Phytochemistry, Vol. 21, No. 5, pp. 1123–1127, 1982.
Printed in Great Britain.

0031-9422/82/051123-05$03.00/0

STRUCTURES OF VERBASCOSIDE AND OROBANCHOSIDE, CAFFEIC ACID SUGAR ESTERS FROM *OROBANCHE RAPUM-GENISTAE*

C. ANDARY, R. WYLDE*, C. LAFFITE*, G. PRIVAT and F. WINTERNITZ*

Laboratoire de Botanique et Cryptogamie, Faculté de Pharmacie, 34000 Montpellier, France; *Ecole Nationale Supérieure de Chimie, 3400 Montpellier, France

(*Received* 4 *June* 1981)

Key Word Index—*Orobanche rapum-genistae*; Orobanchaceae; high field ^{1}H and ^{13}C NMR; caffeic acid sugar esters; 3,4-dihydroxyphenylethanol; 3,4-dihydroxyphenylglycol; glucose; rhamnose; verbascoside; orobanchoside.

Abstract—The complete structural elucidation of the two caffeic acid sugar esters verbascoside and orobanchoside, has been realized by ^{1}H and ^{13}C NMR studies. It has been demonstrated that verbascoside is β-(3′,4′-dihydroxyphenyl)ethyl-*O*-α-L-rhamnopyranosyl(1→3)-β-D-(4-*O*-caffeoyl)-glucopyranoside, and orobanchoside is β-hydroxy-β-(3′,4′-dihydroxyphenyl)-ethyl-*O*-α-L-rhamnopyranosyl(1→2)-β-D-(4-*O*-caffeoyl)-glucopyranoside.

INTRODUCTION

Verbascoside **1** was previously isolated [1] from *Verbascum sinuatum* L. (Schrophulariaceae) but no structural elucidation was given. It is identical [2] with acteoside, a caffeic acid sugar ester from *Syringa vulgaris* (Oleaceae) [3], *Conandron ramoidioides* (Gesneriaceae) [4] and from *Clerodendron myricoides* (Verbenacea) [5]. Since verbascoside has priority over acteoside, the latter name need no longer be used.

We have also found verbascoside to occur in *Orbanche rapum-genistae* Thuill. together with a second glycoside orobanchoside **2**, (the occurrence of which is restricted to the Orobanchaceae [2]). Species of Tubiflorae are particularly rich in caffeic acid esters [6] and recently we isolated verbascoside from other genera of the Acanthaceae, namely *Acanthus*, *Strobilanthes*, *Hemigraphis*, *Ruellia*, *Crossandra* and *Alphelandra* [7].

In this work, verbascoside and orobanchoside were isolated from *Orobanche rapum-genistae*. Preliminary pharmacological screening gave some indications that these two sugar esters act as agonists of the antitremor action of DOPA and as antihypertensive and analgesic agents [2, 8]. These results prompted us to undertake the complete structural elucidation of these two caffeic derivatives.

RESULTS AND DISCUSSION

High resolution ^{1}H and ^{13}C NMR spectra in various solvents were performed throughout this work. Comparative studies with ferulic acid and 1-caffeoyl-glucose were a great help.

Our previous results furnished evidence of the presence in both compounds of a disaccharide unit based on glucose and rhamnose, one caffeic acid molecule and a 3,4-dihydroxyphenylethanol linked to the glucose by an ether function, this aglycone unit being substituted by a secondary alcohol on the carbon of the aromatic side-chain in orobanchoside. Furthermore, the structure of this latter alcohol, which was found for the first time in plants, was proved by a partial synthesis [2, 8]. As can be judged from formulae **1** and **2** there are many questions left open with regard to their detailed structural features.

