Phytochemistry, 1980, Vol. 19, pp. 1481–1484. © Pergamon Press Ltd. Printed in England. 0031-9422/80/0701-1481 $02.00/0

STRUCTURAL CLARIFICATION OF GERMACRANOLIDES FROM *CALEA* SPECIES*

ZENAIDE S. FERREIRA, NÍDIA F. ROQUE, OTTO R. GOTTLIEB, FERNANDO OLIVEIRA† and HUGO E. GOTTLIEB‡

Instituto de Química, Universidade de São Paulo, C. P. 20780, 05508, São Paulo, SP, Brazil

(*Revised received* 7 *November* 1979)

Key Word Index—*Calea pinnatifida*; Compositae; Heliantheae; polyacetylene; germacranolide.

Abstract—The aerial parts of *Calea pinnatifida* contain a 4-glycosyloxybenzoic acid, anisic acid, tetradeca-4*E*,6*E*,12*E*-trien-8,10-diyn-1-ol and arucanolide. The structural elucidation of this germacranolide included comparison with neurolenin B whose stereochemistry had been defined by X-ray crystallographic analysis. The data suggest that the structures of all other known germacranolides of the genus *Calea* should be revised.

INTRODUCTION

Calea pinnatifida Banks (Compositae, Heliantheae), a shrub popularly known as aruca, is widely distributed over São Paulo State. An ethanol extract of the bitter leaves is commercially available for the treatment of amoebic dysentery [1].

RESULTS

Aerial parts of specimens growing near Billings dam, São Paulo, were found to contain chiefly fatty esters, a 4-glycosyloxybenzoic acid, anisic acid, sitosterol, stigmasterol, a polyacetylene, arucanolide and a bitter germacranolide. The glycoside gave a tetraacetate and, after hydrolysis, *p*-hydroxybenzoic acid. This latter compound, anisic acid and the phytosterols were all identified by direct comparison with authentic samples.

The polyacetylene was spectroscopically indistinguishable from **1**, previously isolated from other Heliantheae species [2], particularly with respect to the methyl double doublet (δ1.87, J = 7.0 and 1.5 Hz) in the ^{1}H NMR spectrum. The methyl signal for the isomeric structure **2** would appear as a multiplet [2] or as a broad doublet. Catalytic hydrogenation gave, as expected, *n*-tetradecan-1-ol which was identified by ^{1}H NMR and MS.

High resolution MS revealed the molecular formula $C_{21}H_{26}O_8$ for arucanolide. ^{1}H and ^{13}C NMR permitted the expansion of the formula to

$$C_{15}H_{18}O_4 \cdot OCOMe \cdot OCOC(Me)CH_2,$$

and suggested this compound to be a sesquiterpene possessing α,β-unsaturated ketone and α-methylenelactone moieties. This was confirmed by the presence of 4 carbonyl peaks in the IR spectrum. Irradiation, in turn, of each signal in the ^{1}H NMR spectrum (270 MHz) established several proton sequences, that were accommodated by structure **3a**. The relative location of the acyl groups was based on the formation of two monoacylated methanolysis products **3e** and **3f**. The loss of an acetyl residue upon conversion of **3a** into **3e** was accompanied by a 1.48 ppm diamagnetic shift of one of the carbinolic proton signals which was a double doublet. Decoupling experiments proved this proton to be located on C-8. Alternatively, in monoester **3f** without the methacrylate group, the high-field carbinolic proton signal (shielded by 1.65 ppm relative to **3a**) appeared as a doublet, and can thus only refer to H-9.

The recently described germacranolide neurolenin B (**3b**) differs in constitution only in the nature of the esterifying groups. Its stereochemistry was unambiguously defined by an X-ray crystallographic analysis [3]. The chemical shifts of the carbons and protons (as far as given) of the sesquiterpene moiety, were virtually identical to those obtained for arucanolide (Tables 1–3), which therefore must possess the configuration shown in structure **3a**.

A germacranolide from *Calea urticifolia* was assigned the isomeric structure **4**. The structure was based exclusively on an interpretation of the coupling constants using a Karplus-type relationship and chemical shift comparisons with a reduced derivative formulated as **6** [4]. No evidence was presented to justify the relative location of the acyl residues. Furthermore, the identity of mp, IR, MS and optical rotation (see Experimental), as well as the ^{1}H NMR spectrum (Tables

* Part I in the series "The Chemistry of Brazilian Compositae". Based on part of the Doctorate Thesis presented by Z. S. F. to Universidade de São Paulo (1979). Sponsored by Fundação de Amparo à Pesquisa do Estado de São Paulo.

