Chemistry of Natural Compounds, Vol. 30, No. 3, 1994

TRITERPENE SAPONINS FROM *Thalictrum minus*. VIII. STRUCTURE OF THALICOSIDE D

A. S. Gromova, A. A. Semenov, V. I. Lutskii, S. V. Zinchenko, N. N. Trofimova, and Ya. V. Rashkes*

UDC 547.9:543.422.25:543.51

Thalicoside D – a new triterpene glycoside isolated from Thalictrum minus *L. (Ranunculaceae) – has the structure of oleanolic acid 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside 28-O-[O-β-D-glucopyranosyl-(1→6)-β-D-glucopyranoside]. This structure was established by the use of the results of acid and alkaline hydrolyses, SIMS spectra, and one- and two-dimensional NMR spectroscopy.*

Continuing a study of triterpene glycosides of low meadow rue *Thalictrum minus* L. [1-4], we have isolated a minor triterpenoid, which we have called thalicoside D (**1**). It followed from the ^1H and ^{13}C NMR spectra of this compound that it was a glycoside the genin of which contained fragments substantiating its triterpenoid nature (30 carbon atoms, 7 methyl groups, a trisubstituted double bond, an ester carbonyl, and a glycosylated secondary hydroxyl (Tables 1 and 2)).

The ^1H and ^{13}C NMR spectra of glycoside (**1**) contained the signals of five anomeric carbon atoms (95.9, 101.89, 104.50, 105.11, and 106.54 ppm) and of five anomeric protons (6.05, 6.03, 4.96, 4.84, and 4.58 ppm), while the CS of the carbon atoms of the genin moiety (see Table 1) coincided with those for a 3,28-disubstituted oleanolic acid [5]. These conclusions were confirmed by the SIMS(+) spectrum of thalicoside D. The presence of the peak of the $(M + Na)^+$ ion with *m/z* 1243 and of cluster ions with *m/z* 1097 (1243-146), 1081 (1243-162), 935 (1097-162), 919 (1081-162), and 803 (935-132) and also of the AgI^+ ion with *m/z* 439 showed that a genin with a mass of 456 was glycosylated by five carbohydrate residues (a pentaose, a deoxyhexose, and three hexoses) of which the deoxyhexose and a hexose were terminal.

The position of the signal of one of the anomeric carbon atoms at 95.90 ppm in the ^{13}C NMR spectrum and of an absorption band at 1738 cm^{-1} in the IR spectrum showed the presence of an acyl glycosidic bond with the oleanolic acid. To cleave this bond and establish the structure of the carbohydrate chain in the C-3 position of the genin, glycoside (**1**) was subjected to alkaline hydrolysis (scheme). (see the scheme on the next page)

According to its ^1H and ^{13}C NMR spectra (see Tables 1 and 2) the progenin (**2**) so obtained contained arabinose, glucose, and rhamnose residues, and the carbohydrate moiety was not branched. In view of the fact the ion with *m/z* 935 in the SIMS spectrum of glycoside (**1**) was formed as the result of the splitting out of rhamnose and glucose, the arabinose residue must have been attached to the C-3 atom of the genin. It followed from this that a hexose–hexose block was attached by the acyl glycosidic bond in the thalicoside molecule, as was confirmed by the peaks of ions with *m/z* 347 (two hexoses – H + Na) and 919 (1243 + H – two hexoses).

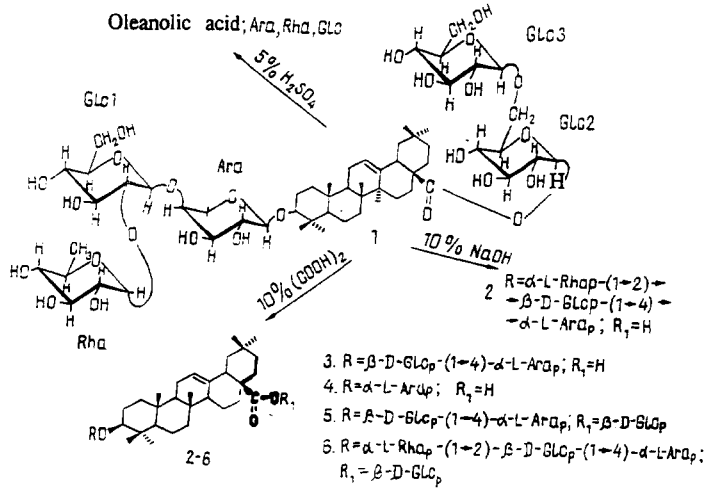
*Deceased.

