	Glc
Capsianoside VIII (**7**)	H	Glc

Fig. 3. Structures of Capsianosides

Experimental

The optical rotations were measured with a JASCO DIP-1000 (l=0.5) automatic digital polarimeter. ^{1}H- and ^{13}C-NMR spectra were measured with JEOL-α-500 and JMX-GX 400 NMR spectrometers, and chemical shifts are given on a δ (ppm) scale with tetramethylsilane as an internal standard. The FAB-MS were measured with a JEOL JMS-DX303HF spectrometer and taken in a glycerol matrix containing NaI. HR-ESI-MS were measured with a JEOL JMS T-100LP spectrometer. TLC was performed on silica gel plates (Kieselgel 60 F_{254}, Merck) and RP C_{18} silica gel plates (Merck). The spots on TLC were visualized by UV light (254/366 nm) and sprayed with 10% H_2SO_4, followed by heating. Column chromatography was carried out on a Diaion HP-20 (Mitsubishi Chemical Ind.), ODS (Wako Pure Chemical Industries, Ltd., Fuji Silysia Chemical, Ltd., Japan), and silica gel 60 (spherical, 40—100 mm, and 230—400 mesh ASTM ; crushed, 40—63 mm, and 230—400 mesh ASTM, Kanto Chemical Co. Inc.).

Plant Material The fruits of Paprika *Capsicum annuum* L. var. *grossum* Bailey and Jalapeño *Capsicum annuum* L. var. *annuum* were harvested at Nagao Farm, Kumamoto prefecture in Japan.

Extraction and Isolation The fresh Paprika (3.15 kg) was extracted successively with 100% MeOH (three times). After evaporation of MeOH *in vacuo*, the residue (179.55 g) was suspended in water and subjected to Diaion HP-20 using H_2O and MeOH (100%). The fraction (5.39 g) eluted with MeOH was subjected to ODS (Wako) column chromatography (eluted with 50—80% MeOH). The fractions eluted with MeOH/H_2O were subsequently subjected to silica gel column chromatography ($CHCl_3$/MeOH/H_2O solvent system, 7 : 3 : 0.3—6 : 4 : 1), and ODS (Chromatorex) column chromatography (MeOH/H_2O solvent system, 55—80% MeOH). From the Paprika, capsianoside II (87.2 mg), IX (183.6 mg), X (25.8 mg), XIII (28.6 mg), XV (28.5 mg) and XVI (16.8 mg) were obtained.

The fresh Jalapeño (6.8 kg) was extracted with 100% MeOH. After evaporation of MeOH *in vacuo*, the residue (252 g) was suspended in water and subjected to Diaion HP-20 using H_2O, MeOH (100%) and acetone. The fraction (5 g from 19.11 g) eluted with MeOH was subjected to ODS (Wako) column chromatography (eluted with 50—100% MeOH). The fractions eluted with MeOH/H_2O were subsequently subjected to silica gel column chromatography ($CHCl_3$/MeOH/H_2O solvent system, 9 : 1 : 0.1—6 : 4 : 1). From the Jalapeño, we obtained capsianoside II (55.8 mg), VIII (42.1 mg), IX (420.5 mg), X (25.9 mg), XIII (87.8 mg).

Capsianoside II (1) An amorphous powder, $[\alpha]_D^{25}$ −31.2° (c=0.19, MeOH), positive FAB-MS m/z: 1108 $[M+Na]^+$, 962, 654; positive HR-FAB-MS $[M+Na]^+$ m/z 1107.5374 (Calcd for $C_{50}H_{84}O_{25}Na$, 1107.5199). ^{1}H-NMR (CD_3OD) δ: 1.26 (3H, d, J=5.5 Hz, Rha H_3-6), 1.27 (3H, d, J=6.1 Hz, Rha H_3-6), 1.39, 1.61×2, 1.77 (each 3H, s, H_3-20, H_3-19, H_3-18, H_3-16), 1.61 (2H, m, H_2-4), 1.95—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.00 (m, sugar), 4.13, 4.32 (each 1H, d, J=11.6 Hz, H_2-17), 4.21, 4.47, 4.56 (each 1H, d, J=7.3 Hz, Glc H-1), 4.71, 4.83 (each 1H, s, Rha H-1), 5.12×2 (each 1H, m, H-6, H-10), 5.22, 5.23 (each 1H, br d, H_2-1), 5.40 (1H, t, H-14), 6.12 (1H, dd, J=11.0, 18.3 Hz, H-2).

