

AMARELLOSIDE, A BITTER TRI-*O*-ACETYL TRI-*O*-BENZOYL TETRASACCHARIDE FROM *POLYGALA AMARELLA*

MARIE-ALETH DUBOIS, ANDRÁS NESZMELYI,* GÜNTHER HEUBL,† MANFRED FIEBIG‡ and HILDEBERT WAGNER‡

Laboratory of Pharmacognosy, Faculty of Pharmacy, Burgundy University, F-21033 Dijon Cedex, France; *Central Research Institute for Chemistry of the Hungarian Academy of Sciences, H-1025 Budapest, Hungary; †Institute of Systematic Botany, University of Munich, D-8000 Munich 2, F.R.G.; ‡Institute of Pharmaceutical Biology, University of Munich, D-8000 Munich 2, F.R.G.

(Received 20 February 1989)

Key Word Index—*Polygala amarella*; Polygalaceae; hexa-*O*-acyl tetrasaccharide; amarelloside; flavonoids; hydroxycinnamoyl esters.

Abstract—A very bitter compound, amarelloside, together with rutin, kaempferol-3-*O*-rutinoside and β -D-(1,3-disinapoyl)-fructofuranosyl- α -D-glucopyranoside have been isolated from the aerial parts of *Polygala amarella*. The structure of amarelloside has been established as β -D-(1,3,4-tri-*O*-benzoyl)-fructofuranosyl- $\{[\beta$ -D-(6-*O*-acetyl)glucopyranosyl(1 \rightarrow 2)][β -D-glucopyranosyl(1 \rightarrow 3)] $\}$ - α -D-(di-*O*-acetyl-4,6)glucopyranoside by NMR spectroscopy, FAB mass spectrometry and GC/MS of the partially methylated alditol acetates.

INTRODUCTION

Plants of the genus *Polygala* are widely distributed in America, Europe and Asia where they are used in traditional medicine for the treatment of parasitical diseases [1]. The presence of several hydroxycinnamoyl esters, saponins and phenolic glycosides has been reported in this genus [2], but *Polygala amarella*, one of the 33 European species, has not been previously studied. A preliminary investigation of the methanolic extract of the aerial parts of this plant led to the isolation of a very bitter tasting principle which we believed could provide a chemotaxonomical marker in the genus *Polygala*. We now report on the isolation and structure elucidation of this new hexa-*O*-acyl tetrasaccharide, which we have named amarelloside (1), together with the isolation and identification of two known flavonoids and one hydroxy cinnamoyl ester.

RESULTS AND DISCUSSION

Amarelloside (1) was obtained as a colourless amorphous powder which gives a positive reaction with Godin reagent (vanillin-H₂SO₄) and has a very bitter taste. Its negative FAB mass spectrum exhibits a strong $[M - H]^-$ ion at m/z 1103 indicating a M_r of 1104, in good agreement with the empirical formula C₅₁H₆₀O₂₇. The IR spectrum (KBr) indicates hydroxyl groups (3400 cm⁻¹, br), carbonyl groups (1730 cm⁻¹) and aromatic rings (1605, 1525, 1500 cm⁻¹). The UV spectrum (MeOH) exhibits typical absorptions of benzoyl derivatives at λ_{max} 203, 230 and 280 (sh) nm. Upon acid hydrolysis, glucose and fructose were identified by TLC and GC of their trimethylsilyl derivatives. The approach adopted here to determine the structure is to make assignments of the protons resonances as complete as possible via COSY experiments [3] and to support this analysis by assigning the ¹³C NMR spectrum through heteronuclear ¹H/¹³C shift correlated experiments [4, 5].

