Pergamon

PII: S0031-9422(97)00574-8

Phytochemistry, Vol. 47, No. 3, pp. 437–440, 1998

Printed in Great Britain
0031-9422/98 $19.00+0.00

THE ELUCIDATION OF THE STRUCTURE OF THALICOSIDE F, A MINOR OLEANANE GLYCOSIDE FROM *THALICTRUM MINUS* L.*

A. S. GROMOVA, V. I. LUTSKY, A. A. SEMENOV, D. LI and N. L. OWEN†

Institute of Organic Chemistry, Siberian Branch of Russian Academy of Sciences, Irkutsk, 664033 Russia; †Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, U.S.A.

(*Received* 3 *December* 1996. *In revised form* 29 *April* 1997)

Key Word Index—*Thalictrum minus* L.; Ranunculaceae; triterpenoid saponins; oleanane; thalicoside F; NMR spectra.

Abstract—A new triterpenoid saponin, thalicoside F was isolated from the above-ground part of *Thalictrum minus* L. (Ranunculaceae family) collected in Eastern Siberia (Russia). The structure of thalicoside F was elucidated by a combination of 1D- and 2D-NMR spectroscopy and mass spectroscopy, and was determined to be $C_{47}H_{74}O_{17}$\{3-β-*O*-[α-L-rhamnopyranosyl-(1 → 2)-β-D-glucopyranosyl-(1 → 4)-α-L-arabinopyranosyl]-11α, 12α-epoxyoleanane-28,13β-olide\}.

INTRODUCTION

It has been previously reported [2] that a chemorace of the Thalictrum genus, producing triterpenoid saponins possessing antitumour and contraceptive activities, grows in the area of Siberia. In particular, seven oleanane and cycloartane oligosides have been isolated from *Th. minus*. The isolation and elucidation of the chemical structure of thalicoside F, a new minor compound from the above-ground part of this plant is described in this paper. Since only a very small amount of the material (12 mg) was available, all structural investigations were conducted using only NMR spectroscopy and mass-spectroscopy.

RESULTS AND DISCUSSION

^{13}C NMR spectra show that glycoside F contains 47 carbon atoms. The molecule also has seven characteristic tertiary methyl signals, a γ-lactone and an epoxide ring and there are 17 other oxygen bound carbon atoms, one on the sterol (C-3) and the rest corresponding to sugar units. These data suggest that the substance examined pertains to a class of triterpenoid glycosides with three sugars. The mass spectrum of the glycoside shows a quasi-molecular ion at *m/z* 933 ($[M+Na]^+$, 100%) and an ion fragment at *m/z* 470 ([M-440], 3%). Analysis of the high-resolution FAB mass spectrum shows a $[M]^+$ of mass 910.4924 corresponding with the molecular formula $C_{47}H_{74}O_{17}$ to within 0.2 ppm. Taking into consideration the data of ^{13}C NMR and mass spectra it can be assumed that the ion at *m/z* 470 belongs to the genol and is formed during loss of the trisaccharide unit (hexose, deoxyhexose and pentose) from the molecule (162 + 146 + 132 = 440).

More detailed examination of the NMR spectra (COSY, TOCSY, ROESY, HETCOR and HMBC) makes it possible to conclude that the glycoside genol belongs to the oleanane group and is 3-hydroxy-11α, 12α-epoxy oleanane-28, 13β-olide as described elsewhere [3, 4]. Two such oleanane lactones enantiomeric in C-3 are known. Once the structure of the genol was determined and the corresponding 30 ^{13}C signals assigned to that structure (Table 1), the chemical shift pattern of the remaining 17 carbons was strongly indicative of the presence of one pentose and two hexoses which agreed with the mass spectral data. Three ^{13}C signals in the range 100–110 ppm (A_1 at 107.16; G_1 at 105.10; and R_1 at 101.97 ppm) were readily recognized as the anomeric carbons of the sugars, and from the HETCOR experiment the corresponding anomeric protons were found to be at 4.74, 5.17 and 6.27 ppm, respectively. The next step was to use COSY data to establish the proton-proton connectivities, and this information together with HETCOR data enabled us to determine the carbon skeletons of the sugars. Starting with the anomeric protons at 5.17 and 6.27 ppm, two hexose proton spin systems were traced out on the COSY spectrum. The only ambiguity encountered was the connectivity between the R2 and R3 protons due to signal overlapping. This ambiguity was later removed and the

* Part IX in the series 'Triterpenoid Saponins from *Thalictrum minus* L.' (For part VIII, see ref. 1).

