Phytochemistry, Vol. 24, No 7, pp 1557–1559, 1985
Printed in Great Britain

0031-9422/85 $3 00 + 0 00

TARAXACOSIDE, A TYPE OF ACYLATED γ-BUTYROLACTONE GLYCOSIDE FROM *TARAXACUM OFFICINALE**†

HANS-WILLI RAUWALD and JAI-TUNG HUANG‡§

Institut für Pharmazeutische Biologie der Universitat, Im Neuenheimer Feld 364, D-6900 Heidelberg 1, West Germany, ‡Institute for Chinese Pharmaceutical Science of the China Medical College, Taichung, Taiwan

(*Received* 21 *September* 1984)

Key Word Index—*Taraxacum officinale*; Compositae; Cichoriaceae; roots; lactone glycoside; butyrolactone acylglycoside; glucose ester; *p*-hydroxyphenylacetic acid; taraxacoside; β-*O*-[4-*O*-(*p*-hydroxyphenylacetyl)-β-D-glucopyranosyl]-β-hydroxy-γ-butyrolactone.

Abstract—In addition to the four new sesquiterpene lactones previously identified, a new acylated γ-butyrolactone glucoside, taraxacoside, was isolated from the roots of *Taraxacum officinale*. Its structure was elucidated mainly by ^1H and ^{13}C NMR studies as β-*O*-[4-*O*-(*p*-hydroxyphenylacetyl)-β-D-glucopyranosyl]-β-hydroxy-γ-butyrolactone This seems to be the first instance of the detection of a monocyclic five-membered, saturated lactone *O*-glycoside. Additionally, *p*-hydroxyphenylacetic acid was identified for the first time as an acylating acid in a sugar ester.

INTRODUCTION

Our previous investigation [1] of the roots and aerial parts from *Taraxacum officinale* afforded, in addition to known compounds, a new eudesmanolide, a tetrahydroridentin B, a eudesmanolide-β-D-glucopyranoside and two germacranolide acids, which are esterified with β-D-glucose. All three of the new glucose derivatives had a strong bitter taste. Our present work resulted in the isolation of taraxacoside, a new type of a γ-butyrolactone glucoside esterified with *p*-hydroxyphenylacetic acid from the bitter fraction of the roots of *T. officinale*. The elucidation of its structure, mainly by ^1H and ^{13}C NMR studies, is described in this paper. The importance of this substance for the identification of *T. officinale*, which is a widely used medicinal plant, will be discussed in another communication.

*Presented at the Annual Congress for Medicinal Plant Research, University of Antwerp, Belgium, July 1984.

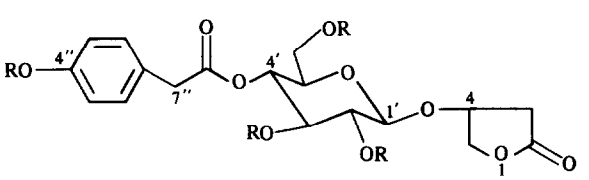
†Dedicated to Professor R Hänsel, Institut für Pharmakognosie und Phytochemie der Freien Universität Berlin, on the occasion of his 65th birthday

§Present address: Sloan Kettering Institute for Cancer Research, Walker Laboratory, Rye, New York 10580, U.S.A

RESULTS AND DISCUSSION

The presence of taraxacoside in the roots of *T. officinale* was detected by TLC of an ethyl acetate extract containing the polar bitter substance fraction [1] on silica gel using methylene chloride–ethanol–water (80:19:1) as solvent. A weak fluorescent extinguishing zone was identified in short-wave UV light at R_f 0.25, giving a grey colour reaction after spraying with vanillin–sulphuric acid and a purple colour with Molisch's reagent. The glycoside was isolated from the aqueous methanolic extract of the dried roots. After elimination of the lipophilic components by petrol and chloroform extraction, the water phase was extracted with ethyl acetate [1]. Further fractionation of the organic phase was carried out by repeated chromatography on silica gel columns using methylene chloride–ethanol–water. Taraxacoside (**1a**) was obtained in a crystalline state (mp 178–180°; determined as $C_{18}H_{22}O_{10}$). Acetylation led to a tetra-acetyl derivative (**1b**) (mp 164–165°; determined as $C_{26}H_{30}O_{14}$).