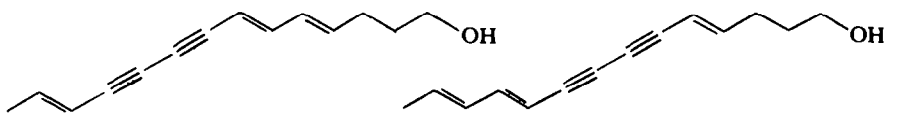
† Faculdade de Ciências Farmacêuticas, Universidade de São Paulo.

‡ Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot, Israel.

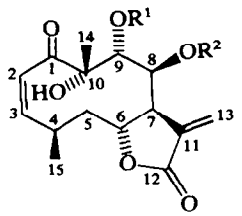
Table 1. ^{1}H NMR chemical shifts (δ) of arucanolide, its derivatives and other reported germacranolides*

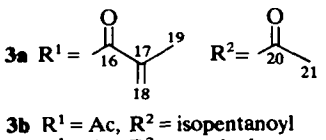
	3a 270 MHz	3a† 270 MHz	3b‡ 100 MHz	3c, 3d§ 100 MHz	3e 270 MHz	3f 60 MHz	7 270 MHz	6† 270 MHz
1	—	—	—	—	—	—	4.49	4.51
2	6.61	6.61	6.61	6.57	6.61	6.67	5.34	5.34
3	6.03	6.03	6.02	5.79	6.00	6.17	5.34	5.34
4	3.13	3.13		3.1	3.08	3.27	3.60	3.57
5	1.83	1.84			1.85		1.7–1.8	1.81
5′	1.44	1.45			1.44		1.7–1.8	1.68
6	4.62	4.62	4.57	4.56	4.50	4.56	4.75	4.74
7	2.66	2.65	2.63	2.62	2.52	2.73	4.34	4.34
8	5.62	5.62	5.57	5.66	4.14	5.43	5.51	5.52
9	5.62	5.62	5.57	5.54	5.54	3.97	5.44	5.44
11	—	—	—	—	—	—	2.86	2.84
13	5.84	5.54‖	5.82	5.70	5.71	5.73	1.09	1.11
13′	6.32	6.33	6.31	6.28	6.31	6.33	—	—
14	1.35	1.36	1.34	1.32	1.32	1.53	1.28	1.30
15	1.14	1.15	1.13	1.11	1.13	1.13	1.04	1.05
18	5.54	5.55	—	—	5.69	—	5.60	5.59
18′	6.01	6.03	—	—	6.25	—	6.07	6.08
19	1.83	1.84	—	—	2.00	—	1.85	1.87
21	2.01	2.02	2.09	1.99	—	2.02	1.93	1.94

* In $CDCl_3$, TMS as internal standard.
† Data from ref. [4].
‡ Data from ref. [3].
§ Data from ref [5].
‖ The anomaly must be due to a misprint; indeed, ref. [4] includes data for a further two germacranolides differing only by the substitution of the acetyl by methacryloyl or angeloyl residues. In both H-13 appears at δ5.83.



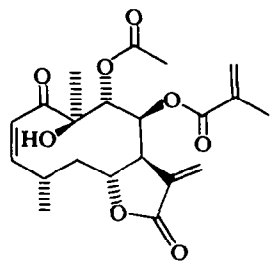
1 2



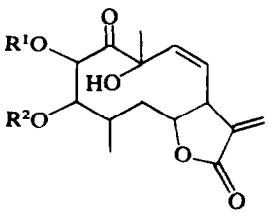


3a R^1 = R^2 =

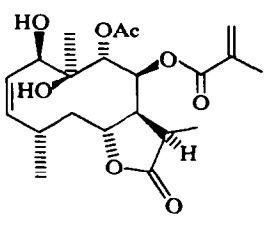
3b R^1 = Ac, R^2 = isopentanoyl
3c R^1 = Ac, R^2 = angeloyl
3d R^1 = angeloyl, R^2 = Ac
3e R^1 = methacryloyl, R^2 = H
3f R^1 = H, R^2 = Ac



