Phytochemistry, Vol 27, No 4, pp 1181 1183, 1988
Printed in Great Britain

0031 9422/88 $3 00+0 00

FURANOCOUMARIN GLUCOSIDES FROM THE SEEDS OF *APIUM GRAVEOLENS*

VINOD K AHLUWALIA, DEREK R BOYD,* ANIL K JAIN, C H KHANDURI and NARAIN D SHARMA

Department of Chemistry, University of Delhi, Delhi 110007, India, *Department of Chemistry, Queen's University of Belfast, Belfast BT9 5AG, U K

(*Received* 17 *August* 1987)

Key Word Index—*Apium graveolens*, Umbelliferae, coumarin glucosides, structural determination

Abstract—Besides celereoside and nodakenin, three new furanocoumarin glucosides have been isolated from the seeds of *Apium graveolens* The new glucosides have been structurally assigned as (+)-2,3,-dihydro-9-hydroxy-2[1-(6-sinapinoyl) β-D-glucosyloxy-1-methylethyl]-7H-furo[3,2*g*] [1]-benzopyran-7-one, (−)-2,3-dihydro-9-*O*-β-D-glucosyloxy-2-isopropenyl-7H-furo[3,2*g*] [1]-benzopyran-7-one, and 5-methoxy-8-*O*-β-D-glucosyloxypsoralen.

INTRODUCTION

In previous phytochemical investigations of *Apium graveolens* L seeds [1–5], we have reported the isolation of both new and known furanocoumarins including two glucosides We now report on the isolation of three new furanocoumarin glucosides (**1**, **3** and **5**) from the seeds.

RESULTS AND DISCUSSION

The concentrated water-soluble portion of the ethanolic extract of the seeds on CC followed by repeated partial crystallizations of the evaporated residues of selected CC fractions yielded five compounds All of the five compounds gave a positive Molisch's test for sugars and showed UV absorption characteristics of linear furanocoumarins [6].

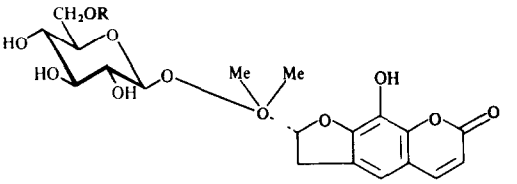
Compound **1**, $[\alpha]_D^{20}+75°$, showed a positive ferric reaction and also gave a positive Gibb's test indicative of an unsubstituted position *para* to a hydroxyl group in the molecule From the ^{1}H NMR spectrum of **1**, the presence of either the rutaretin [1] or the isomeric celereoin unit [5] and a symmetrically trisubstituted styryl unit in compounds **1** were clearly discernible. Two prominent doublets (1H each) at δ6 39 and 7 6 with a large coupling constant ($J=15.9$ Hz) strongly suggested a *trans* orientation for the two olefinic protons of the styryl unit

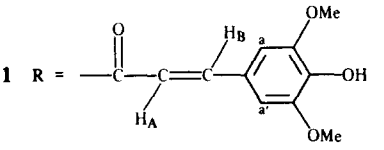
The appearance of another doublet at δ4 7 ($J=7.8$ Hz) assignable to the anomeric proton indicated that the sugar moiety was linked to the coumarin/styryl unit through its C-1 position Moreover, in the ^{1}H NMR spectrum of the pentaacetate (**1a**) of **1**, the presence of two signals integrating in total for nine protons only in the aliphatic acetoxyl region (δ2 01, 2.08) was consistent with the existence of a disubstituted sugar unit in **1**.

Additionally, in the ^{1}H NMR spectrum of **1** the methylene protons on the C-6 position of the sugar group were shifted downfield from their normal value by δ0 4 thereby identifying the C-6 position as the other possible site for linkage [7]

The mass spectrum of the dimethyl ether derivative (**1b**) of **1** (formed by treatment with ethereal CH_2N_2) showed the M^+ peak at *m/z* 658. Accurate mass measurements on the M^+ peak established the composition $C_{33}H_{38}O_{14}$ for **1b**. A further peak at *m/z* 438 was assigned to the glycoside residue left after fission of the acid fragment while the peak at *m/z* 259 arose due to subsequent loss of the glycosyl group. The base peak at *m/z* 217 was derived from the aglycone (*m/z* 276) by the principal fragmentation pathway of hydroxyisopropyldihydrofuranocoumarins [6].

