New triterpene glycosides from *Thalictrum squarrosum* St. ex Willd. The structure of squarrosides B3 and B4

E. A. Khamidullina, A. S. Gromova, V. I. Lutsky, S. B. Zinchenko, and A. A. Semenov*

Irkutsk Institute of Organic Chemistry, Siberian Branch of the Russian Academy of Sciences, 1 ul. Favorskogo, 664033 Irkutsk, Russian Federation. Fax: +7 (395 2) 46 6434

From the above-ground part of *Thalictrum squarrosum* St. ex Willd. (Ranunculaceae), an equilibrium mixture of two triterpene glycosides, squarroside B3 (**1**), 3-*O*-[*O*-(α-L-rhamnopyranosyl)-(1→6)-β-D-glucopyranosyl]-21(*S*),22(ξ),23(*R*)-3β,21,22,30-tetrahydroxy-21,23-epoxycycloarth-24-ene, and squarroside B4 (**2**), its 21(*R*) epimer, was isolated. Their structures were established on the basis of IR and 2D NMR spectra and FAB mass spectra.

Key words: *Thalictrum*, triterpene glycosides, cycloarthane; IR, 2D NMR spectra, FAB mass spectra.

Thalictrum squarrosum (meadow-rue), a Siberian grassy plant, is a source of new, interesting triterpenoids. Previously,[1,2] the isolation of six new triterpene compounds from this plant has been reported. In the present work, the structures of two new saponins isolated as an equilibrium mixture and called squarrosides B3 and B4 (**1** and **2**) are considered.

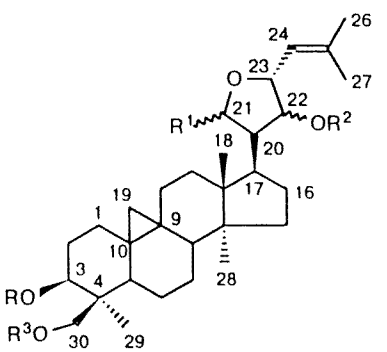
Results and Discussion

A mixture of triterpene glycosides was isolated from a methanolic extract of the above-ground part of meadow-rue, from which squarrosides B3 and B4 were obtained by multiple column chromatography on silica gel and liquid-liquid distribution. Acid hydrolysis of glycosides **1** and **2** results in resinification of the reaction mixture. Using enzymatic cleavage, we obtained 15 mg of one glycoside, progenin (**3**), and 3 mg of a mixture of genins. Acetylation of these glycosides gave a multicomponent mixture, from which only squarroside B3 peracetate (**4**) could be isolated in an individual state. The structure of squarrosides B3 and B4 was established on the basis of NMR and mass spectra of an equilibrium mixture of these compounds, and the spectra of the derivatives prepared.

A thorough study of the ^1H and ^{13}C NMR spectra (Tables 1 and 2) made it possible to conclude that compounds **1** and **2** contain new cycloarthane genins that differ in the stereochemistry of the side chain and are glycosylated with two carbohydrate residues.

The diglycoside character of squarrosides B3 and B4 is indicated by their FAB mass spectrum containing the quasimolecular ion at *m/z* 819 $[M+Na]^+$ (100 %) and the fragment ions at *m/z* 673 $[M+Na-dHex]^+$ and 493 $[M+Na-dHex-Hex-H_2O]^+$, which correspond to the elimination of deoxyhexose and the block of deoxyhexose and hexose.

A comparison of the ^{13}C NMR chemical shifts of the carbohydrate fragments in the spectrum of glycosides **1** and **2** with the literature data[3] allows one to conclude that the carbohydrate chains are identical in both squarrosides and include β-D-glucopyranose and α-L-rhamnopyranose. The downfield shift of the C(6′) signal of glucopyranose (δ 68.34) in **1** and **2** *vs.* that in the spectrum of progenin **3** (see Table 1) is caused by