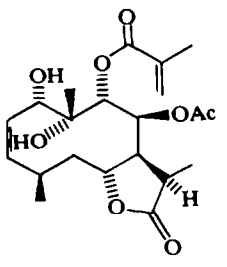
4



5a R^1 = Ac, R^2 = angeloyl
5b R^1 = angeloyl, R^2 = Ac



6



7

Table 2. Coupling constants ($J_{H,H}$ Hz) of arucanolide, its derivatives and other reported germacranolides*

	3a 270 MHz	**3a**† 270 MHz	**3b**‡ 100 MHz	**3c**, **3d**§ 100 MHz	**3e** 270 MHz	**3f** 60 MHz	**7** 270 MHz	**6**† 270 MHz
1, 2	—	—	—	—	—	—	—	3
2, 3	12	12	11	12	12	12		11
3, 4	11.5	11.5	‖	11	11.5	11		12
4, 5	12	12			12.5			12
4, 5′	5	5.5			5.5			5.5
4, 15	7	6.5	7		6.5	6	6.5	6.5
5, 5′	14	13			13.5			13
5, 6	5	4.5	5	5	5	5	6	5.5
5′, 6	12	11	11	11	12	12	11.5	11
6, 13′					1			
7, 8		2		2	1.5	2		2
7, 11	—	—	—	—	—	—	8.5	10
7, 13		1.5	2	1.5	1.5		—	—
7, 13′		1.5	2	1.5	1.5		—	—
8, 9		9.5		10	9.5	9.5	10	9.5
11, 13	—	—	—	—	—	—	7	7
18, 18′		1	—	—	1.5	—		1
18, 19		1	—	—	1.5	—		1
18′, 19		1	—	—	1	—		7

*, †, ‡, § See corresponding footnotes, Table 1.

‖ The H-3 signal is described as a *d*; this is presumably a misprint.

1 and 2) indicates both sets of data to refer to the same compound **3a**. Therefore, not only the compound described as having structure **4** but also two related sesquiterpenes differing only in the nature of the acyl residues [4] belong to the stereochemical series of compounds **3**.

Table 3. ^{13}C NMR chemical shifts (δ) of arucanolide, its reduced derivative and neurolenin B*

	3a 22.6 MHz	**3b**† 25.2 MHz	**7** 22.6 MHz
1	204.1	204.3	79.4
2	125.5	125.3	123.2
3	148.3	147.9	136.0
4	28.3	28.2‡	28.7
5	40.3	40.2	38.9
6	76.4	76.3	78.2
7	41.3	41.2	34.5
8	74.0§	73.8§	69.9
9	74.5§	73.9§	74.6
10	79.3	79.3	76.0
11	134.7	134.8	37.8
12	168.9	168.6‡	179.0
13	126.6	126.2	8.8
14	23.6	23.6	26.7
15	19.7	19.6	21.8
16	165.4	—	166.5
17	135.0	—	135.2
18	127.3	—	128.0
19	18.1	—	17.9
20	170.4	170.0‡	169.7
21	20.4	20.5	20.4

* See corresponding footnote, Table 1; assignments are based in part on a correlation between residual couplings in the single-frequency off-resonance decoupled spectra and the ^{1}H NMR chemical shifts.

† Data from ref. [3].

‡ Literature values for these carbons were reassigned.

§ Signals may be interchanged.

Two other germacranolides, caleins A and B, were isolated from *C. zacatechichi* and assigned structures **5a** and **5b**, but with their configurations undetermined [5]. The proton sequences were established by double irradiation experiments and were identical to those of compounds **3**. The different structural proposal derives from the unlikely assignment of the pair of signals at δ 6.57 and 5.79 to carbinolic protons and of the pair at δ 5.54 and 5.66 to olefinic protons. The overall strong similarity between the H NMR spectra of the caleins with the compounds examined thus far indicated their structural relation and we, therefore, formulate them as **3c** and **3d** respectively.

According to these revisions, the structure of the $NaBH_4$ reduction product **6** [4] must be reformulated. Two new chiral centres are created at C-1 and C-11 and the configuration shown in **7** was deduced as follows. The conformation indicated by the X-ray crystallographic study of neurolenin B [3] is stabilized by the quasi-equatorial position of all the substituents of the decacycle except for the carbonyl oxygen and the vicinal tertiary hydroxyl, which may be hydrogen bonded. As a conspicuous consequence of this arrangement, the C(7)–H bond points towards the centre of the ring system. The dihedral angles predicted by this analysis agree with the H–H vicinal coupling constants, not only for **3a** but also for its tetrahydro derivative **7**, showing that both compounds maintain this same general conformation in solution.

