Phytochemistry, Vol 27, No 11, pp 3541–3544, 1988
Printed in Great Britain

0031-9422/88 $3.00+0.00

STERONES, IRIDOIDS AND A SESQUITERPENE FROM *VERBASCUM THAPSUS*

M. A. KHUROO, M. A. QURESHI, T. K. RAZDAN and P. NICHOLS*

Centre of Research for Development, University of Kashmir, Naseem Bagh, Srinagar 190006, India; *Marine Laboratories CSIRO, GPO 1538 Hobart, Australia

(*Received* 15 *February* 1988)

Key Word Index—*Verbascum thapsus*, Scrophulariaceae, sterones, cholestanones; iridoid glycosides, sesquiterpene acid

Abstract—An unusual sterone distribution has been found in *V. thapsus*. Also two iridoid glycosides and a sesquiterpene acid have been isolated and characterized from the plant.

INTRODUCTION

The herb *Verbascum thapsus* is widely distributed in temperate Himalaya (altitude 5000–12 000 ft) [1]. It has been employed for the treatment of asthma and other pulmonary complaints. The leaves of the plant are considered to be anti-endemic, and the seeds are reported to be aphrodiasic and narcotic in nature [2]. The plant was shown to elaborate oleanane type triterpenoid saponins [3, 4] and iridoid glycosides [5, 6]. The plant lipids are reported [7] to be comprised of sitosterol, ergostan-7-en-3-ol and palmetic, stearic, oleic, linoleic, orchidiac and benzenic acids [7]. This paper, besides describing the isolation and characterisation of two iridoid glycosides and a sesquiterpene acid reveals an unusual distribution of sterones in *V thapsus*

RESULTS AND DISCUSSION

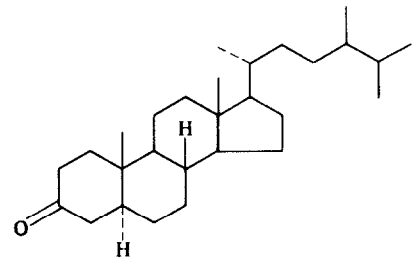
The ethanol extract of *V thapsus* on chromatography over silica gel afforded a mixture which on further chromatography gave a homogeneous fraction which responded positively to the Liebermann–Burchard, TNM and Zimmermann tests for 3-ketosterones [8]. The mass spectrum of these fractions revealed that it was a mixture of three compounds. The mixture defied all efforts for its separation by traditional chromatography, including argentation chromatography. The GC-MS analysis showed that the fraction was actually a composite of seven compounds (Table 1) with close retention times. From the molecular ions (Table 1) and fragmentation patterns the compounds of this fraction were shown to be C_{28} and C_{29}-sterones. The order of elution was deduced by comparison of mobilities of 5α- and 5β-sterols having one or two double bonds at different positions because the same order of mobilities would be expected for the corresponding sterones The mobility order of the sterones was consistent with GC-mobilities of the sterols from the *Curcurbitaceae* family [9].

The mass spectra revealed that the sterones differed by the presence of one or two double bonds and that the plant contained isomeric or epimeric sterones. The sterone composition involved three isomeric sterones with $[M]^+$ at *m/z* 412, other sterones exhibited molecular ion, peaks at *m/z* 400, 410 and 414. The sterones could be differentiated by virtue of their different base peak fragments as well as the intensity of the other prominent peaks in their mass spectra The mass spectral fragmentation of the sterones is well studied [10–13].