Complete acid hydrolysis of **1a** in refluxing 2 M hydrochloric acid–methanol yielded an unknown aglycone moiety plus D-glucose and *p*-hydroxyphenylacetic acid, which were identified by direct comparison (TLC and IR and ^1H NMR spectra) with authentic samples. We had also isolated the corresponding free acid from the aerial parts, as described earlier [1, 2]. Enzymatic hydroly-



	R
1a	H
1b	COMe

1a, b

sis with β-glucosidase yielded neither glucose nor *p*-hydroxyphenylacetic acid, indicating that the anomeric lactol group was linked to a non-aromatic aglycone.

The IR spectrum of **1a** gave major bands at 1770 and 1735 cm^{-1} suggesting a five-membered saturated lactone and an aliphatic ester [3]. The electron impact mass spectrum showed no $[M]^+$, but fragments indicative of a hydroxylactone, a triacetylglucosyl moiety (**1b**: *m/z* 229, 289; $[C_{12}H_{17}O_8]^+$) and *p*-hydroxyphenylacetic acid (**1a**/**1b**: *m/z* 151, $[C_8H_7O_3]^+$, 107, $[C_8H_7O_3-CO_2]^+$) were observed. The lactone moiety (**1a**/**1b**) gave fragments for $[C_4H_6O_3]^+$ and $[C_4H_7O_3]^+$, both of which readily lost water to the major fragments $[C_4H_4O_2]^+$ and $[C_4H_5O_2]^+$, and which underwent further fragmentation typical of five-membered lactones [4]. These data suggest a tetrahydrofuranone substituted with a secondary oxygen function (hydroxyl/alkoxyl). This hypothesis was supported by the ^1H and ^{13}C NMR spectra (Tables 1 and 2), which showed taraxacoside to have aromatic, aliphatic and sugar moieties. In the ^1H NMR spectrum, two signals resonating at δ2.67 and 2.86 suggested the AB part of an ABX system. They were assigned to the non-equivalent C-3 methylene protons of the lactone ring in the neighbourhood of the carbonyl function. The large coupling constants J_{gem} = 17 and 18 Hz (for **1a** and **1b**, respectively) are indicative of a five-membered ring with a neighbouring π-system [3]. Selective decoupling in **1b**, e.g. irradiation of the δ4.40 resonance simplified the two quartets at δ2.28 and 2.60 as two doublets. The signal at δ4.40 (**1b**) (4.89, **1a**) appeared as a broad 1H multiplet and could be assigned to a methine proton at C-4 where it coupled with the four methylene protons of C-3/C-5. The significant Δδ2.3 downfield shift of the C-4 proton signal relative to the corresponding frequency of the unsubstituted γ-butyrolactone (δ2.1 [5]) is in good agreement with the calculated value of δ4.4 due to the additive alkoxy increment, taken from Curphey–Morrison additivity constants [6]. A further distinct resonance was observed at δ4.35 (**1b**) as a doublet for 2H. It was assigned to the isochronous pair of γ-CH_2 protons attributable to coupling with the proton at C-4. Both the resonance position and the shape of this signal correlated with the published data of γ-butyrolactone [5] (triplet at δ4.3). These results were confirmed by the ^{13}C NMR signals of the lactone moiety (Table 2). Atoms C-3–C-5 resonated in good concordance with the tentatively calculated δ-ranges of a C-4 alkoxylation of butyrolactone with the aid of a chemical shift approximation according to ref. [7]: calc. αC/βC/γC: δ36/70/77 (Δδ taken from cyclopentane). On the other hand an α-*O*-alkyl substituent would shift C-5 to ≤ δ65 and, in the case of γ-alkoxylation, the signal of C-5 would be expected to be significantly downfield by up to δ100. On this basis it is established that the glucosyloxy residue is linked to C-4 of the butyrolactone aglycone.

