Chemistry of Natural Compounds, Vol. 30, No. 6, 1994

TRITERPENE AND STEROID GLYCOSIDES OF THE *Melilotus* GENUS AND THEIR GENINS

I. MELILOTOSIDES A, B, AND C FROM THE ROOTS OF *Melilotus albus*

G. V. Khodakov,[a] Yu. A. Akimov,[a] A. S. Shashkov, P. K. Kintya,[b] and V. I. Grishkovets

UDC 547.918:582.738:543.422

Three new triterpene glycosides of the the oleanane series — melilotosides A, B, and C — and the nonglycosylated soyasapogenol B have been isolated from the roots of the plant Melilotus albus Medik. (Leguminosae). The structures of the glycosides have been shown on the basis of chemical transformations and spectral results. Melilotoside A has the structure of soyasapogenol B 3-O-α-L-arabinopyranoside, melilotoside B that of soyasapogenol B 3-O-[O-β-D-galactopyranosyl-(1→2)-α-L-arabinopyranoside], and melilotoside C that of soyasapogenol B 3-O-[O-α-L-rhamnopyranosyl-(1→2)-O-β-D-galactopyranosyl-(1→2)-α-L-arabinopyranoside.

Of the eleven species included in the genus *Melilotus*, four grow in the Crimea. They include two with a wide area of distribution — yellow sweet clover *M. officinalis* (L) Pall. and white sweet clover *M. albus* Medik., one with an area of distribution limited to the Mediterranean region — Neapolitan sweet clover, *M. neapolitanus* Ter., and one endemic species — Crimean sweet clover, *M. tauricus* (Bieb.) Ser. [1, 2]. The most studied in the chemical respect are yellow and white sweet clovers. The epigeal parts of plants of these species have been found to contain coumarins (coumarin, dihydrocoumarin, and dicumarol) [3], acids (*o*-hydrocinnamic, melilotic, coumaric, and *o*-coumaric) [4], coumarin glycosides [5], fatty acids, among which linoleic predominates [6], flavonoids (quercetin and coumestrol) [7], and a polysaccharide — xylan. Screening studies of Crimean representatives of the Leguminosae family have shown the presence of saponins in various species of sweet clover. However, information on them is limited to estimates of their total content in comparison with other species of leguminous plants.

The present paper is devoted to a proof of the structures of glycosides not previously described in the literature from the roots of white sweet clover growing in the Crimea (village of Rybach'e). In a methanolic extract we detected at least eight substances assigned to the triterpene glycosides. We have studied the structures of three of them, which have ben named melilotosides A (**1**), B (**2**), and C (**3**).

Analysis of the product of the complete acid hydrolysis of each of the compopunds studied by PC and TLC (they were characterized by comparing their chromatographic mobilities with those of authentic markers) showed the presence in them of the aglycon soyasapgenol B (olean-12-ene-3β,22β,24-triol). The aglycon was detected in the native form in a chloroform fraction from an extract of the roots of the plant. In the case of melilotoside A the carbohydrate moiety was represented by arabinose, and the ^{13}C NMR spectrum corresponded to one residue of the monosaccharide (one signal in the region of anomeric carbon atoms) (Table 1) [10], while a comparison of the chemical shifts of the carbon atoms of the glycoside with literature figures for the unsubstituted aglycon [11, 12] showed the attachment of the sugar at the C-3 atom of soyasapogenol B (downfield shift from 80.2 to 89.2 ppm).

State Nikitskii Botanical Gardem 334267, Yalta, Botanicheskoe Production Combine; and Institute of Genetics, Academy of Sciences of the Republic of Moldava, Kishinev, ul. Lesnaya, 20. Translated from Khimiya Prirodnykh Soedinenii, No. 6, pp. 756-762, November-December, 1994. Original article submitted March 14, 1994.

