Phytochemistry, Vol. 26, No. 3, pp. 787–791, 1987.
Printed in Great Britain.

0031-9422/87 $3.00 + 0.00
© 1987 Pergamon Journals Ltd.

STRUCTURE OF TWO ANTIVIRAL TRITERPENE SAPONINS FROM *ANAGALLIS ARVENSIS*

M. AMOROS and R. L. GIRRE

Laboratoire de Pharmacognosie, Faculté de Pharmacie, Avenue du Pr. Léon-Bernard, 35043 Rennes Cédex, France

(*Revised received* 23 *July* 1986)

Key Word Index – *Anagallis arvensis*; Primulaceae; triterpenes; saponins; antiviral compounds.

Abstract—From *Anagallis arvensis* two novel saponins were isolated. On the basis of the spectral analysis, the structure proposed for one of the saponins was the 3-*O*-glucose (1 → 3 or 4)[arabinose 1 → 4 or 3]-glucose (1 → 2)-xyloside of 23-hydroxy protoprimulagenin A. The other saponin contained an additional glucose.

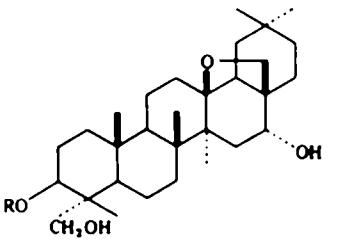
INTRODUCTION

Anagallis arvensis has been reported to possess *in vitro* antiviral activity against herpes simplex virus type 1 and poliovirus [1, 2]. The antiviral compounds were highly haemolytic, gave a copious lather when they were shaken with water and produced positive colour tests for triterpenoids. The isolation of two triterpene saponins from this plant and the spectral evidence leading to the elucidation of their structure are discussed in this paper.

RESULTS AND DISCUSSION

Saponins were extracted from the whole plant by methods described in the Experimental. They are assigned as **1** and **2** in order of increasing polarity on TLC. Both saponins on acid hydrolysis yielded the same aglycone **3** besides glucose, arabinose and xylose as sugar moieties.

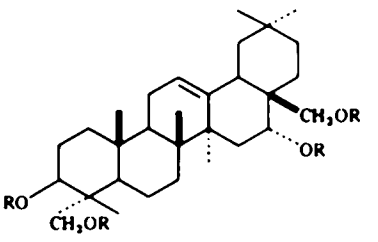
The mass spectrum of compound **3** showed a molecular ion at *m/z* 474 ($C_{30}H_{50}O_4$) and a base peak at *m/z* 219 ($C_{15}H_{23}O$). The IR spectrum exhibited absorptions at ν_{max}^{KBr} cm^{-1} 3400 (–OH), 1640 (CH=CH), 1450 (methyl) and 1380 (*gem*-dimethyl). Acetylation gave the tetraacetate **3a**, M$^+$ at *m/z* 642 ($C_{38}H_{58}O_8$); IR ν_{max}^{KBr} cm^{-1} 1730, 1240, 1360 (alcoholic acetate). ^{13}C NMR chemical shifts of **3** and **3a** are shown in Table 1. The presence of six quaternary carbon atoms and the chemical shifts of C-12 at δ122.3 and C-13 at 144.5 were characteristic of a Δ^{12}-oleanene skeleton [3]. The signal of C-16 at δ73.9 (**3**) and δ75.8 (**3a**) suggested an axial position for the 16-OH [4]. The shifts of C-3 at δ74.1 (**3**) and 74.5 (**3a**) and the shielding of C-24 (δ12.9) were in agreement with a β-equatorial position for the C-3 hydroxyl, the methyl group being β axial [5–8]. The 60 MHz ^{1}H NMR spectrum in C_5D_5N of **3** displayed an ethylenic proton (12-H) at δ5.4 and accounted for six tertiary methyl groups. Five were at δ1.03, 1.08 and 1.12 (15 H). The C-27 methyl group was observed at the lowest fields δ1.77 (**3**) and 1.43 (**3a**), due to its homoallylic disposition, this shift confirming the configuration of the 16α-hydroxy and 14α-methyl groups [9–11]. The signal due to 16β-H was observed as a broad singlet at δ4.59 (**3**) and 5.49 (**3a**). The ^{1}H NMR spectrum of tetraacetate **3a** showed the presence of two