1: R = α-L-Rha$_p$(1→6)-β-D-Glc$_p$, R^1 = β-OH, R^2 = R^3 = H;
2: R = α-L-Rha$_p$(1→6)-β-D-Glc$_p$, R^1 = α-OH, R^2 = R^3 = H;
3: R = β-D-Glc$_p$, R^1 = β-OH, R^2 = R^3 = H;
4: R = 2,3,4-Ac$_3$-α-L-Rha$_p$(1→6)-2,3,4-Ac$_3$-β-D-Glc$_p$, R^1 = β-AcO, R^2 = R^3 = Ac;
5: R = R^2 = R^3 = H, R^1 = β-OH;
6: R = R^2 = R^3 = H, R^1 = α-OH;
7: R = α-L-Rha$_p$(1→6)-β-D-Glc$_p$, R^1 = β-OMe, R^2 = R^3 = H;
8: R = α-L-Rha$_p$(1→6)-β-D-Glc$_p$, R^1 = α-OMe, R^2 = R^3 = H;
9: R = α-L-Rha$_p$(1→6)-β-D-Glc$_p$, R^1 = β-OEt, R^2 = R^3 = H;
10: R = α-L-Rha$_p$(1→6)-β-D-Glc$_p$, R^1 = α-OEt; R^2 = R^3 = H

Translated from *Izvestiya Akademii Nauk. Seriya Khimicheskaya*, No. 6, pp. 1547—1551, June, 1996.

1066-5285/96/4506-1476 $15.00

Table 1. Chemical shifts (δ) in ^{13}C NMR spectra of compounds **1**—**4**, **9**, and **10**

Atom	δ (±0.01)			
	1, 2	**3**	**4**	**9, 10**
C(1)	32.18	32.32	31.41	32.08
C(2)	31.96	32.10	30.81	32.08
C(3)	89.82	89.63	89.12	89.40
C(4)	45.51	45.59	44.18	45.03
C(5)	48.08	48.08	47.68	47.80
C(6)	22.13	22.19	21.09	22.06
C(7)	26.74	26.90	26.47	26.45
C(8)	48.84	48.79	48.15	48.55
C(9)	21.62	21.59	20.89	21.19
C(10)	27.22	26.07	25.61	26.77
C(11)	26.90	26.90	25.93	25.85
C(12)	36.11, 36.32	36.17	35.35	36.04
C(13)	45.75	45.81	45.08	45.24
C(14)	48.92	48.95	48.37	48.93
C(15)	30.39	30.15	28.84	30.29
C(16)	30.18	30.15	27.16	29.86
C(17)	41.28, 45.10	41.37	43.28	40.70, 44.76
C(18)	25.92	26.07	25.47	25.85
C(19)	30.99	30.96	29.51	31.27
C(20)	52.55, 56.72	56.72	51.49	52.35, 54.84
C(21)	98.84, 101.63	101.77	99.57	103.54, 107.38
C(22)	77.04, 77.50	77.22	77.35	75.10, 76.56
C(23)	78.63, 80.53	78.70	78.25	76.94, 80.51
C(24)	—*	—*	138.74	—*
C(25)	—*	—*	118.19	—*
C(26)	21.26	21.43	19.69	21.19
C(27)	19.91	20.02	19.25	19.83
C(28)	18.77	18.77	18.56	18.64
C(29)	18.45	18.77	18.51	18.38
C(30)	63.47	63.51	64.37	63.29
C(1′)	106.15	106.47	102.36	106.19
C(2′)	75.57	75.82	71.83	75.42
C(3′)	79.50	78.91	73.01	78.67
C(4′)	72.08	72.20	69.01	71.85
C(5′)	78.77	78.53	73.03	78.67
C(6′)	68.34	63.25	67.24	67.89
C(1″)	102.96		98.18	102.29
C(2″)	72.35		69.56	72.34
C(3″)	72.86		69.51	72.77
C(4″)	74.16		70.98	73.96
C(5″)	69.64		66.61	69.86
CH_3″	18.61		17.33	18.38
MeCOO			168.87, 169.22, 169.50, 169.53, 169.54, 169.56, 169.91, 169.95, 170.47	
MeCOO			20.36, 20.38, 20.41, 20.42, 20.44 (2 MeCOO), 20.55, 20.64	

* These signals are overlapped by the signals of pyridine.

the α-effect of glycosylation at O(6′). In fact, in the ROESY experiment with acetate **4**, the H(1″) signal of rhamnopyranose gives cross-peaks with the proton signals of the CH_2O group of glucopyranose, and the signal of the anomeric proton of glucose gives a cross-peak with the signal of the H(3) proton of the genin at δ 3.18. In addition, the C(3) signal in the ^{13}C NMR spectrum of compounds **1** and **2** occurs at δ 89.82 due to attachment of the carbohydrate chain at this atom in the genin.

