New triterpenoid glycosides from *Thalictrum minus* L. 10.* Structure of thalicosides G_1 and G_2

N. N. Trofimova,[a] A. S. Gromova,[a] V. I. Lutsky,[a] A. A. Semenov,[a] S. A. Avilov,[b] A. I. Kalinovsky,[b] D. Li,[c] and N. L. Owen[c]*

[a]*Irkutsk Institute of Chemistry, Siberian Branch of the Russian Academy of Sciences, 1 ul. Favorskogo, 664033 Irkutsk, Russian Federation. Fax: +7 (395 2) 35 6046. E-mail: admin@irioch.irk.ru*

[b]*Pacific Institute of Bioorganic Chemistry, Far-Eastern Branch of the Russian Academy of Sciences, 159 prosp. 100-letiya Vladivostoka, 690022 Vladivostok, Russian Federation. Fax: +7 (423 2) 31 4050*

[c]*Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, USA. Fax: (801) 378 5474*

Two triterpenoid diglycosides of the cycloartane series were isolated from the terrestrial part of *Thalictrum minus* L. (Ranunculaceae). Genins of these glycosides are side-chain structural isomers — 3-*O*-β-D-galactopyranosyl-29-*O*-β-D-glucopyranosyl-9β,19-cyclo-20(*S*)-lanost-24(*Z*)-ene-3β,16β,22(*S*),26,29-pentaol and 3-*O*-β-D-galactopyranosyl-29-*O*-β-D-glucopyranosyl-9β,19-cyclo-20(*S*)-lanost-25-ene-3β,16β,22(*S*),24ξ,29-pentaol. The structures of these glycosides were established using 1D and 2D NMR spectroscopy and FAB mass spectrometry.

Key words: *Thalictrum*, Ranunculaceae, triterpenoid glycosides, saponins, cycloartanes; ^1H and ^{13}C NMR spectroscopy and FAB mass spectrometry.

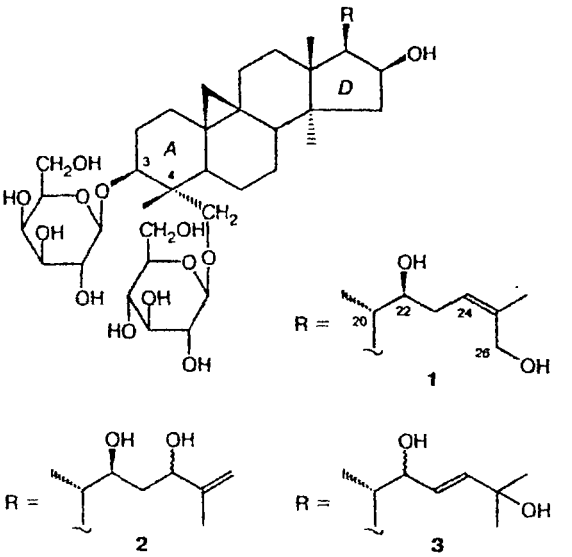
In continuation of the studies of secondary metabolites of meadow rues growing in Siberian biocenoses,[2] we isolated two new minor saponins from *Thalictrum minus* L. and called them thalicosides G_1 (**1**) and G_2 (**2**). The present work is devoted to elucidation of the structures of compounds **1** and **2** by 1D and 2D ^1H and ^{13}C NMR spectroscopy and FAB mass spectrometry.

Results and Discussion

Column chromatography on silica gel repeated many times and drop-counterflow distribution chromatography of the methanolic extract from the terrestrial part of *Thalictrum minus* L. yielded a saponin fraction that consisted, according to HPLC, of two components. The components were separated by preparative HPLC; this gave compounds **1** and **2**.

Compounds **1** and **2** give fast atom bombardment (FAB) mass spectra containing quasimolecular ions with *m/z* 837 $[M+Na]^+$ and belong to the class of diglycosides, as indicated by the fact that their ^1H and ^{13}C NMR spectra exhibit signals for two anomeric protons and two anomeric carbon atoms (Tables 1 and 2).

