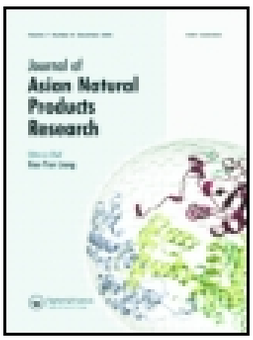


Journal of Asian Natural Products Research

ISSN: 1028-6020 (Print) 1477-2213 (Online) Journal homepage: http://www.tandfonline.com/loi/ganp20

Diterpenoid glycosides and monoterpenoid glycosides from the fruits of *Xanthium chinense*

Yan Cheng, Jiang Fu, Long Chen, Lu-Lu Li & Jing Qu

To cite this article: Yan Cheng, Jiang Fu, Long Chen, Lu-Lu Li & Jing Qu (2018): Diterpenoid glycosides and monoterpenoid glycosides from the fruits of *Xanthium chinense*, Journal of Asian Natural Products Research, DOI: 10.1080/10286020.2018.1526788

To link to this article: https://doi.org/10.1080/10286020.2018.1526788

Published online: 27 Dec 2018.

JOURNAL OF ASIAN NATURAL PRODUCTS RESEARCH
https://doi.org/10.1080/10286020.2018.1526788

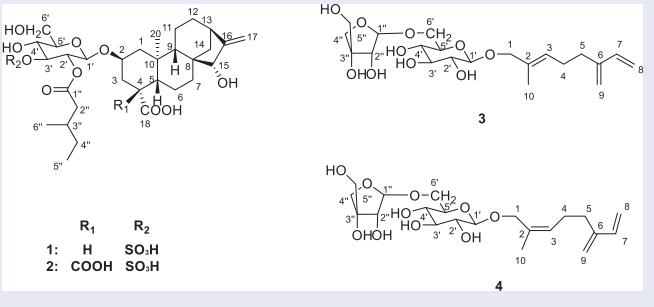
Diterpenoid glycosides and monoterpenoid glycosides from the fruits of *Xanthium chinense*

Yan Cheng, Jiang Fu, Long Chen, Lu-Lu Li and Jing Qu

State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China

ABSTRACT

Two new diterpenoid glycosides, fructusnoids D (**1**) and E (**2**), and two new monoterpenoid glycosides (**3, 4**), together with three known diterpenoid glycosides (**5-7**) and three known monoterpenoid glycosides (**8-10**), were isolated from the fruits of *Xanthium chinense*. Their structures were elucidated by spectrometric analyses.



ARTICLE HISTORY
Received 11 July 2018
Accepted 18 September 2018

KEYWORDS
Compositae; *Xanthium chinense*; diterpenoid glycoside; monoterpenoid glycoside

1. Introduction

The genus *Xanthium* in the family Compositae contains approximately 25 species distributed throughout northern and central America, Europe, Asia, and northern Africa. Chemical investigations of the genus have resulted in the isolation of sesquiterpenoids, diterpenoid glycosides, flavonoids, lignans, thiazinediones, and fatty oils. The fruits of the genus *Xanthium* are commonly used in clinical medicine for the treatment of rheumatism, headache, stuffy nose, runny nose, nasosinusitis, etc. [1]. *Xanthium* has a certain toxicity, and excessive use or improper processing can lead to poisoning and even death [2–5]. Studies have shown that the main toxic components in the genus *Xanthium* are two water-soluble glycosides, carboxyatractyloside and atractyloside, which were controlled by limiting in the Chinese Pharmacopoeia [6–8].

CONTACT Jing Qu qujing@imm.ac.cn State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China.

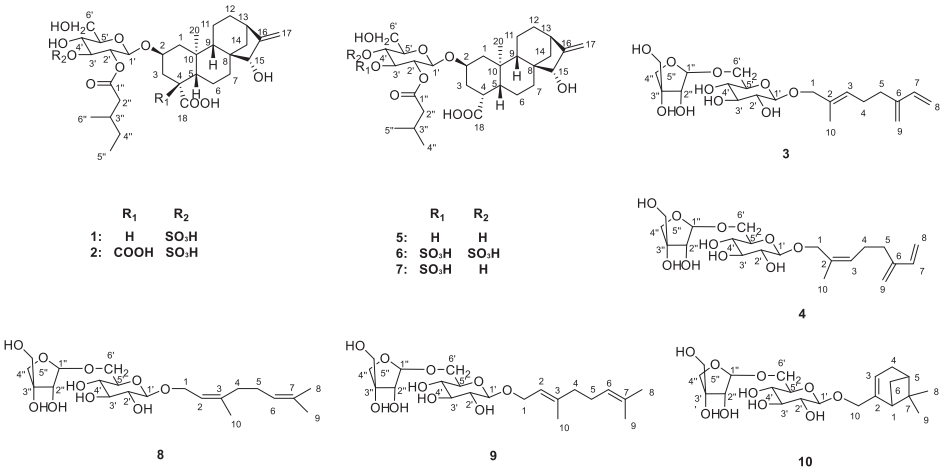


Figure 1. Structures of compounds **1**–**10.**

X. chinense Mill. is a widely distributed species in China. To extend the resources of *Xanthium* in clinical medicine and make use of *Xanthium* more safely, a study on the *n*-BuOH extraction of the fruits of *X. chinense* was carried out. Two new diterpenoid glycosides, fructusnoids D (**1**) and E (**2**), and two new monoterpenoid glycosides (**3**, **4**), together with three known diterpenoid glycosides (**5**-**7**) and three known monoterpenoid glycosides (**8**-**10**), were isolated from the fruits of *X. chinense* (Figure 1). Herein, we report the isolation and structure elucidation of the new diterpenoid glycosides and monoterpenoid glycosides.