The genins of squarrosides B3 and B4, called squarrogenins 3 and 4 (**5** and **6**), were assigned to cycloarthane triterpenoids due to the presence of signals characteristic of a cyclopropane ring (2H(19), δ 0.21 and 0.49, ^2J = 3.9 Hz; C(9), δ 21.62, and C(10), δ 27.22) in the NMR spectra of **1** and **2**. Each genin contains one trisubstituted double bond (C(24), δ 138.74, C(25), δ 118.19 in **4**) and five methyl groups, while three of them are located at sp^3 hybridized carbon atoms and two of them are attached to the double bond. Hence the signals of these protons are shifted downfield (**1**, δ 1.69 and 1.65, **2**, δ 1.66 and 1.64) and are broadened due to allylic splitting by H(24) (^4J = 1.5 Hz).

As can be seen from the ^1H and ^{13}C NMR spectra of compounds **1** and **2**, after subtraction of the signals of the atoms of carbohydrate residues, squarrogenins 3 and 4 contain five carbon atoms each, bonded with oxygen. The side chain contains one epoxy fragment and two hydroxyl groups, and the polycyclic fragment contains two hydroxyl groups.

A comparative study of the ^{13}C NMR spectra of squarrosides B3 and B4 and squarrosides B1 and B2 (**7** and **8**) isolated previously[2] demonstrated the identity of the (C(1)—C(15)) polycyclic fragments of these compounds. Therefore, the secondary hydroxyl group is in the 3β-position, and hydroxymethyl is in the 4β-position.

The spin systems of the side chains of the glycosides were identified by two-dimensional TOCSY and COSY procedures.

For squarroside B3, the following proton-proton interactions were revealed (δ):

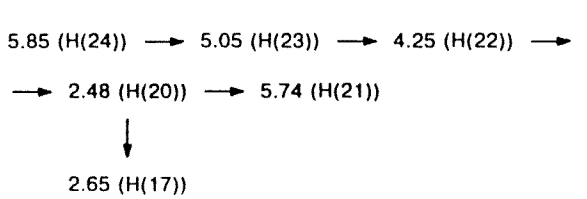


Table 2. ^{1}H NMR spectral data for compounds **1**, **2** and **4**

Atom	δ (*J*/Hz)		
	1	**2**	**4**
2 H(2)	2.1, 2.5 (m)*		1.6, 1.82 (m)
H(3)	3.70 (m)*		3.18 (dd, *J* = 4.5, 12.0)
H(16)	1.55 (m)*		1.22 (m)
H(17)	2.65 (q, *J* = 11.2)	2.82 (q, *J* = 11.2)	2.1 (m)
2 H(19)	0.21, 0.49 (d, *J* = 3.9)*		0.42, 0.33 (d, *J* = 4.2)
H(20)	2.48 (dt, *J* = 4.3, 5.5, 11.2)	2.08 (m)	2.53 (dt, *J* = 4.9, 5.2, 11.7)
H(21)	5.74 (d, *J* = 5.5)	5.60 (d, *J* = 3.9)	6.02 (d, *J* = 4.9)
H(22)	4.25 (dd, *J* = 2.1, 4.3)	4.18 (t, *J* = 3.3)	5.37 (dd, *J* = 5.2, 2.6)
H(23)	5.05 (dd, *J* = 2.1, 7.9)	4.76 (dd, *J* = 3.3, 8.5)	4.73 (dd, *J* = 7.8, 2.6)
H(24)	5.85 (m)*		4.96 (m)
2 H(30)	4.41, 3.70 (d, *J* = 11.2)*		4.12, 4.20 (d, *J* = 11.3)
Me	0.93, 0.95 (both s, CH_3-28); 1.07, 1.36 (both s, CH_3-18); 1.47, 1.48 (both s, CH_3-29); 1.64, 1.65 (both d, CH_3-27, *J* = 1.5); 1.66, 1.69 (both d, CH_3-26, *J* = 1.5)		0.82 (s), 0.93 (s), 0.94 (s), 1.59 (br.s), 1.61 (br.s)
H(1')	4.84 (d, *J* = 7.6)	4.85 (d, *J* = 7.9)	4.48 (d, *J* = 7.8)
H(2')	3.81 (m)*		4.86 (dd, *J* = 7.8, 9.6)
H(3')	4.0 (m)*		5.08 (t, *J* = 9.6)
H(4')	4.47 (m)*		4.82 (t, *J* = 9.6)
H(5')	Not determined		3.62 (m)
H(6'a)	Not determined		3.52 (dd, *J* = 11.7, 7.3)
H(6'b)	Not determined		3.62 (m)
H(1'')	5.35 (br.s)*		4.67 (d, *J* = 1.4)
H(2'')	4.39 (dd, *J* = 1.8, 3.6)*		5.12 (m)
H(3'')	4.32 (t, *J* = 3.6)	4.30 (t, *J* = 3.3)	5.14 (m)
H(4'')	4.02 (m)*		4.96 (m)
H(5'')	4.21 (dq, *J* = 9.4, 6.5)*		3.76 (dq, *J* = 6.3, 9.9)
Me''	1.54 (d, *J* = 6.5)	1.53 (d, *J* = 6.0)	1.1 (d, *J* = 6.3)
MeCOO			1.87 (s), 1.90 (s), 1.91 (s), 1.94 (s), 1.95 (s, 6 H), 1.97 (s), 1.99 (s), 2.30 (s)