Capsianoside XIII (2) An amorphous powder, $[\alpha]_D^{25}$ −38.9° (c=0.5, MeOH), positive FAB-MS m/z: 1108 $[M+Na]^+$; positive HR-FAB-MS $[M+Na]^+$ m/z 1107.5350 (Calcd for $C_{50}H_{84}O_{25}Na$, 1107.5199). ^{1}H-NMR (CD_3OD) δ: 1.26 (3H, d, J=6.1 Hz, Rha H_3-6), 1.27 (3H, d, J=6.7 Hz, Rha H_3-6), 1.37, 1.60×2, 1.77 (each 3H, s, H_3-20, H_3-19, H_3-18, H_3-16), 1.60 (2H, m, H_2-4), 1.95—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.00 (m, sugar), 4.13, 4.32 (each 1H, d, J=11.6 Hz, H_2-17), 4.20, 4.39, 4.41 (each 1H, d, J=7.9 Hz, Glc H-1), 4.71, 4.82 (each 1H, s, Rha H-1), 5.11×2 (each 1H, m, H-6, H-10), 5.20, 5.23 (each 1H, br d, H_2-1), 5.40 (1H, t, H-14), 5.93 (1H, dd, J=11.0, 17.7 Hz, H-2).

Capsianoside XV (3) An amorphous powder, $[\alpha]_D^{26}$ −36.9° (c=0.67, MeOH), positive FAB-MS m/z: 1123 $[M+Na]^-$, 961, 815, 669, 507; positive HR-FAB-MS $[M+Na]^+$ m/z 1123.5298 (Calcd for $C_{50}H_{84}O_{26}Na$, 1123.5149). ^{1}H-NMR (CD_3OD) δ: 1.26 (3H, d, J=6.8 Hz, Rha H_3-6), 1.27 (3H, d, J=5.5 Hz, Rha H_3-6), 1.38, 1.61, 1.77 (each 3H, s, H_3-20, H_3-18, H_3-16), 1.63 (2H, m, H_2-4), 1.99—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.01 (m, sugar), 4.09 (2H, s, H_2-19), 4.13, 4.32 (each 1H, d, J=11.6 Hz, H_2-17), 4.21 (1H, d, J=7.3 Hz, Glc H-1), 4.39, 4.41 (each 1H, d, J=7.9 Hz, Glc H-1), 4.71, 4.82 (each 1H, s, Rha H-1), 5.14 (1H, m, H-10), 5.20, 5.24 (each 1H, br d, H_2-1), 5.29 (1H, t, H-6), 5.40 (1H, t, H-14), 5.94 (1H, dd, J=11.0, 17.6 Hz, H-2).

Capsianoside IX (4) An amorphous powder, $[\alpha]_D^{25}$ −32.7° (c=0.5, MeOH), positive FAB-MS m/z: 1269 $[M+Na]^+$; positive HR-FAB-MS $[M+Na]^+$ m/z 1269.5649 (Calcd for $C_{56}H_{94}O_{30}Na$, 1269.5728). ^{1}H-NMR (CD_3OD) δ: 1.26 (3H, d, J=5.5 Hz, Rha H_3-6), 1.27 (3H, d, J=6.1 Hz, Rha H_3-6), 1.38, 1.61×2, 1.77 (each 3H, s, H_3-20, H_3-19, H_3-18, H_3-16), 1.61 (2H, m, H_2-4), 1.95—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.00 (m, sugar), 4.13, 4.32 (each 1H, d, J=11.6 Hz, H_2-17), 4.20, 4.41, 4.50, 4.58 (each 1H, d, J=7.9 Hz, Glc H-1), 4.71, 4.83 (each 1H, s, Rha H-1), 5.12×2 (each 1H, m, H-6, H-10), 5.21, 5.24 (each 1H, br d, H_2-1), 5.40 (1H, t, H-14), 6.13 (1H, dd, J=11.0, 18.3 Hz, H-2).

Capsianoside XVI (5) An amorphous powder, $[\alpha]_D^{26}$ −35.5° (c=0.88, MeOH), HR-ESI-MS $[M+Na]^+$ m/z 1431.6107 (Calcd for $C_{62}H_{104}O_{35}Na$, 1431.6256). ^{1}H-NMR (CD_3OD) δ: 1.27 (3H, d, J=6.1 Hz, Rha H_3-6), 1.33 (3H, d, J=6.1 Hz, Rha H_3-6), 1.38, 1.61×2, 1.77 (each 3H, s, H_3-20, H_3-19, H_3-18, H_3-16), 1.61 (2H, m, H_2-4), 1.99—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.10 (m, sugar), 4.13, 4.33 (each 1H, d, J=11.6 Hz, H_2-17), 4.20 (1H, d, J=7.3 Hz, Glc H-1), 4.41, 4.51, 4.58×2 (each 1H, d, J=7.9 Hz, Glc H-1), 4.72, 4.84 (each 1H, s, Rha H-1), 5.12×2 (each 1H, m,

H-6, H-10), 5.21, 5.24 (each 1H, br d, H_2-1), 5.40 (1H, t, H-14), 6.13 (1H, dd, J=10.4, 17.7 Hz, H-2).