1 R = —xyl—glc—ara
glc
glc

xyl: xylose
glc: glucose
ara: arabinose

2 R = —xyl—glc—ara
glc



3 R = H
3a R = Ac

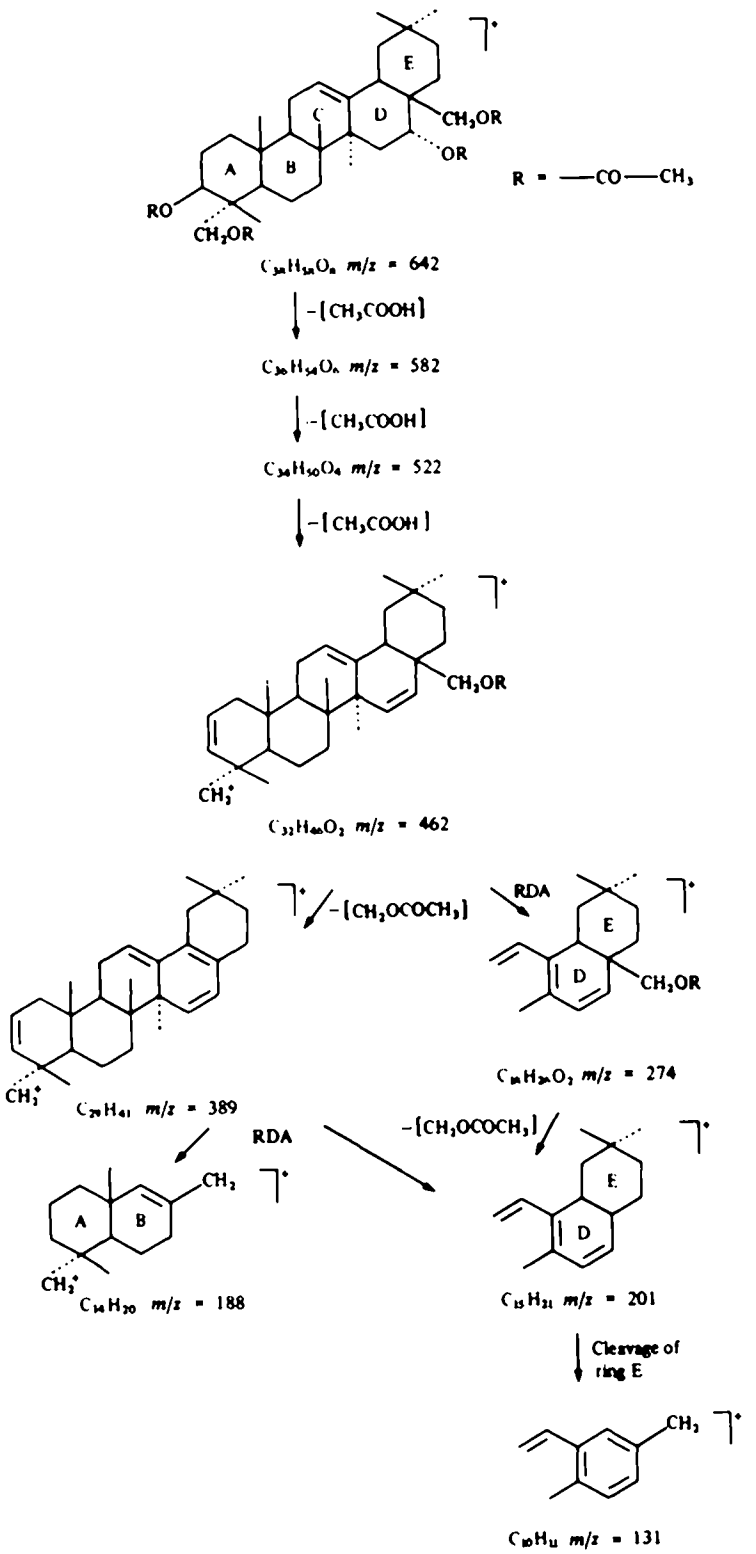
* The linkage positions on the inner glucose of terminal arabinose and terminal glucose were ambiguous and may be reversed.