Note: δ ±0.01, *J* ±0.2; **1**, **2** — C_5D_5N, 80 °C, **4** — $CDCl_3$, 50 °C. * These signals coincide for **1** and **2**.

For squarroside B4 (δ):

5.85 (H(24)) → 4.76 (H(23)) → 4.18 (H(22)) →

→ 2.08 (H(20)) → 5.60 (H(21))

↓

2.82 (H(17))

Thus, the bond sequences in the side chain of both squarrosides are the same. The configurations of the substituents of the side chain were determined using Dreiding models and two-dimensional NOE (ROESY). In Fig. 1, the 2D ROESY NMR spectrum of compounds **1** and **2** is presented, and the proton-proton interactions in the tetrahydrofuran ring are given.

It is known that in lanostane type triterpenoids, H(20) has the β-configuration.[4] Therefore, in connection with this and the presence of cross-peaks of the H(23) and H(20) signals in the spectra of both compounds, the β-configuration of H(23) was established. The consideration of the Dreiding models made it evident that H(21) and H(24) are spatially proximate only for β-H(23) and α-H(21). The presence of the cross-peaks of the H(21) and H(24) signals of **1** in the ROESY spectrum determines the configuration of the substituents at C(21) and C(23) as 21β-OH and 23α-CH=CMe$_2$ in squarroside B3 and 21α-OH in squarroside B4.

The relative configuration of H(22) in **1** and **2** was not determined. The spatial interactions of the H(22) proton in the ROESY spectrum for both isomers are the same, which makes it impossible to determine unequivocally the configuration of this chiral center. The use of ^{13}C NMR spectra to determine the configurations of the substituents in the tetrahydrofuran ring was described in Ref. 5, however, for squarrosides it appeared to be inappropriate, because their chemical shifts did not obey any existing model.

Previously,[6] compounds similar to squarrosides B1 and B2, in which the configuration of H(22) was determined as β, were isolated from *Thalictrum herba*. The absence of experimental data on NOE in that paper made it impossible for us to compare the results obtained.

Thus, for squarroside B3 the structure of 3-*O*-[*O*-(α-L-rhamnopyranosyl)-(1→6)-β-D-glucopyranosyl]-21(*S*),22(ξ),23(*R*)-3β,21,22,30-tetrahydroxy-21,23-epoxycycloarth-24-ene was suggested, and the