0009-3130/94/3006-0704$12.50

TABLE 1. Chemical Shifts of the Carbon Atoms (ppm, 0 — TMS, C_5D_5N) of Melilotosides A (1) and C (3)

Nuclei of the aglycon	1	3	Nuclei of the sugar residues	1	3
1	38.7	38.6	I. Ara_p		
2	26.8	26.6	1'	106.4	105.4
3	89.2	91.2	2'	73.6	77.6
4	44.6	44.4	3'	75.8	76.7
5	56.3	56.1	4'	75.5	74.1
6	19.0	19.0	5'	63.6	63.7
7	33.5	33.4			
8	40.2	40.0	$D\text{-}Gal_p$		
9	48.0	47.8	1''		101.8
10	36.8	36.6	2''		78.5
11	24.3	24.1	3''		75.7
12	122.7	122.5	4''		71.2
13	145.1	144.9	5''		76.5
14	42.6	42.4	6''		61.7
15	27.3	26.6			
16	28.7	28.8	II. Rha_p		
17	38.2	38.1	1''		102.3
18	45.5	45.4	2''		72.5
19	47.0	46.8	3''		72.7
20	31.3	31.0	4''		74.5
21	42.6	42.4	5''		69.5
22	75.8	75.7	6''		18.6
23	23.5	23.1			
24	63.6	63.7			
25	15.7	15.9			
26	17.3	17.1			
27	25.9	25.8			
28	28.7	28.8			
29	33.5	33.4			
30	21.4	21.3			

In the case of melilotoside B, the carbohydrate moiety was represented by arabinose and galactose. The ^1H NMR spectrum with double homonuclear resonance experiments (in the ordinary and the difference variants) confirmed the arabino- and galacto- configurations of the sugar residues in the pyranose forms. The α-configuration of the arabinose and the β-configuration of the galactose followed from the spin-spin coupling constants (SSCCs) of the anomeric protons of the monosaccharides (7.5 Hz in both residues) (Table 2). The application of the nuclear Overhauser effect procedure in the difference variant permitted the observation of changes in the intensities of the signals of the protons at C-2 of the α-*L*-arabinose residue and at C-3 of the aglycon when the proton of the anomeric center of the arabinose was irradiated and also of the protons at the C-2 atoms of the α-*L*-arabinose and of the β-*D*-galactopyranose when the proton of the anomeric center of the β-*D*-galactopyranose was irradiated, which unambiguously showed glycosylation at the C-3 atom of the soyasapogenol B by the α-*L*-arabinopyranose residue and the attachment of the β-*D*-galactopyranose residue at the second carbon atom of the α-*L*-arabinopyranose residue.

Thus, melilotoside A has the structure of soyasapogenol B 3-O-α-*L*-arabinopyranoside, and melilotoside B that of soyasapogenol 3-O-[O-β-*D*-galactopyranosyl-(1→2)-α-*L*-arabinopyranoside].

The presence of galactose, rhamnose, and arabinose in melilotoside C was established. Its ^{13}C NMR spectrum corresponded to a composition with a 1:1:1 ratio of the sugar residues (three signals in the region of anomeric carbon atoms) (Table 1) [10].

To determine the sequence of the carbohydrate residues in the molecule of compound (**3**), we carried out its partial acid hydrolysis, with chromatographic separation of the two progenins formed (**3b** and **3c**), which were subjected to TLC analysis, to a determination of physical constants, and to complete acid hydrolysis. As a result, progenin (**3b**) was

TABLE 2. Chemical Shifts (ppm, 0 — TMS, C_5D_5N) and Spin-Spin Coupling Constants (J, Hz) of the Protons of Melilotosides B (**2**) and C (**3**)

