TRITERPENE SAPONINS FROM Thalictrum minus.

VI. STRUCTURE OF THALICOSIDE C

A. S. Gromova, V. I. Lutskii,
S. V. Zinchenko, N. N. Trofimova,
A. A. Semenov, and N. A. Nakhova

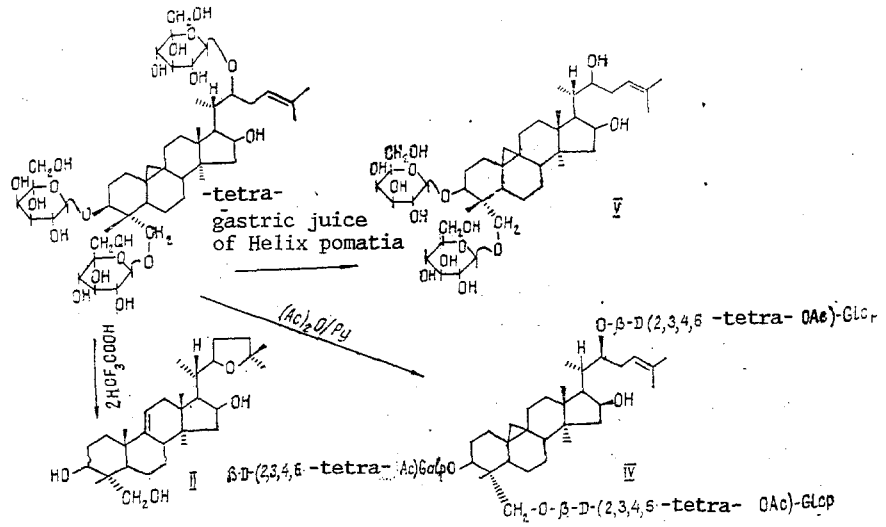
UDC 547.919:543.422.25:543.51

From the epigeal part of Thalictrum minus L. (Ranuncluaceae) we have isolated a minor cycloartane trisdesmoside – thalicoside C – 3β,16β,22(S),29-tetrahydroxy-9,19-cyclo-20(S)-lanost-24-ene 3-O-β-galactopyranoside 22,29-di-O-β-D-glucopyranoside. Its structure was established from the results of chemical transformations and of FAB-mass and NMR spectroscopies ($^1H_1$, ^{13}C, $^1H-^1H_1$, and $^{13}C-^1H$).

In the present paper we consider the structure of a triterpene glycoside – thalicoside C (I) – one of the minor components of a methanolic extract of Thalictrum minus (low meadow rue). We have previously reported the isolation from this species of seven triterpene compounds that are derivatives of cycloartane and oleanane [1-3].

It followed from its ^1H and ^{13}C NMR spectra that the thalicoside C molecule contains a cyclopropane ring (0.39 and 0.65 ppm, d, J = 4.0 Hz; 19.77 and 26.00 ppm), five methyl groups bound to quaternary carbon atoms, three of which are in the state of sp^3-hybridization (1.04 s; 1.13 s; 1.50 s ppm) and two in the state of sp^2-hybridization (1.80 ppm, $2CH_3$, br.s), and one methyl group bound to a tertiary carbon atom (1.32 ppm, d, J = 6.3 Hz) in the state of sp^3-hybridization. There is a trisubstituted double bond in the molecule (5.72 ppm, br.t, J = 6.3 Hz; 123.10 and 131.50 ppm). The facts given are characteristic for triterpenoids of the cycloartane series.

The fact that the compound under investigation was a glycoside was shown by its chromatographic behavior and by the presence in the NMR spectra of the signals of three anomeric protons (5.00, 5.12, and 5.28 ppm, J ≈ 8 Hz) and of three acetal carbon atoms (105.32, 106.37, and 106.63 ppm). Moreover, the FAB-mass spectrum of thalicoside C contained a molecular cluster ion with m/z 983 $(M + Na)^+$ and fragmentary ions with m/z 821 $[(M + Na^+)^+ - 162]$, 803 $[(M + Na)^+ - 162 - H_2O]$, 641 $[(M + Na)^+ - 324 - H_2O]$, 439 $[(M + Na)^+ - 486 - 2H_2O]$, corresponding to the loss of one, two, and three hexose molecules with the simultaneous elimination of water.