Table 1. ^{13}C NMR chemical shifts of aglycone 3, its acetate 3a, saponin 2 and glycosidation shifts $\Delta\delta$ (in parentheses in ppm)

	3	3a	2	$\Delta\delta$
C-1	38.9	38.1	39.2	(+0.3)
C-2	27.4	23.3	25.9	(−1.5)
C-3	74.1	74.5	82.5	(+8.4)
C-4	42.7	41.4	43.6	(+0.9)
C-5	48.9	48.4	47.9	(−1)
C-6	18.6	18.1	17.7	(−0.9)
C-7	33.0	32.7	32.9	(−0.1)
C-8	40.1	40.1	42.4	(+2.3)
C-9	47.2	47.1	50.6	(+3.4)
C-10	37.1	36.9	36.8	(+0.3)
C-11	23.8	23.6	19.3	(−4.5)
C-12	122.3	122.3	36.9	(−85.4)
C-13	145.2	142.3	86.4	(−58.8)
C-14	42.0	41.0	44.6	(+2.6)
C-15	34.7	31.3	34.2	(−0.5)
C-16	73.9	75.8	77.5	(+3.6)
C-17	40.9	38.3	44.6	(+3.7)
C-18	42.6	42.4	51.5	(+8.9)
C-19	48.2	47.4	38.9	(−9.3)
C-20	31.1	30.7	31.8	(+0.7)
C-21	37.1	35.9	37	(−0.1)
C-22	30.2	30.0	31.8	(+1.6)
C-23	68.6	65.6	67.4	(−1.2)
C-24	12.9	13.2	13.2	(+0.3)
C-25	16.2	16.1	17	(+0.8)
C-26	17.1	16.9	19.5	(+2.4)
C-27	27.3	26.7	18.6	(−8.7)
C-28	70.2	71.4	78	(+7.8)
C-29	33.3	33.3	33.7	(+0.4)
C-30	24.9	24.3	24.7	(−0.2)
O-CO-Me		170.8		
		170.6		
		170.5		
		170.0		
O-CO-Me		21.7		
		21.0		
		20.6		
C-1′ (Xyl)			107.4	
C-2′ (Xyl)			80.1	
C-3′ (Xyl)			78.0	
C-4′ (Xyl)			71.4	
C-5′ (Xyl)			64.9	
C-1″ (Glc)			103.9	
C-2″ (Glc)			77.2	
C-3″ (Glc)			85.2	
C-4″ (Glc)			78.3	
C-5″ (Glc)			78.2	
C-6″ (Glc)			62.7	
C-1‴ (Glc)			103.6	
C-2‴ (Glc)			76.0	
C-3‴ (Glc)			78.2	
C-4‴ (Glc)			71.0	
C-5‴ (Glc)			77.7	
C-6‴ (Glc)			61.3	
C-1⁗ (Ara)			104.9	
C-2⁗ (Ara)			73.3	
C-3⁗ (Ara)			75.95	
C-4⁗ (Ara)			70.7	
C-5⁗ (Ara)			64.2	

primary acetoxyl groups, an AB system at δ3.94 and 4.26 (J = 11 Hz, 23 CH_2OAc) [5] and a singlet at 4.01 (28 CH_2OAc). The signal at 5.03 (t-like) in the ^1H NMR spectrum of 3a could be attributed to the 3α-proton geminal to the 3β-hydroxyl group [12–14]. The mass spectrum of 3 and 3a confirmed the Δ^{12}-oleanene structure. The typical retro-Diels-Alder (RDA) fragmentation of ring C [15] resulted in characteristic fragment ions as shown in Scheme 1. Therefore, compound 3 was concluded to be 3β,16α,23,28-tetrahydroxyolean-12-ene. This compound is closely related to sapogenins extracted from other Primulaceae [14, 16] and similar to 23-hydroxyprimulagenin A isolated from *Bupleurum falcatum* (Umbelliferae) [6].