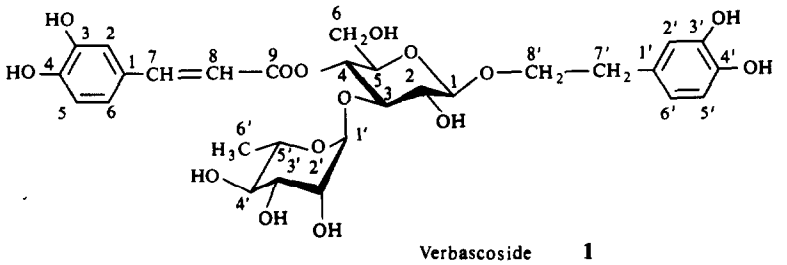
Structure of verbascoside

^{1}H NMR. Aromatic protons are resolved as two ABX systems, one belonging to the caffeic acid substitution, the other to the 3,4-dihydroxyphenylethyl remain (hereafter designated as 'aglycone'). The values summarized in Table 1 are in close agreement with those given by caffeic acid [3] and its derivatives [9, 10].

A triplet at 2.72 ppm contains signals attributable to the α-CH_2 of the aromatic side-chain as compared with phenyl-β-ethanol, but the two protons of the β-CH_2 are non-equivalent, giving rise to quartets at 3.60 and 3.88 ppm. Selective decoupling, e.g. irradiation of the 2.72 resonance, simplifies the two quartets at 3.60 and 3.88 ppm as two doublets ($J = 9$ Hz) and irradiation of the 3.88 ppm quartet modified the 2.72 multiplet.

The anomeric proton H-1′R† and the methyl group of rhamnose were easily recognized: the first at 5.03 ppm with a small coupling constant, the latter at 1.01 ppm as a doublet. The other values (Table 1) are in good agreement only with the α-configuration of

†For the sake of clarity H–R and C–R, H–G and C–G with the suffixed numbering designate the respective protons and carbons of rhamnose and glucose.



Verbascoside 1

Table 1. ^{1}H NMR peaks (ppm) of verbascoside and orobanchoside in DMSO-d_6 + CF_3CO_2H (2 drops) (splitting patterns and J values (Hz) are given in parentheses)

	H	Verbascoside	Orobanchoside
Caffeic acid	2	7.00	7.02
	5	6.74 (*d*, 8)	6.76 (*d*, 8)
	6	6.93 (*d*, 8)	6.94 (*d*, 8)
	7	7.43 (*d*, 16)	7.46 (*d*, 16)
	8	6.18 (*d*, 16)	6.18 (*d*, 16)
Aglycone	2′	6.62	6.75
	5′	6.48 (*d*, 8)	6.60 (*d*, 8)
	6′	6.62 (*d*, 8)	6.68 (*d*, 8)
	7′	2.72 (*m*, 8)	4.57 (*m*, 2.5, 9)
	8′	3.60 (*m*, 8, 9) 3.88 (*m*, 8, 9)	~3.45 (*m*)
Glucose	1	4.33 (*d*, 7.5)	4.55 (*d*, 7.8)
	2	3.22 (*dd*, 7.5, 9)	3.38 (*dd*, 7.8, 9.5)
	3	3.70 (*t*, 9–9.5)	4.05 (*t*, 9.5)
	4	4.72 (*t*, 9.5)	4.92 (*t*, 9.5)
	5	3.45 (*m*)	3.71 (*m*)
	6	~ 3.70, 3.45	3.94, 3.71 (*m*, 12)
Rhamnose	1′	5.03 (*d*, 1)	5.00 (*d*, 1)
	2′	3.70 (*dd*, 1, 2.5)	3.57 (*dd*, 1, 3)
	3′	3.30 (*dd*, 2.5, 9.5)	3.28 (*dd*, 3, 9.5)
	4′	3.12 (*t*, 9.5)	3.12 (*t*, 9.5)
	5′	3.36 (*m*)	3.43 (*m*)
	6′	1.01 (*d*, 6)	1.05 (*d*, 6)