2. Results and discussion

Compound **1** was obtained as a white powder. The HRESIMS $[M\text{-}H]^-$ ion at *m/z* 659.2750 indicated a molecular formula of $C_{31}H_{48}O_{13}S$, suggesting the presence of a sulfate group. The ^{1}H NMR spectrum (Table 1) proved the presence of three methyl groups at δ_H 1.11 (3H, s, H-20), 1.06 (3H, d, $J = 5.6$ Hz, H-6″), 0.85 (3H, t, $J = 6.4$ Hz, H-5″) and two olefinic protons at δ_H 5.47 (1H, s, H-17a) and 5.17 (1H, s, H-17b). The ^{13}C NMR data (Table 2) indicated two carbonyls (δ_C 178.1, 173.1), two olefinic carbons (δ_C 161.2, 108.2), eight oxygenated carbons (δ_C 100.7, 82.9, 82.1, 78.2, 73.5, 73.0, 71.3, 62.3), two sp^3 quaternary carbons (δ_C 48.5, 41.3), five sp^3 methines (δ_C 53.6, 49.9, 44.5, 43.1, 32.3), nine sp^3 methylenes (δ_C 48.2, 42.4, 36.9, 36.2, 35.6, 33.1, 29.9, 26.7, 18.8), and three methyls (δ_C 19.9, 17.3, 11.7). These data implied that the basic skeleton of compound **1** was a diterpene glycoside with an additional substituent group [9]. The correlations in the ^{1}H-^{1}H COSY spectrum (Figure 2(a)) suggested the existence of a sugar and 3-methylpentanoyl [10] in compound **1**, as did the HMBC correlations from H-2″ to C-1″ (δ_C 173.1), C-3″ (δ_C 32.3), and C-4″ (δ_C 29.9); from H-6″ to C-2″ (δ_C 42.4) and C-4″; and from H-5″ to C-3″ and C-4″. The 3-methylpentanoyl group and sulfonic group were attached to C-2′ and C-3′, respectively, indicated by the HMBC correlations from H-2′ to C-1″, together with the downfield shifts of H-3′ (δ_H 5.32) and C-3′ (δ_C 82.1). The rest of the ^{1}H-^{1}H COSY correlations established two spin systems C (1) H_2- C (2) H- C (3) H_2- C (4) H- C

Table 1. ^{1}H NMR spectral data of compounds **1**, **2**, **5-7** (δ in ppm, J in Hz).

	δ_H				
No.	**1**[a]	**2**[b]	**5**[b]	**6**[b]	**7**[a]
1	2.49–2.51, m	2.33–2.35, m	2.31, dd (12.6,3.6)	2.23–2.25, m	2.49, d (11.2)
	1.00–1.02, m	0.81, t (12.0)	0.76, t (12.0)	0.69, t (12.0)	1.00 , overlap
2	4.82–4.84, m	4.16–4.18, m	4.21–4.23, m	4.44–4.46, m	4.81–4.83, m
3	2.93–2.95, m	2.63–2.65, m	2.38–2.40, m	2.42–2.45, m	2.93, d (11.2)
	1.55, overlap	1.30, overlap	1.20–1.22, m	1.07–1.09, m	1.51–1.53, m
4	2.86–2.88, m		2.65–2.67, m	2.46–2.48, m	2.85–2.87, m
5	1.55, overlap	1.80–1.82, m	1.42–1.51, overlap	1.30–1.32, m	1.53–1.55, m
6	2.31–2.33, m	1.85–1.87, m	1.93–1.95, m	1.98–2.00, m	2.31–2.33, m
	1.91, d (12.0)	1.67, overlap	1.63 overlap	1.69–1.71, m	1.91, d (12.0)
7	2.16, overlap	1.67, overlap	1.67–1.69, m	1.93–1.95, m	2.16, d (13.6)
	1.81, overlap	1.44–1.46, m	1.42–1.51, overlap	1.35–1.37, m	1.79–1.81, m
9	1.11, overlap	1.12, d (7.8)	1.05 d (7.8)	1.00, d (7.8)	1.11, overlap
11	1.47, overlap (2H)	1.64, overlap	1.63, overlap	1.60–1.62, m	1.45–1.47, m (2H)
		1.49, overlap	1.42–1.51, overlap	1.47–1.49, m	
12	1.47, overlap	1.64 , overlap	1.63, overlap	1.62–1.64, m	1.42–1.44, m
	1.38–1.40, m	1.49, overlap	1.42–1.51, overlap	1.44–1.46, m	1.38–1.40, m
13	2.68–2.70, m	2.70–2.72, m	2.70–2.72, m	2.68–2.70, m	2.68–2.70, m
14	1.81, overlap	2.87–2.89, m	1.87, d (12.0)	1.65–1.67, m	1.81–1.83, m
	1.59–1.61, m	1.38–1.40, m	1.39, dd (11.4, 4.8)	1.40–1.42, m	1.59–1.61, m
15	4.10, br s	3.78, br s	3.76, s	3.75, br s	4.10, br s
17	5.46–5.48, s	5.18, s	5.18, s	5.17, s	5.48, s
	5.17, s	5.08, s	5.07, s	5.06, s	5.17, overlap
20	1.11, s (3H)	1.01, s (3H)	0.99, s (3H)	1.08, s (3H)	1.11, s (3H)
1′	5.21, d (7.2)	4.69, d (7.8)	4.60, d (7.8)	4.83, overlap	5.17, overlap
2′	5.60–5.62, m	4.77, dd (9.0, 7.8)	4.68, dd (9.6, 8.4)	4.83, overlap	5.60–5.62, m
3′	5.31–5.33, m	4.40, t (9.0)	3.50, t (9.0)	4.56, t (9.6)	5.31–5.33, m
4′	4.33–4.35, m	3.61, t (9.6)	3.38, t (9.0)	4.34, t (9.6)	4.29–4.31, m
5′	3.83–3.85, m	3.39–3.41, m	3.33, overlap	3.51–3.53, m	3.83–3.85, m
6′	4.25–4.37, m	3.86, dd (12.0, 2.4)	3.86, dd (12.0, 2.4)	3.90, d (2H, 3.0)	4.35, d (10.4)
	4.20–4.22, m	3.71, dd (12.0, 5.4)	3.70, dd(12.0,5.4)		4.21, dd (12.0, 5.6)
2″	2.74–2.76, m	2.37–2.39, m	2.26, d (2H, 6.6)	2.27 d (2H, 7.2)	2.59–2.61, m (2H)
	2.52–2.54, m	2.17–2.19, m			
3″	2.16, overlap	1.85–1.87, m	2.10, sep	2.09, sep	2.36–2.38, m
4″	1.55, overlap	1.40–1.42, m	0.98, d (3H, 7.2)	0.96, d (3H, 6.6)	1.05, d (3H, 5.6)
	1.22–1.24, m	1.20–1.22, m			
5″	0.85, t (3H, 6.4)	0.90, t (3H, 7.2)	0.98, d (3H, 67.2)	0.96, d (3H, 6.6)	1.02, d (3H, 5.6)
6″	1.06, d (3H, 5.6)	0.94, d (3H, 6.6)			

