Phytochemistry Letters 44 (2021) 7–13

Contents lists available at ScienceDirect

Phytochemistry Letters

journal homepage: www.elsevier.com/locate/phytol

Four new triterpene glycosides from the aerial parts of *Chenopodium album* and their cytotoxic activity

Phan Thi Thanh Huong [a,d], Do Thi Trang [a,d], Vu Kim Thu [c,d], Nguyen Thi Mai [d], Nguyen Xuan Nhiem [a,b,d], Pham Hai Yen [a,b,d], Bui Huu Tai [a,b,d], Phan Van Kiem [a,b,d,*]

[a] Institute of Marine Biochemistry, Vietnam Academy of Science and Technology (VAST), 18 Hoang Quoc Viet, Caugiay, Hanoi, Viet Nam
[b] Graduate University of Science and Technology, VAST, 18 Hoang Quoc Viet, Caugiay, Hanoi, Viet Nam
[c] Hanoi University of Mining and Geology, 18 Pho Vien, Dong Ngac, Bac Tu Liem, Hanoi, Viet Nam
[d] University of Transport and Communications, Hanoi, Viet Nam

ARTICLE INFO

Keywords:
Chenopodium album
Chenoalbumosides
Triterpene glycoside
Cytotoxicity

ABSTRACT

Six triterpene glycosides including four new ones, chenoalbumosides A–D (**1**-**4**), and two known triterpene glycosides, chikusetsusaponin IVa (**5**) and ginsennoside Ro (**6**), were isolated from the aerial parts of *Chenopodium album* Linn (Chenopodiaceae). Their structures were determined by extensive analyses of HR-ESI-MS and NMR spectral data. The cytotoxic activity of compounds **1**-**6** were evaluated against HT-29, SW480, AGS, and MKN7 cell lines. Unfortunately, our results indicated that none of these compounds were potential cytotoxic agents, showing IC_{50} values over 50 μM.

1. Introduction

The genus *Chenopodium* includes a variety of weedy herbs (more than 200 species) native to Europe, Asia, and both North and South America (Smith, 2006). Phenolics, flavonoids, saponins, ecdysteroids and triterpenoids are the major classes of phytoconstituents of this genus (Zlatina et al., 2009). *Chenopodium album* Linn (Chenopodiaceae) is an herb, that is distributed in many tropical countries in Asia. *C. album* improves appetite and acts as anthelmintic, laxalive, diuretic and tonic agents (Neerja et al., 2007). The ethanolic extract of *C. album* fruits exhibits antipruritic and antinociceptive activities (Dai et al., 2002). A phytochemical investigation of *C. album* has suggested the presence of triterpene saponins (Lavaud et al., 2000), phenols (Chludil et al., 2008; Cutillo et al., 2003, 2004; Cutillo et al., 2006; Horio et al., 1993; Zlatina et al., 2009), ecdysteroids (Dellagreca et al., 2005a, b), and other constituents (Dellagreca et al., 2005a, b; Gonzalez et al., 1998; Jhade et al., 2009). In Vietnam, *C. album* is a common plant in the Red River Delta and the Mekong River Delta. It is commonly used as food and folk medicine for the treatment of snakebites, diarrhea, and gonorrhea (Chi, 2012). In the course of our search for bioactive compounds from nature, we have extensively the performed phytochemical studies on the plant *Chenopodium album*. This paper reports the isolations and structural determinations of four new (**1**-**4**) and two known (**5** and **6**) triterpene glycosides from the aerial parts of *C. album*. In addition, the cytotoxic activity of these compounds was also evaluated in several human cancer cell lines including HT-29 (colorectal adenocarcinoma), SW480 (colon adenocarcinoma), MKN7 (gastric carcinoma), and AGS (stomach gastric adenocarcinoma).

2. Results and discussion

Compound **1** was obtained as a colorless amorphous powder. Its molecular formula was established as $C_{36}H_{56}O_{10}$ by the HR-ESI-MS *quasi*-molecular ion peaks at *m/z* 683.3522 $[M+^{35}Cl]^-$ (Calcd. for $[C_{36}H_{56}O_{10}^{35}Cl]^-$, 683.3562, Δ = -5.9 ppm), *m/z* 685.3511 $[M+^{37}Cl]^-$ (Calcd. for $[C_{36}H_{56}O_{10}^{37}Cl]^-$, 685.3533, Δ = -3.2 ppm), indicating 11 degrees of unsaturation. Its IR spectrum exhibited the presence of hydroxy (3430 cm^{-1}), carbonyl (1712 cm^{-1}) and C—OC— (1090, 1047 cm^{-1}) groups. The ^1H NMR spectrum of **1** contained signals corresponding to six quaternary methyl groups [δ_H 0.90, 1.08, 1.13, 1.17, 1.19, and 1.45 (each, 3H, s)], one methoxy group (δ_H 3.72, 3H, s), one olefinic proton (δ_H 5.60, s), one anomeric proton [δ_H 4.30 (d, *J* = 7.0 Hz)], two oxygenated methylene groups [δ_H 3.04, 3.31, 3.54, 3.86 (each, 1 H)], and five oxygenated methine groups [δ_H 3.20, 3.53, 3.59, 3.82, 4.18 (each, 1 H)]. The ^{13}C NMR spectrum of **1** revealed signals corresponding to 36 carbon atoms which were divided into 9 nonprotonated carbons,

* Corresponding author at: Institute of Marine Biochemistry, Vietnam Academy of Science and Technology, 18-Hoang Quoc Viet, Caugiay, Hanoi, Viet Nam.
E-mail address: phankiem@yahoo.com (P.V. Kiem).

https://doi.org/10.1016/j.phytol.2021.05.004
Received 15 March 2021; Received in revised form 9 May 2021; Accepted 10 May 2021
Available online 15 May 2021

10 methine carbons, 10 methylene carbons, and 7 methyl carbons, identified by HSQC spectra. Of these, two C═O groups (δ_C 202.5 and 178.2), one double bond (δ_C 129.2 and 172.1), six methyl carbons (δ_C 17.0 × 2, 19.1, 23.1, 24.0, 28.5), one methoxy (δ_C 52.3) and one arabinopyranose sugar [δ_C 107.1, 74.3, 72.8, 69.5 (4 x CH), and 66.3 (CH_2)] were indicated (Mizui et al., 1988). The above evidence suggested that compound **1** was a methyl ester of an oleanane triterpene glycoside. All NMR assignments were performed by HSQC, H-H COSY, and HMBC spectra as shown in Table 1 (Fig. 2). In the HMBC spectrum, H-23 (δ_H 1.08) and H-24 (δ_H 0.90) correlated with C-3 (δ_C 90.3)/C-4 (δ_C 40.5)/C-5 (δ_C 56.5), and H-28 (δ_H 3.04/3.31) correlated with C-16 (δ_C 25.5)/C-17 (δ_C 38.0)/ C-18 (δ_C 44.6)/C-22 (δ_C 39.5) confirming that the oxygenated methine carbon and the oxygenated methylene carbon were at C-3 and C-28, respectively. The hydroxy group was attached to C-21, and the methyl carboxylate group was at C-29, as indicated by the ^{1}H-^{1}H COSY cross peaks from H-18 (δ_H 2.20) to H_a-19 (δ_H1.66), from H-21 (δ_H 4.18) to H-22 (δ_H 1.65/1.71), by the HMBC correlations from H-30 (δ_H 1.19) to C-19 (δ_C 37.4)/C-20 (δ_C 49.4)/C-21 (δ_C 70.4)/C-29 (δ_C 178.2), from methoxy protons (δ_H 3.72) to C-29, and from the NOESY cross peaks between H_a-28 (δ_H 3.04)/H-18 (δ_H 2.20)/H-30 (δ_H 1.19). The

Table 1
^{1}H and ^{13}C NMR spectral data for compounds **1**-**4** in CD_3OD.

