Cyclic Peptides from the Seeds of *Annona glauca* and *A. cherimola*

Alassane WELE,*[,a,b] Yanjun ZHANG,[b] Lionel DUBOST,[b] Jean-Louis POUSSET, and Bernard BODO[b]

[a] *Laboratoire de Chimie Organique et Thérapeutique, Faculté de Médecine et de Pharmacie, Université Cheikh; Anta Diop Dakar-Fann, Sénégal: and* [b] *Laboratoire de Chimie des Substances Naturelles; ESA 8041 CNRS, Muséum National d'Histoire Naturelle, 63 rue Buffon, 75005 Paris, France.* Received September 5, 2005; accepted January 6, 2006

From the seeds of *Annona glauca* and *A. cherimola*, two novel cyclic peptides, glaucacyclopeptide B and cherimolacyclopeptide G, have been isolated, respectively. Their structures were elucidated by chemical and spectral methods.

Key words Annonaceae; seed; cyclopeptide; glaucacyclopeptide B; cherimolacyclopeptide G

Preliminary phytochemical studies on *Annona* members have revealed the occurrence of constituents representing cyclopeptides. In preceding papers, we isolated a series of cyclopeptides, glaucacyclopeptide A and cherimolacyclopeptides A, B, C, D, E and F from the methanol extract of the seeds of *Annona glauca* and *A. cherimola*, small trees which grow in Africa, America and Asia and are now cultivated for their edible fruits in Europe.[1—10] This paper reports the isolation and the sequence determination by mass spectrometry and 2D-NMR of two cyclic peptides, glaucacyclopeptide B and cherimolacyclopeptide G.

Results and Discussion

A combination of Sephadex LH-20 gel and silica gel chromatographies, and finally HPLC, gave glaucacyclopeptide B and cherimolacyclopeptide G. IR maximal absorptions at 3320 and 1650 cm^{-1} and a negative ninhydrin reaction indicated that these compounds might be cyclic peptides. Analysis of the total acidic hydrolysate of the compounds, after derivatization, showed the presence of Gly (2), Ile (1), Leu (1), Met (1), Pro (1) and Tyr (1) for glaucacyclopeptide B and Ala (2), Gly (1), Ile (1), Pro (2), Tyr (1) and Val (1) for cherimolacyclopeptide G. The analysis of amino acids by gas chromatography of their methylated derivatives indicated that all the chiral amino acids were L.

The molecular weight and empirical molecular formula of these compounds was determined by an ESI-TOF mass spectrometer. Glaucacyclopeptide B gave a protonated molecular $[M+H]^+$ ion at *m/z* 732 and $[M+Na]^+$ adduct ion at *m/z* 754, corresponding to molecular weight 731. According to the amino acid composition, the empirical molecular formula $C_{35}H_{53}N_7O_8S$ was deduced. Similarly, the molecular weight of 768 for cherimolacyclopeptide G was deduced from the positive ESI-TOF spectrum, which displayed the $[M+Na]^+$ adduct ion at *m/z* 791 and the protonated molecular $[M+H]^+$ ion at *m/z* 769, corresponding to the molecular formula $C_{38}H_{56}N_8O_9$.

The CID experiment on the protonated molecular $[M+H]^+$ ion of glaucacyclopeptide B at *m/z* 732 allowed the sequence determination. The ring opening occurred at the Ile/Leu-Pro amide bond level, and a series of adjacent acylium ions (b_n) at *m/z* 619, 456, 343, 286 and 155 was generated. The successive loss of Ile/Leu, Tyr, Leu/Ile, Gly, Met was observed, yielding to the *N*-terminal dipeptide Pro-Gly (Fig. 1A). A second series of ions was depicted at *m/z* 591, 428, 315, 258 and 127, which were assigned to adjacent (a_n) series related to the above b_n ion series. These results suggested the sequence [H-Pro[1]-Gly[2]-Met[3]-Gly[4]-Ile/Leu[5]-Tyr[6]-Ile/Leu[7]]$^+$ for the linearized peptide ion derived from this compound, but with ambiguity related to the respective location of Leu and Ile.