[a]Data were measured at 800 MHz for ^{1}H and 200 MHz for ^{13}C in C_5D_5N; [b]Data were measured at 600 MHz for ^{1}H and 150 MHz for ^{13}C in CD_3OD.

(5) H- C (6) H_2- C (7) H_2 and C (9) H- C (11) H_2- C (12) H_2- C (13) H- C (14) H_2 (Figure 2(a)). Then, the *ent*-kaurane skeleton of the aglycon was confirmed by the HMBC correlations from H-6, H-11, and H-13 to C-8 (δ_C 48.5); from H-1, H-4, H-5, and H-11 to C-10 (δ_C 41.3); from H-17 to C-13 (δ_C 43.1) and C-15 (δ_C 82.9); and from H-15 to C-9 (δ_C 53.6), C-13, and C-14 (δ_C 36.9). C-18 was oxidized to a carboxyl group and C-19 was degraded, as suggested by the downfield shifted carbon of C-4 (δ_C 44.5) together with the HMBC correlations from H-3, H-4, and H-5 to C-18 (δ_C 178.1). One hydroxyl group was attached to C-15, indicated by the downfield shift of C-15. The HMBC correlation from H-2 to C-1′ indicated that the sugar moiety was located at C-2. Then, the planar structure of compound **1** was elucidated.

The configuration of compound **1** was elucidated based on NOE correlations (Figure 2(b)). The presence of the NOE correlation H-2/H-20 and the absence of H-4/H-2 and H-4/H-20 correlations indicated the relative configurations at C-2, C-4, and C-10, and the NOE correlations H-5/H-9, H-9/H-15, H-13/H-14a, and H-13/H-

Table 2. ^{13}C NMR spectral data of compounds **1**, **2**, **5-7** (δ in ppm).

	δ_C				
No.	**1**[a]	**2**[b]	**5**[b]	**6**[b]	**7**[a]
1	48.2	48.4	48.6	48.9	48.4
2	73.5	73.8	74.1	74.5	73.8
3	35.6	41.2	35.7	37.2	35.9
4	44.5	59.1	44.6	47.6	44.7
5	49.9	52.2	50.4	51.2	50.0
6	26.7	24.1	26.6	26.8	26.9
7	36.2	36.2	36.2	37.2	36.4
8	48.5	49.6	49.6	48.9	48.7
9	53.6	55.0	54.4	54.7	53.8
10	41.3	41.4	41.8	41.9	41.5
11	18.8	19.3	19.2	19.3	19.0
12	33.1	33.6	33.6	33.6	33.3
13	43.1	43.6	43.7	43.8	43.3
14	36.9	37.3	37.2	36.4	37.1
15	82.9	83.4	83.5	83.7	83.1
16	161.2	160.3	160.3	160.6	161.5
17	108.2	109.1	109.0	108.8	108.4
18	178.1	175.2	178.8	182.6	178.3
19		175.5			
20	17.3	17.8	17.2	17.5	17.5
1′	100.7	101.0	101.1	100.1	101.0
2′	73.0	72.6	75.1	73.4	73.2
3′	82.1	82.8	76.3	80.4	82.3
4′	71.3	70.9	71.7	75.3	71.4
5′	78.2	77.6	77.9	76.1	78.5
6′	62.3	62.3	62.6	62.3	62.5
1″	173.1	173.8	173.8	174.2	173.1
2″	42.4	42.6	44.4	44.5	44.5
3″	32.3	33.0	26.8	26.4	26.2
4″	29.9	30.5	22.9	23.1	23.3
5″	11.7	11.7	22.9	23.0	23.3
6″	19.9	19.8			

[a]Data were measured at 800 MHz for ^{1}H and 200 MHz for ^{13}C in C_5D_5N, [b]Data were measured at 600 MHz for ^{1}H and 150 MHz for ^{13}C in CD_3OD.