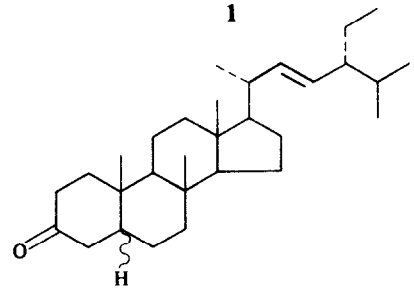
The sterone **1** had $[M]^+$ at *m/z* 400, $C_{28}H_{48}O$, and exhibited a base peak at *m/z* 231, characteristic of 5α-cholestanones [13], the base peak arises by loss of the side chain and $[C_3H_6]^+$ ion from the D-ring. This is characteristic of sterones in which the charge is retained on the oxygen carrying ion. Other prominent peaks are observed at *m/z* 217, 163 and 207 originating from the base peak ion fragment. The mass spectral fragmentation was consistent with the library sample of 24α-methyl-5α-cholestan-3-one

Table 1. Composition of the sterones obtained from *V. thapsus*

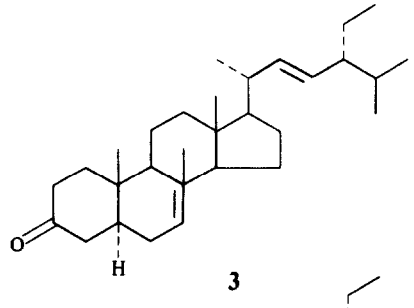
Component	% Composition	(min)	$[M]^+$ at *m/z*
1	2.1	14.36	400
2	24.0	14.73	412
3	42.3	15.06	410
4	17.0	15.32	414
5	3.1	15.43	414
6	9.3	15.63	412
7	2.2	15.75	412



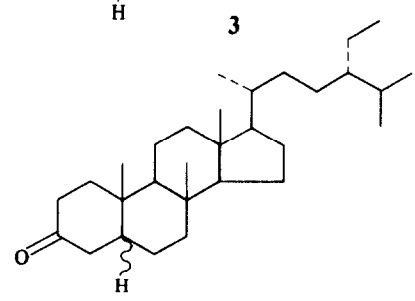
1



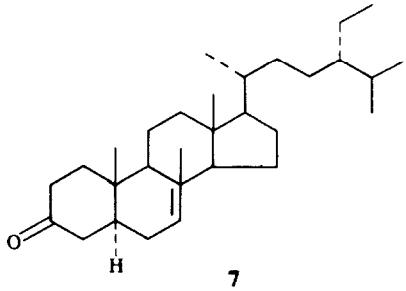
2 H = 5α
6 H = 5β



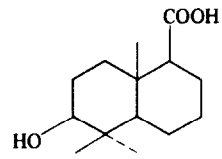
3



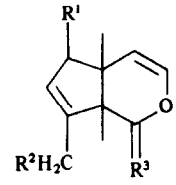
4 H = 5α
5 H = 5β



7



8



9 R^1 = R^2 = OH, R^3 = O-Glu
10 R^3 = OH, R^1 = O-β-Xyl, R^2 = O-Glu

The sterones **2**, **6** and **7** showed molecular ion peaks at *m/z* 412 corresponding to $C_{29}H_{48}O$. The mass spectra of these compounds showed striking differences. The base peak fragment at *m/z* 271 in the mass spectrum of sterone **2** indicated the presence of a double bond in the side chain at C-22 and the presence of a 24-ethyl group. The prominent peaks at *m/z* 229 resulted from the loss of the side chain and D-ring carbons The absence of the $[M-70]^+$ ion in the mass spectrum showed that the sterone was a 5α-cholestanone derivative The rest of the mass