No.	**1**		**2**		**3**		**4**	
	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)
1	40.2	1.04 (m)/2.72 (m)	40.0	1.04 (m)/2.74 (m)	40.2	1.03 (m)/2.74 (m)	40.2	1.03 (m)/2.72 (m)
2	27.0	1.77 (m)/1.86 (m)	27.0	1.78 (m)/1.86 (m)	26.9	1.78 (m)/1.85 (m)	26.9	1.79 (m)/1.87 (m)
3	90.3	3.20 (dd, 11,5, 5.0)	90.6	3.20 (dd, 11,5, 5.0)	90.9	3.21 (dd, 11,5, 5.0)	90.9	3.21 (dd, 11,5, 5.0)
4	40.5	–	40.7	–	40.7	–	40.7	–
5	56.5	0.82 (d, 11.5)	56.3	0.82 (d, 11.5)	56.4	0.82 (d, 11.5)	56.4	0.82 (d, 11.5)
6	18.5	1.48 (m)/1.65 (m)	18.5	1.48 (m)/1.65 (m)	18.4	1.48 (m)/1.62 (m)	18.5	1.48 (m)/1.63 (m)
7	33.9	1.45 (m)/1.75 (m)	33.8	1.46 (m)/1.74 (m)	33.8	1.45 (m)/1.73 (m)	33.7	1.47 (m)/1.76 (m)
8	46.5	–	46.5	–	46.5	–	46.8	–
9	63.1	2.49 (s)	63.0	2.48 (s)	63.0	2.48 (s)	63.1	2.48 (s)
10	38.1	–	38.1	–	38.1	–	38.1	–
11	202.5	–	202.1	–	202.5	–	202.4	–
12	129.2	5.60 (s)	129.2	5.59 (s)	129.1	5.58 (s)	129.1	5.58 (s)
13	172.1	–	171.9	–	172.1	–	171.9	–
14	44.9	–	44.8	–	44.8	–	44.6	–
15	26.9	1.16 (m)/1.76 (m)	26.9	1.17 (m)/1.77 (m)	26.9	1.18 (m)/1.75 (m)	26.9	1.24 (m)/1.81 (m)
16	25.5	1.51 (m)/2.46 (dd, 4.0, 14.0)	25.6	1.52 (m)/2.46 (dd, 4.0, 14.0)	25.5	1.50 (m)/2.44 (dd, 4.0, 14.0)	22.5	1.40 (m)/2.00 (dd, 4.0, 13.5)
17	38.0	–	38.0	–	38.0	–	37.7	–
18	44.6	2.20 (dd, 13.5, 4.0)	44.6	2.21 (dd, 13.5, 4.0)	44.5	2.20 (dd, 13.5, 4.0)	44.9	2.19 (dd, 13.5, 4.0)
19	37.4	1.66 (m)/2.20 (m)	37.4	1.65 (m)/2.20 (m)	37.4	1.65 (m)/2.20 (m)	42.1	1.80 (m)/1.91 (m)
20	49.4	–	49.4	–	49.9	–	45.3	–
21	70.4	4.18 (brt, 3.0)	70.3	4.18 (brt, 3.0)	70.3	4.18 (t, 3.0)	31.6	1.40 (m)/2.20 (m)
22	39.5	1.65 (m)/1.71 (m)	39.4	1.65 (m)/1.72 (m)	39.4	1.65 (m)/1.72 (m)	33.3	1.40 (m)/1.66 (m)
23	28.5	1.08 (s)	28.5	1.10 (s)	28.5	1.09 (s)	28.5	1.09 (s)
24	17.0	0.90 (s)	17.0	0.90 (s)	17.0	0.90 (s)	16.9	0.90 (s)
25	17.0	1.17 (s)	17.1	1.17 (s)	17.0	1.17 (s)	17.0	1.17 (s)
26	19.1	1.13 (s)	19.2	1.13 (s)	19.1	1.13 (s)	19.1	1.15 (s)
27	23.1	1.45 (s)	23.2	1.46 (s)	23.1	1.46 (s)	23.8	1.46 (s)
28	69.3	3.04 (d, 11.0)/3.31 (m)	69.3	3.04 (d, 11.5) 3.31 (d, 11.5)	69.3	3.04 (d, 11.5) 3.30 (d, 11.5)	69.6	3.03 (d, 11.5) 3.38 (d, 11.5)
29	178.2	–	178.1	–	178.1	–	28.5	1.18 (s)
30	24.0	1.19 (s)	24.1	1.20 (s)	24.0	1.19 (s)	178.6	–
30-OCH_3	52.3	3.72 (s)	52.4	3.73 (s)	52.4	3.72 (s)	52.3	3.72 (s)
3-*O*-Ara								
1′	107.1	4.30 (d, 7.0)	105.3	4.54 (d, 6.0)	105.1	4.57 (d, 6.0)	105.1	4.56 (d, 6.0)
2′	72.8	3.59 (dd, 9.0, 7.0)	79.3	3.90 (dd, 9.0, 6.0)	79.1	3.92 (dd, 9.0, 6.0)	79.1	3.91 (dd, 9.0, 7.0)
3′	74.3	3.53 (dd, 9.0, 3.0)	73.6	3.79 (dd, 9.0, 3.0)	73.1	3.90 (dd, 9.0, 3.0)	73.2	3.90 (dd, 9.0, 3.0)
4′	69.5	3.82 (dd, 3.5, 3.0)	68.8	3.86 (dd, 3.5, 3.0)	68.6	3.91 (dd, 3.5, 3.0)	68.7	3.91 (dd, 3.5, 3.0)
5′	66.3	3.86 (dd, 11.5, 3.5) 3.54 (dd, 11.5, 3.5)	65.2	3.85 (dd, 11.5, 3.5) 3.52 (dd, 11.5, 3.5)	64.9	3.86 (dd, 11.5, 3.5) 3.53 (dd, 11.5, 3.5)	65.0	3.85 (dd, 11.5, 3.5) 3.53 (dd, 11.5, 3.5)
2′-*O*-Glc								
1″			104.9	4.61 (d, 7.5)	103.1	4.77 (d, 8.0)	103.2	4.76 (d, 7.5)
2″			76.0	3.22 (dd, 9.0, 7.5)	84.6	3.44 (dd, 9.0, 8.0)	84.7	3.44 (dd, 9.0, 7.5)
3″			77.9	3.37 (dd, 9.0, 9.0)	77.6	3.59 (dd, 9.0, 9.0)	77.7	3.58 (dd, 9.0, 9.0)
4″			71.7	3.28 (dd, 9.0, 9.0)	71.2	3.33 (dd, 9.0, 9.0)	71.3	3.33 (dd, 9.0, 9.0)
5″			78.2	3.28 (m)	78.6	3.33 (m)	78.7	3.34 (m)
6″			62.8	3.67 (dd, 11.5, 5.0) 3.84 (dd, 11.5, 2.5)	62.6	3.72 (dd, 11.5, 5.0) 3.91 (dd, 11.5, 2.5)	62.7	3.72 (dd, 11.5, 5.0) 3.92 (dd, 11.5, 2.5)
2″-*O*-Glc								
1‴					105.7	4.62 (d, 7.5)	105.7	4.61 (d, 7.5)
2‴					76.3	3.30 (dd, 9.0, 7.5)	76.3	3.29 (dd, 9.0, 7.5)
3‴					77.6	3.39 (dd, 9.0, 9.0)	77.6	3.39 (dd, 9.0, 9.0)
4‴					71.6	3.28 (dd, 9.0, 9.0)	71.7	3.28 (dd, 9.0, 9.0)
5‴					77.9	3.30 (m)	78.0	3.28 (m)
6‴					62.9	3.84 (dd, 11.5, 2.5) 3.65 (dd, 11.5, 5.0)	63.0	3.83 (dd, 11.5, 2.5) 3.64 (dd, 11.5, 5.0)