Cherimolacyclopeptide G was characterized by the presence of two proline residues, which gave two specific fragmentations. The protonated molecular formula ion $[M+H]^+$ at *m/z* 769 was subjected to a collisional induced decomposition (CID) experiment. The ring opening at the -Ala[8]-Pro[1]- bond level gave a series of adjacent acylium ions (b_n) at *m/z* 698, 535, 422, 325, 226 and 155 corresponding to the successive loss of Ala, Tyr, Ile, Pro, Val, Ala and yielding the *N*-terminal dipeptide Pro-Gly (Fig. 1B). A second linearized peptide was formed from the $[M+H]^+$ ion at *m/z* 769, due to the cleavage at the Val[4]-Pro[5] amide bond level, and the resulting b'_n and a'_n ions were detected (Fig. 1C). The b'_n ions serie was characterized by ions at *m/z* 670, 599, 445, 374 and 211 corresponding to the successive loss of Val, Ala, Gly-Pro, Ala, Tyr, and yielding to the *N*-terminal dipeptide Pro-Ile/Leu, thus determining the structure *cyclo*(Pro[1]-Gly[2]-Ala[3]-Val[4]-Pro[5]-Ile[6]-Tyr[7]-Ala[8]) for the natural cyclooctapeptide.

The sequence of glaucacyclopeptide B was determined on the basis of the HMBC experiment using the connectivities between the carbonyl of residue *i* with the amide and/or α protons of residue *i*+1. The $^3J_{CH}$, CO(*i*) to NH(*i*+1) and other significant correlations depicted on the HMBC spectrum are shown in Fig. 2A. Thus, the sequence was confirmed using the homonuclear method described by

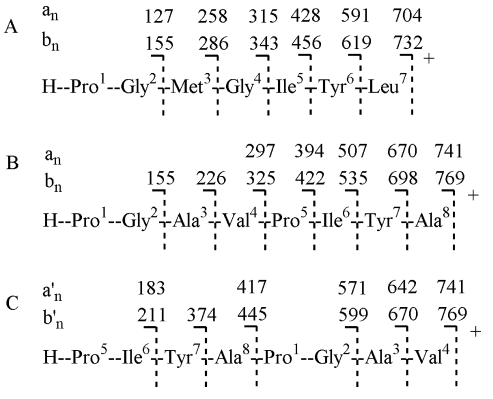

Fig. 1. CID Fragmentation of the Cyclopeptides

(A) Cleavage of the glaucacyclopeptide B at the Leu[7]-Pro[1] amide bond level, (B) cleavage of the cherimolacyclopeptide G at the Ala[8]-Pro[1] amide bond level, (C) cleavage of the cherimolacyclopeptide G at Val[4]-Pro[5].

* To whom correspondence should be addressed. e-mail: alassanewele@yahoo.fr

© 2006 Pharmaceutical Society of Japan

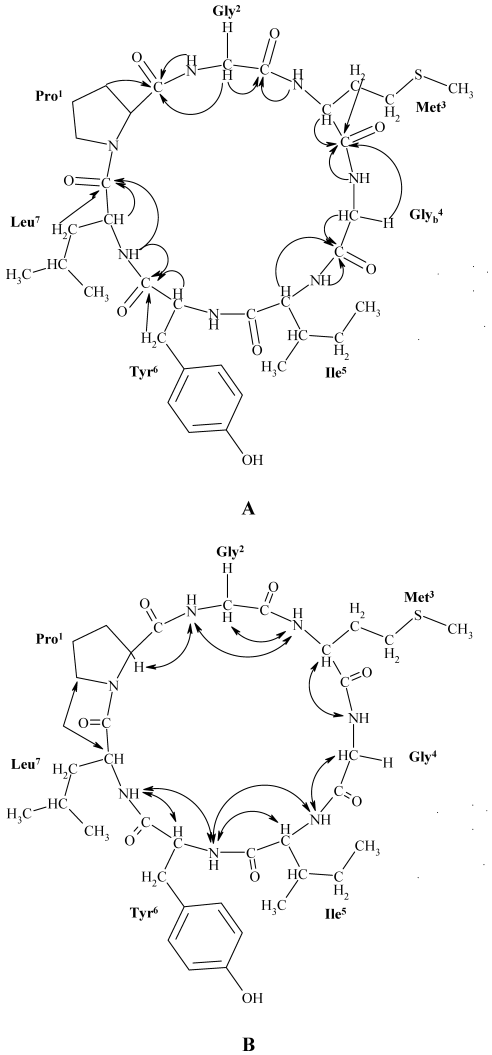

Fig. 2. Significant Correlations in the HMBC (A) and ROESY (B) Spectra of Glaucacyclopeptide B