fragmentation features resembled those of compound **1** The sterone was thus identified 24-ξ-ethyl-5α-cholestan-22-en-3-one. The compound **6** exhibited a base peak at *m/z* 55 in its mass spectrum. The prominent peaks at *m/z* 271 and 273 were indicative of the presence of a Δ^{22}-double bond also indicated by the prominent peak at *m/z* 300 and the characteristic peak at *m/z* 369 $[M-C_3H_7]^+$. Unlike compound **2** the mass spectrum contained well defined peaks at *m/z* 300, 318, 369, 258, 231 and 217 showing that the compound was isomeric with compound **2**. In Δ^{22}-cholestenones the base peak fragment at *m/z* 55 is generally observed with 5β-derivatives; these compounds also exhibit a strong peak at *m/z* 259 Thus the compound **6** was found to be 24-ξ-ethyl-5β-cholestan-22-en-3-one. The sterone **7**, showed a base peak at *m/z* 269 and its mass spectral pattern did not coincide with those of the sterones **2** or **6** Its spectrum contained prominent peaks at *m/z* 397 $[M-Me]^+$ and 312 $[M-Me-C_6H_{13}]^+$ and a base peak at *m/z* 269 arising by the loss of the side chain with two hydrogens from the ring. This feature is generally observed with the steroids carrying a Δ^7-nuclear double bond. The mass spectrum of the compound contained other prominent peaks at *m/z* 244, 229, 185, 122, 129 and 55. The sterone was thus characterized as 24-ξ-ethyl-5α-cholestan 7-en-3-one.

The sterone **3**, $[M]^+$ at *m/z* 410, $C_{29}H_{46}O$, contained two double bonds and was shown to be a 5α-cholestan-one derivative by the lack of a $[M-70]^+$ peak in its mass spectrum The spectrum contained prominent peaks at *m/z* 367 $[M-C_3H_7]^+$ and 298 $[M-C_8H_{16}]^+$ The base peak at *m/z* 269 corresponded to the loss of the side chain with simultaneous transfer of two hydrogens from the nucleus. Other prominent peaks were observed at *m/z* 244, 229, 214, 160, 146, 123 and 55. This fragmentation had the features resembling $\Delta^{7,22}$-diene sterols. The sterone was thus shown to be 24α-ethyl-5α-cholestan-$\Delta^{7,22}$-dien-3-one.

The sterones **4** and **5** were epimeric showing $[M]^+$ at *m/z* 414 corresponding to the molecular formula $C_{29}H_{50}O$. They were thus shown to be saturated sterones. The sterone **4** lacked a prominent peak at $[M-70]^+$ which was present in the mass spectrum of sterone **5** This suggested that the compound **4** was the 5α-epimer while the sterone **5** was the 5β-epimer. Both the sterones showed a prominent peak at *m/z* 273 due to the loss of side chain and D-ring carbons Other mass spectral features resemble 24-ξ-ethyl-cholestanone. The epimeric compounds were thus characterized as 24-ξ-ethyl-5α-cholestan-3-one and 24-ξ-ethyl-5β-cholestan-3-one.

The compound **8** $[M]^+$ at *m/z* 254, $C_{15}H_{26}O_3$ in its ^{1}H NMR spectrum revealed the presence of two methyls at δ 1.10 (6H, *s*) and one downfield methyl at δ 1.65 (3H, *s*) This is consistent with the methyl shifts of dammarane sesquiterpenoids The compound was shown to carry a carbomethoxy function by the presence of a resonance signal at δ 3.25 (3H, *s*). A downfield signal at δ 4.52 (1H, *d*, $J=3$ 6 Hz) was assigned to the proton geminal to the carbomethoxy group. The third oxygen function was presumed to be in the form of a hydroxyl whose resonance signal was submerged under the solvent peak. A single proton resonance signal at δ 4 05 (1H, *br s*) showed that the hydroxyl was secondary in nature. The molecular ion peak at *m/z* 254 in the mass spectrum of the compound was also the base peak of the compound. Additional fragmentation peaks were observed at *m/z* 239 $[M-Me]^+$, 236 $[M-Me-H_2O]^+$ 234, 220, 219, 154 and 95. The spectral data was in agreement with 3α-hydroxy-drimmanyl-8-methanoate. The compounds **9** and **10** were characterised as aucubin and 6-*O*-β-xyloxyl aucubin on the basis of their mp, IR and ^{1}H NMR data which was identical with the reported data [3] for these compounds.