the rhamnopyranose. The coupling constant ($J_{H_1-H_2}$ = 7.5 Hz) of the glucose anomeric proton resonating at 4.33 ppm is in accordance with a β-configuration. The other protons have been identified by selective irradiation. Thus, irradiation of H-1G modifies the triplet at 3.22 ppm, giving the value of H-2G, and irradiation of this latter proton collapses the signal at 3.70 ppm, which is therefore attributed to the H-3G proton. In addition, irradiation of the triplet at 4.72 ppm is reflected by a modification of the H-3G signal and also of the complex signal at 3.45 ppm. Thus, the resonance at 4.72 and 3.45 ppm belongs to H-4G and H-5G. The assignment of H-3G is in agreement with the observed value in β-laminaribiose (β-Glc 1→3 β-Glc) [11]. Furthermore, it is a well-known observation that for glucobioses, only the anomeric proton can be observed above 4 ppm [11]. Thus the higher shift of the 4.72 ppm triplet attributed to H-4G, also reflects the caffeic acid substitution on C-4G. No other effects were observed on the remaining protons by this substitution.

^{13}C *NMR.* Respective values are given in Tables 2–4. In the verbascoside spectrum, we observed the same frequencies as for caffeic acid [12], 1-caffeoylglucose [10] and ferulic acid [12]. In the proton coupled spectrum, the aromatic carbons C-1, C-3 and C-4 are easily recognized by $J_{C-H} = 0$. The signals of C-5 and C-8 are localized as two sharp doublets.

A distinctive difference between **1** and **2** appears clearly as a signal at 36 ppm for the first carbon, in its aglycone side-chain. The aromatic pattern is identical with that of caffeic acid: three singlets for C-1′, C-3′, C-4′ and a fine doublet for C-5′.

Table 2. ^{13}C NMR peaks (ppm) of the caffeoyl moiety of ferulic acid, caffeic acid, 1-caffeoylglycose, orobanchoside and verbascoside (splitting patterns and J values (Hz) are in parentheses)

	Caffeoyl moiety								
	1	2	3	4	5	6	7	8	9
Ferulic acid [12]	127.39	112.28	149.11	149.78	116.89	124.23	146.92	116.22	170.65
Caffeic acid [12]	127.8	114.5	144.1	145.9	116.1	121.4	141.0	121.2	176
1-Caffeoylglucose [10]	128.78	114.93	149.56	150.09	117.42	125.37	146.29	118.20	169.87
Orobanchoside*	127.75	116.14	146.55	149.50	114.90	123.52	148.22	117.23	168.25
Verbascoside*	127.40	115.79	146.35	149.33	114.70	123.36	148.06	117.03	169.5
Orobanchoside†	125.64	114.79	145.69	148.65	113.38	121.61	146.03	115.86	165.55
multiplicity	(*s*)	(*d*, 153)	(*s*)	(*s*)	(*d*, 157)	(*d*, 161)	(*d*, 153)	(*d*, 157)	(*s*)
$^2J_{CH}$ and $^3J_{CH}$ fine structure	(*m*)	(*m*)	(*m*)	(*t*, 5)	(*s*)	(*m*)		(*s*)	
Verbascoside†	125.62	114.70	145.64	148.53	113.68	121.54	145.64	115.86	165.81
multiplicity	(*s*)	(*d*, 155)	(*s*)	(*s*)	(*d*, 161)	(*d*, 161)	(*d*, 157)	(*d*, 153)	(*s*)
$^2J_{CH}$ and $^3J_{CH}$ fine structure	(*t*, 5)	(*t*, 7)	(*m*)	(*t*, 5)	(*s*)	(*tt*)		(*s*)	

*In D_2O + $(CD_3)_2CO$.
†In DMSOd_6 + CF_3CO_2H (2 drops).