† Author to whom correspondence should be addressed.

Table 1. ^1H and ^{13}C NMR assignment of the aglycone moiety of thalicoside F*

Position	Type (DEPT)	^{13}C (ppm)	^1H (ppm)
1	CH_2	38.42	1.05, 1.76
2	CH_2	26.08	1.83, 2.02
3	CH—O	88.71	3.24 *dd* (J = 11.50, 4.55 Hz)
4	C	39.54	—
5	CH	55.00	0.71 *br d*
6	CH_2	17.73	0.88, 1.45
7	CH_2	31.36	1.03, 1.18
8	C	40.89	—
9	CH	51.13	1.64
10	C	36.42	—
11	CH—O	52.74	3.07 *dd* (J = 3.93, 1.51 Hz)
12	CH—O	57.26	3.24 *m* (J = 3.93 Hz)
13	C—O	87.60	—
14	C	41.63	—
15	CH_2	27.05	1.00, 1.34
16	CH_2	21.64	1.28, 2.16
17	C	44.09	—
18	CH	49.76	2.52 *dd*
19	CH_2	38.02	1.72, 1.94
20	C	31.50	—
21	CH_2	34.42	1.19, 1.33
22	CH_2	27.65	1.64, 1.78
23	CH_3	27.83	1.20 *s*
24	CH_3	16.33	0.93 *s*
25	CH_3	17.26	0.90 *s*
26	CH_3	18.38	1.22 *s*
27	CH_3	20.37	1.12 *s*
28	C=O	178.92	—
29	CH_3	33.08	0.89 *s*
30	CH_3	23.48	0.81 *s*

* Assignments were aided by DEPT, COSY, TOCSY and HETCOR experiments. The chemical shift values are presented in ppm and referenced to TMS.

Table 2. ^1H and ^{13}C NMR assignments of the sugar units of thalicoside F*

Compound	^{13}C	^1H
Arabinose		
A1	107.16	4.74, *d* (J = 6.30 Hz)
A2	72.85	4.44
A3	73.77	4.25
A4	78.77	4.38 *br s*
A5	65.34	3.82; 4.67
Glucose		
G1	105.10	5.17, *d* (J = 7.50 Hz)
G2	78.26	4.17
G3	78.89	4.20
G4	71.76	4.17
G5	78.31	3.87
G6	62.68	4.33; 4.46
Rhamnose		
R1	101.97	6.27, *s*
R2	72.09	4.75, *s*
R3	72.37	4.76
R4	74.58	4.22
R5	70.02	4.96
R6	18.65	1.68, *d* (J = 6.00 Hz)

* Chemical shift values are presented in ppm and referenced to TMS.

correct assignment confirmed by a HMBC experiment in which the long-range coupling between the R1 proton and the R3 carbon was unambiguously observed. Thus the proton signals were assigned to specific carbons in the carbohydrates (Table 2), and the hexoses were identified to be α-L-rhamnopyranose and β-D-glycopyranose. The remaining pentose signals could belong to either arabinose or xylose. In the ^1H NMR spectra the H-4 proton of the pentose sugar resonates as a broad singlet (4.38 ppm). The small coupling constants of the H-4 proton with neighboring protons indicates an equatorial orientation for H-4. Hence, we believe the pentose to be arabinose. The site of attachment of the carbohydrate chain to the genol, and the sequence of the sugars were determined using ROESY and HMBC spectra (Figs. 1, 2).

Arabinose is the first sugar in the carbohydrate chain since the HMBC spectrum shows a cross-peak connecting the anomeric proton (4.74 ppm) with the C-3 atom of the sapogenol (88.71 ppm) and in addition, the H-3 of the triterpenoid moiety (3.24 ppm) is linked to the C-1 of arabinose (107.16 ppm). The spatial proximity of the H-3 proton and the H-1 proton of the arabinose can also be seen in the ROESY spectrum (Fig. 1).

