Pergamon

PII: S0031-9422(97)01038-8

Phytochemistry, Vol. 48, No. 6, pp. 1039–1043, 1998

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

MOSCHATINE: AN UNUSUAL STEROIDAL GLYCOSIDE FROM *CENTAUREA MOSCHATA*

SATYAJIT D. SARKER, LAURENCE DINAN,* VLADIMIR ŠIK† and HUW H. REES‡

Department of Biological Sciences, University of Exeter, Hatherly Laboratories, Prince of Wales Road, Exeter, Devon, EX4 4PS, U.K.; † Department of Chemistry, University of Exeter, Stocker Road, Exeter, Devon, EX4 4QD, U.K.; ‡ School of Biological Sciences, The University of Liverpool, Life Sciences Building, Crown Street, Liverpool, L69 7ZB, U.K.

(*Received* 4 *August* 1997)

Key Word Index—*Centaurea moschata*; Compositae; moschatine; steroidal glycoside; pregnane.

Abstract—High-performance liquid chromatographic analysis of the seeds of *Centaurea moschata* has afforded a novel steroidal glycoside: moschatine [(20*S*)-20-{6-*O*-(3-methoxy-4-α-L-rhamnopyranosyloxybenzoyl)-β-D-glucopyranosyloxy}-15α-hydroxy-8β,9α,14α,17α-pregn-4-en-3-one]. The structure of this compound was elucidated by a combination of UV, LSIMS, and a series of 1D and 2D NMR spectral analyses. © 1998 Elsevier Science Ltd. All rights reserved

INTRODUCTION

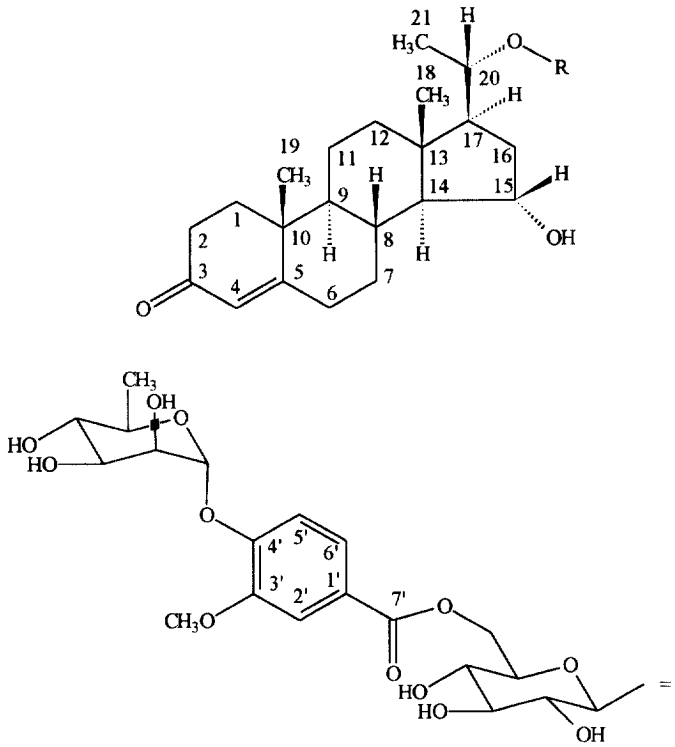
Centaurea moschata L. (common name: 'Sweet Sultan') is an easily grown, green, glabrous, oriental annual from the genus *Centaurea* which comprises ca 500 species, mainly from the Old World [1]. As many species of this genus fall into the category of medicinal plants [2, 3], they have been the object of numerous phytochemical investigations leading to the isolation and identification of a variety of plant secondary metabolites [3, 4]. We have reported earlier, the presence of four indole alkaloids (moschamine, *cis*-moschamine, moschamindole and moschamindolol) [4], and phytoecdysteroids [5] as major constituents, and (20*R*)-15α-hydroxypregn-4-en-3-one-20-*O*-β-D-glucopyranoside as a minor [6] component in the seeds of *C. moschata*. We have re-investigated this species and we now report on the isolation and identification of a further novel steroidal glycoside.