The triterpenoid nature of genins of **1** and **2** is indicated by the fact that the molecule contains 30 carbon atoms (apart from the carbohydrate residues), four tertiary and one secondary methyl groups, a double bond, and five oxygen-containing functions (FAB mass



* For Part 9, see Ref. 1.

Translated from *Izvestiya Akademii Nauk. Seriya Khimicheskaya*, No. 7, pp. 1434—1437, July, 1998.

1066-5285/98/4707-1395 $20.00

Table 1. Chemical shifts (δ) in the ^{13}C NMR spectra of compounds **1** and **2** (C_5D_5N, tetramethylsilane, T = 26 °C)

Atom	δ(±0.01)		Atom	δ(±0.01)		Atom	δ(±0.01)	
	1	**2**		**1**	**2**		**1**	**2**
C(1)	32.1 (t)	32.4 (t)	C(15)	48.7 (t)	48.4 (t)	C(29)	71.1 (t)	71.0 (t)
C(2)	29.4 (t)	29.8 (t)	C(16)	71.7 (d)	71.9 (d)	C(30)	11.7 (q)	11.7 (q)
C(3)	81.7 (d)	81.6 (d)	C(17)	53.0 (d)	53.2 (d)	C(1')[b]	106.5 (d)	106.1 (d)
C(4)	45.0 (s)	44.9 (s)	C(18)	20.6 (q)	20.6 (q)	C(2')	73.3 (d)	73.3 (d)
C(5)	40.7 (d)	40.7 (d)	C(19)	30.5 (t)	30.5 (t)	C(3')	75.5 (d)	75.4 (d)
C(6)	20.7 (t)	20.7 (t)	C(20)	36.0 (d)	36.5 (d)	C(4')	70.4 (d)	70.4 (d)
C(7)	26.5 (t)	26.5 (t)	C(21)	14.7 (q)	15.0 (q)	C(5')	76.2 (d)	76.1 (d)
C(8)	48.4 (d)	48.4 (d)	C(22)	75.1 (d)	72.5 (d)	C(6')	63.0 (t)	63.0 (t)
C(9)	19.8 (s)	19.8 (s)	C(23)	33.0 (t)	39.6 (t)	C(1")	105.0 (d)	104.9 (d)
C(10)	25.9 (s)	25.9 (s)	C(24)	125.5 (d)	72.9 (d)	C(2")[c]	75.3 (d)	75.2 (d)
C(11)	26.3 (t)	26.3 (t)	C(25)	137.5	[a]	C(3")	78.6 (d)	78.5 (d)
C(12)	33.6 (t)	33.6 (t)	C(26)	61.2 (t)	110.0 (t)	C(4")	72.1 (d)	72.0 (d)
C(13)	46.0 (s)	46.0 (s)	C(27)	22.5 (q)	18.8 (q)	C(5")	78.0 (d)	77.9 (d)
C(14)	47.3 (s)	47.3 (s)	C(28)	19.5 (q)	19.8 (q)	C(6")	62.5 (t)	62.4 (t)

[a] Overlapped by the signal of the solvent. [b] The signals of the C(1')—C(6') atoms refer to the 3-*O*-β-D-Gal*p* residue. [c] The signals of the C(1")—C(6") atoms refer to the 29-*O*-β-D-Glc*p* residue.

Table 2. Data of the ^1H NMR spectra obtained in 1D and 2D experiments for compounds **1** and **2** (C_5D_5N, tetramethylsilane, 80 °C)

