Pergamon

PII: S0031-9422(96)00720-0

*Phytochemistry*, Vol. 44, No. 6, pp. 1133–1137, 1997

0031-9422/97 $17.00+0.00

# SESQUITERPENE LACTONES AND LIGNANS FROM *ARTEMISIA ARBORESCENS*

J. ALBERTO MARCO, JUAN F. SANZ-CERVERA, VICENTE GARCÍA-LLISO and JOAN VALLÈS-XIRAU*

Departamento de Química Orgánica, Universidad de Valencia, E-46100 Burjassot, Valencia, Spain; *Laboratori de Botànica, Facultat de Farmàcia, Universitat de Barcelona, E-08028 Barcelona, Spain

(*Received in revised form* 2 *August* 1996)

**Key Word Index**—*Artemisia arborescens*; Compositae; Anthemideae; sesquiterpene lactones; homoditerpene endoperoxide; lignans; azulene derivative.

**Abstract**—The aerial parts of *Artemisia arborescens* yielded, in addition to several known compounds, a new guaianolide, a new homoditerpene endoperoxide, a new lignan of the sesamin type and a new azulene derivative.


## INTRODUCTION

The large genus *Artemisia* has been the subject of numerous chemical studies [1]. *Artemisia arborescens* L., which grows along the Mediterranean coast of European and African countries [2], has been previously investigated. Aerial parts of this species have been shown to contain acetylenes, flavonoids, sesquiterpene lactones and sesamin-type lignans [1, 3–5]. As a part of our recent interest in the chemistry of North African *Artemisia* species [6–8], we have now investigated *A. arborescens* collected in Tunisia. The results of this investigation are presented here.

## RESULTS AND DISCUSSION

From aerial parts of *A. arborescens* we have isolated the following known compounds: **1** [9]; arsanin (**2**) [9, 10]; the acyclic diterpene 7,11,15-trimethyl-3-methylenehexadecane-1,2-diol [11]; the lignans sesamin, yangambin, epiyangambin, aschantin, sesartemin, spinescen, demethoxyexcelsin and pinoresinol [12]; the flavonol artemetin [13]; the coumarins scopoletin, isofraxidin and dracunculin [14]; and costic acid methyl ester [15]. Artemetin and some of the lignans had already been found in previous investigations of the species. Furthermore, we have characterized four new compounds: the guaianolide (**3**), the azulene derivative (**4**), the homoditerpene endoperoxide (**5**) (a mixture of diastereoisomers) and the sesamin-type lignan **6**.

The IR (Experimental) and NMR data (Table 1) for compound **3**, $C_{17}H_{22}O_5$, indicated the presence of a hydroxylated lactone bearing an additional acetate group. Two C=C bonds were also noticeable, one trisubstituted and the other as part of an exocyclic methylene ($^{13}C$ NMR). Spin decoupling results suggested structure **3**, which was definitively supported by 2D heteronuclear correlations. The configurations at C-1, C-5, C-6, C-7 and C-8 were deduced from the values of the coupling constants and confirmed by means of NOE measurements, e.g. the strong NOEs (5–10%) between H-1 and H-5, between H-5 and H-7, between H-6 and H-8 and between H-6 and H-9$\beta$. Furthermore, the NOEs (*ca* 5%) observed between H-2 and H-6 and between H-2 and H-9$\beta$ served to establish the configuration at C-2, which could not be clearly deduced from the value of $J_{1,2}$.

Compound **4** displayed an intense blue–violet colour (UV $\lambda_{max}$ 244, 303 and 398 nm), which suggested an azulene derivative. The FAB mass spectrum showed a protonated molecular peak at *m/z* 199, in agreement with the molecular formula $C_{14}H_{14}O$. The $^{13}C$ NMR spectrum (Experimental) was consistent with this molecular formula as it showed 14 peaks, three of them being methyl signals. The other signals were due to one ketone carbonyl and 10 aromatic carbons. The $^1H$ NMR data (Experimental) showed three methyl singlets around $\delta$ 2.8, an AB system and an AMX system, the latter signals appearing in the aromatic region above $\delta$ 7. With these data to hand, only structure **4** was likely. The locations of the substituents, suggested by biogenetic considerations, were confirmed by both coupling constant values and NOE measurements. Particularly relevant among the latter were those observed (5–10%) between the methyl group at C-1 and both H-2 and H-8, between the methyl group at C-4 and both H-3 and H-5, and between H-8 and both the acetyl methyl and the methyl group at C-1. The isolation of this product is not surprising since azulene derivatives have often been found in the essential oil of the species [1]. Dihy-