RESULTS AND DISCUSSION

RP-HPLC analysis of a methanol extract of the seeds of *Centaurea moschata* L. yielded a new steroidal glycoside: (20*S*)-20-{6-*O*-(3-methoxy-4-α-L-rhamnopyranosyloxybenzoyl) - β - D - glucopyranosyloxy} - 15α-hydroxy-8β,9α,14α,17α-pregn-4-en-3-one (**1**), named moschatine. The structure of this compound was determined by 1D and 2D NMR spectroscopy, notably ^{1}H, ^{13}C PENDANT [7], COSY45, NOESY, HMBC [8] and HMQC.

The UV absorption maxima (log ε) of **1** at 246 (4.10), 255 (4.02), 290 (3.97) nm suggested the presence of an α,β-unsaturated carbonyl group and substituted benzene moiety within the molecule. LSI mass spectrum revealed $[M+H]^+$ (positive ion mode) ion at *m/z* 791, suggesting $M_r = 790$ and solving for $C_{41}H_{58}O_{15}$. Analysis of the ^{1}H NMR, ^{1}H-^{1}H COSY45, and ^{13}C PENDANT spectra (Table 1), along with a ^{1}H-^{13}C HMQC spectrum indicated the presence of a glucose unit, a rhamnose unit, a 3,4-dioxygenated benzoyl group, a methoxy and a 20,15-dioxygenated pregn-4-en-3-one moiety in the molecule. A doublet at δ_H 4.38 ($J = 7.8$ Hz) (δ_C 99.7) integrated for a single proton identified the glucose anomeric proton, from which a ^{1}H-^{1}H COSY45 spectrum enabled identification of the connection of H-2_g to H-6_g of the glucose unit and suggested it as β-D-glucose which was further confirmed from the observed NOE: H-$1_g \leftrightarrow$ H-3_g, H-$1_g \leftrightarrow$ H-5_g and H-$2_g \leftrightarrow$ H-4_g in the ^{1}H-^{1}H NOESY spectrum (Fig. 1). The rhamnose moiety [anomeric signal: δ_H 5.47 ($J = 1.5$ Hz), δ_C 99.4] could easily be identified from the occurrence of a 3H doublet at δ_H 1.22 ($J = 6.2$ Hz) for the methyl group (C-6_r), and from analysis of ^{1}H-^{1}H COSY and NOESY spectra it was further identified as α-L-rhamnose. From the ^{1}H and ^{13}C PENDANT spectra, *ortho*, *ortho-meta* and *meta*-splitting of the aromatic protons, oxygenations at C-3′ and C-4′, and carboxy group at C-1′ could be rationalised. The attachment of the methoxy group at C-3′ was confirmed from the ^{1}H-^{13}C long-range

* Author to whom correspondence should be addressed. Tel: [44]-1392-264605; Fax: [44]-1392-264668; E-mail: L.N.Dinan@exeter.ac.uk