Atom	δ(±0.01) (*J*/Hz)	
	1	**2**
H(1)	1.15 (m)[a], 1.38 (m)[b]	1.15 (m)[a], 1.38 (m)[b]
H(2)	2.01 (m)[a], 2.43 (m)[b]	2.00 (m)[a], 2.43 (m)[b]
H(3)	4.46 (m)	4.46 (m)
H(5)	1.92 (m)	1.91 (m)
H(6)	0.78 (m)[a], 2.07 (m)[b]	0.70 (m)[a], 2.07 (m)[b]
H(7)	1.12 (m)[a], 1.70 (m)[b]	1.10 (m)[a], 1.70 (m)[b]
H(8)	1.98 (m)	1.95 (m)
H(15)	1.72 (m)[a], 2.11 (m)[b]	(1.74 (m)[a], 2.10 (m)[b]
H(16)	4.79 (m)	4.79 (m)
H(17)	2.33 (dd, J = 7.4, 10.5)	2.28 (dd, J = 7.1, 11.1)
H(18)	1.46 (s)	1.46 (s)
H(19)	0.34, 0.59 (both d, J = 3.9)	0.34, 0.59 (both d, J = 3.9)
H(20)	2.60 (m)	2.61 (m)
H(21)	1.18 (d, J = 7.0)	1.19 (d, J = 6.7)
H(22)	4.28 (m)	4.65 (br.d, J = 9.2)
H(23)	2.55 (m), 2.78 (m)	2.08 (m), 2.18 (m)
H(24)	5.72 (t, J = 7.1)	4.81 (m)
H(26)	4.41, 4.53 (both d, J = 12.2)	5.00 (br.s), 5.38 (br.s)
H(27)	1.98 (s)	1.93 (s)
H(28)	0.94 (s)	0.98 (s)
H(29)	4.15 (m)[c], 4.47 (m)[c]	4.11 (m)[c], 4.45 (m)[c]
H(30)	0.98 (s)	0.94 (s)

[a] Data for H(α). [b] Data for H(β). [c] The values of chemical shifts were found from the ^1H—^{13}C correlation spectra.

Table 3. Data of the ^1H NMR spectra of compounds **1** and **2** obtained from 2D experiments (carbohydrate moiety) (C_5D_5N, δ, tetramethylsilane, 26 °C)

Atom	δ(±0.03)		Atom	δ(±0.03)	
	1	**2**		**1**	**2**
	3-*O*-β-D-Gal*p*			29-*O*-β-D-Glc*p*	
H(1')	5.56 (d)[a]	5.39 (d)[a]	H(1")	5.55 (d)[b]	5.39 (d)[b]
H(2')	4.53 (m)	4.49 (m)	H(2")	4.34 (m)	4.18 (m)
H(3')	4.34 (m)	4.28 (m)	H(3")	4.32 (m)	4.29 (m)
H(4')	4.62 (m)	4.59 (m)	H(4")	4.33 (m)	4.30 (m)
H(5')	4.39 (m)	4.35 (m)	H(5")	4.08 (m)	4.06 (m)
H(6'a)	4.42 (m)	Not	H(6"a)	4.39 (m)	4.47 (m)
H(6'b)	4.58 (m)	identified	H(6"b)	Not identified	Not identified

[a] At 80 °C, the signals for the H(1') atoms in **1**: δ 5.31 (d, J = 7.64 Hz), those for the H(1') atoms in **2**: δ 5.31 (d, J = 7.84 Hz). [b] At 80 °C, the signals for the H(1") atoms in **1**: δ 5.18 (d, J = 7.54 Hz), those the H(1") atoms in **2**: δ 5.16 (d, J = 7.64 Hz).

spectra; ^1H and ^{13}C NMR spectra) (see Tables 1 and 2). The presence of two single-proton doublets forming AB systems at δ 0.34 and 0.59, J = 3.9 Hz, typical of a 1,1,2,2-tetrasubstituted cyclopropane ring (CPR) in the ^1H NMR spectra of compounds **1** and **2**, in combination with the data listed above, implies that genins of **1** and **2** can be classified as tetracyclic triterpenoids of the cycloartane series.

The ^{13}C NMR signals of compound **2** were assigned based on the results obtained with incomplete C—H decoupling (see Table 1) and on a comparison of the ^{13}C chemical shifts of compounds **1** and **2** with those observed in the spectrum of a known compound, thalicoside E (**3**) isolated previously.[3] The comparison demonstrated close chemical shifts of signals for the carbon atoms in the polycyclic fragment; this provides grounds for assuming that the structure of this fragment is the same in all the three molecules.

The data of the DQFCOSY and TOCSY 2D experiments in the ^1H NMR spectra confirm the above conclusion by demonstrating the identity of the spin systems of rings *A*—*D* in thalicosides G_1, G_2 (see Table 2), and E.