Protons of the aglycon	2	3
7 CH_3	0.68; 0.89; 0.94; 1.14; 1.19; 1.22; 1.28	0.73; 0.97; 0.98; 1.18; 1.24; 1.27; 1.42
11-12	5.31 br.s $J_{11,12}$=3.0	5.31 br.s $J_{11,12}$=3.5
11-22	3.71 dd $J_{21,22}$=3.2 $J_{21',22}$=6.0	3.69 dd $J_{21,22}$=3.0 $J_{21',22}$=6.7
11-24 A	4.27 d $J_{24A,24B}$=11	
11-24 B	3.37 d	
	L-Arabinose	
1'	4.82 d $J_{1,2}$=7.5	4.90 d $J_{1,2}$=7.0
2'	4.18 dd $J_{2,3}$=7.5	4.41 dd $J_{2,3}$=8.0
3'	4.39 dd $J_{3,4}$=3.0	4.45 dd $J_{3,4}$=2.0
4'		4.31 $J_{4,5a}$=8.0
5e'		4.25
5a'		3.28 m $J_{5a,5e}$=11.4
	D-Galactose	
1''	5.39 d $J_{1,2}$=7.5	5.60 d $J_{1,2}$=7.5
2''	4.41 dd $J_{2,3}$=7.5	4.43 dd $J_{2,3}$=9.0
3''	4.03 dd $J_{3,4}$=3.0	3.99 dd $J_{3,4}$=2.5
4''	4.37 d $J_{4,5}$=3.5	4.29 d $J_{4,5}$=4.0
5''	3.94 t	3.83 m
6A''	4.29 dd $J_{5,6A}$=5.5 $J_{6A,6B}$=12.5	4.28 d
6B''	4.41 dd $J_{5,6B}$=5.3	
	L-Rhamnose	
1'''		6.13 s
2'''		4.65 br.s
3'''		4.58 dd $J_{3,4}$=9.0
4'''		4.20 d $J_{4,5}$=9.0
5'''		4.88 dq $J_{5,6}$=6.5
6'''		1.7 d

identified as melilotoside A, and (**3c**) as melilotoside B. Thus, the carbohydrate component of melilotoside C has a structure with the sequence arabinose, galactose,and rhamnose.

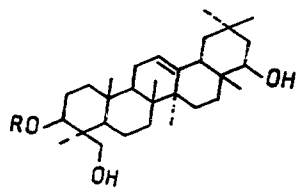
The nature of the splitting of the signals and the SSCCs of all the skeletal protons of the sugar residues of glycoside (**3**) corresponded fully to the arabino-, galacto-, and rhamno- configurations of the monosaccharides in pyranose forms, and the values of the SSCCs showed the α-configuration of the glycosidic center of the arabinose residue (7.5 Hz) and the β-configuration of the anomeric center of the galactose residue (7.5 Hz) (Table 2), while the complete coincidence of the chemical shifts in the ^{13}C NMR spectrum of the rhamnose residue with literature figures for a terminal α-*L*-rhamnopyranose corresponded to the α-configuration of its glycosidic center (Table 1) [10].

A comparison of the chemical shifts of the carbon atoms of the glycoside with literature figures [13, 14] for the unsubstituted fragments it contained showed the attachment of the carbohydrate component at the C-3 atom of the aglycone (downfield shift from 0.2 to 91.2 ppm) and substitution at the C-2 atoms of the α-*L*-arabinopyranose and β-*D*-galactopyranose (downfield shifts from 73.6 to 77.6 ppm and from 72.7 to 78.5 ppm, respectively) (Table 1). This was also confirmed by the PMR spectrum of the acetate of the glycoside, in which it was possible by the use of the double homonuclear resonance method

to assign the signals of the skeletal protons of the monosaccharides and to observe the relatively high position of the signals of the H-2 proton of the arabinose residue and H-2 of the galactose residue, confirming the presence of 1→2-bonds.

A comparative analysis of the glycosylation effects in the ^{13}C NMR spectrum of glycoside C with literature figures was confirmed by the observation of a NOE in the rotating coordinate system (two-dimensional ROESY experiment [15]). The cross-peaks observed recorded spatial propinquity for the H-2 atom of the α-*L*-arabinose and the H-3 atom of the aglycon, for H-1 of the β-*D*-galactopyranose and H-2 of the α-*L*-arabinopyranose, and for H-1 of the α-*L*-rhamnopyranose and H-2 of the β-*D*-galactopyranose.

Thus, melilotoside C is soyasapogenol B 3-O-[O-α-*L*-rhamnopyranosyl-(1→2)-O-β-*D*-galactopyranosyl-(1→2)-α-*L*-arabinopyranoside].



1. R=α-*L*-Ara$_p$ →
2. R=β-*D*-Gal$_p$ $\xrightarrow{2}$ α-*L*-Ara$_p$ →
3. R=α-*L*-Rha$_p$ $\xrightarrow{2}$ β-*D*-Gal$_p$ $\xrightarrow{2}$ α-*L*-Ara$_p$ →
4. R=H

EXPERIMENTAL

General Observations. TLC was conducted on Silufol plates, and CC on silica gel 100-250 μm, (Czechoslovakia)), using the following solvent systems: 1) chloroform—methanol (4:1); 2) chloroform—methanol (7:3); 3) chloroform—methanol—water (65:35:7); 4) chloroform—methanol (9:1); 5) benzene—acetone (4:1); 6) butanol—benzene—pyridine—water (5:1:3:3).