Table 1. $^1$H and $^{13}$C NMR data for sesquiterpene lactones **2** and **3**

| H | 2 | 3 | C | 2 | 3 |
|---|---|---|---|---|---|
| 1 | 3.63 *dd* (11.5, 5.5) | 2.90 *dd* (8.2, 4) | 1 | 76.2 | 57.4 |
| 2α | 2.70 *dd* (15, 5.5) | | 2 | 46.5 | 79.4 |
| 2β | 2.53 *dd* (15, 11.5) | 4.75 *br s* | 3 | 208.4 | 129.9 |
| 3 | | 5.66 *br q* (1.5) | 4 | 44.6 | 145.9 |
| 4 | 2.47 *dq* (12, 7) | | 5 | 50.0 | 55.6 |
| 5 | 1.40 *dd* (12, 10) | 3.00 *br dd* (10, 8.2) | 6 | 82.6 | 80.2 |
| 6 | 3.97 *dd* (10, 10) | 3.96 *dd* (10, 10) | 7 | 53.3 | 52.8 |
| 7 | 1.60 *dddd* (12, 12, 10, 4) | 2.25 *m** | 8 | 22.8 | 75.9 |
| 8α | 1.90 *dddd* (12, 4, 4, 3) | | 9 | 36.5 | 41.5 |
| 8β | 1.52 *dddd* (12, 12, 12, 3) | 4.90 *ddd* (11, 7.5, 5) | 10 | 41.9 | 140.8 |
| 9α | 1.20 *m** | 2.25 *m** | 11 | 40.6 | 40.7 |
| 9β | 2.10 *ddd* (14, 3, 3) | 2.70 *dd* (13, 5) | 12 | 178.9 | 177.8 |
| 11 | 2.26 *dq* (12, 7) | 2.45 *dq* (12, 7) | 13 | 12.4 | 15.6 |
| 13 | 1.20 *d* (7) | 1.28 *d* (7) | 14 | 11.9 | 117.2 |
| 14 | 1.11 *s* | 5.01 *br s*, 4.99 *br s* | 15 | 13.6 | 17.1 |
| 15 | 1.21 *d* (7) | 1.87 *q* (1, 5) | OAc | | 170.1 |
| OAc | | 2.10 *s* | | | 21.1 |

δ in ppm and *J* (parentheses) in Hz; 400 MHz ($^1$H) and 100 MHz ($^{13}$C); $CDCl_3$; 22°.
* Overlapped signal.

droazulenes with the same carbon framework as **4** have been isolated from the essential oil of *Tanacetum annuum* [16]. It is unlikely, however, that these azulenes are true natural products, but may have been formed from guaiane precursors by various structural changes during the isolation procedure.

Compound **5** did not give a molecular peak in either the EI or the FAB mass spectrum. The IR data (Experimental) indicated the absence of relevant functional groups. The NMR data (Table 2) indicated that a mixture of two closely related compounds was at hand here, because some $^1$H and $^{13}$C signals were duplicated. Aside from this signal duplication, the $^{13}$C NMR spectrum showed the presence of 21 carbon atoms, which indicated that the compounds were likely to be homoditerpenes. Two of the signals corresponded to oxygenated quaternary carbon atoms. Taking into account the multiplicity of the carbon signals, the molecular formula $C_{21}H_{34}O_x$ was suggested. The terpenoid nature was further supported by the $^1$H NMR spectrum, which showed the presence of an isopropyl group, a methyl bound to an oxygenated carbon atom and a $CH_3CH$ fragment, as well as 1,1-disubstituted, *cis*-1,2-disubstituted and trisubstituted C=C bonds. The aforementioned molecular formula implied a total of five unsaturations. Since three of them were C=C bonds, the other two consequently were attributable to cycles. In view of this, structure **5**, containing a cyclic endoperoxide was the most likely. Although a molecular ion peak was not observed in the mass spectrum, some key fragment peaks were consistent with the proposed structure (see proposed fragmentation patterns in the formulae scheme). Stereoisomeric endoperoxides of this type are formed from precursor dienes through non-stereospecific cyclooxygenation, which may or may not be enzyme controlled. They usually occur as mixtures which are very difficult to separate [17]. In this particular case, the progenitor diene has already been found in another plant [16], together with its aromatization product. As expected, the NMR data for the diene were in part similar to those for **5**.