Table 3. ^{13}C NMR peaks (ppm) of orobanchoside and verbascoside aglycone and 3-(3',4'-dihydroxyphenyl)-lactic acid (splitting patterns and J values (Hz) are in parentheses)

	Aglycone							
	1'	2'	3'	4'	5'	6'	7'	8'
3-(3',4'-Dihydroxyphenyl)-lactic acid [12]*	130.8	117.2	143.7	142.4	116.1	121.7	39.7	73.5
Orobanchoside*	129.97	114.98	146.04	146.04	117.03	119.17	77.43	72.42
Verbascoside*	131.56	117.54	145.58	144.10	116.84	121.66	35.99	72.07
Orobanchoside†	128.09	113.68	145.26	145.26	115.45	117.29	76.27	71.12
multiplicity	(*s*)	(*d*, 157)	(*s*)	(*s*)	(*d*, 157)	(*d*, 161)	(*d*, 145)	(*t*, 141)
$^2J_{CH}$ and $^3J_{CH}$ fine structure	(*d*, 4)	(*m*)	(*m*)	(*m*)	(*s*)	(*d*, 5)		
verbascoside†	129.31	116.39	145.04	143.58	115.52	119.65	35.07	70.37
multiplicity	(*s*)	(*d*, 161)	(*s*)	(*s*)	(*d*, 155)	(*d*, 160)	(*t*, 125)	(*t*, 145)
$^2J_{CH}$ and $^3J_{CH}$ fine structure	(*d*, 4)	(*m*)	(*d*, 5)	(*t*, 5)	(*s*)	(*d*, 5)		

*In $D_2O + (CD_3)_2CO$.
†In DMSO $d_6 + CF_3CO_2H$ (2 drops).

Table 4. ^{13}C NMR peaks (ppm) of the sugar moiety of verbascoside and orobanchoside (splitting patterns and J values (Hz) are in parentheses)

C	Orobanchoside*	Verbascoside*	Orobanchoside†	Verbascoside†
Glc-1	97.10 (*d*, 161)	102.44 (*d*, 157)	98.63	103.44
2	80.59 (*d*, 141)	74.61 (*d*, 143)	81.39	75.29
3	74.45 (*d*, 143)	79.18 (*d*, 145)	77.70	82.01
4	68.89 (*d*, 145)	69.18 (*d*, 145)	70.36	70.36
5	76.27 (*d*, 145)	74.61 (*d*, 143)	77.94	75.60
6	60.74 (*t*, 141)	60.78 (*t*, 140)	62.01	62.01
Rha-1	100.42 (*d*, 173)	101.25 (*d*, 171)	102.12	102.82
2'	70.52 (*d*, 143)	70.61‡ (*d*, 145)	72.03	71.84‡
3'	70.52 (*d*, 143)	70.52‡ (*d*, 141)	73.03	71.68‡
4'	71.58 (*d*, 141)	71.80 (*d*, 140)	73.47	73.27
5'	68.89 (*d*, 145)	68.82 (*d*, 150)	70.36	70.36
6'	18.01 (*q*, 137)	18.18 (*q*, 127)	18.56	18.60

*In DMSO $d_6 + CF_3CO_2H$ (2 drops).
†In $D_2O + (CD_3)_2CO$.
‡Uncertainty concerning relative assignments.

The $-CH_2-$ carbon (C-8') of the side-chain is clearly apparent between the sugar carbons as a distinctly shaped triplet. It should be observed that C-7' and aromatic carbons values are similar to those of the corresponding carbons in 3-(3',4'-dihydroxyphenyl) lactic acid [12] (Table 3).

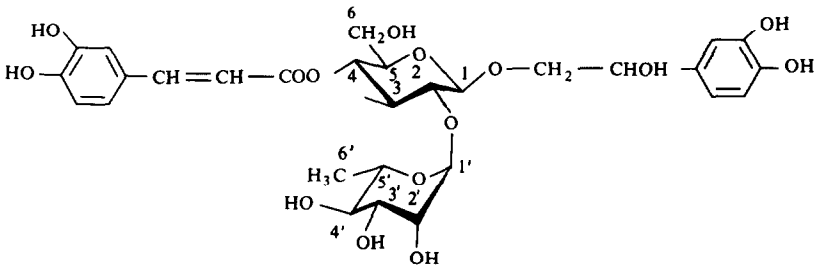
The ^{13}C NMR peaks were assigned on the basis of chemical shift considerations and comparisons with data for glucose and rhamnose. The ambiguity between the two close chemical shifts C-1'R and C-1G was resolved by selective irradiation of the corresponding protons. The resonance of C-1'R was attributed to the 101.25 ppm signal and that of C-1G to 102.4 ppm (Table 4).