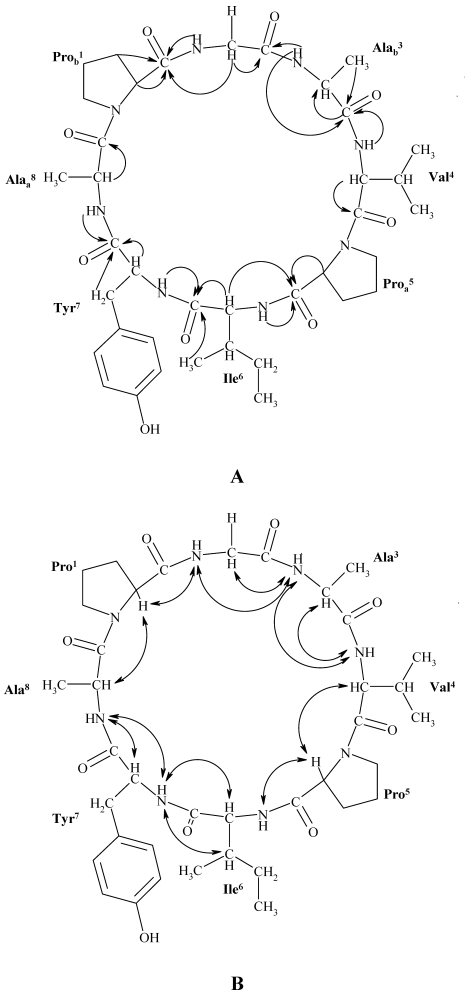

Fig. 3. Significant Correlations in the HMBC (A) and ROESY (B) Spectra of Cherimolacyclopeptide G

Wüthtrich and based on $d_{NN(i,i+1)}$ and $d_{\alpha N(i,i+1)}$ connectivities from the ROESY/NOESY spectra (Fig. 2B).[11,12] $^3J_{CH}$ $CO(i)/\alpha H(i+1)$ correlations between the isoleucine and Tyr^6 and between the leucine and Pro^1 on the HMBC spectrum, indicated unequivocally that the Ile residue is at position 5 and Leu at position 7. Chemical shifts of the β and γ carbons of Pro^1 at 29.9 and 26.4 ppm, respectively, and the strong correlation between the α proton of leucine residue with the $\delta\delta$ protons of the Pro indicates that the Leu^7-Pro^1 amide bond is *trans*.[13] All these connectivities are in full agreement with the structure deduced from the mass spectrometry study, *cyclo*(Pro^1-Gly^2-Met^3-Gly^4-Ile^5-Tyr^6-Leu^7).

The ^{1}H-NMR in pyridine-d_5 of component cherimolacyclopeptide G showed the presence of six amide protons, as well as the presence of eight carbonyl groups in the ^{13}C-NMR spectrum, in agreement with an octapeptide including two prolines. Similarly, this compound was sequenced using HMBC and ROESY spectra (Fig. 3). The α proton of Ala^8 gave a strong correlation with the α proton of Pro^1, indicating that the Ala^8-Pro^1 amide bond was *cis*. Therefore, the Val^4-Pro^5 amide bond was *cis*, because the ROESY spectrum showed a strong correlation between the α proton of Pro^5 and the α proton of Val^4. Further evidence of the presence of two *cis*-Pro amide bonds was done by chemical shifts of the β and γ carbons of Pro^1 at 32.9 and 22.5 ppm, and Pro^5 at 30.8 and 22.4 ppm. The whole data agreed with the cyclic structure for cherimolacyclopeptide G, the sequence of which was thus determined as *cyclo*(Pro^1-Gly^2-Ala^3-Val^4-Pro^5-Ile^6-Tyr^7-Ala^8).