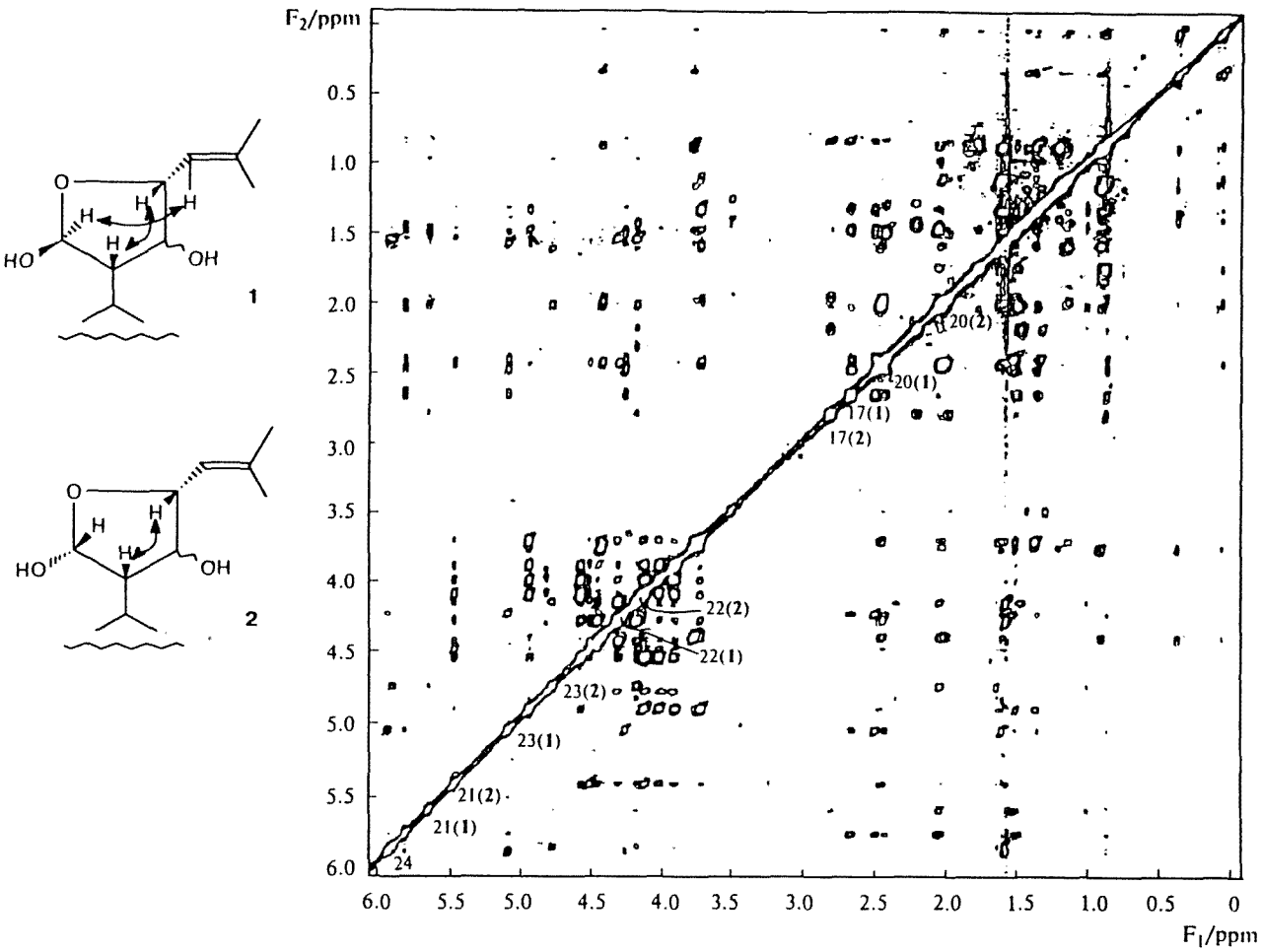


Fig. 1. ROESY spectrum (500 MHz) of squarrosides B3 and B4 (**1** and **2**).

structure of its 21(*R*)-epimer was assigned to squarroside B4.

Squarrosides B1 and B2 isolated previously can be considered as *O*-methyl derivatives of squarrosides B3 and B4. Assuming that the extraction was carried out with methanol and that methanol was widely used in the further individualization of the saponins, it was necessary to prove the native origin of all isolated compounds. Hence, the isolation of squarrosides was repeated, but under conditions avoiding methylation: methanol was replaced by ethanol.

The FAB mass spectral analysis of the isolated total of squarrosides demonstrated the presence of compounds containing the 21-OH group (squarrosides B3 and B4) and the 21-OMe group (squarrosides B1 and B2). The compounds containing the ethoxyl group at C(21) (compounds **9** and **10**, see Table 1) were the major components. Thus, it is evident that squarrosides B1, B2, B3, B4 are native, and in this plant hydroxyl-containing (at C(21)) glycosides are major components and their methylated derivatives are minor components.