Irkutsk Institute of Organic Chemistry, Siberian Branch, Russian Academy of Sciences. Translated from Khimiya Prirodnykh Soedeninii, No. 1, pp. 103-109, January-February, 1993. Original article submitted March 3, 1992.

0009-3130/93/2901-0087$12.50

TABLE 1. Chemical Shifts of the Carbon Atoms of Thalicoside C (I) and Its Derivatives (II), (IV), and (V) [(I), (II), and (V) in C_5D_5N, and (IV) in $CDCl_3$; δ, ppm; TMS – 0]

C-atom of the genin	I	II	IV	V	C-atom of the genin	I	II	IV	V
1	32,07	38,8	31,8	32,07	16	71,85	71,8	75,4	71,78
2	29,53	27,4	29,3	29,53	17	51,85	51,6	50,6	53,06
3	81,73	73,7	82,7	81,62	18	20,62	15,8	20,2	20,74
4	45,08	43,5	44,7	45,03	19	30,50	23,3	30,1	30,62
5	40,72	46,2	40,3	40,70	20	34,51	35,2	32,8	36,07
6	20,78	21,8	19,8	20,67	21	13,10	15,3	11,2	14,65
7	26,62	28,4	25,9	26,50	22	85,28	82,8	82,7	75,40
8	48,45	42,4	48,1	48,60	23	33,36	28,4	33,2	33,60
9	19,77	149,4	18,8	19,72	24	123,10	37,9	121,0	123,87
10	26,00	39,8	25,9	26,00	25	131,50	80,6	133,3	132,11
11	26,00	151,5	25,9	26,44	26	25,94	27,9	26,1	25,81
12	33,50	36,7	32,8	33,60	27	18,24	28,9	18,0	18,14
13	46,03	45,2	46,0	45,97	28	19,67	19,7	19,8	19,65
14	47,34	45,4	47,3	47,33	29	71,38	68,8	70,1	71,31
15	49,67	46,7	46,4	48,88	30	11,80	13,2	11,2	11,73

TABLE 2. ^1H and ^{13}C NMR Spectral Characteristics of the Carbohydrate Residues in Thalicoside C (I) [C_5D_5N; T = 40°C; δ, ppm; TMS – 0]

C atom	3-O-β-D-Galp residue		22-O-β-D-Glcp residue		29-O-β-D-Glcp residue	
	C	H	C	H	C	H
1	106,37	5,30	106,62	4,90	105,32	5,15
2	75,58	4,12[a]	75,43	3,92	75,58	3,88
3	73,42	4,32	78,89	4,05	78,69	4,08
4	70,52	4,45	72,02	4,05	72,02	4,01
5	76,30	4,15[a]	78,03[b]	3,75	78,23[b]	3,86
6	62,50	4,30[a] 4,43[a]	63,06	4,28[a] 4,38[a]	63,17	4,28[a] 4,38[a]

[a]CSs determined from the 2D $^{13}C-^1H$ spectra with an accuracy of ±0.03 ppm.
[b]Alternative assignment.

The acid hydrolysis of thalicoside C (scheme) led to the triterpenoid (II) and to the carbohydrates glucose and galactose (TLC analysis). A comparison of the aglycon parts of the ^1H and ^{13}C NMR spectra of compounds (II) and (I) (Table 1) showed that triterpenoid (II) was not the native genin. Compound (II) was identified as an artefact of thalicogenin [1] by high-performance liquid chromatography using the method of additives.

Enzymatic hydrolysis of thalicoside C with the gastric juice of Helix pomatia led to the progenin (V), which was identified from its ^1H and ^{13}C NMR and mass spectra as thalicoside A [1].

Thus the totality of spectral results for thalicoside C (Tables 1 and 2) and the products of acid (II) and enzymatic (V) hydrolysis permitted the conclusion that the genin of thalicoside C was thalicogenin (III) [1] with galactosyl and glucosyl residues attached at C-3 and C-29, respectively.