Cyclopeptides were always described to have cytotoxic activity *in vitro* against cell cultures. In 3 d, the cytotoxicity activities of the two compounds were evaluated on KB (human nasopharyngeal carcinoma) cell lines. Cherimolacyclopeptide G, with an IC_{50} value of 0.52 μM, was more active *in vitro* than glaucacyclopeptide B (IC_{50} value=16.5 μM). Doxorubicin, with an IC_{50} value of 0.02 μM, was taken as the positive control.

Experimental

General Experimental Procedures IR spectra were obtained using KBr discs, and the melting point was determined on a Büchi melting point B-545 apparatus. Optical rotation was measured with a Perkin-Elmer model 341 polarimeter, and the $[\alpha]_D^{22}$ value is given in deg. cm^2 g^{-1}. ^{1}H- and ^{13}C-NMR spectra were recorded either (1D ^{13}C) on a Bruker AC 300 spectrometer equipped with an Aspect 3000 computer using DISNMR software or (2D spectra) on a Bruker Avance 400 spectrometer operating at 400.13 MHz (2D spectra). Mass spectra were recorded on an API Q-STAR Pulsar i mass spectrometer (Applied Biosystems). For the CID spectra, the collision energy was 40 to 60 eV and the collision gas was nitrogen.

Plant Material Seeds of *A. glauca* were obtained in October 2001 by

DR Modou LO from Dakar. A voucher specimen (AL 301) is deposited at Jardin des plantes utiles, Faculté de Médecine et de Pharmacie, Dakar, Sénégal. Seeds of *Annona cherimola* MILLER (Annonaceae) were collected in the south of Spain in December 2000. Samples were authenticated by Professor P. Boiteau (National Museum of Natural History, Paris) and a voucher specimen (VF 10463) is deposited in the herbarium of the Department of Botany, University of Valencia (Spain).

Extraction and Isolation Details of the methodology of isolation and purification of cyclic peptides are described in previous papers.[6,10] Dried and powdered seeds were extracted successively with cyclohexane and methanol. The methanol extract was partitioned between EtOAc and water. The organic phase was concentrated to dryness and the residue was dissolved in MeOH and chromatographed successively on Sephadex LH-20 and silica gel (Kieselgel 60 H Merck) columns, then finally purified by isochratic reversed phase HPLC (Kromasil C_{18}, 250×7.8 mm, 5 μm, AIT France; flow rate 2 ml/min, detection 220 nm) using MeOH/H_2O: 70/30 to give glaucacyclopeptide B (16.0 min, 11.3 mg) and 75/25 with 1% TFA to yield cherimolacyclopeptide G (t_R 14.3 min, 57.8 mg).

Absolute Configuration of Amino Acids A solution of compounds containing 1 mg of peptide in 6 N HCl (1 ml) was heated at 110 °C for 24 h in sealed tubes. Free amino acid residues were methylated, and the methylated amino acids were analyzed by CPG (for details see ref. 7). Comparison of retention time (min) showed that the values were comparable with those of standards amino acids: DL-Ala (10.6, 11.6), Gly (14.6), DL-Ile (16.2, 16.9), DL-Leu (18.1, 19.2), DL-Met (27.2, 27.9), DL-Pro (18.0, 18.2), DL-Tyr (31.7, 31.9) and DL-Val (13.4, 13.9).

Bioassays Aliquots of the purified cyclopeptides were dissolved in a DMSO solution for cytotoxicity evaluation *in vitro* over 3 d against KB (human nasopharyngeal carcinoma) cells. Details of the assay procedure are described in the literature.[14]