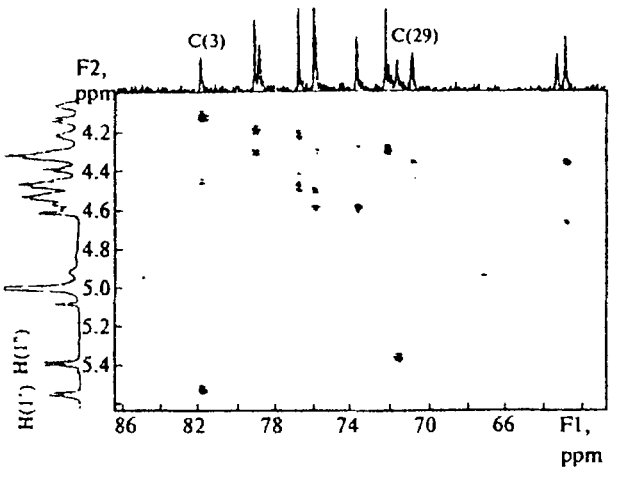


Fig. 1. HMBC NMR spectrum of compound **1**.

The chemical shifts of the signals for the C atoms in the carbohydrate fragments of compounds **1** and **2** correspond to those for terminal β-D-galactopyranoside and β-D-glucopyranoside[4] (see Table 1). The sites of attachment of galactose and glucose to genin, the C(3) and C(29) atoms, respectively, were determined in both cases from an HMBC experiment (Fig. 1).

The protons in the carbohydrate moieties of the genins of **1** and **2** were assigned using the data of the $^1H—^{13}C$ correlation and HMBC spectra (Table 3). Thus, for compounds **1** and **2**, we elucidated the structure of the polycyclic fragments of genins, the relative configurations of the chiral centers (3β-OR, 4α-CH_2OR, 16β-OH), and the structure of the carbohydrate part of the molecule. It was found that these fragments have the same structure in both compounds, and the difference between these two molecules is manifested in the structure of the C(17) side chain of the genin. The side fragments in both molecules contain two hydroxy groups and one double bond (see Tables 1 and 2).

The structure of the side chain in compound **1** was determined in the following way. In the ^1H NMR spectrum of thalicoside G_1, the signal of the H(24) olefinic proton occurs at δ 5.72 and is characterized by a spin-spin coupling constant J = 7.1 Hz. Irradiation of this proton under conditions of nuclear Overhauser effect (NOE) results in an increased intensity of the three-proton signal at δ 1.97, which is manifested as singlet. Thus, one substituent at the double bond is a methyl group, and the other one is a hydroxymethyl group. The ^{13}C NMR spectrum exhibits a signal at δ 61.2, which was assigned to a hydroxymethyl carbon atom that occupies a *trans*-position in relation to the olefinic proton (calculation from the data of Refs. 5, 6 and the data of Ref. 7). In the ^1H NMR spectrum, the same fragments account for two single-proton doublets at δ 4.41 and δ 4.53 (AB system, J = 12.2 Hz) characterized by small allylic constants with the olefinic proton, which were determined by double resonance in the difference variant. The same experiment has revealed allylic coupling with the *cis*-methyl group at δ 1.97.

The splitting of the signal of the H(24) olefinic proton into a triplet implies a methylene group located in the α-position with respect to this proton (at C(23)). Thus, the second hydroxy group in the side chain of compound **1** can be attached only to the C(22) atom, which is not at variance with the whole set of spectral data.

Analysis of the results of DQFCOSY and TOCSY 2D experiments confirmed the reliability of our conclusions; identification of the H(16) proton with δ 4.79 made it possible to identify the spin system of the side chain (δ):

H(16) (4.79) → H(17) (2.33) → H(20) (2.60) →
→ H(21) (1.18);

H(22) (4.28) → H(23α) (2.55), (23β) (2.78) →
→ H(24) (5.72) → H(27) (1.97).

The absolute configuration of the C(22) atom was found by a known procedure,[4] which makes it possible to establish the stereochemistry of secondary hydroxy groups in chiral secondary alcohols using the effect of glycosylation in the ^{13}C NMR spectra, with allowance made for published data.[8] Based on the resulting values, $\Delta\delta_S$(C(1′)) +1.22, $\Delta\delta_A$(C-α) +10.18, and $\Delta\delta_A$[C-β(H)] −1.46, we identified the configuration of the chiral center as 22(*S*).