The majority of higher plants contain 24α-alkyl-Δ^5-sterols. Thus the sterols with a saturated side chain mostly have 24*R*-alkyl groups while the sterols with a Δ^{22}-double bond have usually the 24*S*-alkyl configuration. The Δ^{22}-sterols have the *trans*- or *E*-configuration. On biogenetic grounds similar configurations are believed to exist with the sterones of *V. thapsus*. We have failed to isolate any sterol from this plant The sterones and sesquiterpenes of the plant have not been dealt with so far.

EXPERIMENTAL

Mps uncorr IR spectra were recorded in KBr discs, ^{1}H NMR were run at 60, 90 and 250 MHz, and MS was done at 70 eV.

Extraction and isolation The plant material was obtained from the University Campus, shade-dried, ground and defatted with petrol. The defatted plant material was extracted with EtOH in a Soxhlet and the resulting extract, after drying *in vacuo*, was chromatographed over silica gel using graded solvent systems The development of the column with petrol–$CHCl_3$ (1 1), $CHCl_3$–EtOAc (1 2) and MeOH–EtOAc (4 1) gave the fractions 1, 2 and 3 respectively These fractions were subjected to further chromatography over silica gel for isolation of the compounds and purity was checked with TLC on silica gel G

Identification of sterones Fraction 1 on repeated chromatography gave a TLC homogeneous fraction Its MS showed the fraction to be a mixture which failed resolution by traditional chromatography The mixture was not resolved on a 50 meter BP-1 non-polar capillary column Subsequently, the Biller–Biemann programme showed that the components of the mixture had close retention times The resolution of the mixture was achieved on a GC-4CM instrument with a SCOT OV-17 glass capillary column (30 m × 3 mm internal dia, temp 260°) and by GC-MS, using a Schimadzu LKB-9000 instrument fitted with a 2% OV-17 glass capillary column (12 m × 3 mm internal dia) The order of retention times was determined on the same instrument using standard sterols The R_t values of the seven components are given in Table 1.

Identification of compound **8** $[M]^+$ at *m/z* 254, $C_{15}H_{26}O_3$, IR ν_{max}^{KBr} cm^{-1}. 3450 (OH), 1730 (COOMe), 1370, 1360 (*gem* dimethyl) ^{1}H NMR (DMSO-d_6) δ 1.10 (6H, *s*, H-12 and H-13), 1.65 (4H, *br s*, H-11, H-21, 1.70 (1H, *d*, $J=8$ Hz, H-2), 1 83 (4H, *m*, H-6, H-7), 3.25 (3H, *s*, COOMe), 4 09 (1H, *t*, $J=7$, 3 Hz, H-3), 4.52 (1H, *d*, $J=8$ Hz, H-8), MS *m/z* 254 $[M]^+$ (100), 239 $[M-Me]^+$, 236 $[M-H_2O]^+$, 220, 191, 154, 95

Identification of iridoids **9** and **10**. Compound **9** and **10** responded positively to vanillin and Molischs reagent Compound **9**, mp 185–187°, $[M]^+$ at *m/z* 346, $C_{15}H_{22}O_7$, UV λ_{max}^{MeOH} nm 204; IR ν_{max}^{KBr} cm^{-1}: 3400 (OH), 2920, 1640, 1075 and 1045. ^{1}H NMR δ 2.84 (1H, *br s*, H-5), 3.66 (1H, *pt*, H-9), 4 34 (1H, *br s*, H-10), 4 60 (1H, *br s*, H-6), 5 17 (1H, *dd*, $J=6$, 3 7 Hz, H-4), 5.28 (1H, *d*, $J=7$ 5 Hz, H-1), 5.89 (1H, *br s*, H-7) On hydrolysis with 1N H_2SO_4, the compound **9** gave D-glucose and an aglycone, $[M]^+$ at *m/z* 184, $C_9H_{12}O_4$, ^{1}H NMR· δ 2 35 (1H, *br s*, H-5), 2.80 (1H, *pt*, H-9), 4 35 (1H, *br s*, H-10), 4 70 (1H, *br s*, H-6), 4.82 (1H, *d*, $J=7$ Hz, H-1), 4.96 (1H, *dd*, $J=6$ 5, 2.8 Hz, H-4), 5 84 (1H, *br s*, H-3), 5 85 (1H, *br s*, H-7)