1

HMBC correlation (Table 2) between the methoxy protons (δ_H 3.88) and C-3′ (δ_C 150.1), and also from the strong NOE between this methoxy signal and H-2′ (δ_H 7.63) (Fig. 1). Similarly, the presence of a carboxyl group at C-1′ was confirmed from the HMBC where both H-2′ and H-6′ (δ_H 7.65) showed ^3J correlations to the carbonyl carbon (C-7′, δ_C 166.1). In addition to the signals for glucose, rhamnose and 3-methoxy-4-oxy benzoyl moieties, the ^{1}H NMR provided the distinct signals for three methyls, one of which appeared as a doublet (δ_H 1.07, J = 6.1 Hz), an olefinic proton at δ_H 5.66, two oxymethine protons at δ_H 3.85 and δ_H 3.80, and a number of overlapped complex signals for the methylene and methine protons. The ^{13}C PENDANT spectrum showed signals for 21 additional carbons including a ketonic carbonyl (δ_C 201.1), an olefinic quaternary (δ_C 174.2), two quaternaries, three methyls, seven methylenes and seven methine carbons including an olefinic methine (δ_C 122.40 and two oxymethines (δ_C 72.8 and δ_C 74.2). All the ^{1}H resonances for the steroidal moiety were comparable to those published for the compound (20*R*)-15α,20β-dihydroxypregn-4-en-3-one 20-*O*-β-D-glucopyranoside [6], with the most notable exceptions being the signals for Me-18 (δ_H 0.67) and Me-19 (δ_H 1.00) which were much more shielded than the published values (δ_H 0.89 and 1.27, respectively) which implied *S*-stereochemistry at C-20 [9]. This was further evident from the strong NOE between the Me-18 and H-20 (δ_H 3.80) which confirmed the presence of H-20 at the β-face of the molecule. In the COSY45 spectrum, among the major correlations, the Me-21 coupled to the proton at δ_H 3.80, which again extended the couplings to the protons on C-17 ↔ C-16 ↔ C-15 ↔ C-14 and thus confirmed the presence of a hydroxyl on C-15. The α-orientation of the 15-OH was also determined from the NOE between Me-18 and H-15 (δ_H 3.85) both of which were in β-orientation. The identity of the steroidal moiety was confirmed by ^{1}H-^{13}C HMBC (Table 2). Among the ^{1}H-^{13}C long-range correlations the most notable were: from Me-21 (δ_H 1.07) to C-20 (^2J, δ_C 74.2) and C-17 (^3J, δ_C 53.6); from Me-18 (δ_H 0.67), ^2J to C-13 (δ_C 42.7) and ^3J to C-17, C-12 (δ_C 38.7), C-14 (δ_C 60.8); from H-14 (δ_H 0.95), ^2J to C-15 (δ_C 72.8) and ^3J to C-7 (δ_C 32.1), C-12, C-17, C-18 (δ_C 11.8); from Me-19 (δ_H 1.00), ^3J to C-5 (δ_C 174.2), and from the methylene protons at C-1 and C-2 (δ_H 1.60, 1.80 and 2.25, 2.45, respectively to the carbonyl (C-3, δ_C 201.1). Thus, the steroidal moiety was identified as (20*S*)-20β-oxy-15α-hydroxypregn-4-en-3-one. The positions of linkage between the glucose, rhamnose, 3-methoxy-4-oxybenzoyl and steroid moieties, which led to the identification of this compound as **1**, were also determined by ^{1}H-^{13}C long-range correlations

Table 1. ^{1}H (400 MHz) and ^{13}C PENDANT NMR data of **1**. (Coupling constant J = Hz, in parentheses)

Carbon no.	δ^1H	δ^{13}C
1	1.60*, 1.80*	37.5
2	2.25*, 245*	33.4
3α	—	201.1
4	5.66 *s*	122.4
5	—	174.2
6	2.25*, 2.45*	32.6
7	1.22*, 2.15*	32.1
8β	1.55*	35.0
9α	0.86*	54.2
10	—	38.6
11	0.90*, 1.50*	20.4
12	2.18*, 1.05*	38.7
13	—	42.7
14α	0.95 *dd* (9.4)	60.8
15β	3.85*	72.8
16	1.85*, 1.45*	35.2
17α	1.75*	53.6
18 (Me)	0.67 *s*	11.8
19 (Me)	1.00 *s*	16.3
20β	3.80*	74.2
21 (Me)	1.07 *d* (6.1)	16.6
1′	—	124.6
2′	7.63 *d* (2.0)	113.1
3′	—	150.1
4′	—	149.9
5′	7.23 *d* (8.3)	116.4
6′	7.65 *dd* (2.0, 8.3)	123.0
7′	—	166.1
3′-MeO	3.88 *s*	55.2
1_g	4.38 *d* (7.8)	99.7
2_g	3.15 *dd* (7.8, 9.0)	74.0
3_g	3.38 *bt* (8.9)	76.5
4_g	3.55*	70.1
5_g	3.52*	72.2
6_ga	4.68 *d* (11.7)	62.8
6_gb	4.40 *dd* (11.7, 1.0)	
1_r	5.47 *d* (1.5)	99.4
2_r	4.06 *dd* (1.5, 3.4)	70.6
3_r	3.85*	70.8
4_r	3.48 *t* (9.5)	73.3
5_r	3.70*	69.7
6_r (Me)	1.22 *d* (6.2)	17.6

Spectra obtained in CD_3OD referenced to CH_3OH at δ 3.31 (^{1}H) and δ 49.15 (^{13}C).

^{1}H chemical shifts of methylene protons = former value is for α and the latter is for β orientation.