Irkutsk Institute of Organic Chemistry, Siberian Division, Russian Academy of Sciences. Institute of the Chemistry of Plant Substances, Academy of Sciences of the Republic of Uzbekistan [IKhRV AN RUz], Tashkent, FAX (3712) 62 73 48. Translated from Khimiya Prirodnykh Soedinenii, No. 3, pp. 398-403, May-June, 1994. Original article submitted August 9, 1993.

0009-3130/94/3003-0363$12.50

TABLE 1. Chemical Shifts of the Carbon Atoms of Thalicoside D (**1**) and its Progenins (**2**) and (**6**) (C_5H_5N, δ, ppm, 0-TMS, t°C (**1**, **2**) – 80°C, (**6**) – 30°C)

C Atom	1	6	2	C Atom	1	6	2
	Genin				Carbohydrates		
					Ara		
1	39.28	39.00	39.44	1	106.54	107.06	106.90
2	26.60	26.55	26.82	2	73.03	72.88	73.29
3	89.44	89.16	89.58	3	73.63	73.74	73.91
4	39.73	39.60	39.95	4	77.81	78.34	78.16
5	57.67	56.02	56.56	5	64.94	65.29	65.14
6	18.92	18.68	19.10		Glc-1		
7	32.90	32.68	33.63	1	104.50	104.99	104.91
8	40.36	40.08	40.43	2	78.79	78.98	79.04
9	48.49	48.21	48.68	3	78.88	78.91	79.18
10	37.43	37.14	37.65	4	72.14	72.09	72.57
11	23.85	23.56	24.23	5	78.01	78.58	78.35
12	123.08	123.00	123.04	6	63.11	62.37	63.35
13	144.38	144.23	145.30		Rha		
14	42.58	42.23	42.28	1	101.89	101.96	102.40
15	28.62	28.29	28.82	2	72.07	71.85	72.39
16	24.14	23.84	24.23	3	72.59	72.39	72.88
17	46.77	47.15	47.24	4	74.62	74.56	74.96
18	42.10	41.90	42.62	5	70.14	70.05	70.38
19	46.77	46.38	47.14	6	18.90	18.68	19.10
20	31.00	30.91	31.37		Glc-2		
21	34.46	34.17	34.88	1	95.87	95.83	
22	33.58	33.29	33.86	2	74.13	74.17	
23	28.65	28.39	28.88	3	78.44	78.58	
24	17.12	17.01	17.29	4	71.53	71.26	
25	15.84	15.68	15.89	5	77.94	77.29	
26	17.88	17.62	17.91	6	69.88	62.74	
27	26.29	26.22	26.8		Glc-3		
28	176.60	176.80	181.6	1	105.11		
29	33.33	33.29	33.9	2	75.25		
30	24.07	23.95	24.4	3	78.79		
				4	72.27		
				5	78.19		
				6	63.03		



Scheme of the chemical transformations of thalicoside D

TABLE 2. Chemical Shifts (δ, ppm) and Spin-Spin Coupling Constants (J, Hz) in the PMR Spectra of Compounds (**1**), (**2**) and (**6**) (C_5H_5N, 0 – TMS, t for (**1**) and (**6**), 30°C; for (**2**), 80°C)

Compound	H-3	H-12	CH_3
1	3.07 (11.7; 4.5)	5.27 br.t (3.7)	0.73; 0.75; 0.76; 0.80; 0.94; 1.06; 1.10; 1.52 d (6.1)
2	3.20 dd (11.8; 4.4)	5.40 t (3.5)	0.83; 0.88; 0.90; 0.95; 0.97; 1.12; 1.25; 1.58 d (6.1)
6	3.17 dd (11.6; 4.4)	5.35 t (3.6)	0.78; 0.81; 0.84; 0.86; 1.01; 1.13; 1.16; 1.58 d (6.1)