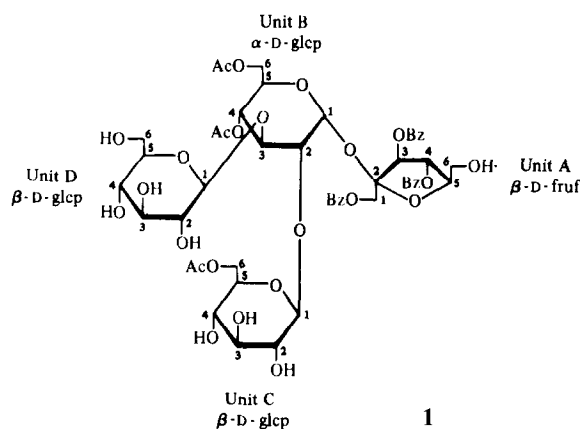
Analysis of the 300 MHz ¹H NMR spectrum revealed three acetyl protons signals as singlets at δ 2.00, 1.98, 1.64, three orthobenzoyl protons signals as doublets at δ 8.18, 8.12 and 7.96, the other benzoyl multiplets between δ 7.80 and 7.40, two separate doublets at δ 5.68 and 5.60, two separate triplets at δ 5.12 and 4.42 and a crowded region between δ 4.60 and 2.50. The 2D homonuclear ¹H NMR (COSY) spectrum at 300 MHz revealed the presence of four monosaccharide units by tracing all closed spin systems in the molecule (Table 1). This conclusion was supported by the presence of four anomeric carbon signals at δ 104.24, 103.83, 103.77 and 92.45 in the ¹³C NMR spectrum (Table 2). The identification of the four monosaccharides (A–D) was achieved by comparing their subspectra to those of the unsubstituted ones [6]. The COSY spectrum revealed the presence of three glucopyranosides. Two of them were characterized by the signals of anomeric protons at δ 4.52 (d , $J = 7.5$ Hz) and 4.31 (d , $J = 7.2$ Hz) indicating the β anomeric configuration. The anomeric proton signal at δ 5.6 (d , $J = 3.5$ Hz) was characteristic of one α -glucopyranoside. The fourth sugar was a β -fructofuranoside where the coupling constants were in good agreement with values reported for sucrose [7]. The sequence of the sugars residues and the position of the six acyl substituents were determined by establishing the connectivities of the proton resonances of each monosaccharide in the region δ 2.8–5.8 of the 2D COSY spectrum [8] and the one-bond carbon–proton connectivities through heteronuclear carbon–proton shifts correlation [5]. By comparing the ¹H and ¹³C NMR data of the four subunits with values reported for sucrose [7, 9, 10], 1,3-disinapoylsucrose [10], methylglycosides [6, 11], methyl *O*-acylglycosides [6, 12] and oligodisaccharides [13, 14], the protons and carbons of the backbone of amarelloside have been identified (Tables 1 and 2). The signals attributed to the protons A-1a, A-1b, A-3 and A-4 (δ 4.58, 4.28, 5.68 and 4.62) appeared downfield (1–2 ppm) of those of the corresponding signals of sucrose and suggested a 1,3,4-trisubstitution of the fructo-

Table 2. ^{13}C NMR data of compound **1** compared with values reported for reference compounds (CDCl_3 , 100 MHz, TMS as int. standard)*

Unit	C	1	Sucrose	Laminaribiose	Sophorose
A	1	66.44	62.15		
	2	103.77	104.07		
	3	80.27	77.26		
	4	73.75	74.44		
	5	83.99	82.56		
	6	62.39	62.21		
B	1	92.45	91.76	91.70	91.37
	2	80.76	71.67	70.89	82.07
	3	78.30	72.96	84.54	71.64
	4	70.57	69.98	68.56	70.17
	5	69.10	72.83	71.80	71.72
	6	63.75	60.64	61.05	60.09
C	1	104.24		104.10	104.92
	2	74.35		73.79	73.81
	3	77.44		76.19	76.34
	4	70.57		70.14	70.10
	5	74.77		76.84	76.73
	6	63.60		60.95	61.17
D	1	103.83			
	2	74.56			
	3	77.32			
	4	70.39			
	5	74.05			
	6	62.79			

*The multiplicity of the signals have been confirmed by DEPT experiments and the attribution of the chemical shifts by using $^1\text{H}/^{13}\text{C}$ heterocorrelated 2D spectrum. For the reference substances, see refs [6, 7, 10, 11, 13].

For the benzoyl carbon atoms the following signals have been attributed: δ 134.57, 134.33, 134.29 (C-1); 130.50, 130.45, 130.29 (C-2, C-6); 130.24, 130.18, 130.20 (C-3, C-5); 129.67, 129.65, 129.65 (C-4), 171.40, 170.97, 170.82 (C-7). For the acetyl carbon atoms: δ 21.40, 21.37, 21.23 (Me-CO); 165.69, 166.04, 165.09 (Me-CO).