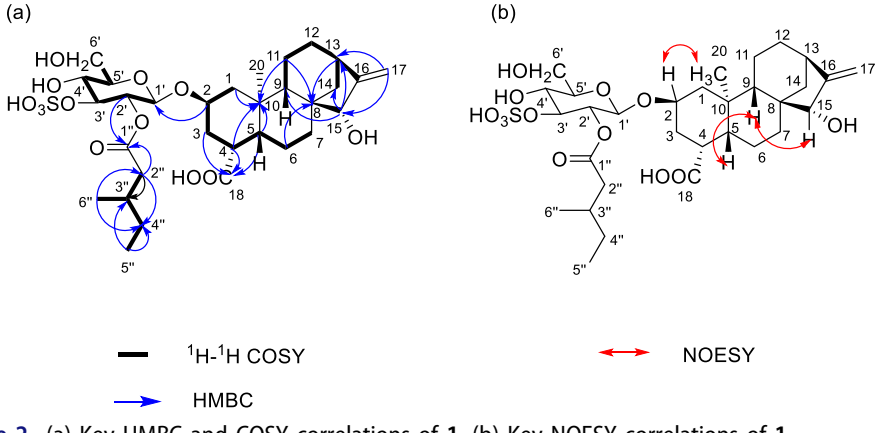


Figure 2. (a) Key HMBC and COSY correlations of **1**. (b) Key NOESY correlations of **1**.

14b indicated the relative configurations at C-5, C-9, C-15, C-8, and C-13, as well as the chair conformation of ring C. The anomeric proton signal (δ_H 5.21, d, $J = 7.2$ Hz) indicated that the sugar of compound **1** was attached via a β-linkage. Basic and acid

Table 3. ^{1}H NMR (C_5D_5N, 800MHz) spectral data of compounds **3**, **4**, **8-10** (δ in ppm, J in Hz).

	δ_H				
No	**3**	**4**	**8**	**9**	**10**
1	4.50, d (12.0) 4.29, overlap	4.54, overlap 5.40–5.42, m	4.66, dd (11.2, 6.4) 4.47. dd (11.2, 8.0)	4.63, dd (11.2, 6.4) 4.47 dd (11.2, 8.0)	2.12–2.14, m
2			5.57–5.59, m	5.56–5.58, m	
3	5.65–5.67, m	5.42–5.44, m			5.60–5.62, m
4	2.22–2.26, overlap (2H)	2.40–2.42, m 3.32, overlap	2.21–2.23, m 2.13, overlap	2.00–2.02, m (2H)	2.27–2.29, m 2.21–2.23, m
5	2.22–2.26, overlap (2H)	2.32, overlap 2.25–2.27, m	2.13, overlap (2H)	2.06–2.08, m (2H)	1.94–1.96, m
6			5.20–5.22, m	5.14–5.16, m	2.34–2.36, m 1.21–1.23, m
7	6.43–6.45, m	6.44–6.46, m			
8	5.27–5.29, m 5.08, overlap	5.37–5.39, m 5.05–5.11, overlap	1.59, s (3H)	1.65, s (3H)	1.18, s (3H)
9	5.08, overlap (2H)	5.05–5.11, overlap (2H)	1.68, s (3H)	1.55, s (3H)	0.83, s (3H)
10	1.76, s (3H)	1.94, s (3H)	1.69, s (3H)	1.71, s (3H)	3.94–3.96, m 4.13, overlap
1′	4.85, d (8.0)	4.81, d (7.2)	4.89, d (8.0)	4.88, d (7.2)	4.86, d (8.0)
2′	4.04–4.06, m	4.00–4.06, overlap	4.04–4.06, m	4.04–4.06, m	4.02–4.06, overlap
3′	4.22–4.24, m	4.23, t (8.8)	4.24, t (8.0)	4.24, t (8.0)	4.23, overlap
4′	4.02, overlap	4.00–4.06, overlap	4.02–4.04, m	4.02, overlap	4.02–4.06, overlap
5′	4.02, overlap	4.00–4.06, overlap	4.03–4.05, m	4.02, overlap	4.02–4.06, overlap
6′	4.58, d (11.2) 4.13–4.15, m	4.54, overlap 4.13–4.15, m	4.59, d (11.2) 4.12, dd (11.2, 6.4)	4.60, d (11.2) 4.12, dd (11.2, 4.8)	4.58, d (11.2) 4.13, overlap
1″	5.72, s	5.69, s	5.68, s	5.70, s	5.68, s
2″	5.11, s	5.05–5.11, overlap	5.10, s	5.10, s	5.10, s
4″	4.31, s (2H)	4.30 s (2H)	4.30, s (2H)	4.29, s (2H)	4.29, s (2H)
5″	4.19, d (11.2) 4.06–4.08, m	4.19, d (12.0) 4.00–4.06, overlap	4.19, d (12.0) 4.05–4.06, m	4.19, d (12.0) 4.05–4.07, m	4.19, overlap 4.02–4.06, overlap

hydrolysis of **1** afforded D-glucose, which was identified by GC analysis. Compound **1** had a similar structure to that of the previously known compound, 4′- desulfated-atractyloside (**7**) [9], in which the 3- methylpentanoyl group takes the place of the 3-methylbuthanoyl in **7**. Thus, compound **1** was assigned as shown and named fructusnoid D.