Table 2. $^1$H and $^{13}$C NMR data for homoditerpenes **5**

| H | | C | |
|---|---|---|---|
| 2 | 6.40 *d* (8.5) | 1 | 74.4 |
| 3 | 6.47 *d* (8.5) | 2 | 136.3 |
| | 6.46 *d* (8.5) | 3 | 133.6, 133.5 |
| 5 | 2.10–1.90 *m** | 4 | 79.9 |
| 6 | 2.10–1.90 *m** | 5 | 25.9[a], 25.8[a] |
| 7 | 1.36 *s* | 6 | 29.5, 29.4 |
| 8 | 1.70 *m** | 7 | 21.4 |
| 9 | 1.60 *m** | 8 | 36.8 |
| 10 | 2.10–1.90 *m** | 9 | 31.4, 31.1 |
| 11 | 5.11 *br t* (7) | 10 | 25.7[a], 25.6[a] |
| 13 | 2.10–1.90 *m** | 11 | 123.9 |
| 14 | 2.10–1.90 *m** | 12 | 135.5 |
| 16 | 2.22 *sept* (7) | 13 | 38.4 |
| | | 14 | 33.0 |
| 17, 18 | 1.02 *d* (7) | 15 | 155.9 |
| 19 | 4.72 *br s* | 16 | 33.8 |
| | 4.66 *br s* | 17, 18 | 21.8 |
| 20 | 1.60 *br s* | 19 | 106.2 |
| 21 | 0.99 *d* (7) | 20 | 16.1 |
| | 0.98 *d* (7) | 21 | 14.0, 13.9 |

δ in ppm and *J* (parentheses) in Hz; 400 MHz ($^1$H) and 100 MHz ($^{13}$C); $CDCl_3$; 22°.
* Overlapped signal.
[a] Signals may be interchanged.

Compound **6** was clearly a sesamin-type lignan, on the basis of the similarities of its spectral data with those for the other isolated lignans of the same structural class. The NMR data (Table 3) showed that