In the ^{13}C NMR spectrum the C-4 atom of arabinose is shifted down-field to 78.77 ppm because of the glycosylation effect. The glucose molecule is attached to this site as indicated in the HMBC spectrum by the long range coupling between $H\text{-}1_{Glc}$ and $C\text{-}4_{Ara}$ (5.17 and 78.77 ppm, respectively) and a distinct cross-peak between $H\text{-}1_{Glc}$ and $H\text{-}4_{Ara}$ (4.38 ppm) is also seen in the ROESY spectrum.

The terminal sugar of the carbohydrate chain is α-L-rhamnopyranose as the chemical shifts of all its carbon atoms correspond to the generally accepted data for methyl-α-L-rhamnopyranoside. The glucose C-2 atom is the site of attachment of the terminal sugar. This is indicated by a down field shift of the glucose C-2 signal (78.26 ppm) and by the corresponding cross-peaks in the HMBC spectra of $C\text{-}1_{Rham}$ and $H\text{-}2_{Glc}$ (101.97 and 4.17 ppm) as well as between $C\text{-}2_{Glc}$ and $H\text{-}1_{Rha}$ (78.26 and 6.27 ppm).

Therefore, thalicoside F is concluded to be 3-β-*O*-[α-L-ramnopyranosyl-(1 → 2)-β-D-glucopyranosyl-(1 → 4)-α-L-arabinopyranosyl]-11α,12α-epoxyoleanan-28,13β-olide (Fig. 3). Two bisdesmosides of oleanolic acid named B and D thalicosides were previously isolated from *Thalictrum minus* L. [1]. In these substances the carbohydrate chain attached to the C-3 atom of the aglycone is similar to that of thalicoside F.