The negative ionization FAB mass spectrum of the saponin 2 showed the molecular peak $[M-H]^-$ at m/z 1061. Glucose, xylose and arabinose were found in an acid hydrolysate and partial hydrolysis on TLC showed glucose and arabinose as terminal sugars. Since the molecular weight of the sapogenin was 474, the relative amounts of the sugars were xylose: 1, arabinose: 1, glucose: 2. This was in agreement with the mass spectral fragmentation pattern with peaks at m/z (rel. int.) 1061 $[M-H]^-$ (100); 929 $[M-Ara]^-$ (16); 899 $[M-Glc]^-$ (10), 767 $[M-Ara-Glc]^-$ (21); 605 $[M-Ara-2\,Glc]^-$ (15); 473 $[M-Ara-2\,Glc-Xyl]^-$ (20). The simultaneous losses of fragment 133 (arabinose) and fragment 162 (glucose) were indicative of two terminal sugars: an arabinose moiety and a glucose moiety, suggesting two sites of attachment or a branched side-chain. The positive ionization FAB MS confirmed this result with peaks at m/z (rel. int.) 1147 $[M+{}_2Na+K]^+$ (1.6), 1108 $[M+{}_2Na]^+$ (15), 1085 $[M+Na]^+$ (100), 953 $[(M+Na)-Ara]^+$ (11), 923 $[(M+Na)-Glu]^+$ (9), 791 $[(M+Na)-Glu-Ara]^+$ (11), 628 $[(M+Na)-Ara-2\,Glu]^+$ (16), 473 $[(M)-Xyl-2\,Glu-Ara]^+$ (5). ^{13}C NMR chemical shifts (250 MHz, C_5D_5N, TMS as internal standard) are shown in Table 1. The 'up-down' ^{13}C NMR spectrum did not give signals at δ122 and δ144 but showed seven quaternary carbons (six for compound 3). The most important changes affected C-12, C-13, C-19 and C-27 which were shifted upfield whereas the C-14, C-16, C-17, C-18 and C-28 signals were shifted downfield. These chemical shifts were very close to those indicated for compounds which possess an oxide ring involving C-28 and C-13, for example androsacenol [17], saikogenin G [4], saikosaponin d [6] and saikosaponin from *Bupleurum longeradiatum* [18]. Moreover, the IR spectrum of 2 showed a sharp band at 890 cm^{-1} characteristic of an ether linkage [13, 17]. This suggested a 13β,28-oxide structure for the aglycone moiety of 2. On acid hydrolysis such a bridged system which has been described in Primulaceae [13, 17] and Umbelliferae [9, 18] would be readily converted to a Δ^{12}-17-CH_2OH moiety [17] which would explain the structure of 3. Thus, the structure of the aglycone moiety of saponin 2 was established as 3β,16α,23-trihydroxy-13β,28-epoxyoleanane or 23-hydroxyprotoprimulagenin A. The shifts of C-3, C-16 and C-23 of saponin 2 were compared with signals available for the corresponding positions in compound 3. The single C-3 signal was found to be shifted downfield by 8.4 ppm, indicating the site of glycosidation. Consequently, saponin 2 would possess only one saccharidic chain which is branched and located at C-3. The assignment of carbon signals due to the common sugar moiety of 2 was carried out by comparison with those


E
C
D
A
B
RO
CH₂OR
CH₂OR
OR
R = —CO—CH₃
C₃₆H₅₈O₈ m/z = 642
-[CH₃COOH]
C₃₄H₅₄O₆ m/z = 582
-[CH₃COOH]
C₃₂H₅₀O₄ m/z = 522
-[CH₃COOH]
CH₂OR
CH₃
C₃₁H₄₈O₂ m/z = 462
-[CH₂OCOCH₃]
RDA
E
D
CH₂OR
C₂₀H₄₁ m/z = 389
C₁₈H₂₆O₂ m/z = 274
RDA
-[CH₂OCOCH₃]
CH₂
A
B
CH₃
C₁₄H₂₀ m/z = 188
E
D
C₁₅H₂₁ m/z = 201
Cleavage of ring E
CH₃
C₁₀H₁₁ m/z = 131