Assignments were done by HSQC, HMBC, H-H COSY, and NOESY experiments.

HMBC correlations from H-27 (δ_H 1.45) to C-14 (δ_C 44.9)/C-13 (δ_H 172.1), from H-18 (δ_H 2.20) to C-13/C-12 (δ_C 129.2), from H-12 (δ_H 5.60) to C-9 (δ_C 63.1), and from H-9 (δ_H 2.49) to C-11 (δ_C 202.5) indicated that the ketone group was at C-11 and the double bond was at C-12/C-13. In addition, H-21 (δ_H 4.18) appeared as a broad triplet with a small coupling constant (J = 3.0 Hz), suggesting its *β/equatorial* orientation, which was further indicated by NOESY cross peak from H-21 to H-30 (δ_H 1.19) (Fig. 3). The larger coupling constant of H-2/H-3 (J = 11.5 Hz) indicated the *α/axial* orientation of H-3, which was further confirmed by NOESY cross peaks from H-3 (δ_H 3.20) to H-23 (δ_H 1.08)/H-5 (δ_H 0.82) and from H-23 to H-5. The arabinopyranose sugar linked to C-3 (Mizui et al., 1988) was indicated by the HMBC correlation from the anomeric proton (δ_H 4.30) to C-3 (δ_C 90.3). Furthermore, NOEs correlations between ara H-1′ (δ_H 4.30) and ara H-3′ (δ_H 3.53)/H-3 (δ_H 3.20) were observed in the NOESY spectrum of **1**, consistent with an *α* configuration for the arabinopyranose sugar (Waltho et al., 1986). Finally, the presence of L-arabinose in the acid hydrolysis product of compound **1** was confirmed by TLC analysis and comparison of the optical rotation with those of authentic L-arabinose (Voutquenne-Nazabadioko et al., 2013). Consequently, the structure of compound **1** was established to be 3-*O*-*α*-L-arabinopyranosyl-3*β*,21*α*,28-trihydroxy-11-oxoolean-12-ene-29-oic acid methyl ester and named chenoalbumoside A (Figs. 1 and S1–S14).

Compound **2** was obtained as colorless amorphous powder. Its molecular formula was determined to be $C_{42}H_{66}O_{15}$ by the *quasi*-molecular ion peaks at *m/z* 845.4084 $[M+^{35}Cl]^-$ (Calcd. for $[C_{42}H_{66}O_{15}^{35}Cl]^-$, 845.4090, Δ = -0.7 ppm) and *m/z* 847.4075 $[M+^{37}Cl]^-$ (Calcd. for $[C_{42}H_{66}O_{15}^{37}Cl]^-$, 847.4061, Δ = +1.9 ppm) in the HR-ESI-MS. Its IR

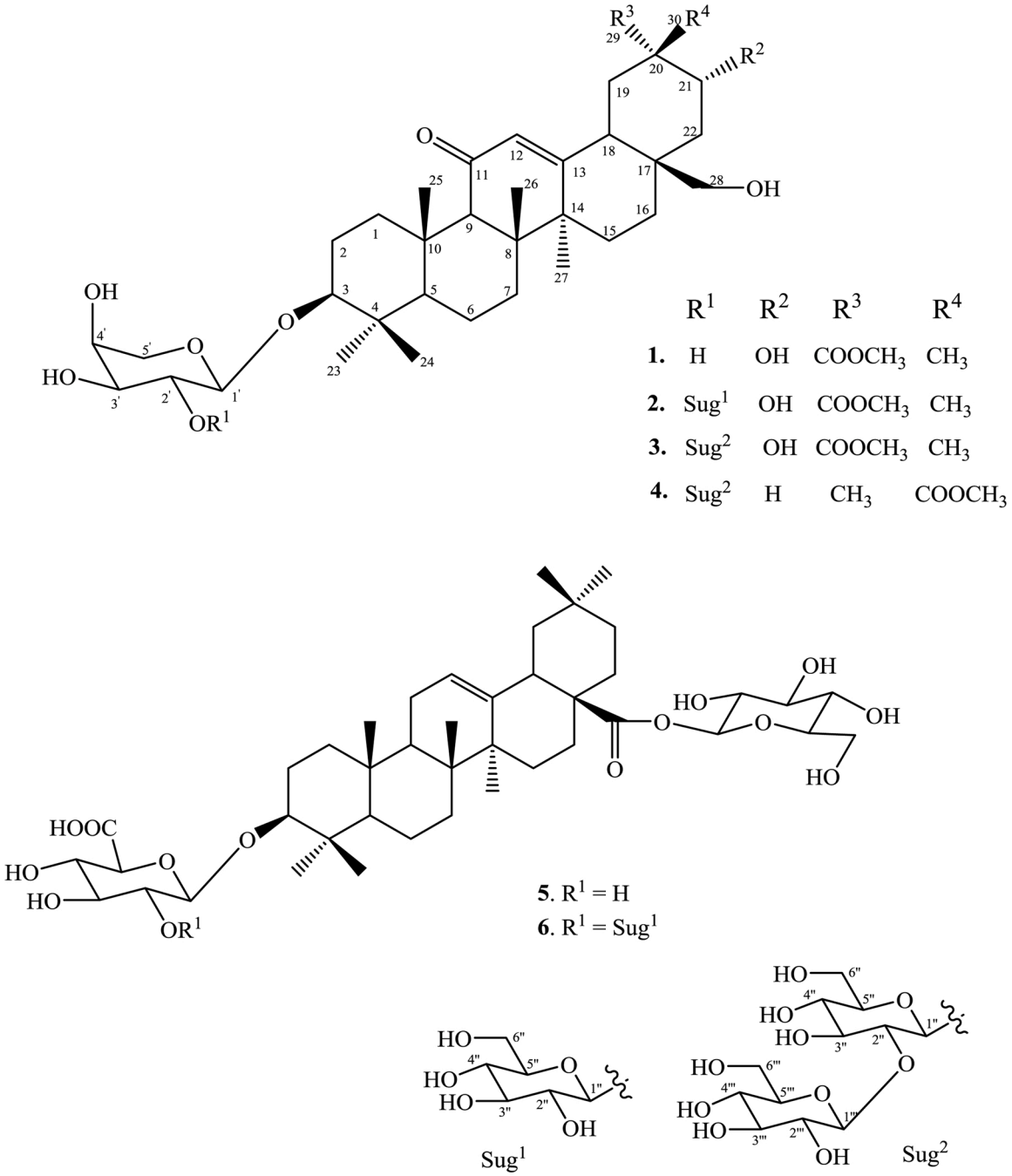

Fig. 1. Chemical structures of compounds **1**-**6**.