of the glucopyranosyl (unit D) and 6-*O*-acetylglucopyranosyl (unit C) residues to their (-2 or -3) positions on unit B was achieved by 2D correlation spectroscopy using the 2D-NOE technique, which establishes connectivities due to the dipole-dipole cross relaxation [18]. Inspection and comparison of the 2D-NOE with the 2D-COSY spectrum revealed through space interactions across the

glycosidic bond observed for the following protons pairs: the anomeric proton of unit C (δ 4.52) and the proton B-2 (δ 3.80), the anomeric proton of unit D (δ 4.31) and the proton B-3 (δ 3.90), indicating the linkages C-1 \rightarrow B-2 and D-1 \rightarrow B-3. On the basis of the foregoing evidence, and if glucose and fructose are assumed to be members of the commonly found D-series, the structure of amarilloside is proposed to be β -D-(1,3,4-tri-*O*-benzoyl)-fructofuranosyl- $\{[\beta$ -D-(6-*O*-acetyl) glucopyranosyl(1 \rightarrow 2)] $[\beta$ -D-glucopyranosyl(1 \rightarrow 3)] $\}$ - α -D-(di-*O*-acetyl-4,6) glucopyranoside.

Analogous derivatives of β -D-fructofuranosyl(2 \rightarrow 1) α -D-glucopyranoside (sucrose) characterized as mono-, di- or triesters with ferulic, sinapic or coumaric acid and as mixed esters of ferulic and acetic acid have been described in a few plants of the families Liliaceae [19, 20], Brassicaceae [21], Polygonaceae [22] and Polygalaceae [10]. Only 6,3'-diferuloyl sucrose and its acetyl derivatives have been isolated as bitter components from *Lilium speciosum* [23] and *Lilium longiflorum* [24]. In amarilloside (**1**), the basic structure of sucrose is substituted by two acetyl and three benzoyl residues via ester linkages and by one glucosyl and one 6-*O*-acetyl glucosyl residues via ether linkages. To our knowledge, the bitter compound **1** represents the first mixed ester of the tri-*O*-acetyl tri-*O*-benzoyl tetrasaccharide type, and as such represents a new natural compound.

Compound **2** was obtained as an amorphous yellow powder. Its spectroscopic data (UV, IR, ^1H NMR, ^{13}C NMR, FAB/MS) were identical with those described for (1,3-disinapoyl)-fructofuranosyl- α -D-glucopyranoside [10].

Compounds **3** and **4** were isolated and identified as rutin and kaempferol-3-*O*-rutinoside respectively by comparison with authentic samples (UV, IR, FAB/MS, ^1H NMR and TLC) [25].

EXPERIMENTAL

General. Mps: uncorr., IR: KBr disc. UV ($\lambda_{\text{max}}^{\text{MeOH}}$). ^1H NMR spectra (δ , TMS as internal standard, *J* Hz) were recorded at 300 MHz ($\text{CDCl}_3 + \text{CD}_3\text{OD}$) and 500 MHz ($\text{Me}_2\text{CO}-d_6$) at room temp., using the selective 5 mm probehead. Standard microprograms of the Bruker library were used for the COSY-45 and NOESY measurements. ^{13}C NMR spectra were recorded at 100 MHz, CDCl_3 (int. standard: TMS) and 75 MHz ($\text{CDCl}_3 + \text{CD}_3\text{OD}$) equipped with a dual probe using the same sample (in a 5 mm diameter NMR tube) utilized for protons studies.

GC/MS: system A: Kratos MS 80, on fused silica capillary column (25 m \times 0.22 mm) with OV-1, gas flow rate for He: 0.8 ml/min., temp. program 120–250 $^\circ$ (+5 $^\circ$ /min.); system B: MAT 312 instrument, on a capillary column (30 m \times 0.25 mm) with DB1, gas flow rate for He: 1 ml/min, temp. program: 150–260 $^\circ$ (+5 $^\circ$ /min.). TLC and HPTLC: precoated plate silica gel 60F₂₅₄. Solvent systems: for compound **1**, (a) EtOAc–MeOH–H₂O (77:15:8); for flavonoids, (b) EtOAc–HCOOH–HOAc–H₂O (100:11:11:26); for monosaccharides, (c) CHCl_3 –MeOH–H₂O (8:5:1). Spray reagents: for oligosaccharides, vanillin–H₂SO₄ (Godin) followed by heating; for flavonoids, Neu + PEG 400; for monosaccharides, diphenylamine–H₃PO₄ followed by heating. CC and flash CC: silica gel Merck, part. Size <0.063 mm, solvent systems (a) and (c). Prep. TLC: silica gel, 0.5 mm, (d) Me_2CO – CH_2Cl_2 –H₂O (16:4:1).

