Isolation and Structure of Gustastatin from the Brazilian Nut Tree *Gustavia hexapetala*[†,1]

George R. Pettit,*[,‡] Qingwen Zhang,[‡] Véronique Pinilla,[‡] Delbert L. Herald,[‡] Dennis L. Doubek,[‡] and James A. Duke[§]

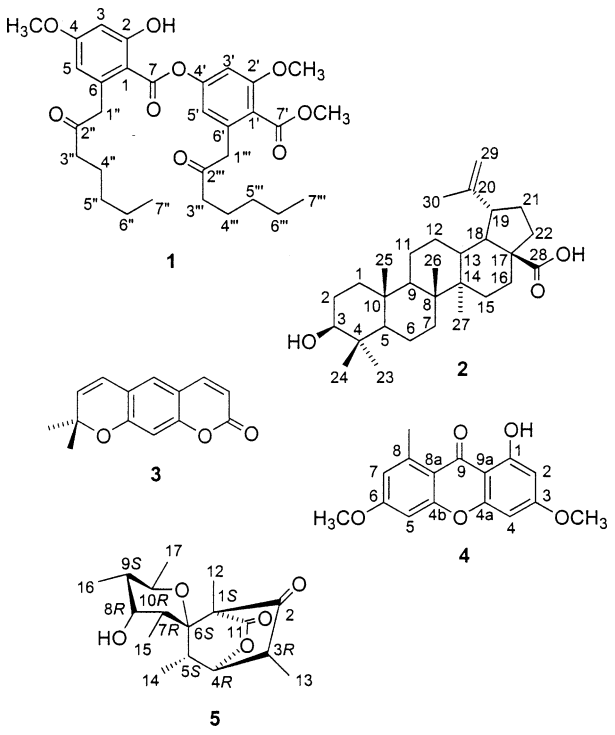
Cancer Research Institute, Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-2404, and U.S. Department of Agriculture, Beltsville, Maryland 20705

Received November 25, 2003

A bioassay-guided investigation of *Gustavia hexapetala* led to the isolation of a new cancer cell growth inhibitor designated gustastatin (**1**) and four previously known cancer cell growth inhibitors that included betulinic acid (**2**). The structures were assigned on the basis of analyses of HRMS combined with 1D and 2D NMR data. The structure of portentol (**5**) was confirmed by an X-ray crystal structure determination.

The Lecythidaceae family is composed of some 325 tropical trees divided among 15 genera. The *Gustavia* genus (Brazil nut trees), represented by 45 species, is one of the better known, albeit very little investigated chemically. One of these, *G. brasiliana* D. C., was reported in 1843[2] to be used in Brazil against liver disease (cancer?). Furthermore, two other Lecythidaceae species, *Barringtonia acutangula* and an unknown *Barringtonia* species, have been used in the traditional treatment of cancer in India and Malaysia.[2] Only *G. augusta*[3] and *G. longifolia*[4] have received a moderate investigation in respect to small molecule constitutents, yielding triterpenes, sterols, and fatty acids including betulinic acid (**2**). Interestingly, the *Gustavia* trees are known for their folk medical use against leishmaniosis[3] and other anthelmintic activities[5] and for their fetid odor believed to arise from emanation of dimethyl disulfide and other volatile sulfur compounds (about 15% of the volatile exudates).[6] These and other components of *Gustavia* provide partial protection against wood-boring longicorn beetles of the Cerambycidae.[6]

As part of the U.S. National Cancer Institute's (NCI) exploratory anticancer drug lead development research, the stems and twigs of *Gustavia hexapetala* (AUBL) Smith (Lecythidaceae) were collected in August 1977 in Peru. A methanol extract of the original collection of *G. hexapetala* led to a 53% life extension at a dosage of 100 mg/kg against murine P388 lymphocyte leukemia. That encouraging result led to a 72 kg re-collection of the stems and twigs of *G. hexapetala* in 1979, which we began to study in May 1980. The NCI in vivo P388 leukemia availability was terminated in 1981. Unfortunately, the early fractions of *G. hexapetala* gave inconsistent results against the P388 lymphocytic leukemia cell line, and pursuit of this lead was delayed. After implementation of our panel of six human solid tumor cell lines, isolation was resumed. A new cancer cell line inhibitor designated gustastatin (**1**) was isolated along with betulinic acid (**2**),[7] the latter being the strongest human cancer cell line inhibitory constitutent of *G. hexapetala*. Other cancer cell growth inhibitors, xanthyletin (**3**),[8] lichexanthone (**4**),[9] and oleanolic/ursolic acids[10,11] that we had previously isolated from *Salvia apiana*[11] as well as the marginally active portentol (**5**)[12] and the inactive β-amyrin,[13] were also isolated.



Results and Discussion

The dry wood (72 kg) of *G. hexapetala* (AUBL) Smith was extracted with CH_2Cl_2−CH_3OH (1:1), followed by a solvent partitioning[14] sequence utilizing 9:1 CH_3OH−H_2O against hexane followed by 3:2 CH_3OH−H_2O against CH_2Cl_2. The CH_2Cl_2 fraction was found to concentrate the cancer cell growth inhibitors, and that fraction was subjected to gel permeation chromatography on Sephadex LH-20 followed by a partitioning sequence employing analogous techniques, yielding compounds **1**−**5**.

Gustastatin (**1**) was isolated as an amorphous powder that resisted crystallization, thus precluding X-ray structural elucidation. Hence, the structure of gustastatin was determined via spectral methods. The molecular formula of gustastatin (**1**) was deduced as $C_{31}H_{40}O_9$ from HR-

[†] In celebration of Professor Carl Djerassi's 80th birthday and his extraordinary contributions to medicinal, natural products, and organic chemistry.