On acetylation with Ac_2O–pyridine, compound **9** gave compound **10** $[M]^+$ at *m/z* 310, $C_{15}H_{18}O_7$ IR ν_{max}^{KBr} cm^{-1} 1720

^{1}H NMR ($CDCl_3$) δ2 95 (1H, *br s*, H-5), 3 17 (1H, *pt*, H-9), 4 74 (1H, *br s*, H-10), 4 98 (1H, *dd*, J = 6 3, 3 3 Hz, H-4), 5.33 (1H, *br s*, H-6), 5 90 (1H, *br s*, H-7), 6.04 (1H, *d*, J = 4 5 Hz, H-1), 6 23 (1H, *dd*, J = 5 7, 2 02 Hz, H-3)

Compound **10**, mp 192–193 , $[M]^+$ at m/z 478, $C_{20}H_{30}O_{13}$, UV λ_{max}^{MeOH} nm 208, IR ν_{max}^{KBr} cm^{-1} 3400 (OH), 2900, 2860, 1655, 1385, 1245, 1165, 1050, 970 ^{1}H NMR δ3 2–2 9 (2H, *m*, H-8 and H-9), 4 40 (2H, *br s*, H-10, 4 58 (1H, *d*, J = 7 4 Hz) and 4 80 (1H, *d*, J = 7 5 Hz), anomeric protons, 4 63 (1H, *br s*, H-6), 5 20 (1H, *dd*, J = 6 0, 3 7 Hz, H-4), 5 29 (1H, *d*, J = 5 Hz, H-1), 5 99 (1H, *br s*, H-7) and 6 40 (1H, *dd*, J = 6 01, 5 1 Hz, H-3)

On hydrolysis with 1N H_2SO_4, **10** gave D-glucose and D-xylose (identified by PC with an authentic sample) and an aglycone whose ^{1}H NMR spectrum was superimposable with the aglycone of the compound **9**

Acknowledgements—The authors are grateful to Dr K L Dhar and Dr S K Koul (RRL, Jammu) for spectral facilities and to the Director CORD, University of Kashmir, Srinagar, for laboratory facilities

REFERENCES

1 Dhar, U (1982) *Alpine Flora of Himalayas*
2 Chopra, R N, Nayar, S L and Chopra, I C (1956) *Glossary of Indian Medicinal Plants*, p 253 CSIR, New-Delhi
3 Degot, A V *Farm Zh (Kiev) Ukrain* **27**, 66
4 De Pascual, T, Diaz, F and Grande, M (1980) *An Quim Ser C* 107
5 Swiatek, L (1984) *Farm Pol* **40**, 415
6 Oswath, K (1982) *Herba Hung* **21**, 141
7 De Pascual, T, Diaz, F and Grande, M (1978) *An Quim* **74**, 566
8 Razdan, T K, Kachroo, V, Harkar, S and Koul, G L (1982) *Tetrahedron* **38**, 991
9 Matsumoto, T (1983) *Phytochemistry* **22**, 183
10 Wyllie, S G and Dgerassi, C (1968) *J Org Chem* **33**, 305
11 Williams, D H, Wilson, J M, Budzikiewicz, H and Djerassi, C (1963) *J Am Chem Soc* **85**, 2091
12 Garg, V K and Nes, W R (1984) *Phytochemistry* **23**, 1919
13 Budzikiewicz, H and Djerassi, C (1962) *J Am Chem Soc* **84**, 1430