Compound **2** was obtained as a white powder. The molecular formula of **2** was established as $C_{32}H_{48}O_{15}S$ based on HRESIMS data at m/z 703.2644 $[M\text{-}H]^-$ in conjunction with the ^{13}C NMR data, containing one more carbon and two more oxygen atoms than compound **1**. The ^{1}H NMR spectrum (Table 1) proved the presence of three methyl groups at δ_H 1.01 (3H, s, H-20), 0.94 (3H, t, $J = 6.6$ Hz, H-6″), 0.90 (3H, d, $J = 7.2$ Hz, H-5″) and two olefinic protons at δ_H 5.18 (1H, s, H-17a) and 5.08 (1H, s, H-17b). The ^{13}C NMR data (Table 2) indicated three carbonyls (δ_C 175.5, 175.2, 173.8), two olefinic carbons (δ_C 160.3, 109.1), eight oxygenated carbons (δ_C 101.0, 83.4, 82.8, 77.6, 73.8, 72.6, 70.9, 62.3), three sp^3 quaternary carbons (δ_C 59.1, 49.6, 41.4), four sp^3 methines (δ_C 55.0, 52.2, 43.6, 33.0), nine sp^3 methylenes (δ_C 48.4, 42.6, 41.2, 37.3, 36.2, 33.6, 30.5, 24.1, 19.3), and three methyls (δ_C 19.8, 17.8, 11.7). These data were similar to those of **1**, with the exception of an additional carbonyl signal (C-19, δ_C 175.5) attached to C-4 (δ_C 59.1), which was confirmed by the downfield shift of C-4 and the HMBC correlations from H-3/H-5 to C-18/C-19. Therefore, compound **2** had the same aglycone as that of 4′-desulfated-carboxyatractyloside [9].

Table 4. ^{13}C NMR (C_5D_5N, 200MHz) spectral data for compounds **3, 4, 8-10** (δ in ppm).

	δ_C				
No	**3**	**4**	**8**	**9**	**10**
1	75.4	67.6	65.8	65.7	43.9
2	133.1	133.1	122.9	121.6	145.7
3	128.6	130.0	140.6	140.8	120.5
4	27.2	26.8	32.7	40.1	31.8
5	31.7	32.1	27.4	26.9	41.4
6	147.1	146.4	124.9	124.9	32.1
7	139.7	139.6	132.1	131.8	38.3
8	114.0	114.1	18.1	26.1	26.6
9	116.8	116.9	26.1	18.0	21.4
10	14.7	22.3	23.9	16.8	72.1
1′	103.3	103.5	103.7	103.3	103.5
2′	75.5	75.3	75.4	75.4	75.4
3′	78.8	78.8	78.8	78.8	78.8
4′	72.0	72.1	72.1	72.2	72.1
5′	77.5	77.3	77.2	77.4	77.4
6′	68.2	67.9	68.1	68.1	68.0
1″	109.4	109.1	109.1	109.1	109.1
2″	86.2	86.2	86.3	86.2	86.2,
3″	92.2	92.0	92.0	92.0	92.0
4″	74.6	74.3	74.3	74.3	74.3
5″	64.7	64.7	64.6	64.6	64.6

The ^1H-^1H COSY, HSQC, and HMBC spectra again allowed unambiguous assignment of the structure of compound **2**, which differed from 4′-desulfated-carboxyatractyloside only in having a 3-methylpentanoyl ester moiety attached to C-2′ of the glycoside, and then, the structure of **2** was established as assigned and named fructusnoid E.

Compounds **3** and **4** were obtained as viscous solids. Their identical molecular formula, $C_{21}H_{34}O_{10}$, was determined by HRESIMS ions at *m/z* 445.2075 and *m/z* 445.2073 $[M\text{-}H]^-$, respectively. The ^1H and ^{13}C NMR data (Tables 3 and 4) for compounds **3** and **4** suggested that they possessed similar structures with the same sugar moiety. The doublet at δ_H 4.85 (1H, d, $J = 8.0$ Hz, H-1′) and a singlet at δ_H 5.72 (1H, s, H-1″) in the ^1H NMR spectrum revealed the presence of two sugar units, with the signals assigned as the anomeric protons of D-glucose and D-apiose, respectively, which was confirmed by further NMR spectral analysis and GC analysis after acid hydrolysis and derivatization. The downfield shift of C-6′ (δ_C 68.2 and 67.9 for **3** and **4**, respectively), and the HMBC correlations from H-1″ to C-6′, C-3″, and C-4″; from H-2″ to C-3″, C-4″, and C-5″; from H-5″ to C-3″ indicated that the apiofuranose was attached to C-6 of the glucopyranose. The β-linkage of the glucopyranose moiety was characterized by the coupling constant of the anomeric proton at δ_H 4.85 (1H, d, $J = 8.0$ Hz, H-1′). As a branched-chain sugar, apiofuranose can occur in four isomeric forms [11]. A β-D-apiofuranose moiety was confirmed by the anomeric proton signal (δ_H 5.72, s) and the NOE correlation between H-2″ and H-5″ [12]. Thus, it was finally determined that compounds **3** and **4** shared the same β-D-apiofuranosyl-(1→6)-β-D glucopyranosyl disaccharide moiety.