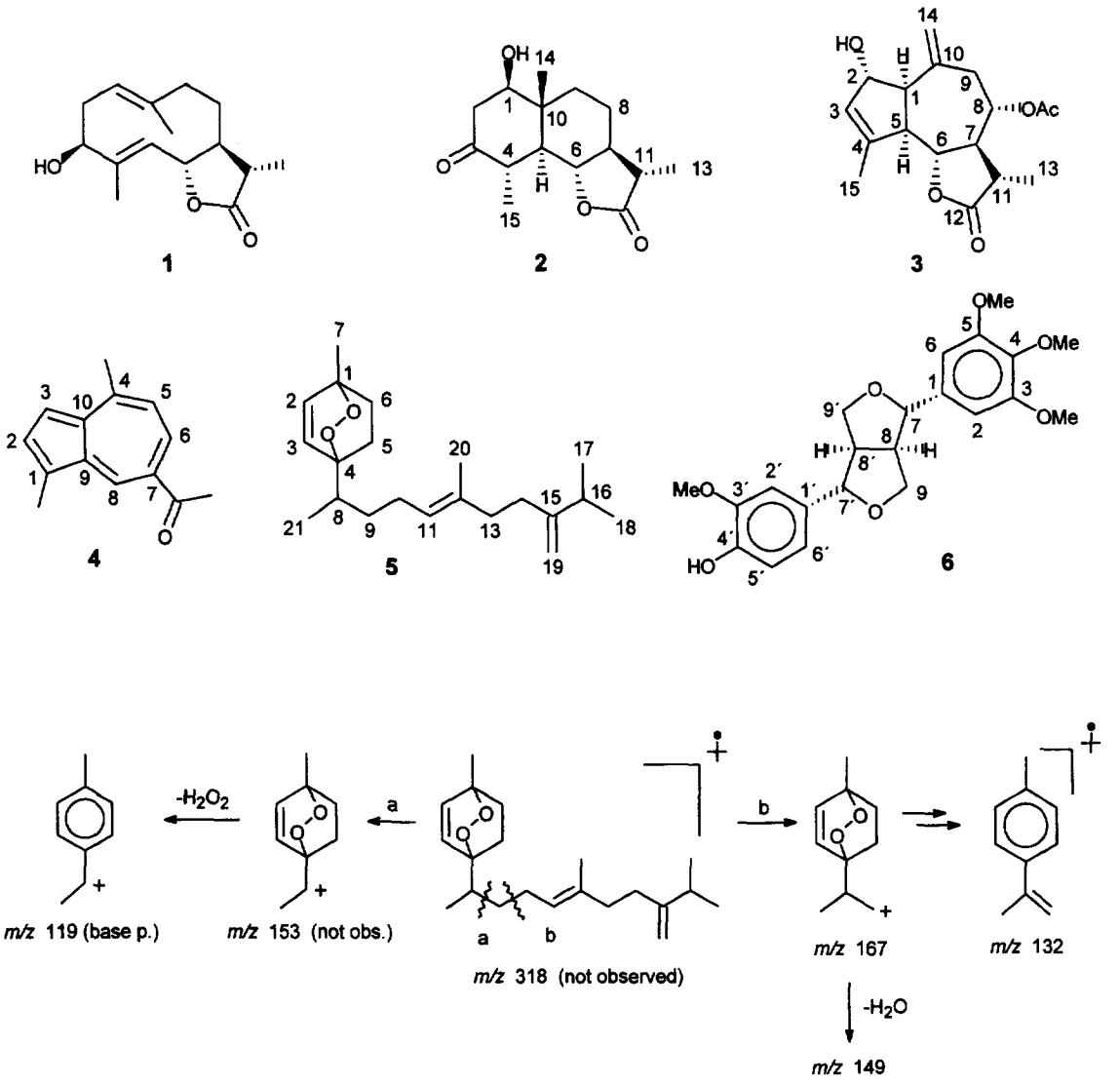


Table 3. $^1H$ and $^{13}C$ NMR data for lignan **6**

| H | | C | |
|---|---|---|---|
| 2, 6 | 6.56 *s* | 1 | 137.3 |
| 2′ | 6.88 *d* (2) | 2, 6 | 102.7 |
| 5′ | 6.88 *d* (8) | 3, 5 | 153.4 |
| 6′ | 6.81 *dd* (8; 2) | 4 | 136.8 |
| 7, 7′ | 4.74 *m** | 1′ | 132.7 |
| 8, 8′ | 3.10 *m** | 2′ | 108.5 |
| 9, 9′ | 3.90 *m** | 3′ | 146.7 |
| | 4.26 *m** | 4′ | 145.2 |
| | | 5′ | 114.2 |
| 3,5-OMe | 3.86 *s* (6H) | 6′ | 118.9 |
| 4-OMe | 3.82 *s* (3H) | 7, 7′ | 86.0, 85.8 |
| 3′-OMe | 3.89 *s* (3H) | 8, 8′ | 54.4, 54.1 |
| | | 9, 9′ | 71.9, 71.7 |
| | | OMe | 60.8, 55.9 |
| | | | 56.1 (×2) |

$\delta$ in ppm and $J$ (parentheses) in Hz; 400 MHz ($^1H$) and 100 MHz ($^{13}C$); $CDCl_3$; 22°.

* Overlapped signal.

one of the benzene rings was 3,4,5-trisubstituted with methoxyl groups, while the other was 3,4-disubstituted with a hydroxyl group and a methoxyl group. The relative position of the latter groups was deduced from NOE measurements, which indicated that the methoxyl group in the disubstituted phenyl ring was contiguous to H-2′. The $^{13}C$ NMR data further confirmed the depicted relative configurations at C-7, C-7′, C-8, C-8′, C-9 and C-9′ [12], which is the same as that displayed by most of the other isolated lignans. The absolute configuration was deduced from the positive sign of the optical rotation [12].

Lactone **2** was isolated for the first time from *A. santolina* and named arsanin [1]. Its structure has been confirmed by chemical correlation with α-santonin [18]. We have now isolated arsanin from *A. arborescens* and determined its structure, including stereochemical aspects, in an unambiguous way by means of exhaustive NMR analyses. In fact, high-resolution NMR data for arsanin have recently been provided [9]. However, the published chemical shifts differ more than expected from our own values. Therefore, we

provide the complete NMR data for arsanin in Table 1.

Aside from the lignans of the sesamin type, the compounds isolated in the present study are different from those found previously [1, 3–5]. This supports previous proposals that there exist different chemotypes of the species [1].

## EXPERIMENTAL

NMR spectra: 400 MHz ($^1$H) and 100 MHz ($^{13}$C) in $CDCl_3$ (22°). NOE measurements: in the 1D difference mode. Optical rotations: 22°. MPCC: Merck silica gel (25–40 μm); gradient elution with the solvent mixts indicated in each case. HPLC: LiChrosorb RP-8 (250 × 8 mm); elution with MeOH–$H_2O$ mixts.