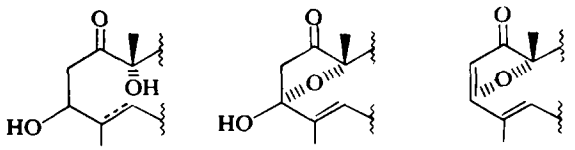
Hydride attack during reduction of **3a** would be expected to occur from the outside of the ring system, thus projecting the resulting hydroxyl towards the middle of the ring, and into close proximity with hydrogens at C-4 and C-7, which as a result will be very strongly deshielded (by 0.47 and 1.68 ppm, respectively). Indeed, a β-OH at C-1 would require a

dihedral angle of *ca* 180° between H-1 and H-2 and a γ-effect between C-1 and C-4. However, the coupling constant of 3 Hz observed between H-1 and H-2 [4] is more compatible with an angle of *ca* 60° (1α-OH), and C-4 is almost invariant relative to **3a**, the C-1 signal appearing at low field (79.4 ppm). The presence of a 1α-OH precludes a γ-effect between these carbons, which would be transmitted through the 1α-H and 4-H atoms in the alternative stereochemistry.

^{1}H and ^{13}C NMR spectroscopy also helped define the configuration at C-11. The large coupling constant between H-7 and H-11, and the presence of a γ-effect between C-8 and C-13 (the former carbon is shielded by *ca* 4 ppm relative to **3a** and the latter appears at high field, 8.8 ppm) indicated the *cis* relationship between H-7 and H-11 and between C-8 and C-13 relative to the pentacycle. Thus the configuration shown in **7** was established.

DISCUSSION

All the hitherto described germacranolides isolated from the genus *Calea* may derive from a common precursor of partial structure **8**. Facile loss of water would lead to structures of type **3**, while oxidation and hemiketal formation would lead to type **9** [6] and, by successive elimination of water, to type **10** [4, 7]. This common biosynthetic origin would imply a common stereochemistry at C-10, as indicated in formulae **9** and **10**. All previous reports [4, 6, 7] on bridged germacranolides of this type from *Calea*, however, feature the epimeric configuration for this centre. In view of the unambiguous establishment of the C-10 stereochemistry reported in the present paper for compounds **3a–3f** a re-examination of the bridged types may be in order.



8 **9** **10**

EXPERIMENTAL

Isolation of the constituents. Air-dried, ground aerial parts (5.3 kg) were percolated successively with petrol and EtOAc. Evapn of the solvents gave respectively extracts A (35 g) and B (25 g). A was chromatographed on Si gel. Elution with C_6H_6–EtOAc (9 : 1 and 8 : 2) gave respectively **1** (800 mg) and **3a** (900 mg). B in MeOH was freed from chlorophyll by filtration through active charcoal. The MeOH was evapd and the residue (12 g) chromatographed on Si gel. Elution with C_6H_6–EtOAc–MeOH (4 : 1 : 0, 7 : 3 : 0, 3 : 7 : 0 and 0 : 4 : 1) gave respectively **3a** (175 mg), sitosterol plus stigmasterol (150 mg), anisic acid (250 mg) and 4-glycosyloxybenzoic acid (850 mg).

Arucanolide (**3a**). Colourless, intensely bitter crystals, mp 172–175° (C_6H_6–petrol), 170–172° (Et_2O–petrol). [Found: C, 66.03; H, 6.35. $C_{21}H_{26}O_8$ requires: C, 62.07; H, 6.40%]. IR ν_{max}^{KBr} cm^{-1}: 3400, 1765, 1740, 1730, 1720, 1690, 1630, 1220, 1150, 815, 760. UV $\lambda_{infl.}^{MeOH}$ nm: 237 (13550). MS *m/e* (rel. int.): 406 (1) M$^+$, 388 (1), 364 (12), 346 (12), 320 (2), 278 (12), 260 (10), 69 (100), 43 (32). $[\alpha]^{25}$ (*c* 0.1 g/ml) −24° (589 nm), −26° (578 nm), −28° (546 nm).