spectrum exhibited the presence of hydroxy (3407 cm^{-1}), carbonyl (1723 cm^{-1}) and C—OC— (1079, 1046 cm^{-1}) groups. The ^{1}H and ^{13}C NMR spectra of compound **2** were similar to those of **1** but differed in the sugar signals in that **2** had one more glucopyranosyl unit. A carefull comparison the NMR data of compounds **1** and **2** found that they had the same aglycon as 3*β*,21*α*,28-trihydroxy-11-oxoolean-12-ene-29-oic acid methyl ester (Table 1), which was further confirmed by HSQC, HMBC, ^{1}H-^{1}H COSY, and NOESY spectra as shown in Table 1 and Figs. 2 and 3. In the HSQC spectra, protons at δ_H 4.61, 4.54, 3.90, 3.20 had cross peaks with carbons at δ_C 104.9, 105.3, 79.3, and 90.6, respectively. In addition, COSY cross peaks H-1′ (δ_H 4.54)/H-2′ (δ_H 3.90)/H-3′ (δ_H 3.79)H-4′ (δ_H 3.86)/H-5′ (δ_H 3.85 and 3.52), as well as HMBC correlations from H-1′ (δ_H 4.54) to C-3 (δ_C 90.6) and from H-1″ (δ_H 4.61) to C-2′ (δ_C 79.3), were observed. The above evidence indicated that the sugar moiety of **2** was 3-*O*-*β*-D-glucopyranosyl-(1→2)-*α*-L-araninopyranoside. The larger coupling constant value ($J = 8.0$ Hz) observed for the anomeric proton at δ_H 4.61 in the ^{1}H NMR spectrum of **2** indicated a *β*-glucoside linkage for the glucose moiety. Furthermore, NOE observations from ara H-3′ (δ_H 3.79) to ara H-1′ (δ_H 4.54) and from H-1′ to H-3 (δ_H 3.20) in the NOESY spectrum of **2** identified an *α* configuration for arabinopyranose (Waltho et al., 1986). Finally, the presence of D-glucose and L-arabinose in the acid hydrolysis product of compound **2** were confirmed by TLC analysis and comparison optical rotations with those of authentic D-glucose and L-arabinose (Voutquenne-Nazabadioko et al., 2013). Therefore, compound **2** was determined to be 3-*O*-*β*-D-glucopyranosyl-(1→2)-*α*-L-arabinopyranosyl-3*β*,21*α*,28-trihydroxy-11-oxoolean-12-ene-29-oic ac id methyl ester and named chenoalbumoside B (Figs. 1 and S15–S29).

Compound **3** was obtained as a colorless amorphous powder. Its molecular formula was determined to be $C_{48}H_{76}O_{20}$ by the *quasi*-molecular ion peaks at *m/z* 1007.4591 $[M+^{35}Cl]^-$ (Calcd. for $[C_{48}H_{76}O_{20}^{35}Cl]^-$, 1007.4618, $\Delta = -2.7$ ppm), *m/z* 1009.4562 $[M+^{37}Cl]^-$ (Calcd. for $[C_{48}H_{76}O_{20}^{37}Cl]^-$, 1009.4589, $\Delta = -2.7$ ppm) in the HR ESI MS. A comparison of the NMR spectra of **3** with the corresponding spectra of **1** and **2** (data shown in Table 1) found that they had the same aglycone, 3*β*,21*α*,28-trihydroxy-11-oxoolean-12-ene-29-oic acid methyl ester (Table 1), but differed in the sugar moieties. The additional NMR signals of **3** [(δ_H 4.62 (d, $J = 7.5$ Hz))/δ_C (105.7), δ_H (3.30)/δ_C (76.3), δ_H (3.39)/δ_C (77.6), δ_H (3.28)/δ_C (71.6), δ_H (3.30)/δ_C (77.9), δ_H (3.84 and 3.65)/δ_C (62.9)] compared to **2** suggested that **3** had more than one glucose sugar. This suggestion was consistent with the molecular formula of $C_{48}H_{76}O_{20}$ obtained from the HR-ESI-MS results. Its IR spectrum exhibited the presence of hydroxy (3417 cm^{-1}), carbonyl (1727 cm^{-1}) and C—OC— (1075, 1027 cm^{-1}) groups. The similarity of NMR spectra of **3** with **1** and **2** indicated that they are glycosyl derivatives of each other. Therefore, the assignments of NMR data of **3** were first done by direct comparison with those of compounds **1** and **2** (Table 1) and further elucidated by HSQC, ^{1}H-^{1}H COSY, and HMBC spectra. ^{1}H-^{1}H COSY cross peaks from H-1′ (δ_H 4.57) to H-2′ (δ_H 3.92) and from H-1″ (δ_H 4.77) to H-2″ (δ_H 3.44) were observed. On the other hand, the carbon chemical shifts corresponding to these protons were determined from HSQC spectra. HMBC correlations from H-1′ to C-3 (δ_C 90.9), from H-1″ to C-2′ (δ_C 79.1), and from H-3‴ (δ_H 4.62) to C-2″ (δ_C 84.6) were observed, indicating the sugar moiety as 3-*O*-glucopyranosyl-(1→2)-glucopyranosyl-(1→2)-arabinopyranosyl. As with compound **2**, the sugar