Glaucacyclopeptide B Colorless solid, mp 232—233 °C (MeOH), $[\alpha]_D^{22}$ −47° (c=0.1, MeOH); IR (KBr) cm^{-1}: 3320 and 1650. UV λ_{max} (MeOH) nm (ε): 242, 254, 260; ^{1}H- and ^{13}C-NMR (methyl alcohol-D_3, OH, 400 Mz, 298 K, TMS): 4.28 (1H, m, Pro^1-Hα), 2.22 (1H, m, Pro^1-Hβ), 1.96 (1H, m, Pro^1-Hβ'), 1.15 (1H, m, Pro^1-Hγ), 1.96 (1H, m, Pro^1-Hγ'), 3.80 (1H, m, Pro^1-Hδ), 3.59 (1H, m, Pro^1-Hδ'), 4.31 (1H, dd, 17.2, 8.6, Gly^2-Hα), 3.54 (1H, dd, 17.2, 3.9, Gly^2-Hα'), 8.95 (1H, dd, 8.6, 3.9, Gly^2-NH), 4.74 (1H, ddd, 9.4, 7.5, 7.5, Met^3-Hα), 2.20 (1H, m, 14.2, 7.5, 7.5, Met^3-Hβ), 1.83 (1H, m, 14.2, 7.5, 7.5, Met^3-Hβ'), 2.49 (2H, dd, 7.7, 7.7, Met^3-H$\gamma\gamma'$), 2.07 (1H, s, Met^3-S-CH$_3$), 8.54 (1H, d, 9.4, Met^3-NH), 4.18 (1H, dd, 16.3, 4.6, Gly^4-Hα), 3.65 (1H, dd, 16.3, 5.4, Gly^4-Hα'), 8.41 (1H, dd, 5.4, 4.6, Gly^4-NH), 3.97 (1H, dd, 6.3, 5.2, Ile^5-Hα), 1.72 (2H, m, Ile^5-$\beta\beta'$H), 0.99 (2H, m, Ile^5-$\gamma\gamma'$H), 0.75 (3H, dd, 7.3, 7.3, Ile^5-δCH$_3$), 0.67 (3H, d, 6.9, Ile^5-γCH$_3$), 8.18 (1H, d, 6.3, Ile^5-NH), 4.45 (1H, m, Tyr^6-Hα), 3.21 (1H, dd, 14.1, 4.6, Tyr^6-Hβ), 2.96 (1H, dd, 14.1, 11.4, Tyr^6-Hβ'), 7.10 (2H, m, 8.5, Tyr^6-2′6′), 6.71 (2H, m, 8.5, Tyr^6-3′5′), 7.89 (1H, d, 7.7, Tyr^6-NH), 10.20 (1H, s, Tyr^6-OH), 4.76 (1H, ddd, 10.0, 6.7, 6.7, Leu^7-Hα), 1.54 (2H, dd, 6.7, 6.7, Leu^7-H$\beta\beta'$), 1.63 (1H, m, 13.2, 6.5, 6.3, Leu^7-γH), 0.95 (3H, d, 6.5, Leu^7-δCH$_3$), 1.00 (3H, d, 6.3, Leu^7-δ'CH$_3$), 7.32 (1H, d, 10.0, Leu^7-NH). ^{13}C-NMR (methyl alcohol-D_3, OH, 400 MHz, 298 K, TMS): 175.2 (Pro^1-CO), 62.8 (Pro^1-Cα), 29.9 (Pro^1-Cβ), 26.4 (Pro^1-Cγ), 48.5 (Pro^1-Cδ), 171.6 (Gly^2-CO), 44.0 (Gly^2-Cα), 173.6 (Met^3-CO), 53.0 (Met^3-Cα), 33.0 (Met^3-Cβ), 31.3 (Met^3-Cγ), 15.4 (Met^3-S-CH$_3$), 172.2 (Gly^4-CO), 44.2 (Gly^4-Cα), 174.0 (Ile^5-CO), 61.0 (Ile^5-Cα), 37.1 (Ile^5-Cβ), 25.2 (Ile^5-Cγ-CH$_2$), 15.6 (Ile^5-Cγ-CH$_3$), 11.9 (Ile^5-Cδ-CH$_3$), 173.6 (Tyr^6-CO), 58.4 (Tyr^6-Cα), 37.7 (Tyr^6-Cβ), 128.9 (Tyr^6-C1′), 130.9 (Tyr^6-C2′6′), 116.4 (Tyr^6-C3′5′), 157.5 (Tyr^6-C4′), 171.9 (Leu^7-CO), 50.7 (Leu^7-Cα), 43.6 (Leu^7-Cβ), 25.7 (Leu^7-Cγ), 23.7 (Leu^7-Cδ-CH$_3$), 22.5 (Ile^5-Cδ'-CH$_3$); ESI-QTOF, m/z: 770 [M+K]$^+$, 754 [M+Na]$^+$, 732 [M+H]$^+$; ESI-QTOF MS/MS on m/z 732 [M+H]$^+$ (ce 40 eV) m/z (%): 732 (100), 704 (18), 669 (7), 612 (9), 619 (25), 591 (6), 546 (12), 428 (37), 343 (22), 327 (5), 315 (3), 286 (5), 258 (19), 155 (11), 127 (3).