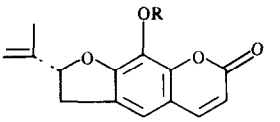
Rutaretin [1] was isolated from the ethyl acetate extract of the acid hydrolysate of **1**, and D-glycose was detected in the aq phase. Graded hydrolysis with barium hydroxide produced isorutarin (**2**) [8] and 4-hydroxy-3,5-dimethoxycinnamic acid (sinapinic acid). The latter was identified by GC-MS comparison of its volatile di-TMSi derivative with an authentic sample Formation of 2,3,4-

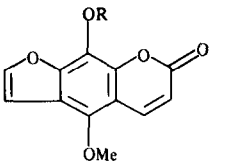




1 R =

2 R = H





3 R = β-D-Glcpyr
4 R = H

5 R = β-D-Glcpyr
6 R = H

tri-*O*-methyl-D-glucopyranose on methanolysis of the permethylate of **1** [9] established the (1→6) linkage of glucose with rutaretin and sinapinic acid respectively The β-linkage of the glucose unit was confirmed by enzymatic hydrolysis Thus compound **1** was identified as (+)-2,3-dihydro-9-hydroxy-2[1-(6-sinapinoyl)-β-D-glucosyloxy-1-methylethyl]7H-furo-[3,2*g*] [1]-benzopyran-7-one

Compound **3**, $C_{20}H_{22}O_9$, $[\alpha]_D^{20}-60$, gave negative ferric chloride and Gibb's tests and did not form any methyl derivative on treatment with ethereal diazomethane The ^{1}H NMR spectrum of **3** revealed the presence of an apiumetin nucleus (**4**) as the major component Acetylation of **3** gave the tetraacetate **3a**, $C_{28}H_{30}O_{13}$, the ^{1}H NMR spectrum of which showed that only aliphatic acetoxyl groups are present Enzymatic hydrolysis of compound **3** gave apiumetin (**4**) [1] but acid hydrolysis with 7% refluxing H_2SO_4 seemed to decompose the aglycone and only D-glucose could be identified in the aq hydrolysate Compound **3** did not show the molecular ion peak on EI mass spectrometric analysis presumably because of its thermal degradation before ionization [10] The mass spectrum was quite similar to that of the aglycone. Based on the above data the structure of compound **3** was established as (−)-2,3-dihydro-9-*O*-β-D-glucosyloxy-2-isopropenyl-7H-furo[3,2*g*] [1]-benzopyran-7-one

Compound **5**, gave negative ferric chloride and Gibb's tests did not yield any methyl derivative on treatment with ethereal diazomethane, but readily formed a tetraacetate (**5a**), $C_{26}H_{26}O_{14}$, on reaction with acetic anhydride in pyridine The ^{1}H NMR spectrum of acetate **5a** showed three upfield signals at δ2 00, 2 07, 2 25 and a downfield signal at δ4 22 assignable to four aliphatic acetoxyl groups and to a methoxyl group respectively In addition the NMR spectrum contained a set of signals attributable to a 5,8-disubstituted psoralen nucleus Acid hydrolysis of compound **5** produced 8-hydroxy-5-methoxypsoralen (**6**) [2] and D-glucose Finally, the β-linkage of glucose was confirmed by enzymatic hydrolysis Compound **5** was thus assigned the structure 5-methoxy-8-*O*-β-D-glucosyloxypsoralen

The other two glycosides were identified as celeroside [3] and nodakenin [11] This is the first report of the occurrence of nodakenin in *A graveolens*

EXPERIMENTAL

Mps uncorr, ^{1}H NMR 250 MHz, TMS as internal standard, A VG analytical 12–250 mass spectrometer system was used for mass spectra of TMSi derivatives introduced from a directly coupled Hewlett Packard Model 5790 gas chromatograph equipped with 25 m capillary column (BP1) 180→230° at 5°/min MS (70 eV) Accurate mass measurement was obtained using the peak matching method TLC and CC silica gel using the solvent system (i) EtOH–MeOH–H_2O (100 16 5 13 5), (ii) $CHCl_3$–MeOH (3 1), (iii) $CHCl_3$–MeOH (1 1), (iv) C_6H_6–Me_2CO (3 1), (v) $CHCl_3$–MeOH (9 1), (vi) C_6H_6–Me_2CO (9 1)