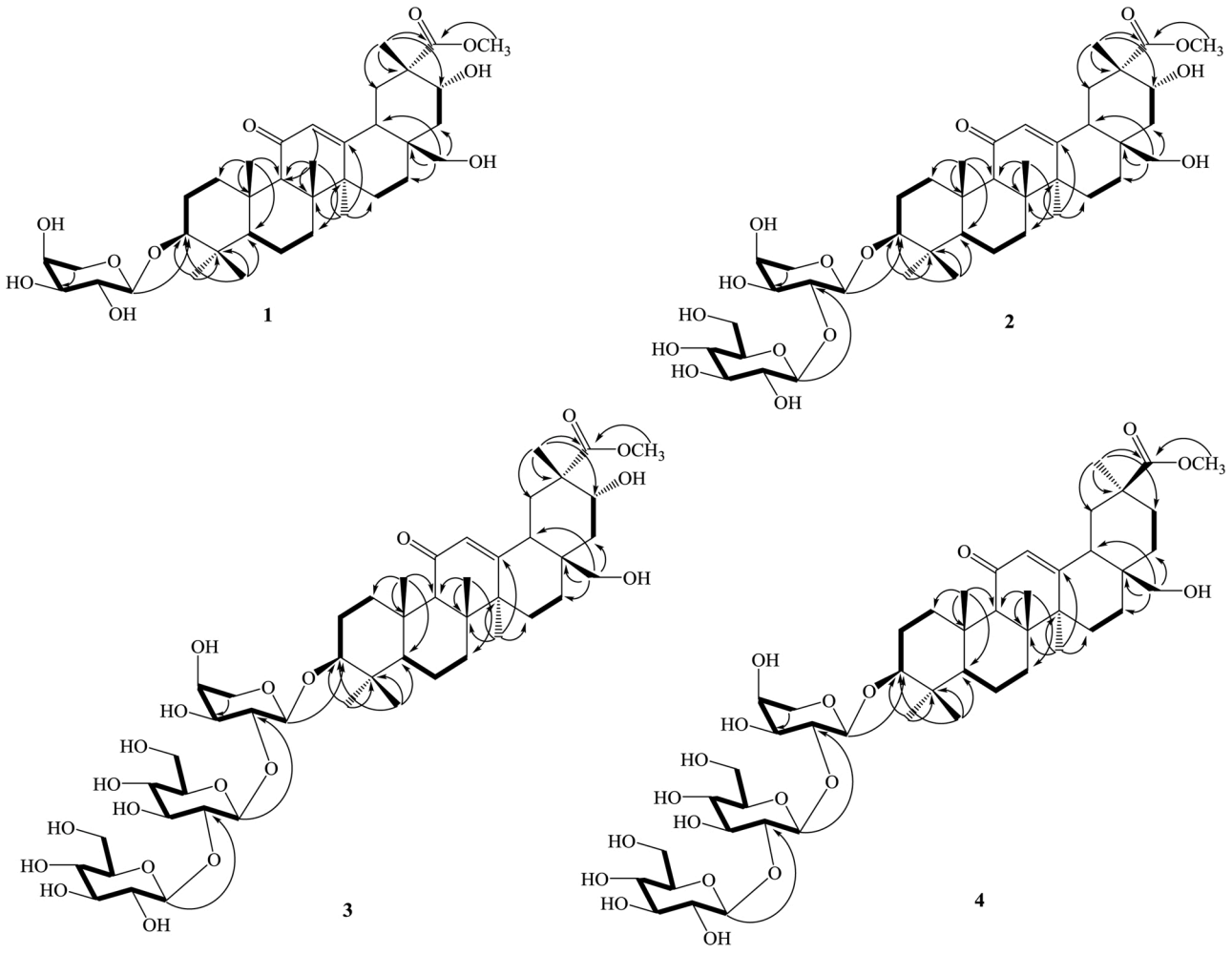

Fig. 2. Important HMBC (H→C) and H-H COSY correlations of compounds **1**-**4**.

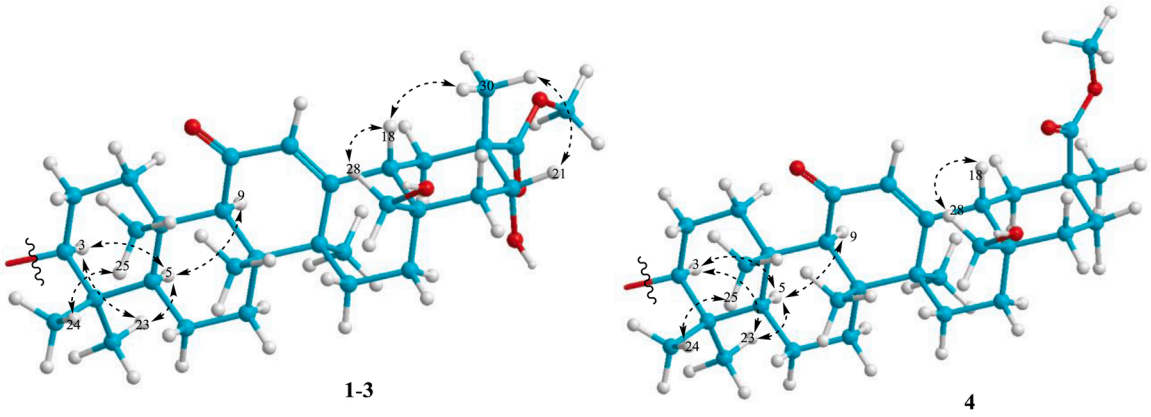

Fig. 3. Key NOESY correlations of compounds **1-4**.

linkages must be in the β-form for the two glucose moieties ($J_{H\text{-}1/H2}$ = 7.5–8.0 Hz) and in the α-form for the arabinose moiety ($J_{H\text{-}1/H2}$ = 7.0 Hz). In addition, the NOE observation from ara H-1′ (δ_H 4.57) to ara H-3′ (δ_H 3.90)/H-3 (δ_H 3.21) in the NOESY spectrum of **3** further confirmed an α configuration for the arabinopyranosyl sugar (Waltho et al., 1986). Finally, the presence of D-glucose and L-arabinose in the acid hydrolysis product of compound **3** were confirmed by TLC analysis and comparison optical rotations with those of authentic D-glucose and L-arabinose (Voutquenne-Nazabadioko et al., 2013). Thus, the structure of compound **3** was determined to be 3-*O*-β-D-glucopyranosyl-(1→2)-β-D-glucopyranosyl-(1→2)-α-L-arabinopyranosyl-3β,21α,28-trihydroxy-11-oxoolean-12-ene-30-oic acid methyl ester and named chenoalbumoside C (Figs. 1 and S30–S44).