Furthermore, downfield in the ^1H NMR four aromatic protons are resolved as an AA'BB' system, belonging to the *p*-hydroxy- and acetoxy phenylacetyl moieties. Its CH_2-7″ protons were non-equivalent, giving rise to an AB quartet at δ3.84 and 3.76 (**1a**). These values, as well as the ^{13}C NMR shifts summarized in Tables 1 and 2, are in good agreement with those given by the authentic free acid and its published data [7]. The coupling constants ($H_{1',2'}$ = 8 Hz) of the glucose anomeric proton resonating at δ4.98 (**1a**) and 4.49 (**1b**) are in accordance with a β-configuration. The other glucosyl protons have been identified by selective irradiation. Thus, irradiation of H-1′ in **1a** modifies the triplet at δ4.15, giving the value of H-2′, and irradiation of this latter proton collapses the signal at δ4.61 which is, therefore, attributed to H-3′. In addition, irradiation of the triplet at δ5.53 is reflected by a modification of the H-3′ signal and also of the complex signal at δ3.98. Thus, the resonances at δ5.53 and 3.98 belong to H-4′ and H-5′, respectively. The significant downfield shift of the δ5.53 triplet in **1a**, attributed to H-4′, established the phenylacetic acid substitution of

Table 1. ^1H NMR data of taraxacoside (**1a**) and its tetra-acetate (**1b**) [250 MHz; shifts in δ-values relative to TMS; splitting pattern and *J* values (Hz) are given in parentheses]

Proton(s) at C atom	**1a** ($C_5D_5N+D_2O$)	**1b** ($CDCl_3$)
Aglycone		
$3_{1,2}$	2.67 (*dd*; 17.8, 1.8)	2.28 (*dd*; 18.2, 2.6)
	2.86 (*dd*; 17 8, 6.2)	2.60 (*dd*; 18.2, 6 4)
4	4 89 (*m*, 1H)	4.40 (*m*)
$5_{1,2}$	4 46 (*m**, 2H)	4.35 (*d*; 3 8)
Glucose		
1′	4.98 (*d*; 8 1)	4.49 (*d*; 8)
2′	4 15 (*t*; 8.1)	4 96 (*dd*; 8, 9 5)
3′	4.61 (*dd*; 8, 8 7)	5.20 (*t*; 9 5)
4′	5.53 (*t*, 8.7)	5.03 (*t*; 9.5)
5′	3.98 (*m*)	3.66 (*ddd*; 9.5, 4.5, 2.3)
$6'_{1,2}$	4.27, 4.31 (*m*)	4.24 (*dd*; 12, 4 5)
		4 18 (*dd*; 12, 2.3)
p-Hydroxyphenyl acetic acid		
2″, 6″	7.39, 7.18 (AA'BB';	7.26, 7.06 (AA'BB';
3″, 5″	$J_{AB} = J_{A'B'} = 8.5$)	$J_{AB} = J_{A'B'} = 8.7$)
7″	3 84 (*d*, 15 3)	3.57 (*s*)
	3 76 (*d*, 15 3)	
OH-4″	11 53 (*s*)	OAc: 2.28 (*s*, ar);
		2.08; 2.01;
		1.85 (3 × *s*, al)

*Signal pattern unclear due to overlapping.
†Signal before D_2O exchange.

Table 2. ^{13}C NMR data of taraxacoside tetra-acetate (**1b**) in deuterochloroform (shifts in δ-values relative to TMS at 20 4 MHz)

Aglycone*		*p*-Acetoxyphenylacetic acid	
C-2	174.41	C-1″	131.22
C-3	34.79	C-2″, C-6″	130.42
C-4	73.50†	C-3″, C-5″	122.13
C-5	74.66	C-4″	150.48
		C-7″	40.70
		C-8″	169 51
Glucose			
C-1′	100 03		
C-2′	72.48†		
C-3′	72.48†	MeCO	169 51; 169.61;
C-4′	68 59		170 34; 170.62
C-5′	71 46†	MeCO	20.31; 20.40;
C-6′	61.97		20.54; 20.95