Compound	Anomeric protons of the carbohydrates				
	Ara	Rha	Glc-1	Glc-2	Glc-3
1	4.58 (5.8)	6.03 d (1.8)	4.96 d (7.7)	6.05 d (7.9)	4.84 d (7.7)
2	4.66 d (6.2)	6.02 d (1.6)	4.99 d (7.4)		
6	4.66 d (6.2)	6.16 d (1.8)	5.06 d (7.8)	6.24 d (8.0)	

d – doublet, s – singlet, br.t – broadened triplet. The signals of the methyl groups had a singlet nature, with the exception of the rhamnose CH_3, which had doublet splitting.

TABLE 3. Chemical Shifts of the Protons in the PMR Spectrum of Progenin (**2**) (carbohydrate moiety)

Atom	Rha		Glc		Ara	
	δ, ppm	J, Hz	δ, ppm	J, Hz	δ, ppm	J, Hz
H-1	6.02	d 1.6	4.99	d 7.4	4.66	d 6.2
H-2	4.52	dd 1.6; 3.2	3.97	t 7.4	4.24	dd 6.2; 7.5
H-3	4.46	dd 3.2; 8.5	4.01	m	4.09	dd 7.5; 5.0
H-4	3.98	t 8.5	3.94	m	4.27	m
H-5a	4.69	m 8.5; 6.5	3.67	ddd 2.5; 5.0; 8.0	3.73	dd 11.0; 7.5
H-5e					4.45	dd 11.0; 4.0
H-6	1.56	d 6.5	4.26	dd 2.5; 11.0		
H-6			4.12	dd 11.0; 5.0		

The positions of the links between the monosaccharide residues in progenin (**2**) were determined with the aid of NMR spectroscopy. By employing the procedures of two-dimensional homonuclear resonance (TOCSY, COSY, RELAY) we first made a complete assignment of the proton magnetic resonance signals in the 3.6-6.0 ppm interval (Tables 2 and 3). Then, using the HETCOR program, we determined the values of the chemical shifts of the carbon atoms of the monosaccharide residues forming the carbohydrate chain of progenin (**2**) (see Table 1). The results showed that the terminal rhamnose was attached to the C-2 position of β-D-glucose (glycosylation effect, $\Delta\delta$,3-4 ppm [6]. In its turn, this disaccharide was linked with the C-4 atom of the α-L-arabinose residue ($\Delta\delta$ = +5-6 ppm) [7], i.e., the progenin (**2**) had the structure of oleanolic acid 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside.

The partial acid hydrolysis of thalicoside D with 10% oxalic acid gave, together with the monodesmosides (**2**)-(**4**), two bisdesmosides (**5** and **6**). Compound (**6**) differed from progenin (**2**) by the presence of an additional O-glycosyl fragment, as followed from a comparison of the ^{13}C NMR spectra of these two substances (see Table 1). The chemical shifts of the carbon atoms of the additional glucosyl residue in the spectrum of (**6**) corresponded to that of a 1-O-acyl glycoside [8]. Consequently, progenin (**6**) had the structure of oleanolic acid 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside] 28-O-β-D-glucopyranoside.

Progenins (**3**)-(**5**) were identified by comparison with authentic specimens [4].

Three carbohydrates (arabinose, glucose, and rhamnose) were identified qualitatively in the products of the total acid hydrolysis of glycoside (**1**). Thus, the terminal hexose in the acyloside chain must have been glucose. As compared with progenin (**6**), thalicoside D contained one more glucose residue. The position of attachment of this carbohydrate at C-6 of the acyl glucoside was determined with the aid of ^{13}C NMR spectra. Thus, on passing from progenin (**6**) to glycoside (**1**), apart from the newly appearing signals of the carbon atoms of the terminal glucose residue, changes were observed only for the following atoms of the acyloside glucose: C-6 $\Delta\delta$(6-1) +7.14 ppm and C-5 $\Delta\delta$(6-1) −1.35 ppm (see Table 1).