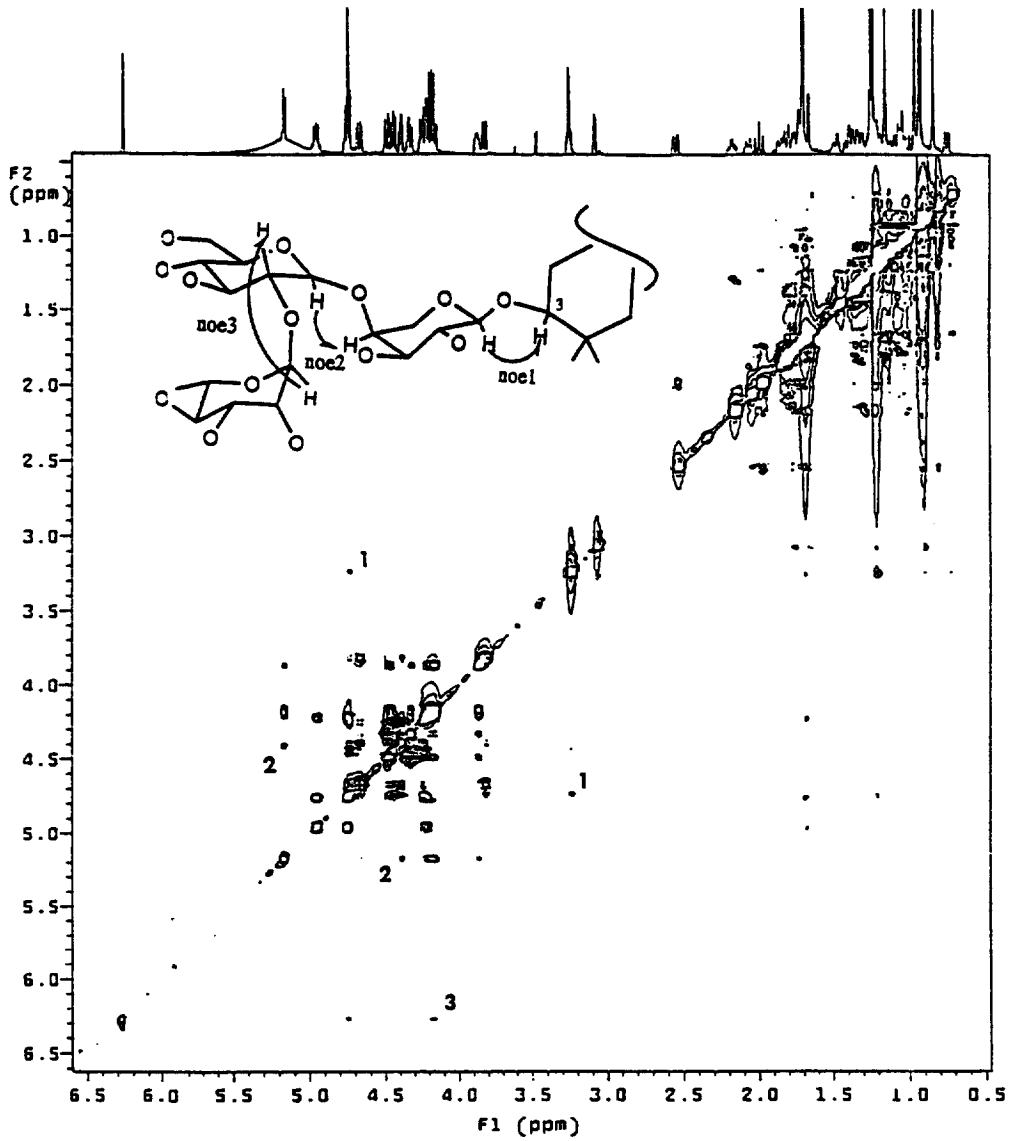

Fig. 1. Through-space 1H-1H nOe Responses in the ROESY NMR Spectrum and the Corresponding Glycosidic Bonds.

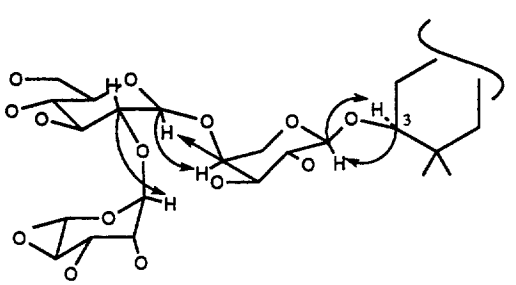

Fig. 2. Long-range ^{13}C-^{1}H couplings of the glycosidic bonds observed in HMBC experiment.

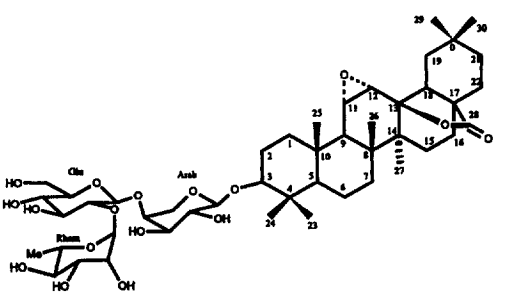

Fig. 3.

EXPERIMENTAL

Mass spectra. MS spectra were obtained using both a LKB-2091/PDB-11/34 mass spectrometer, and a JEOL SX102A double focussing, reverse geometry, spectrometer. High resolution spectra were obtained with the latter using a FAB source with xenon as the ionizing gas and thioglycerol (plus 1% sodium acetate) as a matrix. Polyethylene glycol was used for calibration purposes.

NMR spectra. All NMR spectra were obtained using a Varian VXR-500S spectrometer equipped with a SUN SPARC 20 work-station. Samples were run at 26° in 5 mm NMR tubes at concentrations of about