* Assignments obtained from COSY45.

observed in the HMBC spectrum as follows: ^3J correlation from H-1_g (δ_H 4.38) to C-20 (δ_C 74.2) confirmed the attachment of the glucose moiety at C-20, ^3J connectivity between H-6_g (δ_H 4.68, 4.40) to the carbonyl carbon of the benzoyl moiety (C-7′, δ_C 166.1) established that the benzoyl moiety was linked to C-6_g; similarly, ^3J correlation from H-1_r (δ_H 5.47) to C-4′ (δ_C 149.9) confirmed the attachment of the rhamnose at C-4′ of the benzoyl group. All these attachments were also supported from the following NOEs: H-$1_g \leftrightarrow$ H-20, H-$1_g \leftrightarrow$ 21-Me, and H-$1_r \leftrightarrow$ 3′-MeO (weak). Apart from these NOE features, a NOESY spectrum established unambiguously the different spatial orientations of the protons in the molecule as a whole. Thus, the structure of this compound was unequivocally identified as (20*S*)-20-{6-*O*-(3-methoxy-4-α-L-rhamnopyranosyloxybenzoyl)-β-D-glucopyranosyloxy}-15α-hydroxy-8β,9α,14α,17α-pregn-4-en-3-one (**1**).

The *R*-form of the aglycone of **1**, (20*R*)-15α,20β-dihydroxypregn-4-en-3-one, was first isolated from the defence secretion of the water beetle *Platambus maculatus* [10]. A glucoside of (20*R*)-15α,20β-dihydroxypregn-4-en-3-one was isolated from *Carthamus tinctorus* [11] and *Centaurea moschata* [6]. Most plant steroids have a hydroxyl group at C-3 like **2**, whereas compound **1**, having a 3-keto-Δ^4 structure, is more closely related to animal steroids e.g. progesterone analogues. However, progesterone has also been found in plants and, although not very common, there are reports concerning other plant steroids possessing a 3-keto group [11, 12]. The extract from which this steroidal glycoside was isolated was very active in a bioassay [13] for ecdysteroid agonists. However, moschatine is not responsible for this activity.

EXPERIMENTAL

The UV spectrum was in MeOH. NMR spectra were obtained on a Bruker AVANCE DRX400 instrument using standard Bruker microprograms. The chemical shifts are expressed in ppm. LSIMS (positive ion mode); glycerol matrix using a Cs^+ primary ion beam on a VG Quattro triple quadrupole mass spectrometer (VG Biotech. Altrincham, U.K.). HPLC separations were performed on a Gilson model 802C HPLC coupled with Gilson UV-Visible detector. RP stands for reversed-phase. Preparative RP-HPLC was performed on Technoprep 10 C_8 column and final purification was done on Spherisorb C_6 semi-preparative column. HPLC separations were monitored at 242 nm.

Plant material

Seeds of *Centaurea moschata* L. were purchased from Chiltern Seeds, Ulverston, U.K.

Extraction

Ground seeds (11 g) were extracted in a Soxhlet with, successively, *n*-hexane, $CHCl_3$ and MeOH, 500