In addition to the signals of sugar, the ^1H NMR (Table 3) data for compound **3** showed one methyl proton at δ_H 1.76 (3H, s, H-10); three methylenes at δ_H 2.22-2.26 (4H, overlap, H-4, 5), 4.50 (1H, d, $J = 12.0$ Hz, H-1a), and 4.29 (1H, overlap, H-1b);

and six olefinic protons at δ_H 6.43-6.45 (1H, m, H-7), 5.65-5.67 (1H, m, H-3), 5.27-5.29 (1H, m, H-8a), 5.08 (1H, overlap, H-8b), and 5.08 (2H, overlap, H-9). The ^{13}C NMR data (Table 4) indicated six olefinic carbons (δ_C 147.1, 139.7, 133.1, 128.6, 116.8, 114.0), one oxygenated carbon (δ_C 75.4), two sp^3 methylenes (δ_C 31.7, 27.2), and a methyl (δ_C 14.7). The HMBC correlations from H-1 to C-2, C-3, and C-10; from H-5 to C-4, C-6, and C-9; from H-7 to C-5; and from H-9 to C-4 and C-7, and the absence of H-3/H-10 NOESY correlations indicated the *E* configuration at C-3/C-4, suggested that (*E*) -2- methyl-6- methyleneocta-2, 7- dien-1-ol [13] was present as the aglycone in compound **3**.

Compound **4** differed from compound **3** in a C-10 signal at δ_C 22.3 (Table 4), and the presence of H-3/H-10 NOESY correlations indicated the *Z* configuration at C-3/C-4, suggesting that (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol [14] was present as the aglycone in compound **4**.

The six known compounds **5-10** were identified as 3′,4′-desulfated-atractyloside (**5**) [15], atractyloside (**6**) [16], 4′-desulfated-atractyloside (**7**) [9], neryl 6-*O*-*β*-D-apiofuranosyl-*β*-D- glucopyranoside (**8**), geranyl 6-*O*-*β*-D-apiofuranosyl-*β*-D-glucopyranoside (**9**) [17], and myrtenol 6-*O*-*β*-D-apiofuranosyl-*β*-D-glucopyranoside (**10**) [18] by comparison of their NMR (Tables 1–4) and MS data with reported values.

3. Experimental

3.1. General experimental procedures

Optical rotations were measured on a P2000 automatic digital polarimeter (JASCO Corporation, Tokyo, Japan). IR spectra were recorded on a Nicolet 5700 FT-IR spectrometer (Thermo Electron Corporation, Waltham, MA, USA). ESIMS were measured on an Agilent 1100 Series LC/MSD trap mass spectrometer (Agilent Technologies Inc, Santa Clara, CA, USA). HRESIMS were measured on an Agilent Technologies 6250 Accurate-Mass Q-TOF LC/MS instrument (Agilent Technologies Inc, Santa Clara, CA, USA). NMR spectra were recorded on Bruker AV-III-500, Bruker AV IIIHD-600, or Bruker AV IIIHD-800 (Bruker Daltonics Inc, Berlin, German), spectrometers for ^1H, ^{13}C, ^1H-^1H COSY, HSQC, HMBC, and NOESY. Column chromatography was performed with polyamide (30–60 mesh, Jiangsu Changfeng Chemical Co., Jiangsu, Ltd, China), macroporous resin (D101, Tianjin Nankai Hecheng Science & Technology Co., Ltd, Tianjin, China), Sephadex LH-20 (Amersham Pharmacia Biotech AB, Uppsala, Sweden), and ODS (50 μm, 120 Å, Silicycle Inc., Quebec, Canada). Preparative HPLC was carried out on a Shimadzu LC-6AD instrument with an SPD-10A detector (Shimadzu Scientific Instruments Co. Ltd., Kyoto, Japan) using YMC-Pack ODS-A columns (250 × 20 mm, 5 μm, or 250 × 10 mm, 5 μm, YMC Co., Ltd, Kyoto, Japan). Two-dimensional high-throughput chromatography was carried out on a Sepiatec Sepbox-2D-2000 (Sepiatec GmbH, Berlin, Germany) with two preparative HPLC pumps (a gradient pump and a water pump) and two detectors (a UV detector and an ELSD detector) using 1 first separation column (150 × 32 mm), 6 second separation columns (250 × 16 mm), and 18 trap columns (30 × 32 mm).

3.2. Plant material

The fruits of *Xanthium chinense* (110 kg) were collected in Xining, Guangxi Province, China, in September 2013. The plant was identified by Associate Prof. Lin Ma (Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College). A voucher specimen (No. ID-S-2522) was deposited in the herbarium of the Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, China.

3.3. Extraction and isolation

A large amount of material (110 kg) was crushed and extracted with 95% EtOH (150 L; 2 h × 3) under reflux. After removal of the solvent under vacuum, the extractum was suspended in H_2O and then extracted with petroleum ether, CH_2Cl_2, EtOAc, and *n*-BuOH in turn. Each extract was evaporated under reduced pressure to result in petroleum ether-soluble (6 kg), CH_2Cl_2-soluble (450 g), EtOAc-soluble (250 g), and *n*-BuOH-soluble (450 g) extracts. The *n*-BuOH-soluble extract (250 g) was fractionated by macroporous resin (D101) column chromatography, eluting with H_2O, 10%, 30%, 50%, 70%, and 95% EtOH, to afford six fractions, A–F.