Compound **4** was obtained as a colorless amorphous powder. Its molecular formula was determined to be $C_{24}H_{36}O_5$ by a *quasi*-molecular ion peaks at *m/z* 991.4654 $[M+^{35}Cl]^-$ (Calcd. for $[C_{48}H_{76}O_{19}^{35}Cl]^-$, 991.4669, Δ = -1.5 ppm) and 993.4625 $[M+^{37}Cl]^-$ (Calcd. for $[C_{48}H_{76}O_{19}^{37}Cl]^-$, 993.4640, Δ = -1.5 ppm) in the HR-ESI-MS. Its IR spectrum exhibited the presence of hydroxy (3420 cm^{-1}), carbonyl (1701 cm^{-1}) and C—OC— (1083, 1042 cm^{-1}) groups. The NMR spectra of **4** were similar to those of **3**, suggesting an oleanane triterpene glycoside. Three sugars were identified by the anomeric signals at δ_H 4.56 (d, J = 6.0 Hz)/ δ_C 105.1, δ_H 4.76 (d, J = 7.5 Hz)/ δ_C 103.2, and δ_H 4.61 (d, J = 7.5 Hz)/ δ_C 105.7. A ketone at C-11 (δ_C 202.4), double bond at C-12 (δ_H 5.58 (s)/ δ_C 129.1) and C-13 (δ_C 179.1), one methyl carboxylate group at δ_C 178.6 and δ_H 3.72 (3H, s)/ δ_C 52.3, one oxygenated methylene group at C-28 [δ_H 3.03 and 3.38 (each, d, J = 11.5 Hz)/ δ_C 69.6, and one oxygenated methine group at C-3 [δ_H 3.21 (dd, J = 11.5, 5.0 Hz)/ δ_C 90.9] were observed. The complete similarity of the sugar NMR data of **4** and **3** (Table 1) suggested that they had the same sugar moiety of 3-*O*-β-D-glucopyranosyl-(1→2)-β-D-glucopyranosyl-(1→2)-α-L-arabinopyranosyl, but their aglycones were different. The hydroxyl signals at C-21 were absent in the NMR spectra of **4** and the CH_3-20 signals changed from δ_C 24.0 (in **3**) to δ_C 28.5 (in **4**). The value of δ_C 28.5 coincided highly with those reported for $\underline{C}H_3$-29/$COOCH_3$-30 as $\delta_{C\text{-}29}$ 28.0–28.5 (Yi et al., 1995; Bandara et al., 1989; Harkar et al., 1984) and was quite different from $\underline{C}H_3$-30/$COOCH_3$-29 as $\delta_{C\text{-}30}$ 18.8–21.2 (Yahara et al., 1989; Rogers and Subramony, 1988; Sakamoto et al., 1992). In addition, HMBC correlations from the methoxy group (δ_H 3.72) to C-30 (δ_C 178.6), from H-29 (δ_H 1.18) to C-19 (δ_C 42.1)/C-20 (δ_C 45.3)/C-21 (δ_C 31.6)/C-30 (δ_C 178.6), from H-28 (δ_H 3.03 and 3.38) to C-16 (δ_C 22.5)/C-17 (δ_C 37.7)/C-18 (δ_C 44.9)/C-22 (δ_C 33.3), from H-12 (δ_H 5.58)/H-9 (δ_H 2.48) to C-11 (δ_C 202.4), from H-27 (δ_H 1.46) to C-13 (δ_C 171.9), and from H-23 (δ_H 1.09)/H-24 (δ_H 0.90) to C-3 (δ_C 90.9) were observed, confirming that the methyl carboxylate, hydroxy, ketone, double bond, and methine carbinol groups were at C-30, C-28, C-11, C-12/C-13, and C-3, respectively. Furthermore, the sugar moieties were further indicated by HSQC, ^1H-^1H COSY, HMBC, and NOESY spectra. The HMBC correlations from ara H-1′ (δ_H 4.56) to C-3 (δ_C 90.9), from glc H-1″ (δ_H 4.76) to ara C-2′ (δ_C 79.1), and from glc H-1‴ (δ_H 4.61) to glc C-2″ (δ_C 84.7) were observed, confirming the sugar moieties as shown in Figs. 1 and 2. The sugar linkages must be in the β-form for the two glucose moieties ($J_{H\text{-}1/H2}$ = 7.5 Hz) and in the α-form for the arabinose moiety ($J_{H\text{-}1/H2}$ = 6.0 Hz). NOE observation from ara H-1′ (δ_H 4.56) to ara H-3′ (δ_H 3.90)/H-3 (δ_H 3.21) further confirmed an α configuration for the arabinopyranosyl sugar of **4** (Waltho et al., 1986). Finally, the presence of D-glucose and L-arabinose in the acid hydrolysis product of compound **4** were confirmed by TLC analysis and comparison optical rotations with those of authentic D-glucose and L-arabinose (Voutquenne-Nazabadioko et al., 2013). Consequently, compound **4** was established to be 3-*O*-β-D-glucopyranosyl-(1→2)-β-D-glucopyranosyl-(1→2)-α-L-arabinopyranosyl-3β,28-dihydroxy-11-oxoolean-12-ene-29-oic acid methyl ester and named chenoalbumoside D (Figs. 1, S45–S60).

The two known compounds were identified to be chikusetsusaponin IVa (**5**) and ginsennoside Ro (**6**) (Mizui et al., 1990), by consistency of their NMR spectral data with those reported in the literature.

The cytotoxic activities of compounds **1-6** were assessed in HT-29 (colorectal adenocarcinoma), SW480 (colon adenocarcinoma), AGS (stomach gastric adenocarcinoma), and MKN7 (gastric carcinoma) human cancer cell lines (Skehan et al., 1990). As shown in the Table S1, compounds **1-4** exhibited weak cytotoxic effects on all four tested cell lines with IC_{50} values ranging from 46.5 ± 1.3 to 71.6 ± 2.9 μM. Meanwhile, compounds **5** and **6** exhibited inactivity (IC_{50} > 100 μM). Ellipticine was used as a positive control, showing IC_{50} values of 1.54 ± 0.08, 1.74 ± 0.17, 1.67 ± 0.12, and 1.59 ± 0.08 μM against HT-29, SW480, AGS, and MKN7 cell lines, respectively. A previous review indicated that oleanane-type saponins with oleanolic acid as an aglycone had more potent cytotoxic effects than hederagenin saponins (23-hydroxy oleanolic acid derivative). Additionally, carboxylic groups at C-28 are suggested to be important for cytotoxicity (Podolak et al., 2010). The esterification at position C-28 generally decreased cytotoxicity. Compounds **5** and **6** were inactive, which could be due to the esterification at position C-28. Compounds **1-4** had a methocarboxy group (C-29 or C-30). However, the carboxylic group at C-28 in those compounds was replaced by a hydroxymethyl group. These changes in the chemical structure could decrease the cytotoxic of compounds **1-4**.

3. Experimental

3.1. General experimental procedures

Optical rotation was recorded on a Jasco P2000 polarimeter. HR-ESI-MS was acquired on an Agilent 6530 Accurate Mass Q-TOF LC/MS (ESI source voltage of 3.2 kV). NMR spectra were recorded on an Bruker AM500 FT NMR spectrometer with standard pulse sequences using deterated solvent as standard reference (Gottlieb et al., 1997) and

operating at 500 MHz for ^{1}H NMR and 125 MHz for ^{13}C NMR spectra. Column chromatography (CC) was performed using silica gel (particle size 40−63 μm) or reversed phase C-18 resins (particle size 120−150 μm) as stationary phase. Thin layer chromatography (TLC) was carried out on pre-coated silica gel 60 F_{254} and/or RP-18 F_{254S} plates. Preparative HPLC were run on an Agilent 1100 system including quaternary pump, autosampler, DAD detector, and preparative HPLC column YMC J'sphere ODS-H80 (4 μm, 20 × 250 mm). Isocratic mobile phase with the flow rate of 3 mL/min was used in pre-HPLC. The compound was monitored at wavelengths of 205, 230, 254, and 280 nm. The compounds were detected under UV irradiation (254 and 365 nm) or by spraying with 5% H_2SO_4 solution followed by heating with a heat gun.

3.2. Plant material

The aerial parts of Chenopodium album L. were collected at Xuan Thuy District, Nam Dinh Province, Vietnam in Jun 2019 and identified by Dr. Nguyen The Cuong, Institute of Ecology and Biological Resources. A voucher specimen (NCCT-P121) was deposited at the Institute of Marine Biochemistry, VAST.