Capsianoside X (6) An amorphous powder, $[\alpha]_D^{25}$ −16.7° (c=0.18, MeOH), positive FAB-MS m/z: 1286 $[M+Na]^+$; positive HR-FAB-MS $[M+Na]^+$ m/z 1285.5880 (Calcd for $C_{56}H_{94}O_{31}Na$, 1285.5677). ^{1}H-NMR (CD_3OD) δ: 1.27 (3H, d, J=6.1 Hz, Rha H_3-6), 1.38, 1.61×2, 1.79 (each 3H, s, H_3-20, H_3-19, H_3-18, H_3-16), 1.99—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.00 (m, sugar), 4.22 (1H, d, J=11.6 Hz, H-17), 4.29 (1H, d, J=11.0 Hz, H-17), 4.36, 4.42 (each 1H, d, J=7.9 Hz, Glc H-1), 4.50, 4.58 (each 1H, d, J=7.3 Hz, Glc H-1), 4.62 (each 1H, d, J=7.9 Hz, Glc H-1), 4.75 (1H, s, Rha H-1), 5.12×2 (each 1H, m, H-6, H-10), 5.21 (1H, br d, H-1), 5.24 (1H, br d, H-1), 5.39 (1H, t, H-14), 6.12 (1H, dd, J=11.0, 18.3 Hz, H-2).

Capsianoside VIII (7) An amorphous powder, $[\alpha]_D^{25}$ −20.4° (c=0.19, MeOH), positive FAB-MS m/z: 1124 $[M+Na]^+$; positive HR-FAB-MS $[M+Na]^+$ m/z 1123.5348 (Calcd for $C_{50}H_{84}O_{26}Na$, 1123.5149). ^{1}H-NMR (CD_3OD) δ: 1.26 (3H, d, J=6.1 Hz, Rha H_3-6), 1.37, 1.60, 1.61, 1.79 (each 3H, s, H_3-20, H_3-19, H_3-18, H_3-16), 1.60 (2H, m, H_2-4), 1.99—2.20 (10H, m, H_2-5, H_2-8, H_2-9, H_2-12, H_2-13), 3.20—4.00 (m, sugar), 4.21 (1H, d, J=11.6 Hz, H-17), 4.29 (1H, d, J=11.0 Hz, H-17), 4.35, 4.39, 4.41, 4.63 (each 1H, d, J=7.9 Hz, Glc H-1), 4.75 (1H, s, Rha H-1), 5.12×2 (each 1H, m, H-6, H-10), 5.21 (1H, br d, H-1), 5.25 (1H, br d, H-1), 5.38 (1H, t, H-14), 5.93 (1H, dd, J=11.0, 17.7 Hz, H-2).

Acid Hydrolysis of 2 and 3 with 1 N HCl A solution of **2** (4.4 mg) and **3** (2.5 mg) in 1 N HCl was heated under reflux for 2 h. The reaction mixture eluted with H_2O and MeOH successively was subjected to Amberlite IRA-400. The aqueous layer was subjected to HPLC analysis under the following conditions: HPLC column, COSMOSIL Sugar-D, 4.6 mm i.d.×250 mm (Nacalai Tesque, Inc., Tokyo, Japan); detector, JASCO OR-2090; pump, JASCO PU-2080; solvent, 80% CH_3CN; column oven, Co-2060 plus; column temperature, 30 °C. Identification of D-glucose and L-rhamnose in the aqueous layer was carried out by comparison of retention time with those of an authentic sample: D-glucose, t_R 12.30 min; L-rhamnose, t_R 6.27 min.

References

1) Banga I., Szent-Gyorgyi A., *Biochem. J.*, **28**, 1625—1628 (1934).
2) Molnár P., Kawase M., Satoh K., Sohara Y., Tanaka T., Tani S., Sakagami H., Nakashima H., Motohashi N., Gyémánt N., Molnár J., *Phytotherapy Res.*, **19**, 700—707 (2005).
3) Maoka T., Mochida K., Kozuka M., Ito Y., Fujiwara Y. Hashimoto K., Enjo F., Ogata M., Nobukuni Y., Tokuda H., Nishino H., *Cancer Lett.*, **172**, 103—109 (2001).
4) Ochi T., Takaishi Y., Kogure K., Yamauti I., *J. Nat. Prod.*, **66**, 1094—1096 (2003).
5) Izumitani Y., Yahara S., Nohara T., *Chem. Pharm. Bull.*, **38**, 1299—1307 (1990).