Fraction C (95.3 g) was fractionated by polyamide (30–60 mesh) column chromatography, eluting with H_2O, 50% and 95% EtOH, to afford three fractions, C1-C3. Fraction C1 (70 g) was separated by chromatography on an MCI column, eluting with EtOH-H_2O to yield six major fractions (C1-a–C1-f). Fraction C1-c (6.0g) was separated by an ODS column using a gradient of MeOH-H_2O as the eluent to obtain seven subfractions, C1-c-1–C1-c-7. Subsequent separation of C1-c-6 (768 mg) by a Sephadex LH-20 column (H_2O: MeOH, 5: 1) afforded subfraction C1-c-6-2 (142 mg), and then, *p*-HPLC (25% CH_3CN/H_2O, $\upsilon = 4$ ml/min, $\lambda = 210$ nm) of C1-c-6-2 gave compound **5** (t_R =35 min, 10 mg).

Fraction D (35.6 g) resulted in four subfractions (D1–D4) after being chromatographed over a silica gel column. Fraction D2 (2g) was subjected to 2D Sepbox high-throughput preparation to give 23 subfractions, D2-a–D2-w. Subsequent separation of D2-j (10 mg) by *p*-HPLC (30% CH_3CN/H_2O, $\upsilon = 4$ ml/min, $\lambda = 210$ nm) gave compounds **1** (t_R= 20 min, 2 mg) and **7** (t_R= 17 min, 3 mg). D2-i (20 mg) was purified by *p*-HPLC (25% CH_3CN/H_2O, $\upsilon = 4$ ml/min, $\lambda = 210$ nm) to afford compounds **8** (t_R= 23 min, 2 mg), **9** (t_R= 26 min, 2 mg), and **10** (t_R= 48 min, 2 mg). D2-r (10 mg) was purified by *p*-HPLC (23% CH_3CN/H_2O, $\upsilon = 4$ ml/min, $\lambda = 210$ nm) to afford compounds **3** (t_R= 23 min, 2 mg) and **4** (t_R= 21 min, 2 mg).

Fraction D3 (6 g) was fractionated by an ODS column using a gradient of MeOH-H_2O as the eluent to obtain 5 subfractions, D3-a-D3-e. D3-a was afforded as compound **6** (250 mg). D3-c was purified by *p*-HPLC (17% CH_3CN/H_2O, $\upsilon = 4$ ml/min, $\lambda = 210$ nm) to afford compound **2** (t_R= 21 min, 2 mg).

3.3.1. Fructusnoid D (1)

$C_{31}H_{48}O_{13}S$, white powder; $[\alpha]_D^{20}$ +64.0 (*c* 0.00169, H_2O); UV (H_2O) λ_{max} (log ε) 192 (4.02), 255 (3.14) nm; IR ν_{max} 3410, 2931, 1700, and 1076 cm^{-1}; for ^1H NMR

(C_5D_5N, 800MHz) and ^{13}C NMR (C_5D_5N, 200MHz) spectral data (see Tables 1 and 2); HRESIMS: *m/z* 659.2750 $[M\text{-}H]^-$ (calcd for $C_{31}H_{47}O_{13}S$, 659.2743).

3.3.2. Fructusnoid E (2)

$C_{32}H_{48}O_{15}S$, white powder; $[\alpha]_D^{20}$ +66.9 (*c* 0.00171, H_2O); UV (H_2O) λ_{max} (log ε) 192 (3.98), 255 (3.47) nm; IR ν_{max} 3441, 2981, 1718, 1250, 1079, 1453, 1379, 1042, and 1004 cm^{-1}; for ^1H NMR (CD_3OD, 600MHz) and ^{13}C NMR (CD_3OD, 150MHz) spectral data (see Tables 1 and 2); HRESIMS: *m/z* 703.2644 $[M\text{-}H]^-$ (calcd for $C_{32}H_{47}O_{15}S$, 703.2641).

3.3.3. (E)-2-Methyl-6-methyleneocta-2,7-dien-1-ol β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside (3)

$C_{21}H_{34}O_{10}$, viscous solid; $[\alpha]_D^{20}$ +54.8 (*c*, 0.00153 H_2O); UV (H_2O) λ_{max} (log ε) 192 (3.96), 220 (3.59), 255 (3.38) nm; IR ν_{max} 3394, 2936, 1684, 1436, 1207, 1139, 842, 803, and 724 cm^{-1}; for ^1H NMR (C_5D_5N, 800MHz) and ^{13}C NMR (C_5D_5N, 200MHz) spectral data (see Tables 3 and 4); HRESIMS: *m/z* 445.2075 $[M\text{-}H]^-$ (calcd for $C_{21}H_{33}O_{10}$, 445.2079).

3.3.4. (Z)-2-Methyl-6-methyleneocta-2,7-dien-1-ol β-D-apiofuranosyl-(1→6)-β-D glucopyranoside (4)

$C_{21}H_{34}O_{10}$, viscous solid; $[\alpha]_D^{20}$ +76.7 (*c* 0.00215, H_2O); UV (H_2O) λ_{max} (log ε) 192 (3.98), 220 (3.71), 255 (3.28) nm; IR ν_{max} 3406, 2938, 1682, 1433, 1201, 1137, 840, 802, and 723 cm^{-1}; for ^1H NMR (C_5D_5N, 800MHz) and ^{13}C NMR (C_5D_5N, 200MHz) spectral data (see Tables 3 and 4); HRESIMS: *m/z* 445.2073 $[M\text{-}H]^-$ (calcd for $C_{21}H_{33}O_{10}$, 445.2079).