Isolation Air-dried *Apium graveolens* seeds (3 5 kg) were extracted successively with petrol (bp 60–80°), Et_2O and 98% EtOH under reflux The EtOH extract was concd under red pressure, the residue was suspended in H_2O and extracted with EtOAc using a liquid–liquid extractor The H_2O soluble portion was again concentrated under reduced pressure and the gummy residue (*ca* 150 g) was chromatographed on silica gel (1.5 kg) using a $CHCl_3$→50% $CHCl_3$–MeOH gradient Fractions eluted with $CHCl_3$–MeOH (9 1) were pooled together and rechromatographed on silica gel using a $CHCl_3$→10% MeOH–$CHCl_3$ gradient Fractions eluted with 5 10%, MeOH–$CHCl_3$ were combined and the solvent evapd The residue was dissolved in warm EtOH and left for several days at ambient temperature (*ca* 28°) for crystallization Four crops of crystals were collected TLC examination (solvent i) of these showed that the first two crops were richer in compound **1** while the subsequent two had compound **3** as their major component Recrystallization from EtOH of the combined first and second crop of crystals yielded pure compound **1** Similar recrystallization of the combined third and fourth crops of crystals gave pure compound **3**

The earlier fractions eluted with $CHCl_3$ MeOH (17 3) on repeated CC followed by partial crystallization of the purified material ($CHCl_3$–MeOH) yielded compound **5** and nodakenin [11] From the latter fractions celeroside [3] was isolated by careful crystallization from EtOAc Compounds **1**, **3** and **5** were obtained in *ca* 5–10 mg quantities

Compound **1** Recrystallization from EtOH yielded compound **1** as colourless shiny plates, mp 166 68°, $[\alpha]_D^{20}+75$ [MeOH, *c* 0 38], R_f 0 58 (system i), 0 45 (system ii) UV λ_{max}^{MeOH} nm (log ε) 246 sh (4 06), 265 (3 73), 330 (4 35) IR ν_{max}^{KBr} cm^{-1} 3580, 1705, 1630, 1586, 1425 and 1158, ^{1}H NMR [$(CD_3)_2CO$] δ1 35 and 1 37 (s, 3H each, *gem* dimethyl), 3 00–3 92 (*m*, 6H, 4 glucose protons and Ar–C$\underline{H}_2$ CH<), 3 87 (s, 6H, 2× –OMe) 4 20 (*m*, 2H, H-6 of glucose), 4 70 (*d*, 1H, *J* = 7 8 Hz anomeric H), 4 86 (*m*, 1H, Ar–CH_2–C$\underline{H}$<), 6 1 (*d*, 1H, *J* = 9 5 Hz, H-6), 6 39 (*d*, 1H, *J* = 15 9 Hz, H_A), 6 92 (s, 1H, H-4), 6 96 (s, 2H, $H_a$$H_a$), 7 6 (*d*, 1H, *J* = 15 9 Hz, H_B) and 7 71 (*d* 1H, *J* = 9 5 Hz, H-5)

Methylation of **1** by dropwise addition of ethereal CH_2N_2 to its methanolic soln yielded the dimethyl ether **1b**, which crystallized from $CHCl_3$–*n*-hexane as colourless needles, mp 145°, $[\alpha]_D^{20}-38$ 1 ($CHCl_3$, *c* 0 47), (Found M$^+$ 658 2257 $C_{33}H_{38}O_{14}$ requires M 658 22612) UV λ_{max}^{MeOH} nm 276 314 The tetraacetate **1a** (prepared *via* Ac_2O-pyridine) recrystallized from $CHCl_3$–MeOH as white powder, mp 139–140°, $[\alpha]_D$ −25 [$CHCl_3$, *c* 0 77] UV λ_{max}^{MeOH} nm (log ε) 240 sh (3 96), 258 (3 84) 298 (4 36), 310 sh (4 33), ^{1}H NMR ($CDCl_3$) δ1.25 and 1 34 (s, 3H each, *gem* dimethyl), 2 01 (s, 6H, 2× –OCOMe), 2 08 (s, 3H, –OCOMe), 2 35 (s, 3H, –OCOMe), 2 45 (s, 3H –OCOMe) 3 35 (*m*, 2H, Ar–C$\underline{H}_2$–CH<), 3 57–3 98 (*m*, 3H glucose protons), 3 86 (s, 6H, 2× –OMe), 4 88 (*d*, 1H *J* = 8 0 Hz, anomeric H) 4 74–5 27 (*m*, 4H, 3 glucose protons and Ar–CH_2 C$\underline{H}$<) 6 21 (*d*, 1H, *J* = 9 6 Hz H-6), 6 33 (*d*, 1H *J* = 16 0 Hz, H_A), 6 77 (s, 2H, H_aH_a), 7 11 (s, 1H, H-4), 7 56 (*d*, 1H, *J* = 16 0 Hz, H_B) and 7 59 (*d*, 1H, *J* = 9 6 Hz, H-5)