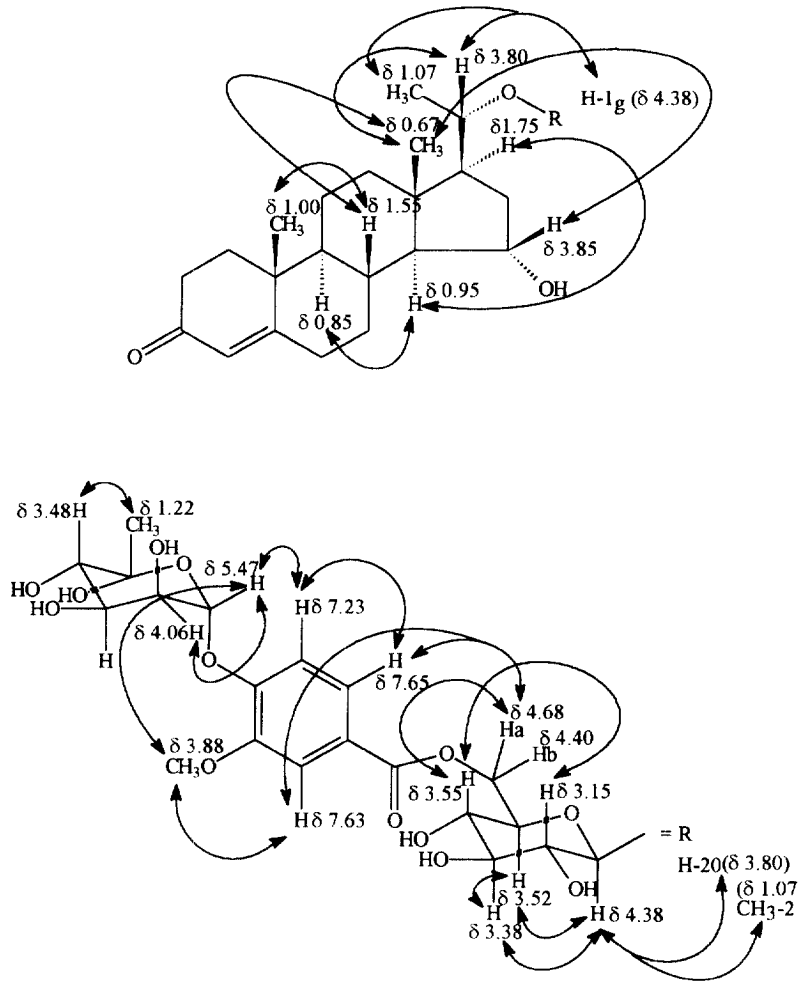


Fig. 1. Key NOE from ^{1}H-^{1}H NOESY spectrum.

ml each. The MeOH extract was concentrated using a rotary evaporator at a maximum temperature of 45 .

Isolation of compounds

Preparative RP-HPLC with 55% MeOH–H_2O at 5 ml min^{-1} yielded **1** (26–36 min) as a broad UV-absorbing peak. Compound **1** was further purified on C_6 semipreparative column, eluted with 45% MeOH–H_2O, 2 ml min^{-1} (R_t = 35 min).

Bioassay

Ecdysteroid agonist and antagonist activities of **1** were assessed with a microplate-based bioassay using the *Drosophila melanogaster* B_{II} cell line [13]. Compound **1** was tested at concentrations ranging from 10^{-8} M–10^{-4} M. for the antagonist assay, a concentration of 20-hydroxyecdysone of 5×10^{-8} M was used.

Moschatine A [(20*S*)-20{6-O-(3-*methoxy*-4-α-L-*rhamnopyranosyloxybenzoyl*)-β-D-*glucopyranosyloxy*}-15α-*hydroxy*-8β,9α,14α,17α-*pregn*-4-*en*-3-*one*] (**1**). (1.2 mg). Gum. $UV\lambda^{MeOH}_{max}$ nm (log ε): 218 (4.30), 246 (4.10), 255 (4.02), 290 (3.97). ^{1}H NMR (Table 1). ^{13}C NMR (Table 1). LSIMS (positive ion mode) *m/z* 791 $[M+H]^+$.

Acknowledgements—This research was supported by a grant from the Biotechnology and Biological Sciences Research Council. Dr Tamara Savchenko and Pensri Whiting are thanked for valuable assistance with the bioassay.

Table 2. ^{1}H-^{13}C HMQC direct correlation (^{1}J) and major ^{1}H-^{13}C HMBC long-range correlations (^{2}J and ^{3}J) in **1**