Thus, thalicoside D has the structure of oleanolic acid 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside 28-O-[O-β-D-glucopyranosyl-(1→6)-β-D-glucopyranoside].

Progenin (**6**) and the products of its degradation were identical with the thalicoside B previously isolated from *Thalictrum minus* and its progenins [4]. However, in the paper cited, this glycoside was ascribed the structure of oleanolic acid 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→3)-α-L-arabinopyranoside 28-O-β-glucopyranoside. On the basis of the results obtained in the determination of the structure of thalicoside D, we consider it necessary to revise the structure of thalicoside B as (**6**), where there is a 1→4 bond between the glucose and arabinose residues. Consequently thalicoside B has the structure of oleanolic acid 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside 28-O-β-glucopyranoside.

It was possible to find confirmation of this in the NOESY spectra of thalicoside D. The signal of the anomeric proton of glucose 1 had, in addition to cross-peaks with H-2, H-3, and H-5 of its own residue (4.02, 4.97, and 3.70 ppm), an intense cross-peak at 4.28 ppm where the H-4 proton of arabinose resonates.

EXPERIMENTAL

For column chromatography we used type L 40/100 silica gel and the anion-exchange cellulose Servacel DEAE 23 SS. TLC was conducted on silica gel L 5/40 with the following solvent systems: 1) ethyl acetate–methanol–water (10:3:2); 2) chloroform–methanol–water (70:23:4), 3) n-butanol–ethanol–water (5:3:2), 4) ethyl acetate–hexane–methanol (6:40:5), 5) chloroform–methanol–water (70:23:1).

In TLC the substances were revealed by 0.5% vanillin in 50% orthophosphoric acid with heating to 120°C. Melting points were determiend on a Boetius stage, and angles of rotation on a Polamat A polarimeter.

Mass spectra were taken by Yu. M. Mil'grom (IKhRV AN RUz, Tashkent) on a MKh-1310 instrument, with a St. Petersburg NTO [Scientific and Technical Co.] secondary ion source and bombardment with cesium ions having an energy of 7 keV, the matrices being glycerol and glycerol + NaCl. IR spectra were recorded on a Specord 75 IR instrument in a KBr tablet, and PMR spectra under the following conditions: Varian VXR 500 S, ^1H – 499.843 MHz, ^{13}C – 125.697 MHz. Solvent C_5H_5N. Internal standard – TMS. 2D NMR spectra were recorded under conditions described previously [2, 3].

Isolation of Thalicoside D (1). A fraction containing saponins A, B, C, and D (40 g) was obtained as described previously [1]. An eluent containing saponins C and D was isolated by repeated chromatography on silica gel in systems 2 and 5, and the residue from its evaporation was dissolved in methanol and reprecipitated with acetone. The precipitate (5.06 g) was chromatographed in system 2. This led to the isolation of 1.3 g of thalicoside D contaminated with flavonoids. The compound was purified further on anion-exchange cellulose (in the OH form) with elution by water–methanol mixtures having proportions of the latter increasing from 0 to 10%. This yielded 950 mg of thalicoside D (**1**).

Thalicoside D (1). $C_{59}H_{96}O_{26}$, mp. 218-220°, $[\alpha]_{546}^{20}$–0.8° (s 5.0; water). SIMS mass spectrum, *m/z* (%): 1243 [(M + Na)$^+$, 80], 1259 (M + K)$^+$, 1225(6), 1097(4), 1081(5), 1079(5), 935(7), 919(24), 873(15), 803(7), 671(7), 595(13), 481(15), 479(14), 439(38), 393(20), 365(16), 363(15), 347(100).

Complete Acid Hydrolysis of Thalicoside D (1). A mixture of 10 mg of (**1**) and 2 ml of 5% sulfuric acid was heated in a sealed tube for 5 h. The precipitate that had formed was separated off and washed with water, and it was identified by TLC in system 4 with an authentic specimen as oleanolic acid. The mother solution was neutralized with AV-17 anion-exchange resin, and the carbohydrates arabinose, glucose, and rhamnose were identified in it by TLC in system 3.