1 (6 mg) was hydrolysed with H_2SO_4 (7%, w/v) for 2 hr under reflux The soln was thoroughly extracted with EtOAc and the residue left after evapn of the solvent was purified by prep TLC (system v) The purified aglycone, crystallized from EtOAc–*n*-hexane, was found to be identical with rutaretin [1] (mp co-TLC, UV and co-IR) D-Glucose was identified by PC in the neutralized aq phase

Compound **1** (2 mg) was hydrolysed with 0 1 N $Ba(OH)_2$ soln at ambient temp The soln was carefully neutralized with cold dil H_2SO_4, filtered and the aq filtrate extracted with Et_2O The $BaSO_4$ ppt was washed with Et_2O and the combined extract evapd The residue obtained was shown to contain sinapinic acid by GC-MS analysis of its di-TMSi derivative, prepared by treatment with *N*-*O*-bis(trimethylsilyl)acetamide in pyridine (2 1) An authentic sample of di-TMSi derivative of sinapinic acid was also analysed by GC-MS for comparison The gas chromatograms of both the samples showed a prominent peak corresponding to a major component which eluted at *ca* 16 min

post-injection The mass spectrum of this component (*m/z* 368) of each sample, which corresponded to the incorporation of two Me_3Si groups in the acid, was indistinguishable from the other

Isorutarin [8] was identified in the residue obtained after lyophilization of the extracted aq hydrolysate by its UV spectrum, λ_{max}^{MeOH}, 265, 335 nm Permethylation of compound **1** by Hakomari's method [9] and subsequent methanolysis of the product with MeOH–HCl (5%) gave 2,3,4-tri-*O*-methyl-D-glucopyranose which was identified by direct comparison with an authentic sample [TLC (system IV) and PC (*n*-BuOH–EtOH–H_2O, 5 1 4)] The β-configuration of the glucose linkage was established by the hydrolysis of **1** with emulsin

Compound **3** crystallized from EtOH as colourless needles, mp 165°, $[\alpha]_D^{20} - 60°$ [$CHCl_3$, *c* 0 16], R_f 0 60 (system I), 0 45 (system II) UV λ_{max}^{MeOH} (log ε) 250 (3 41), 262 (3 35), 333 (3 87, IR ν_{max}^{KBr} cm^{-1} 3400, 1725, 1600, 1440, 1140 and 860 (Found C, 59 0, H 5 2 $C_{20}H_{22}O_9$ requires C 59 1, H, 5.4%). ^{1}H NMR (DMSO-d_6 + D_2O) δ1 74 (*s*, 3H, –Me), 4 94 and 5 12 (*s*, 1H each, $=CH_2$), 5.38 (*m*, 1H, Ar–CH_2–$CH<$), 3 10–5 35 (7 glucose protons and Ar–CH_2–$CH<$), 6 26 (*d*, 1H, *J* = 9 5 Hz, H-6), 7 23 (*s*, 1H, H-4), 7 94 (*d*, 1H, *J* = 9 5 Hz, H-5)