Scheme 1.

reported for the saccharide chain of related saponins [7, 19–22], chemical shifts of pure sugars and previous knowledge of glycosidation effects. The anomeric carbon signals at δ107.4, 103.9, 103.6 and 104.9 confirmed the presence of four monosaccharide units which are substituted on C-1. Thus, the carbon chemical shifts of the terminal arabinose and of the terminal glucose could be attributed. The inner glucose unit must not be linked at C-6 since the signal at δ62 which would correspond to the $-CH_2OH$ group of an unsubstituted glucopyranoside was present. C-3″ and C-4″ of the inner glucose were shifted downfield by *ca* 7 ppm indicating glycosidic linkages but we did not discriminate which one of the two terminal sugars was 1 → 3 and 1 → 4 linked. In the same way, the signal at δ80.1 was attributed to C-2′ of the inner xylose bonded to the anomeric carbon of the inner glucose. For the four sugars, the distribution of chemical shifts was in agreement with a pyranoside structure. Therefore, the saponin isolated from *Anagallis arvensis* would be the 3-*O*-glucose (1 → 3 or 1 → 4)-[arabinose (1 → 4 or 1 → 3)]-glucose (1 → 2)-xyloside of 23-hydroxyprotoprimulagenin A.

Incubation of compound **1** with β-glucosidase afforded glucose together with saponin **2**. Negative FAB-MS of saponin **1** gave a well-defined molecular ion at *m/z* 1223. The fragmentation pattern shown in Table 2 involved successive expulsions of sugar molecules. The first signals at *m/z* 1091 (12%) and 1062 (12%) corresponded to the simultaneous losses of a terminal arabinose and a terminal glucose. The other important ions were at *m/z* 767 (20%) and 605 (80%) showing the further elimination of two glucose and one xylose units to give aglycone peak $[M-H]^-$ at *m/z* 473. By comparison with **2**, this fragmentation mechanism permitted the assignment of the site of the additional glucose on the glucosyl moiety but not on the arabinosyl moiety. Positive FAB-MS was in agreement with this result, the molecular peak $[M+Na]^+$ appeared at *m/z* 1247. In a comparison of the ^{13}C NMR spectra of compounds **1** and **2**, signals due to the aglycone moiety of **1** were superimposable on those of **2**, which confirmed the identity of the aglycone moiety of both saponins. Thus, **1** could be formulated as the 3-*O*-glucose-glucose-[arabinose]-glucose-xyloside of 23-hydroxyprotoprimulagenin A.

EXPERIMENTAL

NMR spectra were taken in C_5D_5N using TMS as an internal standard: ^1H NMR at 60 MHz, ^{13}C NMR at 20 MHz for sapogenin and its acetyl derivative (**3** and **3a**), ^{13}C NMR at 250 MHz for saponins **1** and **2**. Mass spectra of **3** and **3a** were recorded by direct inlet at 70 eV ionization. Mass spectra of saponins **1** and **2** were measured by FAB-MS using a ZAB HF combined with a data system. The solvent used was PEG 400. Spectra in positive ionization and in negative ionization were recorded.

Plant material. Anagallis arvensis was collected in the suburbs of Rennes, France, in Summer.

Extraction and isolation of saponins. Dried and milled whole plant was exhaustively defatted and depigmented with $CHCl_3$, then repeatedly extracted with MeOH at room temp. The MeOH extract was evaporated to dryness and a soln of the residue (yield 19.4%) in H_2O was washed with EtOAc and then extracted × 3 with 1-BuOH saturated with H_2O. The BuOH was removed *in vacuo* affording crude saponin extract (yield 4.1%) which was subjected to CC on silica gel (70–230 mesh Merck) with $CHCl_3$ MeOH (1:1) as the eluant. The major saponins **1** and **2** were further submitted to droplet counter current chromatography (DCCC, Rikakikay Co Ltd, DCCC-A); $CHCl_3$ MeOH–*n*-PrOH–H_2O (9:12:1:8) ascending mode. Eluates chromatographed on silica gel TLC ($CHCl_3$–MeOH H_2O 65:25:4) gave analytical material.