Thus, for thalicoside G_1, we propose the structure of 3-*O*-β-D-galactopyranosyl-29-*O*-β-D-glucopyranosyl-9β,19-cyclo-20(*S*)-lanost-24(*Z*)-ene-3β,16β,22(*S*),26,29-pentaol (**1**).

The structure of the side chain in thalicoside G_2 was established in the following way. The ^1H NMR spectrum of compound **2** contains two broadened single-proton singlets at δ 5.00 and δ 5.38 and a three-proton singlet for the methyl group at δ 1.93. By the COSY experiment, we found that the signals of olefinic protons give cross-peaks with each other, while in the TOCSY experiment, both exhibit cross-peaks with the signal of the methyl group. Thus, we identified this fragment as a terminal double bond bearing a methyl group (CH_2=C(CH_3)—).

In the ^{13}C NMR spectrum of compound **2**, the signal corresponding to the C atom of the double bond[9] is manifested at δ 110.0 (C(26)), and the signal for the second carbon atom of the methylidene group (C(25)) is overlapped by the signal of pyridine.

The side chain of thalicoside G_2 contains two secondary hydroxy groups, whose geminal protons are responsible for a doublet at δ 4.65 with spin-spin coupling constant J = 9.2 Hz and a multiplet at δ 4.81 in the ^1H NMR spectrum of compound **2**.

The COSY spectrum contains no cross-peaks for the signals at δ 4.65 and δ 4.81. Hence, the two —CHOH-

groups are not bound to each other by a covalent bond. The two hydrogen atoms at δ 2.08 and δ 2.18 (2 H(23)) in the same spectrum exhibit cross-peaks both with the signal at δ 4.81 and with the signal at δ 4.65. It is obvious that the hydroxy groups are separated by the C(23) methylene bridge.

The signals for H(22) and H(24) were identified based on the presence of a cross-peak for the signal of the methylidene proton at δ 5.38 and the signal at δ 4.81 in the TOCSY spectrum of **2**. Hence, the latter can be assigned to the H(24) proton, and the signal at δ 4.65 belongs to H(22).

The signals at δ 72.5 and δ 72.9 (doublets) in the ^{13}C NMR spectrum of compound **2** correspond to the C atoms in the hydroxymethine fragments; based on the data of the HETCOR spectrum, they were assigned to C(22) and C(24), respectively.

The absolute configuration of the C(22) chiral center was determined to be 22(*S*) based on the values $\Delta\delta_S$(C(1′)) −1.22, $\Delta\delta_A$(C-α) +12.78, and $\Delta\delta_A$[C-β(H)] −1.99, found by the known procedure.[4]

Thus, for thalicoside G_2, we propose the structure of 3-*O*-β-D-galactopyranosyl-29-*O*-β-D-glucopyranosyl-9β,19-cyclo-20(*S*)-lanost-25-ene-3β,16β,22(*S*),24ξ,29-pentaol (**2**).

The metabolites of triterpenoid saponins of the cycloartane series isolated from *Thalictrum minus* L. are characterized by the presence of a relatively stable polycyclic fragment in the genin part of the molecule, whereas the side chains in the molecules differ by the number and positions of substituents (hydroxyl groups, carbohydrates) at the double bond, which demonstrates the large potential of secondary metabolism of the *Thalictrum* family.

Experimental

General procedures are presented in previous publications.[2,3] The IR spectra were recorded on an IFS25 Fourier Transform spectrometer.

^1H NMR spectra were recorded on a Varian VXR-500 S instrument (499.843 MHz) equipped with a SUN 3/50 computer with the standard VNMR software. The procedures for recording 2D DQFCOSY and TOCSY spectra were presented in a known paper.[2] The procedures for registration of the HETCOR and HMBC 2D spectra were also reported previously.[1] ^{13}C NMR spectra were measured on a Bruker WM-250 spectrometer (62.5 MHz).

The FAB mass spectra were run on a Jeol SX102A instrument with double focusing (using thioglycerol as the matrix and xenon as the ionizing gas).