Stepwise double-resonance techniques allowed the attribution of a singlet at 79.18 to C-3G to 1, the chemical shifts depending on the solvent [13] (see Table 4). As can be expected, this deshielding on C-3G, with respect to the glucose, must be attributed to the Rha (1→3) Glc bond (+ δ effect 8 ppm), and to the β-caffeoyl effect (− δ effect 2 ppm).

Selective irradiation of H-4G allows the identification of C-4G as the location of the caffeic ester function. It may be noted that there is no appreciable change in the chemical shift of the C-4G going from the hydroxyl substituent to the ester. Finally, coupling constants of the anomeric carbons of glucose and rhamnose, respectively at 171 and 157 Hz are completely consistent with a β-configuration for the former and α- for the latter [14]. Additional evidence was obtained from the C-5 signals, *ca* 75 ppm for that of glucose and 69 ppm for that of rhamnose.

Structure of orobanchoside

^1H *NMR*. The spectra of the caffeic and of the aromatic moiety of the aglycone are very similar to those of verbascoside. The H-7' protons, i.e. the methine proton of –CHOH–, is readily located, giving a signal at 4.57 ppm as a doublet of a doublet with $J = 9$ and 2.5 Hz. It is obvious that these values reflect a preferential conformation of the hydroxyl



Orobanchoside 2

opposite (exo) to the glycosyl moiety. Irradiation of this proton (H-7′) allows the recognition of H-8′ at 3.45 ppm. In the most deshielded sugar region, the anomeric protons (H-1G, H-1′R) are observed with a small coupling constant as for that of rhamnose, and also a triplet, the coupling constant of which determines the position of the caffeic group on C-4′R, C-3G or C-4G as in **1**.

The resonances of H-2′R, H-3′R and H-6′R can be definitely assigned from the 250 MHz spectrum. Here these signals are shifted outside the glucose proton area. H-5′R on the other hand was identified by irradiation of the methyl protons. Here too, the β-configuration of the glucose is consistent with the signal of the H-1G proton at 4.55 ppm and $J = 7.8$ Hz. By double irradiation the resonance at 3.38 ppm can be unambiguously assigned to H-2G. It should be mentioned that this signal is perturbed by irradiation of a triplet at 4.05 ppm and at the same time the shape of another triplet at 4.92 ppm is modified. This observation gives definitive evidence for the linking of the caffeic ester function to C-4G. Irradiation of H-4G readily allows the location of the H-5G signal.

^{13}C *NMR.* This spectrum, as expected, corroborated very well with the above structure deduced from the ^1H NMR. Actually, in spite of the dissimilarity of the aglycone moieties, the aromatic carbon resonances are superimposable. From the coupled spectrum the C-8′ signal appeared as a triplet at 71.12 ppm and C-7′ as a doublet at 76.27 ppm (Table 3). The anomeric carbons are identified by irradiating the corresponding protons. In the present case, a larger shielding effect was observed for C-1G. The values for the coupling constants, 161 Hz for glucose and 173 Hz for rhamnose are in agreement with the respective β- and α-configurations. Furthermore, selective irradiation of the H-2G signal interacts with that at 80.59 ppm of C-2G. From this strongly deshielded value, it could be inferred that C-2G is involved in the glycosidic bond. In the same manner, irradiation of H-3G allowed the assignment of C-3G at 74.45 ppm (this value is indeed very close to that of C-7′).