*Plant material.* Aerial parts of *A. arborescens* were collected in Tunisia, 11 km west of Sejnane (May 1992). A voucher specimen (BCF-37204) has been deposited in the Herbarium of the Laboratory of Botany, Faculty of Pharmacy, University of Barcelona, Spain (Prof. J. Vallès-Xirau).

*Extraction and chromatography.* Plant material (935 g air-dried aerial parts) was processed according to the described protocol [19]. The defatted extract was prefractionated by CC on silica gel (A, hexane–$Et_2O$, 1:1; B, $Et_2O$; C, $Et_2O$–MeOH, 6:1). The 3 frs were subjected to further chromatographic sepns as described below.

MPCC of fr. A [elution with hexane–$Et_2O$ (2:1) to $Et_2O$], followed where necessary by CC, prep. TLC or HPLC, allowed isolation of costic acid Me ester (25 mg), sesamin (880 mg), aschantin (1.2 g), yangambin (440 mg), epiyangambin (85 mg), sesartemin (125 mg), spinescen (67 mg), demethoxyexcelsin (7 mg), 7,11,15-trimethyl-3-methylenehexadecane-1,2-diol (10 mg), artemetin (115 mg), dracunculin (5 mg), **1** (475 mg), **4** (7 mg) and **5** (19 mg). In the same way, MPCC of fr. B [elution with hexane–$Et_2O$ (1:2) to $Et_2O$] allowed isolation of more aschantin (37 mg), yangambin (300 mg), spinescen (13 mg), artemetin (250 mg) and **1** (96 mg), as well as pinoresinol (10 mg), scopoletin (15 mg), **2** (24 mg), **3** (11 mg) and **6** (10 mg). Finally, fr. C afforded only artemetin (90 mg) and isofraxidin (6 mg).

*8α-Acetoxy-2α-hydroxy-1α,5α,6β,7α,11βH-guaia-3,10(14)-dien-12,6-olide* (**3**). Oil, $[\alpha]_D$ +35° ($CHCl_3$; *c* 0.7); IR $\nu_{max}^{film}$ cm$^{-1}$: 3450 (*br*, OH), 3060, 1770 (lactone C=O), 1740 (acetate C=O), 1450, 1375, 1240, 1175, 1120, 1030. EIMS (probe) *m/z* (rel. int.): 246 $[M-HOAc]^+$ (12), 228 $[M-HOAc-H_2O]^+$ (12), 164 (100), 91 (77). NMR: Table 1.

*7-Acetyl-1,4-dimethylazulene* (**4**). Blue–violet oil, IR $\nu_{max}^{film}$ cm$^{-1}$. 3090, 2980, 2970, 2875, 1675, 1600, 1570, 1525, 1450, 1380, 1365, 1245. UV $\lambda_{max}^{EtOH}$ nm: 244, 293sh, 303, 324sh, 398. FABMS *m/z* (rel. int.): 199 $[M+H]^+$ (45), 198 $[M]^+$ (60), 154 (100), 136 (90), 119 (75). $^1$H NMR ($CDCl_3$, 400 MHz, 22°): 8.98 (1H, *d*, *J* = 2 Hz; H-8), 8.20 (1H, *dd*, *J* = 11 and 2 Hz; H-6), 7.66 (1H, *d*, *J* = 3.8 Hz; H-2), 7.51 (1H, *d*, *J* = 3.8 Hz; H-3), 7.06 (1H, *d*, *J* = 11 Hz; H-5), 2.87 (3H, *s*; Me-$C_4$), 2.74 (3H, *s*; COMe), 2.73 (3H, *s*; Me-$C_1$). $^{13}$C NMR ($CDCl_3$, 100 MHz, 22°): 127.9 (C-1), 136.9 (C-2), 119.1 (C-3), 149.5 (C-4), 124.4 (C-5), 136.1 (C-6), 137.2 (C-7), 133.4 (C-8), 133.9, 134.3 (C-9, C-10), 13.3 (Me-$C_1$), 24.3 (Me-$C_4$), 27.0 (MeCO), 198.9 (ketone CO).

*6,10-Dimethyl-9-methylene-2-(4-methyl-1,2-dioxabicyclo[2.2.2]oct-5-en-1-yl)undec-5-ene* (**5**) (*mixt. of 2 diastereoisomeric peroxides*). Oil, IR $\nu_{max}^{film}$ cm$^{-1}$: 3040, 2975, 2960, 2860, 1640, 1450, 1375, 1260, 1110, 880, 725, 690. EIMS (probe) *m/z* (rel. int.): 167 (28), 149 (62), 132 (63), 119 (100), 109 (36), 95 (33), 81 (38), 69 (32), 55 (49). NMR: Table 2.