Cherimolacyclopeptide G Colorless solid, mp 277—278 °C (MeOH), $[\alpha]_D^{22}$ −52° (c=0.1, MeOH); IR (KBr) cm^{-1}: 3320 and 1650. UV λ_{max} (MeOH) nm (ε): 242, 254, 260; ^{1}H- and ^{13}C-NMR (pyridine-d_5, 400 Mz, 298 K, TMS): 4.80 (1H, dd, 7.1, 1.5, Pro^1-Hα), 2.30 (1H, m, Pro^1-Hβ), 2.15 (1H, m, Pro^1-Hβ'), 1.70 (2H, m, Pro^1-H$\gamma\gamma'$), 3.84 (1H, m, Pro^1-Hδ), 3.61 (1H, m, Pro^1-Hδ'), 4.61 (1H, dd, 16.9, 6.3, Gly^2-Hα), 4.13 (1H, dd, 16.9, 5.9, Gly^2-Hα'), 9.37 (1H, dd, 6.3, 5.9, Gly^2-NH), 5.15 (1H, dq, 8.5, 6.8, Ala^3-Hα), 1.46 (3H, d, 6.8, Ala^3-βCH$_3$), 8.67 (1H, d, 8.5, Ala^3-NH), 4.30 (1H, dd, 8.6, 5.1, Val^4-Hα), 2.35 (1H, m, Val^4-Hβ), 1.34 (3H, d, 6.6, Val^4-γCH$_3$), 1.04 (3H, d, 6.9, Val^4-γ'CH$_3$), 9.93 (1H, d, 5.1, Val^4-NH), 4.90 (1H, d, 6.1, Pro^5-Hα), 2.75 (1H, m, Pro^5-Hβ), 1.75 (1H, m, Pro^5-Hβ'), 1.85 (1H, m, Pro^5-Hγ), 1.65 (1H, m, Pro^5-Hγ'), 3.75 (1H, m, Pro^5-Hδ), 3.59 (1H, m, Pro^5-Hδ'), 4.76 (1H, dd, 9.1, 9.5, Ile^6-Hα), 2.33 (1H, m, Ile^6-Hβ), 1.65 (1H, m, Ile^6-γH), 1.30 (1H, m, Ile^6-γ'H), 1.00 (3H, d, 6.8, Ile^6-δCH$_3$), 0.83 (3H, d, 7.4, Ile^6-γCH$_3$), 9.16 (1H, d, 9.1, Ile^6-NH), 5.51 (1H, ddd, 8.3, 7.0, 6.4, Tyr^7-Hα), 3.36 (1H, dd, 7.0, 13.8, Tyr^7-Hβ), 3.12 (1H, dd, 6.4, 13.8, Tyr^7-Hβ'), 7.40 (2H, m, 8.4, Tyr^7-2′6′), 7.12 (2H, m, 8.4, Tyr^7-3′5′), 8.12 (1H, d, 8.3, Tyr^7-NH), 11.40 (1H, s, Tyr^7-OH), 5.00 (1H, m, Ala^8-Hα), 1.30 (3H, d, 7.1, Ala^8-βCH$_3$), 10.37 (1H, d, 8.0, Tyr^7-NH). ^{13}C-NMR (pyridine-d_5, 400 MHz, 298 K, TMS): 173.8 (Pro^1-CO), 61.5 (Pro^1-Cα), 32.9 (Pro^1-Cβ), 22.5 (Pro^1-Cγ), 46.7 (Pro^1-Cδ), 169.0 (Gly^2-CO), 45.0 (Gly^2-Cα), 175.7 (Ala^3-CO), 48.4 (Ala^3-Cα), 16.4 (Ala^3-Cβ), 171.7 (Val^4-CO), 60.0 (Val^4-Cα), 30.5 (Val^4-Cβ), 19.3 (Val^4-Cγ), 19.0 (Val^4-Cγ'), 171.9 (Pro^5-CO), 61.8 (Pro^5-Cα), 30.8 (Pro^5-Cβ), 22.4 (Pro^5-Cγ), 47.4 (Pro^5-Cδ), 170.5 (Ile^6-CO), 61.3 (Ile^6-Cα), 36.3 (Ile^6-Cβ), 25.5 (Ile^6-Cγ), 16.1 (Ile^6-Cδ-CH$_3$), 10.3 (Ile^6-Cγ-CH$_3$), 172.3 (Tyr^7-CO), 54.8 (Tyr^7-Cα), 39.7 (Tyr^7-Cβ), 127.9 (Tyr^7-C1′), 130.9 (Tyr^7-C2′6′), 116.4 (Tyr^7-C3′5′), 157.8 (Tyr^7-C4′), 172.4 (Ala^8-CO), 47.4 (Ala^8-Cα), 17.5 (Ala^8-Cβ). ESI-QTOF, m/z: 807 [M+K]$^+$, 791 [M+Na]$^+$, 769 [M+H]$^+$; ESI-QTOF MS/MS on m/z 769 [M+H]$^+$ (ce 40 eV) m/z (%): 769 (100), 741 (20), 698 (13), 670 (55), 642 (13), 599 (10), 571 (2), 535 (15), 507 (7), 445 (10), 422 (10), 417 (21), 396 (34), 394 (3), 374 (63), 325 (56), 297 (31), 226 (16), 211 (16), 183 (14), 155 (3).