In conclusion, the ^1H NMR spectra have enabled us to establish the nature of the second aromatic residues in **1** and **2**, the presence of the hydroxyl group at C-7′ in **2** and the esterification by the caffeoyl moiety at C-4G in both compounds. Selective irradiation of protons in the ^{13}C NMR spectra lead to the identification of the linkage between glucose and rhamnose in **1** and **2**. At the same time, the following observations can be made. The strongly shielded value of the anomeric carbon C-1G (97.10 ppm) in **2** reflects the nature of the (1→2) glycoside bond. The deshielding effect on the other chemical shifts of the inter-glycosidic carbon is the same as that expected from our preliminary studies [15]. In fact the deshielding effect (8 ppm) is depressed here by the acyl neighbouring effect, giving the value of 82 ppm, in the case of **1**. Likewise, C-2G in **2** supports the deshielding effect of the linkage, and the shielding of β-substitution. Thus verbascoside is β - (3′,4′ - dihydroxyphenyl) - ethyl - *O* - α - L - rhamnopyranosyl(1→3) - β - D - (4 - *O* - caffeoyl) - glucopyranoside and orobanchoside is β - hydroxy - β - (3′,4′ - dihydroxyphenyl) - ethyl - *O* - α - L - rhamnopyranosyl(1→2) - β - D - (4 - *O* - caffeoyl) - glucopyranoside.

EXPERIMENTAL

NMR spectra were measured on a CAMECA instrument at 250 MHz for ^1H and at 62.86 MHz for ^{13}C; chemical shifts are given on the δ (ppm) scale with TMS as int. standard. Solvents used are given in each table. Real coupling constants are obtained by a technique called gated decoupling.

Isolation of principals. **1** and **2** have been crystallized from an *Orobanche* phenolic extract [2, 8]. (Yield: **1** = 6% and **2** = 3%, from the dried plant.) Physico-chemical constants of **1** and **2** (PF, $[\alpha]_D$, UV IR, CCM) have been reported in [2, 8].

Acknowledgement—The authors are indebted to Mrs. M. Bon of the NMR Laboratory, Université Paul Sabatier, Toulouse, France for NMR measurements.

REFERENCES

1. Scarpati, M. L. and Monache, D. (1963) *Ann. Chim. (Rome)* **53**, 356.
2. Andary, C. (1975) Thèse Doct. (Etat) Pharm., Montpellier, France.
3. Birkofer, L., Kaiser, C. and Thomas, U. (1968) *Z. Naturforsch. Teil B* **23**, 1051.
4. Nonaka, G. and Nishioka, I. (1977) *Phytochemistry* **16**, 1265.
5. Cooper, R., Solomon, P. H., Kubo, I., Nakanishi, K., Shoolery, J. N. and Occolowitz, J. L. (1980) *J. Am. Chem. Soc.* **102**, 7953.
6. Harborne, J. B. (1966) *Z. Naturforsch. Teil B* **21**, 604.
7. Andary, C., Pellecuer, J., Soediro, I. and Privat G. (1978) *Pharmacia Mediterr.* **12**, 237.
8. Andary, C., Privat, G., Chevallet, P., Orzalesi, H., Serrano, J. J. and Boucard, M. (1980) *Il Farmaco, Ed. Sci.* **35**, 3.
9. Kelley, C. J., Mahajan, J. R., Brooks, L. C., Neubert, L. A., Brenneman, W. R. and Carmack, M. (1975) *J. Org. Chem.* **40**, 1804.

10. Pratviel-Sosa, F., Sosa, A. and Wylde, R. *Phytochemistry* (in press).
11. De Bruyn, A., Anteunis, M. and Verhegge, G., (1975) *Bull. Soc. Chim. Belg.* **84**, 721.
12. Kelley, C. J., Harruff, R. C. and Carmack, M. (1976) *J. Org. Chem.* **41**, 449.
13. Dutton, G. G. S., Laffite, C., Pratviel-Sosa, F. and Wylde, R. (to be published).
14. Bock, K. and Pedersen, C. (1974) *J. Chem. Soc.* 293.
15. Laffite, C., Nguyen Phuoc Du, A. M., Winternitz, F., Wylde, R. and Pratviel-Sosa, F. (1978) *Carbohydr. Res.* **67**, 105.