Plant material. *Polygala amarella* was cultivated in the Munich Botanical Garden and provided by Dr G. Heubl. A sample is

deposited in the herbarium of the Institute of Pharmaceutical Biology, Munich.

Isolation and purification of 1–4. Powdered air-dried aerial parts of *P. amarella* (500 g) were treated with 3 l petrol in a Soxhlet apparatus for 2 days, dried and extracted with 3 l MeOH (Soxhlet, 2 days). The MeOH extract was concd under red. pres. (89 g), dissolved in H₂O (1.5 l) and shaken successively with petrol, CH₂Cl₂, EtOAc and *n*-BuOH (3 × 400 ml for each solvent). Petrol extracts containing chlorophylls were discarded. CH₂Cl₂ and EtOAc extracts having the same composition on TLC (a) were collected together and concd *in vacuo* to afford a bitter residue (6.4 g). HPTLC: mixture of 6 Godin positive spots in the *R_f* region 0.3–0.7 (a). This extract was chromatographed over 300 g of silica gel (a) in two portions of 3.4 g each. Ten fractions were collected and the isolation of pure compound 1 (174 mg) was achieved by prep. TLC on silica gel (d) of the fractions II and III. The BuOH extract was concd to dryness to afford a brown residue (22 g). HPTLC: mixture of 8 Neu positive spots in the *R_f* region 0.19–0.40 (b). This extract was submitted to flash CC on silica gel (300 g). Elution with the solvent system EtOAc–MeOH–H₂O (21:4:3), yielded 7 fractions. Two of them (IV and V) were further separated by CC on silica gel (c) to afford pure compounds 3 (50 mg) and 4 (30 mg). The isolation of pure compound 2 (90 mg) was achieved by prep. TLC (CH₂Cl₂–MeOH, 1:1) of fraction V. Compounds were purified over a Sephadex LH-20 column prior to spectral analysis.

Acid hydrolysis. The compound (2 mg) was treated with 2 M CF₃COOH–MeOH (1:1) at 100° in a sealed tube for 1 hr. After repeated evapns of the solvent, the residue was taken up in H₂O and extracted with Et₂O. The aglycone was identified in the Et₂O extracts of compounds 3 and 4 by TLC (silica gel, C₆H₅Me–HCOOEt–HCOOH, 5:4:1) as quercetin and kaempferol respectively. Sugars were identified in the aq. phase by TLC (c) as rhamnose and glucose for compounds 3 and 4, glucose and fructose for compounds 1 and 2. Then 1-(trimethylsilyl)imidazole (0.2 ml) was added and the reaction mixture was heated at 70° in a stoppered test tube for 1 hr and then analysed by GC [3% Silicone OV-17 on Shimalite W (80–100 mesh) column: 3 mm × 2 m; gas pressure 1.51 kg/cm² air; column temp. 120°].

Acetylation. The compound (2 mg) was dissolved in Ac₂O–pyridine (1:1) and stirred at room temp. for 36 hr. The soln was evapd with toluene and the crude extract purified on Sephadex LH 20 with MeOH.

Permethylation according to Hakomori's method and preparation of the partially methylated alditol acetates. In a 10 ml injection bottle containing an atmosphere of N₂, 5 mg of compound 1 was dissolved in DMSO and Na-methylsulphinylmethanide in DMSO (2 ml) added dropwise. The mixture was sonicated for 30 min at 25° in a Bransonic 12 ultrasound bath (60 W, 50 kHz) and kept at room temp. Me I (1.5 ml) was added to this soln under cooling, and the soln further sonicated at room temp. for 1 hr. The excess of Me I was distilled off at 40°, 5 ml H₂O added and the soln extracted (4 × 5 ml) with CH₂Cl₂. The CH₂Cl₂ phases were washed with H₂O and then evapd to dryness. The residue was hydrolysed with 90% HCOOH (1 ml) and 1M CF₃COOH (1 ml) for 2 hr at 100°, the soln evapd to remove the acid, the residue dissolved in H₂O (2 ml) and 25 mg NaBD₄ added. After standing at room temp. for 2 hr the mixture was acidified with HOAc to pH 3.5 and the solvent evapd to dryness. H₃BO₃ in the residue was removed by three cycles of codistillation with MeOH. The partially methylated alditols were acylated with Ac₂O (1 ml) for 1 hr at 100°. The Ac₂O was removed by codistillation in the presence of toluene, and the residue dissolved in CH₂Cl₂ (3 ml), washed with H₂O (3 × 1 ml) and then evapd to dryness. The mixture thus obtained was

subjected to GC/MS (system A).