Experimental

The general procedures have been published previously.[1,2] The ^{1}H NMR spectra were recorded with a Varian VXR-500 S instrument (for ^{1}H 499.843 MHz, and for ^{13}C 125.699 MHz), equipped with a SUN 3/50 computer and VNMR conventional software. The spectra of compounds **1** and **2** were recorded in C_5D_5N (22 mg/0.6 mL), and the spectra of **5** and **6** were recorded in DMSO-d_6 (3 mg/0.6 mL). The ^{1}H and ^{13}C NMR spectra of compounds **4** were recorded in $CDCl_3$ (30 mg/0.6 mL) in a 5-mm tube using Me_4Si as the internal standard. The ^{13}C NMR spectra were obtained in a 10-mm tube using Me_4Si as the internal standard, in C_5D_5N with a Jeol FX-90Q spectrometer (22.49 MHz): for compounds **1** and **2** the concentration was 22 mg/1.8 mL, for **3** 15 mg/1.8 mL, for **9** and **10**, 25 mg/1.8 mL.

To obtain the two-dimensional ^{1}H—^{1}H spectra, the following procedures were used: COSY, the conventional RELAYH program, the dimension of a matrix 1 K×0.5 K, spectral width 3000 Hz. In the pulse sequence two 90° pulses (COSY-90), with a delay of 1.8 s were used. Before the Fourier transform, the FID was multiplied by the bell-shaped function with zero shift. For TOCSY, the conventional TOCSY program in a

phase-sensitive version was used. The matrix dimension was 2 K×1 K; spectral width 3000 Hz. The values for 90° and trim-pulse were 22 μs; mixing time 0.05 ms; the delay of relaxation 1.8 s. To equilibrate the spin system, the pulse sequence sspul : 90°—HS—90° was used. Before the Fourier transform, a Gaussian function with zero shift in both dimensions was used for weighting. For ROESY, the conventional ROESY program in a phase-sensitive version was used. The dimension of a matrix was 2 K×1 K; spectral width 3557 Hz; mixing time 1 s. To equilibrate the spin system, the pulse sequence sspul : 90°—HS—90°. A Gaussian function with zero shift in both dimensions was used for weighting.

The mass spectra were measured with a LKB-2091/PDP-11/34 instrument equipped with an Iontech Ltd. Feddington FAB ion source. Ionization was carried out with a beam of accelerated Xe atoms with 6 keV energy and a 1.2 mA discharge current. Glycerol with a NaCl or KCl additive was used as the matrix.

Isolation of squarrosides B3 and B4 was carried out as described for squarrosides B1 and B2.[2]

Squarrosides B3 and B4 (1 and 2), $C_{42}H_{68}O_{14}$, m.p. 187—189 °C (methanol), IR (KBr), ν_{max} 3350—3500 cm^{-1} (OH). FAB MS, *m/z*: 835 $[M+K]^+$, 819 $[M+Na]^+$, 804 $[M+Na-CH_3]^+$, 801 $[M+Na-H_2O]^+$, 673 $[M+Na-146]^+$, 515 $[M+2Na-H-308-H_2O]^+$, 493 $[M+Na-308-H_2O]^+$. The ^{13}C NMR spectrum is given in Table 1.

Acetylation of squarrosides B3 and B4. A total of 600 mg of squarrosides B3 and B4 was dissolved in pyridine (4 mL) and acetic anhydride (2 mL) was added. The reaction mixture was kept at 20 °C for 24 h. After the usual workup, the reaction products were chromatographed on a column with silica gel to afford 200 mg of 2',3',4',2",3",4",21,22,30-*nona*-*O*-acetylsquarroside B3 (**4**), m.p. 148—150 °C, $[\alpha]_{546}^{20}$ +7.75° (*c* 2.71, acetone). The ^1H and ^{13}C NMR spectra are given in Tables 1 and 2.