3.3. Extraction and isolation

The dried powder of aerial parts of *Chenopodium album* L. (10 kg) were sonicated with methanol (3 times, each 15 L MeOH). After removal of solvent, the MeOH extract (300 g) was suspended with water and then partitioned with *n*-hexane, dichloromethane and then ethyl acetate to give corresponding residues, *n*-hexane (HCA1, 70 g), dichloromethane (HCA2, 90 g), ethyl acetate (HCA3, 8 g), and water layer (HSB4). The HCA3 (8 g) was chromatographed on a RP-18 column eluting with MeOH/water (1/2, v/v) to give five fractions, HCA3A → HCA3E. HCA3C (106 mg) was chromatographed on a sephadex LH-20 column, eluting with MeOH/water (1.5/1, v/v) to give compound **1** (14.5 mg). The water layer (HCA4) was chromatographed on a Diaion HP-20 column, first eluting with water to remove sugar components, then increasing concentration of MeOH in water (50 and 100 %) to obtain two fractions, HCA4A (7.2 g) and HCA4B (9.8 g). The HCA4B fraction was subjected on a silica gel CC eluting with gradient solvent of CH_2Cl_2/MeOH (30/1, 10/1, 5/1, 0/1 v/v) to give four fractions, HCA4B1 → HCA4B4. The HCA4B3 (956 mg) was re-chromatographed on a silica gel CC eluting with CH_2Cl_2/MeOH/H_2O (2/1/0.05, v/v) to give a smaller subfractions, HCA4B3C (109 mg). The HCA4B3C fraction was chromatographed on HPLC (J'sphere H-80 column, 250 mm length x 20 mm ID, eluting with 65 % MeOH in water, a flow rate of 3 mL/min) to give **4** (35 mg). HCA4B1 (1.6 g) was re-chromatographed on a silica gel CC eluting with CH_2Cl_2/MeOH (5/1/, v/v) to give four fractions, HCA4B1A → HCA4B1D. The HCA4B1B fraction (126 mg) was re-chromatographed on HPLC (J'sphere H-80 column, 250 mm length x 20 mm ID, eluting with 25 % acetonitrile in water, a flow rate of 3 mL/min) to give compound **2** (20.8 mg). HCA4B1D (173 mg) was chromatographed on HPLC (J'sphere H-80 column, 250 mm length x 20 mm ID, eluting with 25 % acetonitrile in water, a flow rate of 3 mL/min) to give compound **3** (60 mg). The HCA4A fraction (7.2 g) was on a RP-18 column eluting with acetone/water: 1/1.8 (v/v) to give three fractions, HCA4A1 → HCA4A3. HCA4A3 (124 mg) was re-chromatographed on HPLC (J'sphere H-80 column, 250 mm length x 20 mm ID, eluting with 26 % acetonitrile in water, a flow rate of 3 mL/min) to give compound **6** (31.3 mg). The HCA4B4 (1.5 g) was chromatographed on a RP-18 column eluting with MeOH/water (1.5/1, v/v) to give five compounds, HCA4B4A → HCA4B4E. The HCA4B4E fraction (58 mg) was chromatographed on HPLC (J'sphere H-80 column, 250 mm length × 20 mm ID, eluting with 60 % MeOH in water, a flow rate of 3 mL/min) to give compound **5** (9 mg).

3.3.1. Chenoalbumoside A (1)

Colorless amorphous powder, $[\alpha]_D^{25}$ = +43.0 (*c* 0.1, MeOH). UV (MeOH) λ_{max} nm (logε): 250 (2.95). IR (KBr) ν_{max}: 3430 (broad), 2940, 1712, 1644, 1090, 1047 cm^{-1}. HR-ESI-MS: *m/z* 683.3522 [M+^{35}Cl]- (Calcd. for [$C_{36}H_{56}O_{10}^{35}$Cl]-, 683.3562, Δ = -5.9 ppm), *m/z* 685.3511 [M+^{37}Cl]- (Calcd. for [$C_{36}H_{56}O_{10}^{37}$Cl]-, 685.3533, Δ = -3.2 ppm). ^{1}H and ^{13}C NMR spectral data are given in the Table 1.

3.3.2. Chenoalbumoside B (2)

Colorless amorphous powder, $[\alpha]_D^{25}$ = +34.5 (*c* 0.1, MeOH). UV (MeOH) λ_{max} (logε): 250 (3.23). IR (KBr) ν_{max}: 3407 (broad), 2936, 1723, 1644, 1079, 1046 cm^{-1}. HR-ESI-MS: *m/z* 845.4084 [M+^{35}Cl]- (Calcd. for [$C_{42}H_{66}O_{15}^{35}$Cl]-, 845.4090, Δ = -0.7 ppm), *m/z* 847.4075 [M+^{37}Cl]- (Calcd. for [$C_{42}H_{66}O_{15}^{37}$Cl]-, 847.4061, Δ = +1.9 ppm). ^{1}H and ^{13}C NMR spectral data are given in the Table 1.

3.3.3. Chenoalbumoside C (3)

Colorless amorphous powder, $[\alpha]_D^{25}$ = +27.4 (*c* 0.1, MeOH). UV (MeOH) λ_{max} (logε): 250 (3.25). IR (KBr) ν_{max}: 3417 (broad), 2936, 1727, 1651, 1075, 1027 cm^{-1}. HR-ESI-MS: *m/z* 1007.4591 [M+^{35}Cl]- (Calcd. for [$C_{48}H_{76}O_{20}^{35}$Cl]-, 1007.4618, Δ = -2.7 ppm), *m/z* 1009.4562 [M+^{37}Cl]- (Calcd. for [$C_{48}H_{76}O_{20}^{37}$Cl]-, 1009.4589, Δ = -2.7 ppm). ^{1}H and ^{13}C NMR spectral data are given in the Table 1.

3.3.4. Chenoalbumoside D (4)

Colorless amorphous powder. $[\alpha]_D^{25}$ = +23.5 (*c* 0.1, MeOH). UV (MeOH) λ_{max} (logε): 250 (3.25). IR (KBr) ν_{max}: 3420 (broad), 2949, 1701, 1654, 1083, 1042 cm^{-1}. HR-ESI-MS: *m/z* 991.4654 [M+^{35}Cl]- (Calcd. for [$C_{48}H_{76}O_{19}^{35}$Cl]-, 991.4669, Δ = -1.5 ppm), *m/z* 993.4625 [M+^{37}Cl]- (Calcd. for [$C_{48}H_{76}O_{19}^{37}$Cl]-, 993.4640, Δ = -1.5 ppm). ^{1}H and ^{13}C NMR spectral data are given in the Table 1.

3.4. Cytotoxic assay

AGS (human stomach gastric adenocarcinoma), MKN7 (human gastric carcinoma), SW480 (human colon adenocarcinoma), and HT-29 (human colorectal adenocarcinoma) cell lines were obtained from Milan University, Italy and Long Island University, USA. The cells were maintained and cultured in DMEM supplemented with FBS, trypsin-EDTA, L-glutamine, sodium piruvat, $NaHCO_3$, and penicillin/streptomycin at 37 °C in a humidified atmosphere of 5% CO2. Cytotoxic effects of compounds were determined using Sulforhodamine B (SRB) assay as previously described by Skehan et al. (1990). In brief, the cells were incubated with/without compounds for three days in 96-well culture plate. After incubation, cells were stained with sulforhodamine B and measured optical density (OD) at 540 nm. Difference of OD between samples and vehicle well during experiments indicated cells situation induced by compounds. Results are expressed as percentage of cells death in comparison with vehicle well. Dose-response curves of compounds were generated to determine IC_{50} values of the compounds corresponding with each cell line. Ellipticine was used as the positive control throughout experiments (Skehan et al., 1990).