Permethylation using methyl trifluoromethanesulphonate in trimethyl phosphate. Compound 1 (5 mg) was dried over P₂O₅ and suspended in trimethyl phosphate (1 ml, Aldrich Chemical Co). To the clear soln 2,6-di-(*tert*-butyl)pyridine (150 μl) and methyltrifluoromethanesulphonate (100 μl) were added and allowed to react for 2 hr at 50°. The soln was then distributed between CHCl₃ (5 ml) and H₂O (20 ml). The CHCl₃ layer was separated by centrifugation at 1700 g for 5 min, concd. by evapn and applied to a Sephadex LH-20 column (MeOH). The recovered methylated carbohydrate was then treated according to the procedure described above. The partially methylated alditol acetates obtained were subjected to GC-MS analysis (system B).

β -D-(1,3,4-tri-*O*-benzoyl)-Fructofuranosyl- $\{[\beta$ -D-(6-*O*-acetyl)-glucopyranosyl(1 → 2)][β -D-glucopyranosyl(1 → 3)] $\}$ - α -D-(di-*O*-acetyl-4,6)-glucopyranoside. Amarelloside (1), Amorphous powder; mp 137–142° with decompn; $[\alpha]_D^{25} -12^\circ$ (MeOH; *c* 0.1); TLC (silica gel), (a): *R_f* 0.60; blue grey spots by spraying with Godin reagent; IR ν_{\max}^{KBr} cm⁻¹: 3500–3300 (OH), 1725 (C=O ester), 1630 (C=C), 1600, 1450, 1380; UV $\lambda_{\max}^{\text{MeOH}}$ nm: 203, 230, 280 (sh); FABMS (negative ion, thioglycerol) *m/z*: 1103 [M–H]⁻, 1061, 999, 957, 899, 795, 753, 691, 633, 591, 529, 467, 387, 367; ¹H NMR (300 MHz, CDCl₃ + CD₃OD): see Table 1; ¹³C NMR (100 MHz, CDCl₃): see Table 2. After GC-MS analysis of the partially methylated alditol acetates (methylation according to Hakomori's method) the following carbohydrates could be assigned for 1: system A: terminal fructose (2,5-di-*O*-acetyl-1,3,4,6-tetra-*O*-methylhexitol, *R_t* = 2.31 min., *m/z* 43, 45, 71, 87, 101, 129, 145, 161, 205); terminal glucose (1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methylhexitol, *R_t* = 3.19 min, *m/z* 43, 71, 75, 101, 117, 118, 129, 145, 161, 162, 205); 1,2,3-linked glucose (1,2,3,5-tetra-*O*-acetyl 4,6-di-*O*-methylhexitol, *R_t* = 5.45 min., *m/z* 43, 71, 85, 87, 101, 129, 161, 187, 201, 261). After methylation analysis (according to the method with methyl trifluoromethanesulphonate in trimethyl phosphate) the following carbohydrates could be assigned for 1: system B: terminal glucose (*R_t* = 23.50 min.), terminal 6-*O*-acetylglucose (1,5,6-tri-*O*-acetyl-2,3,4-tri-*O*-methylhexitol, *R_t* = 25.80 min, *m/z* 23, 71, 75, 87, 99, 101, 102, 118, 129, 162, 173, 189, 233); 4,6-di-*O*-acetyl-1,2,3-linked-glucose (hexitolhexaacetate, *R_t* = 30.10 min, *m/z* 43, 70, 85, 103, 115, 128, 145, 171, 187, 217, 289, 290, 361, 362); terminal 1,3,4-tri-*O*-benzoyl-fructose (1,3,4-tri-*O*-benzoyl-2,5-di-*O*-acetyl-6-*O*-methylhexitol, *R_t* = 44 min, *m/z* 43, 45, 57, 71, 77, 101, 105, 129, 161, 190, 207, 250, 385).

β -D-(1,3-disinapoyl)-Fructofuranosyl- α -D-glucopyranoside (2). The experimental results were identical with those described in ref. [10].

Rutin (3) and kaempferol-3-*O*-rutinoside (4) were identified by comparison of UV, chromatographic properties and MS data with those of standard compounds [25].