3.4. Hydrolysis of compounds 1 and 3, and sugar identification

Compound **1** (1.5 mg) was added in 2 M NaOH (2 ml) and stirred for 5 h under room temperature. The solution was regulated by 0.5M HCl-H_2O to pH 2-3, and was purified by *p*-HPLC (25% CH_3CN/H_2O) to remove the side chain on sugar, then the compound that was removed the side chain on sugar was added in 2 M HCl (2 ml) and refluxed for 12 h under 90 °C. The solution was extracted three times with EtOAc. The water layer was dried to yield a residue. Then the residue was dissolved in pyridine (1 ml) and the L-cysteine methyl hydrochloride (2 mg) was added. The mixture was refluxed for 2 h under 60 °C, and dried by nitrogen and heated for 0.5 h under 80 °C. Then the N-trimethylsilylimidazole (1 ml) was added, and heated for 2 h under 60 °C. Finally, the mixture was added in H_2O (2 ml) and partitioned with n-hexane (2 ml) three times. The organic layer was combined and concentrated, used for GC analysis. The conditions of GC experiments: capillary column, HP-5 (60 m × 0.25 mm, with a 0.25 μm film, Dikma); detector, FID; injection temperature, 300 °C; detection temperature, 300 °C; initial temperature, 200 °C, risen to 280 °C at the rate of 10 °C/min, and sustained for 35 min, then declined to 200 °C at the rate of 40 °C/min and then sustained for 1 min; carrier, N_2. The D-glucose (t_R = 29.6 min) was determined by comparing with the standard sugar.

Compound **3** (1.5 mg) was added in 2 M HCl (2 ml) and conducted as the same way as compound **1**. D-glucose (t_R = 29.5 min) and D-apiofuranose (t_R = 20.3 min) were determined by comparing with the standard sugars.

Acknowledgements

We are grateful to the Department of Instrumental Analysis, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College for measuring the IR, NMR, MS, and HRESIMS.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This project was financially supported by the National Natural Science Foundation of China [grant number 81573321], the CAMS Innovation Fund for Medical Science [grant number 2017-I2M-3-010], Young Elite Scientists Sponsorship Program by CAST [YESS, grant number 2016QNRC001], and National Science and Technology Project of China [No. 2014ZX09304306-001].

References

[1] National Pharmacopoeia Commission, *Pharmacopoeia of the People's Republic of China* (China Medical Science Press, Beijing, 2010), Part **I**. pp. 151.
[2] J. Wang, C.L. Mo, L. Que, G. Zhaori, and J. Yu, *Chin. Tradit. Herb. Drugs* **42**, 613 (2011).
[3] T.T. Zhang, L.C. Yan, J.N. Zhao, and L. Cheng, *Drug Evaluation Res.* **33**, 361 (2010).
[4] X.M. Zhang and C.H. Zhang, *Chin. J. Integr. Med.* **1**, 71 (2003).
[5] Q. Song and X.H. Meng, *Guid. J. Tradit. Chin. Med. Pharm.* **10**, 26 (2004).
[6] E. Pebay-Peyroula, C. Dahout-Gonzalez, R. Kahn, V. Trezeguet, G.J.M. Lauquin, and G. Brandolin, *Nature* **426**, 39 (2003).
[7] R. Duo, Y. Chen, Y.H. Liu, Z.F. Huang, Y.H. Liu, and J.H. Yi, *Chin. J. Chin. Mater. Med.* **37**, 2313 (2012).
[8] M.E. Krejci and D.A. Koechel, *Toxicology* **72**, 299 (1992).
[9] P. Sonia, P. Cosimo, N. De Tommasi, and F. De Simone, *Phytochemistry* **41**, 1357 (1996).
[10] H. Jiang, G.X. Ma, L. Yang, X.D. Xing, M.L. Yan, Y.Y. Zhang, Q.H. Wang, H.X. Kuang, and X.D. Xu, *Phytochem. Lett.* **18**, 192 (2016).
[11] T. Ishii and M. Yanagisawa, *Carbohydr. Res.* **313**, 189 (1998).
[12] K. Zhou, F. Zhao, Z. Liu, Y.L. Zhuang, L.X. Chen, and F. Qiu, *J. Nat. Prod.* **72**, 1563 (2009).
[13] S.N. Garg, R. Charles, and S. Kumar, *Fitoterapia* **70**, 472 (1999).
[14] R. Granger, J. Passet, and J.P. Girard, *Phytochemistry* **11**, 2301 (1992).
[15] H. Richter and G. Spiteller, *Chem. Ber.* **111**, 3506 (1978).
[16] R.J. Cole, B.P. Stuart, J.A. Lansden, and R. H. Cox, *J. Agric. Food Chem.* **28**, 1330 (1980).
[17] Y. Sekiwa, A. Kobayashi, K. Kubota, and M. Takenaka, *Nat. Prod. Lett.* **15**, 267 (2001).
[18] T. Kikuchia, J. Zhang, Y. Huang, K. Watanabe, K. Ishii, A. Yamamoto, M. Fukatsu, R. Tanaka, and T. Akihisa, *Chem. Biodivers.* **9**, 1221 (2012).