The third carbohydrate residue in glycoside (I) was glucose having the β-configuration of the glycosidic center and the pyranose form of the oxide ring. These conclsuions followed from the CSs in the ^{13}C NMR spectra of thalicoside C and its acetate (IV) [4]. To determine the position of attachment of the second glucose residue we made use of the following facts: the signals of the C-22 atom of the genin, C-2 of galactose, and C-2 of glucose lay in the 75.4-75.6 ppm region, and any of these positions could be glycosylated, since in the ^{13}C NMR spectrum of thalicoside C one of the signals mentioned had undergone a shift to 85.28 ppm.

The choice from the three alternative structures was based on two-dimensional NMR spectra. By making use of the 2D $^{13}C-^1H$ correlation spectrum for (I) (Fig. 1) we found an interrelationship of the signals at 85.28 ppm (F2 scale) and 4.50 ppm (F1 scale). It was

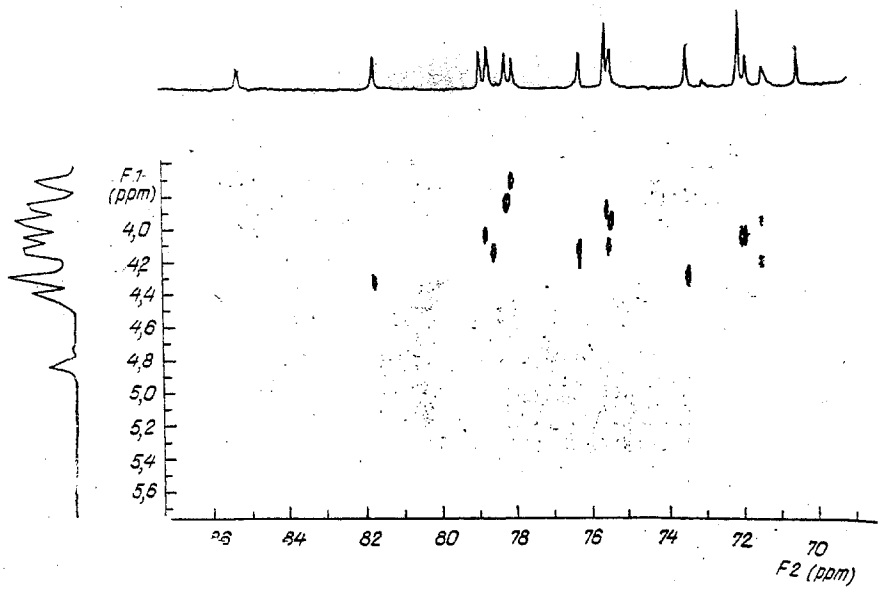


Fig. 1. Fragment of the two-dimensional $^{13}C-^{1}H$ correlation spectrum of compound (I) (T = 40°C).

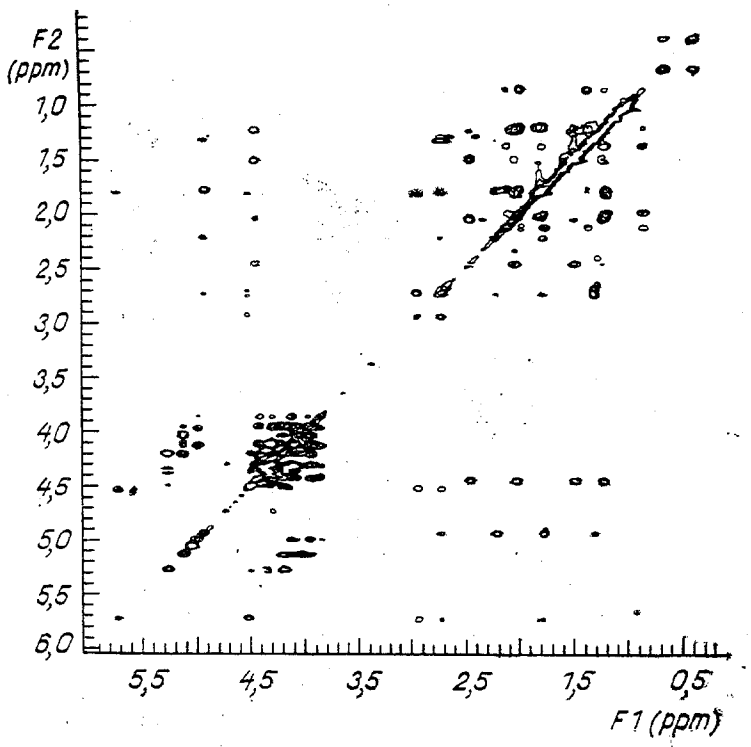


Fig. 2. Two-dimensional NMR $^{1}H-^{1}H$ (TOCSY) spectrum of compound (I) (T = 90°C).