*δ_c of γ-butyrolactone [7] (C-2–C-5): 177.9; 27 7; 22.2, 68.6.
†Resonance assignments are interchangeable.

glucose on C-4′. Additionally, before deuterated water exchange, only the H-4′ signal was observed as a sharp triplet, whereas H-2′, H-3′ and H-6′ further split due to the vicinal couplings with the hydroxyl groups. According to some esterified phenylpropanoid glycosides, such as raphanusol A [8] and verbascoside [9], no other effects were observed on the remaining protons of this substitution. Furthermore, esterification of the anomeric hydroxyl by the acyl moiety could be excluded, since the signals of both anomeric protons do not resonate further downfield than *ca* δ5.5 [10]. The ^{1}H NMR of the acetylated taraxacoside (**1b**) revealed the presence of four acetyl groups (one aromatic and three aliphatic) and confirms the above results.

In conclusion, the ^{1}H and ^{13}C NMR data of taraxacoside and its tetra-acetate have enabled us to establish the nature of the aglycone moiety, its position of *O*-glucosylation and the esterification by the hydroxyphenylacetic acid at C-4′ of glucose. On the basis of the above evidence the isolated compound, taraxacoside, could be identified as 4-[4-*O*-(*p*-hydroxyphenylacetyl)-β-D-glucopyranosyloxy]-tetrahydrofuran-2-one. Taraxacoside represents a previously unknown structural type of *O*-glycosylated monocyclic five-membered, saturated lactone. Additionally, *p*-hydroxyphenylacetic acid was found for the first time as an acylating acid in a sugar ester [11].

EXPERIMENTAL

Plant material. As described in ref. [1].

Extraction and isolation of taraxacoside. The powdered dried roots of *T. officinale* (5 kg) were extracted with cold 80% MeOH using an Ultra Turrax. The filtrate after MeOH removal under red. pres. was extracted with petrol and pure $CHCl_3$. The aq. phase was extracted with EtOAc. The EtOAc fraction (*ca* 50 g brown oil) was chromatographed on a column of silica gel (800 g) eluting with CH_2Cl_2–EtOH (9:1) (**1**) and finally with CH_2Cl_2–EtOH–H_2O (85:14.1), from which fractions of pure taraxacoside (0.48 g), were collected TLC identification: see Results and Discussion.

Taraxacoside (**1a**). Colourless crystals ($CHCl_3$–EtOH–Me_2CO), mp 178–180°. IR ν_{max}^{KBr} cm^{-1}: 3400 (br OH), 3000 (ar CH), 1770 (γ-lactone), 1735 (RCOOR), 1620, 1600, 1520 (ar C=C), 1190, 1170 (C–O st). EIMS 70 eV, *m/z* (rel. int.): 151 $[C_8H_7O_3]^+$ (23.6), 133 $[151-H_2O]^+$ (2.4), 107 $[151-CO_2]^+$ (100), 103 $[C_4H_7O_3]^+$ (1.4), 102 $[C_4H_6O_3]^+$ (1.0), 85 $[C_4H_5O_2]^+$ (4.4), 84 $[C_4H_4O_2]^+$ (29.4). [Found. C, 54.1; H, 5 7. $C_{18}H_{22}O_{10}$ (MW 398) requires: C, 54 2; H, 5.52%.]

Taraxacoside tetra-acetate (**1b**). Acetylation of **1a** following usual methods gave colourless needles (MeOH), mp 164–166° IR ν_{max}^{KBr} cm^{-1}. 3050, 2960–2920 (ar CH), 1765 (γ-lactone), 1750 (COOR). UV λ_{max}^{MeCN} nm: 240, 260. EIMS 70 eV, *m/z* (rel. int): 524 $[M-COCH_2]^+$ (2.6), 464 $[M-C_4H_6O_3]^+$ (0.8), 422 $[524-C_4H_6O_3]^+$ (10), 289 $[C_{12}H_{17}O_8]^+$ (0.9), 229 $[289-MeCOOH]^+$ (0.9), 151 $[C_8H_7O_3]^+$ (6.0), 133 $[151-H_2O]^+$ (59.6), 107 $[151-CO_2]^+$ (100), 103 $[C_4H_7O_3]^+$ (4.4), 102 $[C_4H_6O_3]^+$ (4.8), 85 $[C_4H_5O_2]^+$ (10.6), 84 $[C_4H_4O_2]^+$ (14.2). [Found: C, 54.8; H, 5.4. $C_{26}H_{30}O_{14}$ (MW 566) requires: C, 55.1; H, 5.3%.]