3.5. Acid hydrolysis and confirmation of monosaccharide

Compounds **1**-**4** (each, 5 mg) were dissolved in 1 mL of HCl 1 M solution in dioxan/water (1/1, v/v). Solution was refluxed at 80 °C in water bath for 2 h. Reaction mixture was cooled down to room temperature and extracted twice times with ethyl acetate (each 1 mL). Water layer was neutralized by amberlite IRA400 resin (OH form) and then dried in *in vacuo* to give sugar residue. Monosaccharide in hydrolysis products was purified by preparative TLC (dichloromethane/methanol/water, 3/1/0.1, v/v/v) and then immediately measured optical rotation. The presence of D-glucose and/or L-arabinose in the acid hydrolysis products of compounds **1**-**4** were confirmed by TLC analysis and

comparison their optical rotations with those of authentic D-glucose [R_f 0.36, $[\alpha]_D^{25} = +16.4$ (*c* 0.1, H_2O)] and L-arabinose [R_f 0.50, $[\alpha]_D^{25} = +30.7$ (*c* 0.1, H_2O)] as reported (Voutquenne-Nazabadioko et al., 2013).

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 104.01-2019.18. The authors would like to thank Dr. Nguyen The Cuong, Institute of Ecology and Biological Resources for defining the scientific name of the plant.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.phytol.2021.05.004.

References

Bandara, B.M.R., Jayasinghe, L., Karunaratne, V., Wannigama, G.P., Kraus, W., Bokel, M., Sottheewaran, S., 1989. Diploclisin, a bidesmosidic triterpenoid saponin from *Diploclisia glaucescens*. Phytochemistry 28, 2783–2785.

Chi, V.V., 2012. Dictionary of Vietnamese Medicinal Plants, 2. Medicine Publishing House, Hanoi, pp. 526–527.

Chludil, H.D., Corbino, G.B., Leicach, S.R., 2008. Soil quality effects on Chenopodium album flavonoid content and antioxidant potential. J. Agric. Food Chem. 56, 5050–5056.

Cutillo, F., D'Abrosca, B., Dellagreca, M., Di, M.C., Golino, A., Previtera, L., Zarrelli, A., 2003. Cinnamic acid amides from *Chenopodium album*: effects on seeds germination and plant growth. Phytochemistry 64, 1381–1387.

Cutillo, F., D'Abrosca, B., Dellagreca, M., Zarrelli, A., 2004. Chenoalbicin, a novel cinnamic acid amide alkaloid from *Chenopodium album*. Chem. Biodivers. 1, 1579–1583.

Cutillo, F., Dellagreca, M., Gionti, M., Previtera, L., Zarrelli, A., 2006. Phenols and lignans from *Chenopodium album*. Phytochem. Anal. 17, 344–349.

Dai, Y., Ye, W.C., Wang, Z.T., Matsuda, H., Kubo, M., But, P.P.H., 2002. Antipruritic and antinociceptive effects of *Chenopodium album* L. in mice. J. Ethnopharmacol. 81, 245–250.

Dellagreca, M., D'Abrosca, B., Fiorentino, A., Previtera, L., Zarrelli, A., 2005a. Structure elucidation and phytotoxicity of ecdysteroids from *Chenopodium album*. Chem. Biodivers. 2, 457–462.

Dellagreca, M., Previtera, L., Zarrelli, A., 2005b. A new xyloside from Chenopodium album. Nat. Prod. Res. 19, 87–90.

Gonzalez, J.A., Gallardo, M., De, I.L.A., 1998. Leaf flavonoids in Chenopodium hircinum Schrad. and Chenopodium album L. (Chenopodiaceae). Phyton 63, 279–281.

Gottlieb, H.E., Kotlyar, V., Nudelman, A., 1997. NMR chemical shifts of common laboratory solvents as trace impurities. J. Org. Chem. 62, 7512–7515.

Harkar, S., Razdan, T.K., Waight, E.S., 1984. Further triterpenoids and 13C NMR spectra of oleanane derivatives from *Phytolacca acinose*. Phytochemistry 23, 2893–2898.

Horio, T., Yoshida, K., Kikuchi, H., Kawabata, J., Mizutani, J., 1993. A phenolic amide from roots of *Chenopodium album*. Phytochemistry 33, 807–808.

Jhade, D., Paarakh, P.M., Gavani, U., 2009. Isolation of phytoconstituents from the aerial parts of *Chenopodium album* Linn. J. Pharm. Res. 2, 1192–1193.

Lavaud, C., Voutquenne, L., Bal, P., Pouny, I., 2000. Saponins from *Chenopodium album*. Fitoterapia 71, 338–340.

Mizui, F., Kasai, R., Ohtani, K., Tanaka, O., 1988. Saponins from brans of quinoa, *Chenopodium quinoa* WILLD. I. Chem. Pharm. Bull. 36, 1415–1418.

Neerja, Y., Neeru, V., Sumitra, S., Surendra, K.S., 2007. Medicinal properties of genus *Chenopodium* Linn. Nat. Prod. Radiance 6, 131–134.

Podolak, I., Galanty, A., Sobolewska, D., 2010. Saponins as cytotoxic agents: a review. Phytochem. Rev. 9, 425–474.

Rogers, C.B., Subramony, G., 1988. The structure of imberbic acid, a 1α-hydroxy pentacyclic triterpenoid from *Combretum imberbe*. Phytochemistry 27, 531–533.

Sakamoto, S., Kofuji, S., Kuroyanagi, M., Ueno, A., Sekita, S., 1992. Saponins from *Trifolium repens*. Phytochemistry 31, 1773–1777.

Skehan, P., Storeng, R., Scudiero, D., Monks, A., McMahon, J., Vistica, D., Warren, J.T., Bokesch, H., Kenney, S., Boyd, M.R., 1990. New colorimetric cytotoxic assay for anticancer-drug screening. J. Natl. Cancer Inst. 82, 1107–1112.

Smith, B.D., 2006. Eastern North America as an independent center of plant domestication. Proc. Natl. Acad. Sci. U. S. A. 103, 12223–12228.

Voutquenne-Nazabadioko, L., Gevrenova, R., Borie, N., Harakat, D., Sayagh, C., Weng, A., Thakur, M., Zaharieva, M., Henry, M., 2013. Triterpenoid saponins from the roots of *Gypsophila trichotoma* Wender. Phytochemistry 90, 114–127.

Waltho, J.P., Williams, D.H., Mahato, B.S., Bikas, C., Barna, J.C.J., 1986. Structure elucidation of two triterpenoid tetrasaccharides from *Androsacea saxifragifolia*. J. Chem. Soc. Perkin Trans. 1, 1527–1531.

Yahara, S., Emura, S., Feng, H., Nohara, T., 1989. Studies on the constituents of the bark of *Dalbergia hupeana*. Chem. Pharm. Bull. 37, 2136–2138.

Yi, Y., Zhang, J., Zheng, H., Zhang, J., Su, Z., 1995. New saponins from the roots of *Phytolacca polyandra*. J. Nat. Prod. 58, 1880–1882.

Zlatina, K.N., Paraskev, T.N., Stefan, D.N., 2009. The genus *Chenopodium*: phytochemistry, ethnopharmacology and pharmacology. Pharmacogn. Rev. 3, 280–306.