Alkaline Hydrolysis. A solution of 120 mg of (**1**) in 7 ml of 10% KOH in 70% ethanol was heated at 90°C for 3 h, the reaction mixture was neutralized with KU-1 cation-exchange resin and was then evaporated until the ethanol had been eliminated completely, and it was diluted with water, and extracted with butanol. The butanolic extracts were evaporated and chromatographed on silica gel in system 5. A progenin (60 mg) was isolated that, from its ^1H and ^{13}C NMR spectra and the absence of a depression of the melting point with an authentic specimen, was identified as the progenin (**2**), which we had isolated previously from the same source [4].

Partial Hydrolysis of Thalicoside D. A solution of 500 mg of (**1**) in 100 ml of a 10% solution of oxalic acid was heated at 75°C for 4 h, and the hydrolysate was extracted with *tert*-amyl alcohol. The extracts were washed with water and evaporated. The residue (300 mg) was chromatographed on silica gel in system 1. By rechromatography in systems 2 and 5, 30 mg of progenin (**2**) and 25 mg of progenin (**6**), and also progenins (**3**)-(**5**), were isolated.

Oleanolic Acid 3-O-α-L-Arabinopyranoside (4) from (1). Identical with an authentic sample according to TLC in systems 1 and 5.

Oleanolic Acid 3-O-[O-α-*L*-Rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside] (2) from (1). mp 248-250°C (methanol). ν_{max} (KBr, cm^{-1}): 1698 (acid C=O), 3310-3450 (OH) [4]. For the ^1H and ^{13}C NMR spectra, see Tables 1 and 2.

Oleanolic Acid 3-O-[O-β-D-Glucopyranosyl-(1→4)-α-L-arabinopyranoside] (3) from (1). Identified with an authentic specimen by TLC in systems 2 and 5 [4].

Oleanolic Acid 3-O-[O-β-D-Glucopyranosyl-(1→4)-α-L-arabinopyranoside] 28-O-β-D-Glucopyranoside (5) from (1). Identified with an authentic specimen by TLC in systems 2 and 5 [4].

Oleanolic Acid 3-O-[O-α-L-Rhamnopyranosyl-(1→2)-O-β-D-glucopyranosyl-(1→4)-α-L-arabinopyranoside] 28-O-β-D-Glucopyranoside (6) from (1). Vitreous mass, mp 212-216°C, $[\alpha]_{max}^{20}$+10.10° (*c* 3.58; pyridine) ν_{max} (KBr, cm^{-1}): 1740 (ester C=O), 3400-3500 (OH). For the ^1H and ^{13}C NMR spectra, see Tables 1 and 2.

REFERENCES

1. M. N. Mats, V. V. Korkhov, V. I. Lutskii, A. S. Gromova, T. V. Ganenko, and A. A. Semenov, Rast. Res., No. 4, 570 (1988).
2. A. S. Gromova, V. I. Lutskii, S. V. Zinchenko, N. N. Trofimova, A. A. Semenov, and N. A. Nakhova, Khim. Prir. Soedin., 103 (1993).
3. A. S. Gromova, V. I. Lutskii, S. V. Zinchenko, T. V. Ganenko, and A. A. Semenova, Khim. Prir. Soedin., 567 (1993).
4. A. S. Gromova, V. I. Lutskii, A. A. Semenov, R. B. Valeev, G. A. Kalabin, and Yu. N. El'kin, Khim. Prir. Soedin., 670 (1985).
5. K. Tori, A. Shimiako, S. Seo, and Y. Tomito, Tetrahedron Lett., 4227 (1974).
6. Y. Kurihara, K. Ookubo, H. Tasaki, H. Kodama, Y. Akiyama, A. Yagi, and B. Halpern, Tetrahedron, **44**, 61 (1986).
7. N. K. Kochetkov, A. Ya. Ott, and A. S. Shashkov, Izv. Akad. Nauk SSSR, Ser. Khim., 200 (1986).
8. V. U. Ahmad, N. Bano, I. Fatima, and S. Bano, Tetrahedron, **44**, 247 (1988).