shown with the aid of the 2D TOCSY procedure [5, 6] that the proton signal at 4.50 ppm had no cross-peaks with the signals of the carbohydrate protons and, consequently, the third carbohydrate residue was attached to the genin (Fig. 2). In actual fact the proton with a signal at 4.50 ppm gave a cross-peak only with the H-24 proton (5.72 ppm) and, in the strong field, with the H-20 proton (2.75 ppm) and with the two H-23 protons (2.70 and 2.91 ppm). This enabled the signals at 4.50 and 85.28 ppm to be assigned to the H-22 and C-22 atoms, respectively, and permitted the conclusion that the third carbohydrate residue was attached at the C-22 position of the genin.

Thus, thalicoside C has the structure 3β,16β,22(S),29-tetrahydroxy-9,19-cyclo-20(S)-lanost-24-ene 3-O-β-galactopyranoside 22,29-di-O-β-D-glucopyranoside.

EXPERIMENTAL

For column chromatography we used type L 40/100 silica gel and the anion-exchange cellulose derivative Servacel DEAE 23SS. TLC was conducted on silica gel L5/50 with the following solvent systems: 1) chloroform–methanol–water (70:23:1); 2) chloroform–methanol–water (70:23:4); 3) chloroform–methanol–water (80:35:6); 4) chloroform–methanol–water (60:35:8); 5) hexane–acetone (3:2); and 6) n-butanol–thanol–water (5:3:2). For revealing the saponins we used a 0.5% solution of vanillin in 50% orthophosphoric acid with heating to 120°C.

HPLC analysis was performed by A. G. Gorshkov (IrIOKh) on a Milikhrom 1 microcolumn liquid chromatograph with UV detection at a wavelength of 200 nm. Chromatographic conditions: 2 × 64 nm column filled with the reversed-phase sorbent Nucleosil-5 C18. The mobile phase was acetonitrile–water (65:35). Rate of elution 200 μl/min, sensitivity of the detector 1.6 A.

Melting points were determined on a Boëtius stage, and angles of rotation on Polamat A polarimeter. The mass spectra of (I) and (V) were taken by A. L. Vereshchagin (IrIOKh) on a LKB-2091/PDP-11/34 instrument with an ion source from Iontech Ltd., Teddington, UK. Ionization was brought about by a beam of accelerated xenon atoms with an energy of 6 keV at a discharge current of 1.2 mA. The matrix used was glycerol with added NaCl. IR spectra were recorded on a Specord 75 IR instrument in paraffin oil.

NMR spectra were recorded under the following conditions: Bruker WP-200 (^{1}H – 200.13 MHz) and Jeol FX-90 Q (^{13}C – 22.49 MHz) instruments for compounds (II) and (IV); Varian VXR 500S (^{1}H – 499.483 MHz; ^{13}C – 125.697 MHz) instrument for compounds (I) and (V). The solvents were C_5D_5N and $CDCl_3$, and the internal standard TMS.

Two-Dimensional NMR Spectroscopy. 2D NMR spectra were recorded on a Varian VXR 500S spectrometer fitted with a SUN 3/50 computer and the standard VNMR equipment. The following procedures were used to obtain the two-dimensional ^{1}H–^{1}H and ^{13}C–^{1}H spectra:

TOCSY – the standard TOCSY program in the phase-sensitive variant. Size at the matrix 2K × 1K; width of the spectrum 3000 Hz. Length of the 90° reading and trim pulses 22 μsec. Mixing time 0.05 msec. Relaxation delay 1.8 sec. To bring the spin system into equilibrium we used the sequence sspul: 90°–H–90°. We used a Gaussian function with zero shift for weighting before Fourier transformation.