The isolation of the saponin fraction was described previously.[3] The fraction was analyzed using a Milikhrom A-02 liquid chromatograph equipped with a 2×70 mm analytical column with Nucleosil 5-C18 using MeOH : 0.0005 *M* KH_2PO_4, 60 : 40 as the eluent. UV detection was performed at 200 and 210 nm.

Isolation of compounds 1 and 2. Preparative HPLC of the saponin fraction (0.11 g) was carried out on a Yanako-2000L instrument with a refractometry detector using a 10×250 mm column with Silasorb 12-C18. A 55 : 45 MeOH—H_2O mixture, 3 mL min^{-1}, served as the mobile phase. TLC analysis was carried out on silylated plates using a 42 : 58 EtOH—H_2O mixture. Compounds **1** (0.011 g) and **2** (0.021 g) were obtained.

Thalicoside G_1 (**1**), $C_{42}H_{70}O_{15}$, m.p. 296—298 °C (pyridine), $[\alpha]^{24}_{546}$ +11.1° (*c* 0.18, pyridine). IR (KBr), ν/cm^{-1}: 3408 (O—H), 3050 (CPR), 1069 (C—O). MS (FAB), *m/z* (I_{rel} (%)): 837 $[M+Na]^+$ (5). The NMR spectra are given in Tables 1—3.

Thalicoside G_2 (**2**), $C_{42}H_{70}O_{15}$, m.p. 292—294 °C (pyridine), $[\alpha]^{24}_{546}$ +12.3° (*c* 0.41, pyridine). IR (KBr), ν/cm^{-1}: 3395 (O—H), 3050 (CPR), 1070 (C—O). MS (FAB), *m/z* (I_{rel} (%)): 837 $[M+Na]^+$ (68). The NMR spectra are given in Tables 1—3.

The authors are grateful to S. V. Zinchenko (Irkutsk State University) for recording the NMR spectra and to G. F. Fedorova (Irkutsk Limnological Institute, Siberian Branch of the RAS) for performing HPLC analysis.

The authors also wish to thank V. A. Stonik for assistance in this study.

This work was supported by the Russian Foundation for Basic Research (Project No. 97-03-96100).

References

1. A. S. Gromova, V. I. Lutsky, A. A. Semenov, D. Li, and N. L. Owen, *Phytochemistry*, 1997, **47**, 437.
2. A. S. Gromova, V. I. Lutskii, A. A. Semenov, V. A. Denisenko, and V. V. Isakov, *Khim. Prir. Soedin.*, 1984, 213 [*Chem. Nat. Compd.*, 1984 (Engl. Transl.)].
3. A. S. Gromova, V. I. Lutskii, S. V. Zinchenko, T. V. Ganenko, and A. A. Semenov, *Khim. Prir. Soedin.*, 1993, 568 [*Chem. Nat. Compd.*, 1993 (Engl. Transl.)].
4. S. Seo, J. Tomita, K. Tori, and Y. Yashimura, *J. Am. Chem. Soc.*, 1978, **100**, 3331.
5. G. T. Emmons, W. K. Wilson, and G. J. Schroepfer, *Magn. Reson. Chem.*, 1989, **27**, 1012.
6. A. A. Kicha, A. I. Kalinovskii, E. V. Levina, V. A. Stonik, and G. B. Elyakov, *Bioorgan. Khim.*, 1983, **9**, 975 [*Sov. J. Bioorg. Chem.*, 1983, 2 (Engl. Transl.)].
7. R. Riccio, O. S. Greco, L. Minale, D. Laurent, and D. Duhet, *J. Chem. Soc., Perkin Trans. 1*, 1986, 665.
8. A. S. Gromova, V. I. Lutskii, S. V. Zinchenko, N. N. Trofimova, A. A. Semenov, and N. A. Nakhova, *Khim. Prir. Soedin.*, 1993, 103 [*Chem. Nat. Compd.*, 1993 (Engl. Transl.)].
9. S. Yahara, K. Kaji, and O. Tanaka, *Chem. Pharm. Bull.*, 1979, **27**, 88.

Received November 17, 1997;
in revised form January 12, 1998