Acknowledgments We are indebted to the French "Ministère de la Coopération" (EGIDE) and the "Région Ile-de-France" for its generous contribution to the funding of the 400 MHz NMR and the ESI-QTOF mass spectrometers. We also thank Miss C. Caux for the 400 MHz NMR spectra, Mr. L. Dubost for the mass spectra, and Mrs. Christiane Gaspard (ICSN-CNRS, Gif-sur-Yvette) for the cytotoxicity bioassays.

References

1) Li C.-M., Tan N.-H., Mu Q., Zheng H.-L., Hao X.-J., Wu Y., Zhou J., *Phytochemistry*, **45**, 521—523 (1997).
2) Li C.-M., Tan N.-H., Mu Q., Zheng H.-L., Hao X.-J., Liang H.-L., Zhou J., *Phytochemistry*, **47**, 1293—1296 (1998).
3) Li C.-M., Tan N.-H., Zheng H.-L., Mu Q., Hao X.-J., He Y.-N., Zhou J., *Phytochemistry*, **48**, 555—556 (1998).
4) Li C.-M., Tan N.-H., Zheng H.-L., Mu Q., Hao X.-J., He Y.-N., Zhou J., *Phytochemistry*, **50**, 1047—1052 (1999).
5) Morita H., Shishido A., Matsumoto T., Itokawa H., Takeya K., *Tetrahedron*, **55**, 7509—7518 (1999).
6) Wélé A., Ndoye I., Zhang Z., Brouard J.-P., Pousset J.-L., Bodo B., *Phytochemistry*, **66**, 1154—1157 (2005).
7) Wélé A., Landon C., Labbé H., Vovelle F., Zhang Y., Bodo B., *Tetrahedron*, **60**, 405—414 (2002).
8) Wélé A., Zhang Z., Ndoye I, Brouard J.-P., Pousset J.-L., Bodo B., *J. Natural Prod.*, **67**, 1577—1579 (2004).
9) Wélé A., Zhang Z., Brouard J.-P., Pousset J.-L., Bodo B., *Phytochemistry*, **66**, 693—696 (2005).
10) Wélé A., Zhang Z., Brouard J.-P., Pousset J.-L., Bodo B., *Phytochemistry*, **66**, 2376—2380 (2005).
11) Wüthrich K., Wider G., Wagner G., Braun W., *J. Mol. Biol.*, **155**, 311—319 (1982).
12) Wüthrich K., Billeter M., Braun W., *J. Mol. Biol.*, **180**, 715—740 (1984).
13) Douglas D. E., Bovey F. A., *J. Org. Chem.*, **38**, 2379—2383 (1973).
14) Chang F.-R., Chen J.-L., Chiu H.-F., Wu M.-J., Wu Y.-C., *Phytochemistry*, **47**, 1057—1061 (1998).