The PC of the sugars was conducted on FN-11 paper. Aniline phthalate was used to reveal the sugars, and 25% tungstophosphoric acid, followed by the heating of the chromatograms, for the glycosides, their acetates, and the aglycon.

IR spectra were taken on a UR-20 instrument in KBr tablets.

PMR spectra were obtained on a Bruker WM-250 instrument with working frequencies of 250 MHz for ^1H atoms and 62.9 MH for ^{13}C at 70°C in pyridine-d_5.

Isolation of the Glycosides. The comminuted air-dry roots of white sweet clover gathered in the phase of the flowering of the plant (8 kg) were extracted with 70% ethanol. The extract, evaporated to 1/3 volume, was treated successively with chloroform and butanol. The butanolic fraction was evaporated, giving a total of 100 g of extractive substances. Chromatography, using systems 1, 2, and 3 in succession, enabled fractions enriched with compounds (**1**), (**2**), and (**3**) to be obtained.

The Full Acetates of (1), (2), and (3) — (1a, 2a, and 3a). The fractions were acetylated with acetic anhydride in pyridine (1:1, 20°C, 48 h), followed by dilution with water and extraction with chloroform. The chloroform layer was washed and evaporated to dryness. The residue was deposited om a column and eluted with system 5. This gave: 108 mg of (**1a**) ($C_{45}H_{68}O_{12}$), mp 136-138°C; 72 mg of (**2a**) ($C_{57}H_{84}O_{20}$), mp 137-140°C; and 1.9 g of (**3a**) ($C_{67}H_{98}O_{20}$), mp 148-151°C (from the solvent system). The IR spectra of all the products lacked the absorption of OH groups.

PMR spectrum of (**3a**) (C_5D_5N, δ, ppm, 0 — TMS): 4.50 (dd, $J_{2,3}$ = 10.0 Hz, H-2′), 5.42 (d, $J_{1,2}$ = 7.5 Hz, H-1″), 4.42 (dd, $J_{2,3}$ = 10.0 Hz, H-2″), 5.55 (dd, $J_{3,4}$ = 3.0 Hz, H-3″), 5.86 (br. d, $J_{4,5}$ = 4.0 Hz, H-1‴), 5.50 (d, $J_{1,2}$ = 2.0 Hz, H-1⁗), 5.58 (dd, $J_{2,3}$ = 3.0 Hz, H-2‴), 5.79 (dd, $J_{3,4}$ = 10.0 Hz, H-3‴), 5.63 (t, $J_{4,5}$ = 10.0 Hz, H-4‴), 4.68 (m, H-5‴), 1.55 (H-6‴), 3.48 (H-3), 4.85 (t, $J_{22e,21a}$ = $J_{22e,21e}$ = 4.0 Hz, H-22).

Deacetylation of (1a), (2a), and (3a). Solutions of 100 mg of (**1a**), 70 mg of (**2a**), and 1.5 g of (**3a**) in 15, 10, and 20 ml, respectively, of 5% ethanolic KOH were boiled in the water bath under reflux for 5 h. Each reaction mixture was diluted with water and the ethanol was evaporated off. The aqueous solution was extracted repeatedly with water-saturated butanol. The combined butanolic extract was washed to neutrality, evaporated to dryness, and chromatographed on silica gel with elution by the appropriate system (1 for compound (**1**), 2 for compound (**2**), and 3 for compound (**3**)). In this way we

isolated: 80 mg (0.001% on the weight of the air-dry raw material) of melilotoside A ($C_{35}H_{58}O_7$); 42 mg (0.0005% on the weight of the air-dry raw material) of melilotoside B ($C_{41}H_{68}O_{12}$), decomp. p. 219-220°C (from system 2); and 1.15 g (0.014% on the weight of the air-dry raw material) of melilotoside C ($C_{47}H_{78}O_{16}$), decomp. p. 239-240°C (from system 3); all (KBr, ν, cm^{-1}) 3580-3200 (OH), 1630 (C=C).