Enzymolysis of squarrosides B3 and B4. A total of 100 mg of squarrosides B3 and B4 was suspended in water (20 mL), and an enzymatic preparation from grape snail *Helix pomatia* (10 mg) and several drops of toluene (10 mg) were added. The reaction mixture was kept in a thermostat at 36 °C with continuous stirring for 14 days and periodic addition of the enzyme (5 mg portions). The reaction was monitored by TLC. The reaction products were extracted with chloroform. The chloroformic extracts were washed with water and concentrated. A dry extract (30 mg) was obtained. The aqueous layer was extracted with butanol. The butanolic extract was concentrated and the residue was digested as described above. After repeating enzymolysis, the weight of the chloroformic extract was 30 mg.

The reaction products were chromatographed on a column with silica gel, and 3 mg of a mixture of genins **5** and **6** was isolated.

Squarrogenins 3 and 4 (5 and 6). ^1H NMR (DMSO-d_6), δ: 3.23* (dd, H-3); 0.45, 0.31* (d, 2 H(19), ^2J = 4 Hz); 1.94* (m, H(20)); 5.02 (d, H(21)); 5.08 (d, H(21)); 3.75 (br.s, H(22)); 3.81 (br.s, H(22)); 4.52 (br.d, H(23)); 4.43 (dd, H(23)); 5.28 (br.d, H(24)); 5.38 (br.d, H(24)); 3.95, 3.40* (d, 2 H(30), ^2J = 11 Hz); 1.70* (d, 3 H, CH_3(26), ^4J = 1.0 Hz); 1.64* (d, 3 H, CH_3(27), ^4J = 1.0 Hz); 1.04* (s, 3 H, CH_3); 1.0* (s, 3 H, CH_3); 0.85* (s, 3 H, CH_3).

The subsequent elution afforded 15 mg of squarroside progenin (**3**). $C_{36}H_{58}O_{10}$, FAB MS, *m/z*: 698 $[M+K]^+$, 673 $[M+Na]^+$, 695 $[M+2Na-H]^+$, 493 $[M+Na-162-H_2O]^+$, 475 $[M+Na-162-2H_2O]^+$. The ^{13}C NMR spectrum is given in Table 1.

Isolation of squarrosides under conditions avoiding methylation. Air-dried plant material (1.2 kg) was extracted with ethanol five times. The extract was concentrated using circulation and rotor evaporators. The dry extract (72 g) was dissolved in a minimum amount of ethanol and a double excess of acetone was added. The precipitate was filtered off. The solution was concentrated and a second precipitation was carried out. The solution (after the second precipitation) containing squarrosides (TLC) was concentrated and the residue (44.8 g) was chromatographed on a column with silica gel using a chloroform—ethanol (0 to 20 % of ethanol) system. A fraction containing squarrosides was obtained when the column was eluted with a system containing 14 % ethanol. Repeat chromatography of the squarrosides under the same conditions afforded compounds **1** and **2** (21-OH); FAB MS, *m/z* 835 $[M+K]^+$, **7** and **8** (21-OCH_3); FAB MS, *m/z* 833 $[M+Na]^+$, **9** and **10** (21-OCH_2CH_3); FAB MS, *m/z* 847 $[M+Na]^+$.

The ^{13}C NMR spectrum of compounds **9** and **10** is given in Table 1.

The authors thank A. L. Vereshchagin for recording the FAB mass spectra.

* These signals coincide for **5** and **6**.

References

1. V. I. Lutsky, E. A. Khamidullina, A. S. Gromova, and A. A. Semenov, *Khim. Prir. Soedin.*, 1989, 510 [*Chem. Nat. Compd.*, 1989 (Engl. Transl.)].
2. E. A. Khamidullina, A. S. Gromova, V. I. Lutsky, A. L. Vereshchagin, A. A. Semenov, and M. F. Larin, *Khim. Prir. Soedin.*, 1989, 516 [*Chem. Nat. Compd.*, 1989 (Engl. Transl.)].
3. S. Seo, J. Tomita, K. Tori, and Y. Yashimura, *J. Am. Chem. Soc.*, 1978, **100**, 3331.
4. P. Pant and R. P. Rastogi, *Phytochemistry*, 1979, **18**, 1095.
5. H. Achenbach and D. Frey, *Phytochemistry*, 1992, **31**, 4263.
6. H. Yashimitsu, K. Hagoshi, M. Kumabe, and T. Nohara, *Chem. Pharm. Bull.*, 1993, **47**, 786.

Received August 2, 1995;
in revised form March 22, 1996