2D ^{13}C–^{1}H – the standard HETCOR program. Size of the matrix 2K × 0.25K. Width of the spectrum: for ^{13}C: 2500 Hz; for ^{1}H: 1000 Hz. Length of the ^{13}C 90° pulse 12.4 μs, and of the ^{1}H 90° pulse 18 μsec. Relaxation delay 2 sec. The delay τ was tuned to the value J = 140 Hz. For weighting, in the F2 region we used a bell-shaped function with zero shift. In a number of cases we made use of the suppression of the signal of residual protons by presaturation for one second.

Isolation of Thalicoside C (I). A fraction containing saponins A, B, C, and D (40 g) was obtained as described previously [1]. By repeated column chromatography on silica gel in systems 1-3 we isolated a mixture of saponins B and C which, after evaporation, was dissolved in the minimum volume of methanol and was precipitated with acetone. The residue (3.96 g) was chromatographed in system 3, and 668 mg of thalicoside C contaminated with flavonoids was isolated. The further purification of the compound was achieved on an anion-exchange cellulose derivative (in the OH form) with elution by water containing increasing concentrations (from 0 to 10%) of methanol. This face 650 mg of thalicoside C (I).

Thalicoside C (I). $C_{48}H_{80}O_{19}$, mp 205-207°C, $[\alpha]_{546}^{24}$ +50° (c 1.0; pyridine). FAB mass spectrum, m/z: 983 $(M + Na)^+$, 965 $[(M + Na)^+ - H_2O]$, 821 $[(M + Na)^+ - 162]$, 803 $[(M + Na)^+ - 162 - H_2O]$, 641 $[(M + Na)^+ - 324 - H_2O]$, 636 $(M - 324)^+$, 439 $[(M + H)^+ - 486 - 2H_2O]$, 421 $[(M + H)^+ - 486 - 3H_2O]$.

^{1}H NMR spectrum (500 MHz C_5D_5N, 90°C, δ, ppm): 0.39; 0.65 (2H·19, d, ^{2}J = 4.0 Hz); 1.04 (s, CH_3); 1.13 (s, CH_3); 1.32 (d, ^{3}J = 6.3 Hz, CH_3-21); 1.50 (s, CH_3); 1.80 (br.s, CH_3-26, 27. The CSs for the protons given below were determined from 2D ^{1}H–^{1}H (TOCSY) spectra with an accuracy of ±0.01 ppm: 2.49; 2.01 (2H-2); 1.30; 1.69 (2H-15); 2.20 (H-17); 2.70 (H-23); 2.75 (H-20); 2.91 (H-23); 4.45 (H-3); 4.50 (H-22); 3.96; 4.20 (2H-29, the CS_S were determined from the 2M ^{13}C–^{1}H spectra with an accuracy of ±0.03 ppm); 4.94 (H-16); 5.00 (H_{anomer}, d, ^{3}J = 7.8 Hz); 5.12 (H_{anomer}, d, ^{3}J = 7.5 Hz); 5.28 (H_{anomer}, d, ^{3}J = 7.8 Hz); 5.72 (H-24, br.t, ^{3}J = 6.3 Hz). The NMR spectra of (I), which were taken at T = 40°C, are given in Tables 1 and 2.

Acid Hydrolysis of Thalicoside C (I). A solution of 32 mg of compound (I) in 3 ml of 2 N CF_3COOH was heated in the water bath at 90°C for 1.5 h and then the reaction mixture was diluted with a twofold volume of methanol and was evaporated. The operation was repeated until the acetic acid had been completely eliminated, and the evaporated residue was treated with water. Glucose and galactose were detected in the lyophilic fraction of the hydrolysate by TLC in systems 2-4. The lyophobic residue was analyzed by HPLC, the retention volumes of a standard (the artefact from thalicogenin) and of the hydrolysis product being compared: $V_{R\,stand} = V_{R\,hydr.prod} = 1255$ µl.

Enzymatic Hydrolysis of Thalicoside C (I). A solution of 286 mg of thalicoside C in 50 ml of H_2O was treated with 100 mg of the unpurified enzyme of Helix pomatia and 1 ml of toluene, and the mixture was subjected to constant stirring in the dark at 30°C for 7 days with the addition of a further 7 mg of enzyme to the reaction mixture each day. Monitoring was carried out by TLC in system 4. The reaction product was extracted with mixtures of chloroform and methanol (1:1 and 1:2). The residue was exhaustively extracted with butanol. The chloroform–methanolic and butanolic extracts were combined, washed with water and evaporated, and the residue was chromatographed on a silica gel column in systems 1-4. This gave 47.2 mg of a progenin, and this was identified as thalicoside A (V) [1].