Complete Acid Hydrolysis of the Glycosides. Solutions of compounds (**1**), (**2**), and (**3**) in 10 ml of 5% aqueous sulfuric acid were boiled in a sand bath under reflux for 24 h. Then each reaction mixture was diluted with water and extracted repeatedly with chloroform. The combined extract was washed with water to neutrality, evaporated to a dry residue, and chromatographed on silica gel in system 4. This gave 10 mg of the genin of compound (**3**) with mp 253-255°C (from methanol). The TLC analysis of the genins of compounds (**1**) and (**2**) enabled soyasapogenol B to be identified.

The hydrolysate was neutralized with barium carbonate and evaporated to dryness, and the minimum volume of methanol was added. The following sugars were identified by PC in system 6: for glycoside (**1**), arabinose; for (**2**), arabinose and galactose; and for (**3**), arabinose, galactose, and rhamnose.

Partial Acid Hydrolysis of Glycoside (3). A solution of 100 mg of (**3**) in 15 ml of 2.5% aqueous sulfuric acid was boiled on a sand bath under reflux for 24 h. Then the reaction mixture was diluted with water and was extracted repeatedly with water-saturated butanol. The combined butanolic extract was washed to neutrality, evaporated to a dry residue, and chromatographed successively in systems 4, 1, and 2. This led to the isolation of 20 mg of the progenin (**3a**) and 27 mg of the progenin (**3b**) with decomp. p. 219-220°C (from system 2).

Complete Acid Hydrolysis of the Progenins. Compounds (**3b**) and (**3c**) (5 mg) were each dissolved in 5 ml of 5% aqueous sulfuric acid and the solutions were boiled on a sand bath under reflux for 24 h. The reaction mixtures were neutralized with barium carbonate, filtered, evaporated to dryness, and treated with the minimum volume of methanol. The following sugars were identified in the residue by PC in system 6: for (**3b**), arabinose; for (**3c**), arabinose and galactose.

REFERENCES

1. N. I. Rubtsov, Handbook of Higher Plants of the Crimea [in Russian], Leningrad (1972), p. 254.
2. P. D. Sokolov [editor-in-chief], Plant Resources of the USSR. Flowering Plants, their Composition and Uses. The Families Hydrangeaceae—Haloragaceae [in Russian], Leningrad (1987), p. 159.
3. G. A. Kuznetsova, Natural Coumarins and Furocoumarins [in Russian], Leningrad (1967), p. 246.
4. W. R. Akeson, H. J. Gorz, and F. A. Haskins, Crop Sci., **3**, No. 2, 167 (1963).
5. J. B. Harborne, The Biochemistry of Phenolic Compounds, Academic Press, New York (1964) [Russian translation: Moscow (1968), p. 257].
6. A. S. Akramov, A. N. Umarov, and A. L. Markman, Khim. Prir. Soedin., 314 (1968).
7. M. Torck, L. Bezanger-Beauquesne, and M. Pinkas, Ann. Pharm. Fr. **27**, No. 6., 419 (1969). [Chem. Abstr., **72**, 63577 (1970)].
8. E. I. Kozarets, M. S. Dudkin, and I. S. Kazanskaya, Khim. Prir. Soedin., 249 (1980).
9. Yu. M. Fadeev and V. N. Golubev, Sb. Nauchn. Trudov. Nikit. Botan. Sad, **95**, 54 (1985).
10. V. I. Grishkovets, N. V. Tolkacheva, A. S. Shashkov, and V. Ya. Chirva, Khim. Prir. Soedin., 686 (1991).
11. A. I. Kalinovskii, Khim. Prir. Soedin., 3 (1992).
12. T. Fukunaga, K. Nishiya, K. Takeya, and H. Itokawa, Chem. Pharm. Bull., **35**, 1610 (1987).
13. G. V. Shevchuk, Yu. S. Vollerner, A. S. Shashkov, and V. Ya. Chirva, Khim. Prir. Soedin., 678 (1991).
14. V. I. Grishkovets, A. A. Poloiko, A. S. Shashkov, and V. Ya. Chirva, Khim. Prir. Soedin., 779 (1990).
15. A. A. Bothner-By, R. L. Stephens, J. Lee, C. D. Warren, and R. W. Jeanloz, J. Am. Chem. Soc, **106**, No. 3, 811 (1984).