Acid hydrolysis of the saponins. Saponin **1** or **2** (1 g) in 0.5 N HCl (100 ml) was heated on a steam bath for 90 min, cooled, diluted with H_2O and extracted × 3 with $CHCl_3$ MeOH (9:1). The H_2O layer was lyophilized and monosaccharides analysed by TLC on cellulose, solvents: EtOAc C_5H_5N–CH_3COOH–H_2O (36:36:7:21) and H_2O PhOH (1:5). Glucose, xylose and arabinose were identified. The $CHCl_3$ extracts were combined, washed with H_2O and dried. Isolation of the sapogenin (compound **3**) was carried out using a silica gel column and eluting with $CHCl_3$ MeOH (9:1).

Compound **3**. MS *m/z* (rel. int.): 474 $[M]^+$ (12), 456 (8), 250 (32), 232 (20), 223 (28), 219 (100), 201 (60), 189 (20), 131 (24); IR ν_{max}^{KBr} cm^{-1}: 3400, 2940, 1640, 1450, 1380, 1045, 1000, 760; ^1H NMR (C_5D_5N): δ1.03, 1.08, 1.12 (15 H, *s*, 5 × Me); 1.77 (3 H, *s*, Me) 2.59 (1 H, *br s*, 18-H), 3.69 (1H, *s*, 3α-H), 3.69, 3.74, 4.08, 4.22 (4H, *m*, 23-H), 4.59 (2H, *br s*, 16β-H), 5.4 (1H, *s*, 12-H). ^{13}C NMR data: see Table 1.

Acetylation of **3**. The aglycone **3** (300 mg) was refluxed for 2 hr with Ac_2O (10 ml) in pyridine (10 ml). The reaction mixture was evaporated to dryness. Column chromatography of the tetra-acetate on silica gel eluting with $CHCl_3$ MeOH (95:5) gave pure compound **3a**. MS *m/z* (rel. int): 642 $[M]^+$ (2), 582 (6), 522 (20), 462 (4), 389 (8), 274 (22), 201 (100), 188 (18), 131 (20); IR ν_{max}^{KBr} cm $^{-1}$: 2940, 1760, 1450, 1380, 1240, 1045, 780; ^1H NMR: δ0.84, 0.92, 0.98 (5 × Me), 1.43 (Me), 2.03, 2.13 (4 × $_3$H, *s*, CH_3COO) 3.94, 4.26 (AB system, J = 11 Hz, 23 CH_2OAc) 4.01 (2H, *s*, 28-CH_2OAc), 5.03 (1H, *t*-like, 3α-H) 5.35 (1H, *br s*, 12-H) 5.49 (1H, *br s*, 16β-H). ^{13}C NMR data: see Table 1.

Partial acid hydrolysis on TLC [23]. Saponins **1** or **2** were applied on silica gel TLC and left in an HCl atmosphere at room temp. for 60 min. HCl vapour was eliminated under hot ventilation then authentic sugars were applied to the plate. The plate was developed with the solvent system $CHCl_3$ MeOH H_2O (6.4:4:0.8) and spots detected by spraying:

Table 2. Negative ion FAB-mass spectrum of saponin **1**

m/z		Rel. int.	Interpretation
1223		100	$[M-H]^-$ (for $C_{58}H_{96}O_{27}$)
1091	(1223 − 132)	12	$[M-Ara]^-$
1062	(1223 − 162)	12	$[M-Glc]^-$
767	(1223 − 132 − 2 × 162)	20	$[M-Ara-2\,Glc]^-$
605	(1223 − 132 − 3 × 162)	80	$[M-Ara-3\,Glc]^-$
473	(1223 − 2 × 132 − 3 × 162)		$[M-Ara-3\,Glc-Xyl]^-$ = aglycone

aniline–diphenylamine–H_3PO_4–MeOH (2:2:10:96), followed by heating. For both saponins two terminal sugars appeared which were identified as glucose and arabinose.