Acknowledgements—The authors thank Dr O. Seligmann, (Institute of Pharmaceutical Biology, University of Munich) for recording the FAB mass spectra, Dr W. Schäffer (Max-Planck-Institut für Biochemie, Martinsried, Munich) and Dr C. Tsaonas, (Faculty of Medicine, Dijon) for recording GC/MS spectra, Dr Dabrowsky (University of Budapest for the possibility of measuring the ¹H NMR spectrum at 500 MHz on his facility, Dr B. Hanquet, (CEREMA, Faculty of Sciences, Dijon) for the ¹³C NMR spectrum and DEPT experiments, Dr C. Lavaud, (Faculty of Pharmacy, Reims) for recording the two dimensional heterocorrelated ¹H/¹³C NMR spectrum and Prof. F. Tillequin, (Faculty of Pharmacy, Paris V), for helpful discussions.

REFERENCES

1. Haerdi, F., Kerharo, J. and Adam, J. G. (1964) in *Afrikanische Heilpflanzen, Acta Tropica Suppl.* **8**, 77.
2. Hamburger, M. and Hostettmann, K. (1986) *J. Nat. Prod.* **49**, 557.
3. Bax, A., Freeman, R. and Morris, G. (1981) *J. Magn. Res.* **42**, 164.
4. Freeman, R. and Morris, G. (1978) *J. Chem. Soc. Chem. Commun.*, 684.
5. Lanzetta, R., Laonigro, G., Parilli, M. and Breitmaier, E. (1984) *Can. J. Chem.* **62**, 2874.
6. Bock, K. and Thogersen, H. (1982) *Ann. Rep. NMR Spectrosc.* **13**, 1.
7. De Bruyn, A., Van Beeumen, J., Anteunis, M. and Verhegge, G. (1975) *Bull. Soc. Chim. Belg.* **84**, 799.
8. Massiot, G., Lavaud, C., Le Men-Olivier, L., Van Binst, G., Miller, S. P. F. and Fales, H. M. (1988) *J. Chem. Soc., Perkin Trans I*, 3071.
9. Christofides, J. C. and Davies, D. B. (1984) *J. Chem. Soc., Perkin Trans II*, 481.
10. Hamburger, M. and Hostettmann, K. (1985) *Phytochemistry* **24**, 1793.
11. Gorin, R. A. J. and Mazureck, M. (1975) *Can. J. Chem.* **53**, 1212.
12. Yoshimoto, K., Itatani, Y. and Tsuda, Y. (1980) *Chem. Pharm. Bull.* **28**, 2065.
13. Usui, T., Yamaoka, N., Matsuda, K., Tuzimura, K., Sugiyama, H. and Seto, S. (1973) *J. Chem. Soc., Perkin Trans I*, 2425.
14. Hoffman, R. E., Christophides, J. C., Davies, D. B. and Lawson, C. J. (1986) *Carbohydr. Res.* **153**, 1.
15. Hakomori, S. I. (1964) *J. Biochem. (Tokyo)* **55**, 205.
16. Jansson, P. E., Kenne, L., Liedgren, H., Lindberg, B. and Lönngren, J. (1976) *Chem. Commun. (Stokholm Univ.)* **8**, 1.
17. Prehm, P. (1980) *Carbohydr. Res.* **78**, 372.
18. Kumar, A., Ernst, R. R. and Wüthrich, K. (1980) *Biophys. Res. Commun.* **95**, 1.
19. Nakano, K., Murakami, K., Takaishi, Y. and Tomimatsu, T. (1986) *Chem. Pharm. Bull.* **34**, 5005.
20. Strack, D., Sachs, G., Römer, A. and Wiermann, R. (1981) *Z. Naturforsch. C.* **36**, 721.
21. Linscheid, M., Wendisch, D. and Strack, D. (1980), *Z. Naturforsch. C.* **35**, 907.
22. Fukuyama, Y., Sato, T., Miura, I., Asakawa, Y. and Takemoto, T. (1983) *Phytochemistry* **22**, 549.
23. Shimomura, H., Sashida, Y. and Mimaki, Y. (1986) *Phytochemistry* **12**, 2897.
24. Shoyama, Y., Hatano, K., Nishioka, T. and Yamagishi, T. (1987) *Phytochemistry* **26**, 2965.
25. Mabry, T. J., Markham, K. R. and Thomas, M. B. (1970) *The Systematic Identification of Flavonoids*. Springer, Berlin.