Reduction of arucanolide. To a soln of **3a** (200 mg) in MeOH (5 ml), $NaBH_4$ (400 mg) was added in small portions. The mixture was stirred at room temp. (2 hr), diluted with H_2O (100 ml), acidified with HCl and extracted with $CHCl_3$. The $CHCl_3$ soln was dried (Na_2SO_4), filtered and evapd. The residue (120 mg) was chromatographed on Si gel. Elution with C_6H_6–EtOAc (7 : 3) gave **7**, colourless crystals, mp 78–80° (C_6H_6). IR ν_{max}^{KBr} cm^{-1}: 3500, 1770, 1750, 1720, 1635, 1240, 1110, 1020, 950, 760. MS *m/e* (rel. int.): 410 (1) M$^+$, 350 (1), 324 (4), 282 (17), 264 (7), 69 (87), 43 (100).

Hydrolysis of arucanolide. To a soln of **3a** (300 mg) in MeOH (10 ml), conc HCl (4 ml) was added drop by drop. The mixture was left at room temp. (72 hr), diluted with H_2O and extracted with $CHCl_3$. The $CHCl_3$ soln was washed, dried (Na_2SO_4), filtered and evapd. The residue (180 mg) was purified by TLC (Si gel, C_6H_6–EtOAc, 1 : 1) to a mixture of **3e** plus **3f**. Crystallization from C_6H_6–MeOH gave **3e**, mp 183–185°. MS *m/e* (rel. int.): 364 (1) M$^+$, 282 (10), 169 (70), 71 (25), 55 (10), 43 (100). The data of **3f** (Tables 1, 2) were determined by difference.

4-Glycosyloxybenzoic acid. Colourless needles, mp 213–214° (MeOH) [Found: C, 51.98; H, 5.41. $C_{13}H_{16}O_8$ requires: C, 52.00; H, 5.37%]. IR ν_{max}^{KBr} cm^{-1}: 3450, 1680, 1620, 1600, 1500, 1260, 1190, 1090, 1060, 1000, 860, 790. *Tetraacetate*, mp 178–180° (MeOH + C_6H_6). IR ν_{max}^{KBr} cm^{-1}: 3130, 2690, 1700, 1610, 1500, 1230, 1090, 1040, 990, 860, 780. ^{1}H NMR ($CDCl_3$, 60 MHz): δ2.1 (*s*, 4-OAc), 3.7–4.5 (*m*, CH_2O), 5.35 (*m*, CHO), 7.05 and 8.1 (AA'BB' system, $J = ca$ 9 Hz). MS *m/e* (rel. int.): 332 (4) M$^+$, 331 (7), 289 (2), 271 (2), 229 (2), 211 (2), 169 (35), 139 (4), 138 (4), 127 (13), 121 (7), 115 (7), 109 (38), 93 (5), 43 (100). *Hydrolysis.* A soln of the glycoside (100 mg) in 2N HCl (5 ml) was heated on the steam bath (3 hr), cooled and extracted with Et_2O. The Et_2O soln was dried and evapd. The residue was crystallized from C_6H_6 to *p*-hydroxybenzoic acid (40 mg).

Hydrogenation of tetradeca-4,6,12-trien-8,10-diyn-1-ol (80 mg) in MeOH (10 ml) over Pd/C (10 mg) gave tetradecanol, IR ν_{max}^{film} cm^{-1}: 3300, 1460, 1380, 1070, 735. ^{1}H NMR (CCl_4, D_2O, 60 MHz): δ0.88 (*t*, Me), 1.25 (*s*, 12-CH_2) 3.61 (*t*, $J = 7$ Hz, CH_2O). MS *m/e* (rel. int.): 196 (19) M$^+$ −H_2O, 168 (19), 140 (11), 139 (10), 126 (19), 125 (21), 112 (15), 111 (38), 98 (35), 97 (67), 84 (50), 83 (98), 82 (56), 70 (71), 69 (92), 57 (31), 55 (100), 49 (88), 41 (77).

REFERENCES

1. Malhado Filho, J. (1950) *Arq. de Biol. (São Paulo)* **34,** 95.
2. Bohlmann, F. and Kleine, K. M. (1965) *Chem. Ber.* **98,** 872.
3. Manchand, P. S. and Blount, J. F. (1978) *J. Org. Chem.* **43,** 4352.
4. Bohlmann, F. and Jakupovic, J. (1979) *Phytochemistry* **18,** 119.
5. Quijano, Z., Romo de Vivar, A. and Rios, T. (1978) *Rev. Latinoam. Quim.* **9,** 86.
6. Bohlmann, F. and Zdero, C. (1977) *Phytochemistry* **16,** 1065.
7. Ortega, A., Romo de Vivar, A., Diaz, E. and Romo, J. (1970) *Rev. Latinoam. Quim.* **1,** 81.