Enzymatic hydrolysis of **1**. A mixture of **1** (3 mg), β-glucosidase (3 mg) and phosphate buffer pH 5 (0.75 ml) was left to stand at 37° for 24 hr. Control TLC using BuOH–HOAc–H_2O (6:1:2) as eluant showed that compound **1** gave saponin **2** together with glucose.

Acknowledgements We wish to thank Dr. Fraisse, C.N.R.S. Vernaison for recording FAB mass spectra of saponins and Dr. Tabet, Ecole Polytechnique d'Orsay for their helpful suggestions; Dr. Guenno and Dr. Sinbandhit, Faculty of Sciences, Rennes, for measurements of ^{1}H NMR, ^{13}C NMR and mass spectra and useful discussions. This work has been supported by a grant from Fondation Langlois which is gratefully acknowledged.

REFERENCES

1. Amoros, M., Fauconnier, B. and Girre, L. (1977) *Ann. Pharm. Fr.* **35**, 371.
2. Amoros, M., Fauconnier, B. and Girre L. (1979) *Pl. Med. Phytoth.* **XIII**, 122.
3. Doddrell, D. M., Khong, P. W. and Lewis, K. G. (1974) *Tetrahedron Letters* **27**, 2381.
4. Tori, K., Yoshimura, Y., Seo, S., Sakurawi, K., Tomita, Y. and Ishii, H. (1976) *Tetrahedron Letters* **46**, 4163.
5. Sousa, M. P., Matos, M. E. O., Machado, M. I. L., Braz Filho, R., Vencato, I. and Mascarenhas, Y. P. (1984) *Phytochemistry* **23**, 2589.
6. Tori, K., Seo, S., Yoshimura, Y., Nakamura, M., Tomita, Y. and Ishii, H. (1976) *Tetrahedron Letters* **46**, 4167.
7. Kizu, H. and Tonimori, T. (1982) *Chem. Pharm. Bull.* **30**, 3340.
8. Sakakibara, J., Kaiya, T. and Fukuda, H. (1984) *Phytochemistry* **23**, 627.
9. Kitagawa, I., Yoshikawa, M. and Yosioka, I. (1974) *Tetrahedron Letters* **5**, 469.
10. Alexander, R., Croft, K. D., Kagi, R. I. and Shea, S. (1978) *Aust. J. Chem.* **31**, 2741.
11. Becchi, M., Bruneteau, M., Trouilloud, M., Combier, H., Sartre, J. and Michel, G. (1979) *Eur. J. Biochem.* **102**, 11.
12. Woo, W. S., Kang, S. S., Wagner, H., Seligmann, O. and Chari, M. (1978) *Planta Med.* **34**, 87.
13. Heitz, S., Billet, D. and Raulais, D. (1971) *Bull. Soc. Chem. Fr.* **6**, 2320.
14. Kitagawa, I., Matsuda, A. and Yosioka, I. (1972) *Chem. Pharm. Bull.* **20**, 2226.
15. Budzikiewicz, H., Wilson, J. M. and Djerassi, C. (1963) *J. Am. Chem. Soc.* **85**, 3688.
16. Boiteau, P., Pasich, B. and Ratsimamanga, A. R. (1964) Gauthier-Villars, Paris.
17. Pal, B. C., Roy, G. and Mahato, S. B. (1984) *Phytochemistry* **23**, 1475.
18. Kimata, H., Rasai, R. and Tanaka, O. (1982) *Chem. Pharm. Bull.* **30**, 4373.
19. Kitagawa, I., Wang, H. K. and Yoshikawa, M. (1983) *Chem. Pharm. Bull.* **31**, 716.
20. Kitajima, J., Komori, T., Kawasaki, T. and Schulten, H. R. (1982) *Phytochemistry* **21**, 187.
21. Oshima, Y., Ohsawa, T. and Hikino, H. (1984) *Planta Med.* **3**, 254.
22. Mizutani, K., Ohtani, K., Wei, J. X., Kazai, R. and Tanaka, O. (1984) *Planta Med.* 327.
23. Le Turdu, M. (1984) Thèse Doctorat 3eme Cycle, Université de Nantes.