Acid hydrolysis. Taraxacoside (15 mg) was dissolved in MeOH, and refluxed with 2 M HCl for 2 hr, diluted with H_2O and extracted with EtOAc. The EtOAc soln was evaporated to dryness; the residue afforded *p*-hydroxyphenylacetic acid [TLC, silica gel; CH_2Cl_2–MeOH–H_2O (85:14:1); R_f 0.3; grey-brown spot with vanillin–H_2SO_4]. The aq. layer was neutralized with basic ion exchange resin (Amberlite 400) and evaporated to dryness. The presence of D-Glc was shown by TLC [1].

Enzymatic hydrolysis. With β-glucosidase, as described in ref [1].

p-Hydroxyphenylacetic acid. Isolated from aerial parts, as described in ref. [1]. Colourless crystals, mp 148–150°. EIMS 70 eV, *m/z* (rel. int.): 152 $[M]^+$ (26), 134 $[M-H_2O]^+$ (17), 107 $[M-COOH]^+$ (100), 77 (24). ^{1}H NMR (60 MHz, DMSO-d_6): δ3.55 (2H, *s*, Ar–C<u>H</u>$_2$–CO), 5.25 (1H, *s*, OH), 6 69 and 7.07 (4H Ar, AA′BB′, J_{AB} = 8.7 Hz), 9.21 (1H, *s*, OH-4). (Found: C, 63.3; H, 5.2. Calc. for $C_8H_8O_3$: C, 63.2; H, 5.3%.)

Acknowledgements—We thank Dr. D. Rosenberg, Schering AG Berlin and Mrs. G. Dreke, Freie Universität Berlin, for the recording of ^{1}H and ^{13}C NMR spectra and Professor Dr. W. D. Clark, Heidelberg/Arizona, for reading the manuscript.

REFERENCES

1. Hänsel, R., Kartarahardja, M., Huang, J T. and Bohlmann, F. (1980) *Phytochemistry* **19**, 857.
2. Power, F. B and Browning, H (1914) *J. Chem. Soc.* **105**, 1829.
3. Williams, D. H. and Flemming, I (1971) *Spektroskopische Methoden in der organischen Chemie*, Vol. 2. Georg Thieme, Stuttgart
4. Silverstein, R. M., Bassler, C. G. and Morill, T. C. (1974) *Spectroscopic Identification of Organic Compounds*. John Wiley, New York.
5. Clerc, T. and Pretsch, E. (1970) *Kernresonanzspektroskopie*, p. 34. Akadem. Verlagsgesellschaft, Frankfurt.
6. Fuchs, P. L and Bunell, C. A. (1979) *Carbon 13 NMR Based Organic Spectral Problems* p. 307. John Wiley, New York.
7. Kalinowski, H. O., Berger, S. and Braun, S. (1984) *^{13}C-NMR-Spektroskopie*. Georg Thieme, Stuttgart.
8. Hase, T. and Hasegawa, K (1982) *Phytochemistry* **21**, 1021.
9. Andary, C., Wylde, R. Laffite, C , Privat, G. and Winternitz, F. (1982) *Phytochemistry* **21**, 1123.
10. Lehmann, J. (1976) *Chemie der Kohlenhydrate*, p. 59. Georg Thieme, Stuttgart.
11 Harborne, J. B. and Williams, C. A (1982) in *The Flavonoids Advances in Research* (Harborne, J B. and Mabry, T J , eds) p. 270. Chapman & Hall, London