The Progenin (Thalicoside A). $C_{42}H_{70}O_{14}$, mp 255-257°C (methanol), $[\alpha]_{546}^{22}$ + 8.94° (c 2.0; pyridine). FAB mass spectrum, m/z: 798 $(M + H)^+$; 677 $[(M + H)^+ - 162]$; 457 $[(M + H)^+ - 162 - 18]$. ^1H NMR spectrum (500 MHz, C_5D_5N, δ, ppm): 0.25; 0.46 (2H-19, d, ^2J = 4.0 Hz); 0.89 (s, CH_3); 0.92 (s, CH_3); 1.20 (d, ^3J = 6.5 Hz); 1.44 (s, CH_3); 1.66 (br.s, (CH_3); 1.65 (br.s, CH_3-26); 2.07; 1.70 (2H-15 m); 2.34 (H-17, m); 2.48; 2.58 (2H-23, m); 2.58 (H-20, m); 4.20-4.40 (2H-29, m); 4.31 (H-22, m); 4.50 (H-3, m); 4.83 (H-16, m; 5.31 ($H_{anomer.glucose}$, d, ^3J = 7.8 Hz); 5.47 ($H_{anomer.galactose}$, d, ^3J = 8.0 HZ); 5.58 (H-24, br.t, ^3J = 6.0 Hz). The ^{13}C NMR spectrum is given in Table 1.

Acetylation of Thalicoside C (I). A solution of 89.6 mg of thalicoside C in 1.2 ml of pyridine was treated with 0.7 ml of acetic anhydride, and the reaction mixture was kept at room temperature for 8 h. The reaction product was worked up by the scheme given above and was chromatographed on a silica gel column in system 5. This yielded 60 mg of thalicoside C acetate (IV).

Thalicoside C Pentadecaacetate (IV). $C_{74}H_{106}O_{32}$, mp 160-162°C, $[\alpha]_{546}^{24}$ + 13.33° (c 0.1; $CHCl_3$). IR spectrum, cm^1 (CCl_4): 3430; 1735. ^1H NMR spectrum (200 MHz, $CDCl_3$, δ, ppm): 0.28; 0.42 (2H-19, d, ^2J = 4.0 Hz); 0.64 (s, CH_3); 0.82 (d, ^3J = 6.5 Hz); 0.88 (s, CH_3); 0.92 (s, CH_3); 1.27 (d, ^3J = 6.5 Hz); 1.43 (s, CH_3); 1.62 (s, CH_3); 1.65 (s, CH_3); 3.29 (H-22, m); 3.30; 3.46 (2H-29, d, ^2J = 10.0 Hz); 3.5-3.7 (H-3, 2H of glucose); 3.9-4.35 (H-5, 2H-6 of galactose, 4H-6 of glucose); 4.44 (H_{anomer}, ^3J = 7.8 Hz); 4.47 (H_{amomer}, ^3J = 8.0 Hz); 4.51 (H_{anomer}, ^3J = 7.8 Hz); 4.9-5.25 (H-16, H-24 of the genin, 3H-2, 3H-3, 2H-4 of carbohydrates); 5.33 (H-4, d, ^3J = 4.0 Hz of galactose). The ^{13}C NMR spectrum is given in Table 1.

LITERATURE CITED

1. A. S. Gromova, V. I. Lutskii, A. A. Semenov, V. A. Denisenko, and V. V. Isakov, Khim. Prir. Soedin., 213 (1984).
2. 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).
3. A. S. Gromova, V. I. Lutskii, A. L. Vereshchagin, and A. A. Semenov, Khim. Prir. Soedin., 107 (1987).
4. S. Seo, G. Tomita, K. Tori, and G. Goshimura, J. Am. Chem. Soc., 100, 3331 (1978).
5. A. Bax and D. G. Davis, J. Magn. Reson., 65, 355 (1985).
6. M. Levitt, R. Freeman, and T. Frenkell, J. Magn. Reson., 47, 328 (1982).