	δ^{13}C		
Proton	^{1}J	^{2}J	^{3}J
H_2-1	37.5 (C-1)		201.1 (C-3), 54.2 (C-9),
H_2-2	33.4 (C-2)	201.1 (C-3)	
H-4	122.4 (C-4)		33.4 (C-2), 32.6 (C-6), 38.6 (C-10)
H_2-6	32.6 (C-6)		
H_2-7	32.1 (C-7)		
H-8	35.0 (C-8)		
H-9	54.2 (C-9)		37.5 (C-1), 38.7 (C-12), 16.3 (C-19), 60.8 (C-14)
H_2-11	20.4 (C-11)	54.2 (C-80	42.7 (C-13)
H_2-12	38.7 (C-12)	20.4 (C-11)	54.2 (C-8), 60.8 (C-14), 53.6 (C-17)
H-14	60.8 (C-14)	35.0 (C-8), 42.7 (C-13), 72.8 (C-15)	32.1 (C-7), 38.7 (C-12), 53.6 (C-17), 11.8 (C-18)
H_2-15	72.8 (C-15)		53.6 (C-17)
H_2-16	35.2 (C-16)		
H-17	53.6 (C-17)		
Me-18	11.8 (C-18)	42.7 (C-13)	38.7 (C-12), 60.8 (C-14), 53.6 (C-17)
Me-19	16.3 (C-19)	38.6 (C-10)	174.2 (C-5)
H-20	74.2 (C-20)	53.6 (C-17)	
Me-21	16.6 (C-21)	74.2 (C-20)	53.6 (C-17)
H-2′	113.1 (C-2′)		150.1 (C-3′), 123.0 (C-6′), 166.1 (C-7′)
H-5′	116.4 (C-5′)		124.6 (C-1′), 150.1 (C-3′)
H-6′	123.0 (C-6′)		113.1 (C-2′), 149.9 (C-4′), 166.1 (C-7′)
MeO-3′	55.2 (MeO-3′)		150.1 (C-3′)
H-1_g	99.7 (C-1_g)		74.2 (C-20), 76.5 (C-3_g), 72.2 (C-5_g)
H-2_g	74.0 (C-2_g)	99.7 (C-1_g), 76.5 (C-3_g)	70.1 (C-4_g)
H-3_g	76.5 (C-3_g)	70.1 (C-4_g)	99.7 (C-1_g), 72.2 (C-5_g)
H-4_g	70.1 (C-4_g)	76.5 (C-3_g)	74.0 (C-2_g)
H-5_g	72.2 (C-5_g)		76.5 (C-3_g)
H_2-6_g	62.8 (C-6_g)		70.1 (C-4_g), 166.1 (C-7′)
H-1_r	99.4 (C-1_r)	70.6 (C-2_r)	69.7 (C-5_r), 149.9 (C-4′), 70.8 (C-3_r)
H-2_r	70.6 (C-2_r)	70.8 (C-3_r)	73.3 (C-4_r)
H-3_r	70.8 (C-3_r)	70.6 (C-2_r)	
H-4_r	73.3 (C-4_r)	70.8 (C-3_r)	70.6 (C-2_r), 17.6 (C-6_r)
H-5_r	69.7 (C-5_r)		70.8 (C-3_r)
Me-6_r	17.6 (C-6_r)	69.7 (C-5_r)	73.3 (C-4_r)

Spectra obtained in CD_3OD.

REFERENCES

1. Bailey, L. H., *Manual of Cultivated Plants*, Macmillan, New York, 1977, p. 1027.
2. Monya, M. and Racz, G., *Plant Med. et Phytother.*, 1974, **8**, 126.
3. Kaij-a-Kamb, M., Amoros, M. and Girre, L., *Pharm. Acta Helv.*, 1992, **67**, 178.
4. Sarker, S. D., Savchenko, T., Whiting, P., Šik, V. and Dinan, L. N., *Nat. Prod. Letts.*, 1997, **9**, 189.
5. Sarker, S. D., Savchenko, T., Whiting, P., Šik, V., Lafont, R. and Dinan, L., *Biochem. Syst. Ecol.*, 1997, **25**, 367.
6. Sarker, S. D., Girault, J.-P., Lafont, R. and Dinan, L., *Int. J. Pharmacognosy*, (in press).
7. Homer, J. and Perry, M. C., *J. Chem. Soc., Chem. Commun.*, 1994, 373.
8. Bax, A. and Summers, M. F., *J. Am. Chem. Soc.*, 1986, **108**, 2093.
9. Lee, H. and Wolf, M. E., *J. Org. Chem.*, 1967, **32**, 192.
10. Schildknecht, H., Tachei, H. and Maschwitz, U., *Die Naturwissenschaften*, 1969, **56**, 37.
11. Palter, R., Lundin, R. E. and Fuller, G., *Phytochemistry*, 1972, **11**, 822.
12. Geuns, J. M. C., *Phytochemistry*, 1978, **17**, 1.
13. Clément, C. Y., Bradbrook, D. A., Lafont, R. and Dinan, L., *Insect Biochem. Mol. Biol.*, 1993, **23**, 187.