*(1R,2S,5R,6S)-2-(3,4,5-Trimethoxyphenyl)-6-(4-hydroxy-3-methoxyphenyl)-3,7-dioxabicyclo[3.3.0]octane* (**6**). Oil, $[\alpha]_D$ +18° ($CHCl_3$; *c* 0.88). IR $\nu_{max}^{film}$ cm$^{-1}$: 3400 (*br*, OH), 3060, 1600, 1510, 1450, 1420, 1370, 820, 730. EIMS (probe) *m/z* (rel. int.): 402 $[M]^+$ (100), 371 $[M-OMe]^+$ (10), 240 (10), 224 (15), 207 (30), 195 (32), 181 (40), 151 (42), 137 (28). NMR: Table 3.

*Acknowledgement*—One of us (J. V.-X.) thanks the Ministry of Education and Science for financial support (CICYT project 88-PR-033).

## REFERENCES

1. Marco, J. A. and Barberá, O., in *Studies in Natural Products Chemistry*, Vol. 7A, ed. Atta-ur-Rahman. Elsevier, Amsterdam, 1990, pp. 201–264.
2. Tutin, T. G., Persson, K. and Gutermann, W., in *Flora Europaea*, Vol. 4, eds T. G. Tutin, V. H. Heywood, N. A. Burges, D. M. Moore, D. H. Valentine, S. M. Walters and D. A. Webb. Cambridge University Press, London, 1976, p. 180.
3. Grandolini, G., Casinovi, C. G., Betto, P., Fardella, G., Menichini, F., Gabriele, R., Barbetti, P., Kajtar-Peredy, M. and Radics, L., *Phytochemistry*, 1988, **27**, 3670.
4. Betto, P., Casinovi, C. G., Gabriele, R., Grandolini, G. and Menichini, F., *Revista Latinoamer. Química*, 1988, **19**, 40.
5. Zarga, M. A., Qauasmeh, R., Sabri, S., Munsoor, M. and Abdalla, S., *Planta Medica*, 1995, **61**, 242.
6. Marco, J. A., Sanz-Cervera, J. F., Ocete, G., Carda, M., Rodríguez, S. and Vallès-Xirau, J., *Journal of Natural Products*, 1994, **57**, 939.
7. Marco, J. A., Sanz-Cervera, J. F., Pareja, J. M., Sancenón, F. and Vallès-Xirau, J., *Phytochemistry*, 1994, **37**, 477.
8. Marco, J. A., Sanz-Cervera, J. F., Sancenón, F. and Vallès-Xirau, J., *Phytochemistry*, 1994, **37**, 1095.
9. Barrero, A. F., Sánchez, J. F., Barrón, A. and Ramírez, A., *Phytochemistry*, 1992, **31**, 332.
10. Moiseeva, G. P., Kasymov, S. Z., Yagudaev, M. R. and Sidyakin, G. P., *Khimiya Prir. Soedin.*, 1980, 343.

11. Urones, J. G., De Pascual-Teresa, J., Marcos, I. S., Fernández-Moro, R., Basabe-Barcala, P. and Sexmero-Cuadrado, M. J., *Phytochemistry*, 1987, **26**, 1113.
12. Ayres, D. C. and Loike, J. D., *Lignans. Chemical, Biological and Clinical Properties*. Cambridge University Press, Cambridge, U.K., 1990, Ch. 2.
13. Gottlieb, O. R., in *The Flavonoids*, Vol. 1, eds J. B. Harborne, T. J. Mabry and H. Mabry. Academic Press, New York, 1975, Ch. 7.
14. Murray, R. D. H., Méndez, J. and Brown, S. A., *The Natural Coumarins*. John Wiley, Chichester, U.K., 1982.
15. Herz, W., Chikamatsu, H. and Tether, L. R., *Journal of Organic Chemistry*, 1966, **31**, 1632.
16. Barrero, A. F., Sánchez, J. F., Altarejos, J. and Zafra, M. J., *Phytochemistry*, 1992, **31**, 1727.
17. Bohlmann, F. and Abraham, W.-R., *Phytochemistry*, 1979, **18**, 1754.
18. Yamakawa, K. and Nishitani, K., *Chemical and Pharmaceutical Bulletin*, 1976, **24**, 2810.
19. Marco, J. A., Sanz-Cervera, J. F. and Manglano, E., *Phytochemistry*, 1993, **33**, 875.