The acetate **3a** (prepared via Ac_2O–pyridine) recrystallized from aq EtOH as colourless needles, mp 139–142° (Found C, 58 3, H, 5 5 $C_{28}H_{30}O_{13}$ requires C, 58 5, H, 5 2%) ^{1}H NMR ($CDCl_3$) δ1 74 (*s*, 3H, Me) 2 02 (*s*, 9H, 3 × –OCOMe), 2 18 (*s*, 3H, –OCOMe), 3.38 (*m*, 2H, Ar-CH_2–$CH<$), 3 42–4 50 (*m*, 3H, glucose protons), 4 75–5 30 (*m*, 5H, 4 glucose protons and Ar–CH_2–$CH<$), 4 88 and 5 08 (*s*, 1H each $=CH_2$), 6 18 (*d*, 1H, *J* = 10 Hz, H-6), 7 02 (*s*, 1H, H-4) and 7 54 (*d*, 1H, *J* = 10 Hz, H-5) Compound **3** (3 mg) was hydrolysed with emulsin in Pi buffer (0 02 M, pH 7) at ambient temp The aglycone isolated from the EtOAc extract of the enzyme digest was identified as apiumatin (**4**) [1] by direct comparison with an authentic sample (mmp, TLC, UV and co-IR). Compound **3** (2 mg) was hydrolysed with H_2SO_4 (7%, w/v) for 1 hr under reflux. The hydrolysate was extracted with EtOAc and D-glucose was identified by PC in the neutralized aq extract

Compound **5** Recrystallization from $CHCl_3$–MeOH yielded compound **5** as colourless needles mp 230°, R_f 0.42 (system I), 0 57 (system III), UV λ_{max}^{MeOH} nm (log ε) 241 (4.04), 266 (4.13), 271 (4 17) and 313 (4 04), IR ν_{max}^{nujol} cm^{-1} 3375, 1680, 1590, 1450, 1170 and 823. The tetraacetate **5a** (Ac_2O-pyridine) recrystallized from $CHCl_3$–C_6H_6 as colourless needles mp 129° (Found· C 55 9, H 4 4 $C_{26}H_{26}O_{14}$ requires C, 55 5, H, 4 6%) ^{1}H NMR (90 MHz, $CDCl_3$) δ2 00 (*s*, 3H, –OCOMe), 2 07 (*s*, 6H, 2 × –OCOMe), 2 25 (*s*, 3H, –OCOMe), 4 22 (*s*, 3H, –OMe), 3.63–4.20 (*m*, 3H, glucose protons) 4 80–5 20 (*m*, 4H, glucose protons) 6 26 (*d*, 1H, *J* = 10 Hz, H-3), 7 01 (*d*, 1H, *J* = 2 Hz, H-3′ furan), 7 63 (*d*, *J* = 2 Hz, H-2′ furan) and 8 10 (*d*, 1H, *J* = 10 Hz, H-4) Compound **5** (2 mg) was hydrolysed with H_2SO_4 (7%, w/v) for 2 hr under reflux The hydrolysate was thoroughly extracted with EtOAc and the aglycone obtained by evapn of solvent was crystallized, as pale yellow needles mp 218°, from EtOAc–*n*-hexane after purification by TLC (system VI) The aglycone was identified as 8-hydroxy-5-methoxypsoralen [3] by direct comparison with an authentic sample (mmp, UV and co-IR) D-Glucose was identified by PC in the neutralized aq phase The aglycone (**6**) was also obtained by hydrolysis of compound **5** with emulsin

Acknowledgement—The authors are grateful to Dr J M Webber, Department of Chemistry, University of Birmingham, U K for a sample of 2,3,4-tri-*O*-methyl-D-glucopyranose.

REFERENCES

1 Garg, S K, Gupta, S R and Sharma, N. D (1978) *Phytochemistry* **17**, 2135
2 Garg, S K, Gupta, S R and Sharma, N D (1979) *Phytochemistry* **18**, 352
3 Garg, S K, Gupta, S R and Sharma, N. D (1980) *Planta Med* **38**, 363
4 Garg, S K, Sharma, N D and Gupta, S R (1981) *Planta Med* **43**, 306
5 Jain, A K, Sharma, N D, Gupta, S. R and Boyd, D R (1986) *Planta Med* **3**, 246
6 Murray, R D H, Mendez, J and Brown, S A (1982) *The Natural Coumarins*, p 31–49 Wiley, Chichester
7. Shimomura, H, Sashida, Y and Adachi, T (1987) *Phytochemistry* **26**, 249
8 Varga, E, Szendrei, K, Novak, I, Sohar, P and Horvath, Gy (1976) *Fitoterapia* **47**, 107
9. Hakomari, S. (1964) *J Biochem* **55**, 205
10 Prox, A (1968) *Tetrahedron* **24**, 3697
11 Spath, E and Kamrath, P (1936) *Ber Dtsch Chem Ges* **69B**, 2062