12 mg 0.7 ml^{-1} of pyridine-d_5. ^1H NMR chemical shifts were measured with respect to TMS whereas ^{13}C NMR chemical shifts were measured with reference to the pyridine signal at 123.5 ppm. In COSY experiments, both the f_1 and f_2 spectral widths were 5 kHz, and the (t_1, t_2) data matrix was zero-filled to 2048 by 2048 data points to give a final resolution of 5 Hz per point. For ROESY and absolute value TOCSY experiments, the same spectral width was used in both f_1 and f_2 dimensions. The mixing time used for the ROESY experiment was 300 ms. For HETCOR experiments the spectral width was 4 kHz in the f_1 (^1H) dimension and 20 kHz in f_2 (^{13}C). Sixty-four scans in the f_2 dimension were accumulated for each of the 400 increments in the f_1 dimension. The total acquisition time was 12 hr. For the HMBC experiment, a Nalorac indirect-detection 5 mm probe was used with appropriate re-cabling of the NMR spectrometer. The acquisition recycle time was 1.21 s, and data acquisition took 11 hr to complete.

Plant material. Thalictrum minus L. was collected at the time of its flowering (July) in the upper Kitoi river (a left tributary of the Angara river, Eastern Siberia, Russia). A herbarium sample is deposited in the M.G. Popov Herbarium of the Siberian Institute of Plant Physiology and Biochemistry of the Russian Academy of Sciences, Irkutsk.

Isolation. Analytical grade solvents were used throughout the isolation procedure. The air-dried, above-ground part of the plant (4.5 kg) was extracted with 80% aq. MeOH (3 × 50 l). After removal of the MeOH, the H_2O soln was extracted with $CHCl_3$ (3 × 20 l), followed by *n*-BuOH (3 × 20 l). The *n*-BuOH extract was taken to dryness, yielding 299 g of residue which was dissolved in MeOH (1 l), and stored in a refrigerator for a day at 1–3°. The ppt. which formed was filtered, and the filtrate concd to obtain a viscous liquid, which was slowly added to 3 l of Me_2CO with vigorous stirring. The resulting ppt., which comprised the more polar saponins [2] was sepd from the filtrate containing the less polar saponins. The Me_2CO was removed and the residue was treated with EtOAc (3 × 100 ml), followed by *n*-BuOH (3 × 100 ml). This *n*-BuOH extract yielded 18.3 g of solid which was chromatographed [Si-gel, 40–100 (Lachema); $CHCl_3$–MeOH–H_2O (80:30:7)] to give 54 frs, each of 400 ml. The first 10 frs were combined and passed through a short Al_2O_3 column to eliminate phenols. The solvent was removed to give a residue (1.1 g) which was re-chromatographed [Si-gel, 5–40 (Lachema); $CHCl_3$–MeOH–H_2O (70:23:4 and 70:23:1)] and sepd into 30 50 ml frs. The compound of interest was contained in frs 19–21. These frs were pooled following removal of the solvent and the resulting 30 mg of material was again subjected to silica gel chromatography [$CHCl_3$–MeOH–H_2O (70:23:4)] to give 15 20 ml frs. The material in frs 9– = 13 of this last set was repeatedly chromatographed [$CHCl_3$–MeOH–H_2O (70:23:1)] until pure. The final product was recrystallized from MeOH to yield 12 mg of pure thalicoside F. [mp 268–270°; FAB-MS, *m/z*: $[M + Na]^+$ 933.4822].

Acknowledgements—We express our appreciation to Dr A. Vereschagin (Institute of Limnology of the Siberian Branch of the Russian Academy of Sciences, Irkutsk), and Mr Bruce Jackson (Brigham Young University) for measurements of MS spectra, and to Dr Steven G. Wood (BYU) for helpful discussions.

REFERENCES

1. Gromova, A. S., Semenov, A. A., Lutsky, V. I., Zinchenko, S. V., Trofimova, N. N. and Rashekes, Ja. V., *Khimia Prirodnykh Soedinenii*, 1994, 398.
2. Semenov, A. A., Lutsky, V. I., Gromova, A. S., Ganenko, T. V., Khamidullina, E. A. and Trofimova, N. N., in *Saponins: Chemistry and Biological Activity*, ed. G. R. Waller, Plenum Press, New York, 1996, p. 193.
3. Melek, F. R., Radwan, A. S., Ahmed, A. A. and Aboutabl, E. A., *Pharmazie*, 1989, **44**, 735.
4. Ikuta, A. and Morikawa, A., *Journal of Natural Products*, 1992, **55**, 1230.