* To whom correspondence should be addressed. Tel: (480) 965-3351. Fax: (480) 965-8558. E-mail: bpettit@asu.edu.

[‡] Cancer Research Institute.

[§] U.S. Department of Agriculture (USDA).

10.1021/np030509i CCC: $27.50
Published on Web 05/20/2004

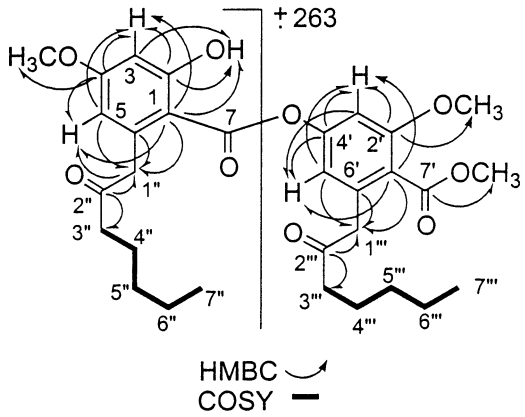


Figure 1. Gustastatin (1) HMBC and COSY correlations and HRFAB cleavage.

FABMS (*m*/*z* 557.2750 $[M + H]^+$, Δ 0 mmu), implying 12 degrees of unsaturation. The HMBC spectrum (see Figure 1) of phenol **1** displayed a chelated proton signal at δ 11.28 (2-OH), which showed a correlation with the signal at δ 104.3 (C1), corresponding to the aromatic carbon linked to the carbonyl group. The 2-OH proton also exhibited connectivity to the signals at δ 166.5 (C2) and 100.1 (C3). The C4 (δ 164.9) carbon showed correlations with 3-H (δ 6.46), 5-H (δ 6.29), and methoxy protons (δ 3.85), confirming that C4 was substituted with a methoxy group. Another 2′-methoxy group was deduced from the correlation of the quaternary carbon at δ 158.4 (C2′) with methoxy protons (δ 3.83) and an aromatic proton at δ 6.89 (3′-H). The last methoxy was identified as part of a methyl ester by the cross-peak between the methoxy proton (δ 3.88) and a carbonyl carbon at δ 167.5 (C7′). Two carbonyl carbons at δ 207.4 (C2″) and 206.5 (C2‴) exhibited correlations with 1″-H and 3″-H, and 1‴-H and 3‴-H, respectively. That suggested both side chains to be 2-oxoheptyl groups. The two side chains were determined to link to C6 and C6′ by the cross-peaks of (C1, C5, and C6)/1″-H and (C1′, C5′, and C6′)/1‴-H, respectively. From the carbon chemical shifts of C1′ (121.8) and C4′ (151.3), the $COOCH_3$ unit was deduced to be at C1′ and the ring-A/ring-B core to be linked by an ester bond at C4′. The ester bond linkage of ring-A and -B was confirmed by interpretation of the mass spectral fragmentation results, completing the structure assignment for gustastatin (**1**). The previously known constituents, betulinic acid (**2**),[7] xanthyletin (**3**),[8] lichexanthone (**4**),[9] oleanolic/ursolic acids,[10,11] portentol (**5**),[12] and β-amyrin,[13] were identified by analogous spectral techniques.

Portentol (**5**) was previously isolated from the lichens *Roccella luteola*,[12a] *R. portentosa*,[12b] *R. fuciformis*,[12c] *Dirina repanda*,[12d] *D. stenhammari*,[12e] and *Lobodirina cerebriformes*.[12f] Aberhart reported the proton NMR data at 100 MHz in 1970,[12g] but not ^{13}C NMR data. We elucidated the structure by 1D and 2D NMR analyses and confirmed the structure by single-crystal X-ray structure determination. The X-ray data for portenol (**5**) were consistent with the structure previously assigned.[12] Each asymmetric portion of the unit cell was found to contain a single molecule of **5**. Intermolecular hydrogen bonding was observed between the O-8 hydroxyl hydrogen and the O-11 carbonyl oxygen of an adjacent molecule (d(D···A) = 2.779 Å).

The new depside (phenol-ester linkage) gustastatin (**1**) represents a new human cancer cell growth inhibitor that is potentially available by synthesis, which will be undertaken to advance biological evaluations. The isolation of betulinic acid as the most significant cancer cell growth inhibitory constituent of *G. hexapetala* underscores the potential utility of this relatively common triterpene, which is now in preclinical development.

Table 1. Murine P388 Lymphocytic Leukemia Cell Line and Human Cancer Cell Line Inhibition Values (GI_{50} in μg/mL) for Compounds **1**−**5**[a]

cancer cell line[b]	**1**	**2**	**3**	**4**	**5**
P388	>10	>10	>10	>10	>10
BXPC-3	>10	0.85	5.5	2.5	7.0
MCF-7	6.4	1.1	5.9	4.1	>10
SF-268	7.9	0.86	11.2	3.0	4.7
NCI-H460	5.1	4.7	9.0	7.1	4.7
KM20L2	>10	3.4	6.3	2.0	>10
DU-145	5.2	5.0	2.9	3.7	1.3

[a] In DMSO. [b] Cancer type: P388 (lymphocytic leukemia); BXPC-3 (pancreas adenocarcinoma); MCF-7 (breast adenocarcinoma); SF268 (CNS glioblastoma); NCI-H460 (lung large cell); KM20L2 (colon adenocarcinoma); DU-145 (prostate carcinoma).

Experimental Section

General Experimental Procedures. The general experimental procedures applied in the following isolation and structural elucidation techniques have been summarized in the Introduction to the Experimental Section of our recent contribution on the isolation and structure of palstatin.[15]

***Gustavia hexapetala* (AUBL) Smith.** The wood, bark, and twigs of this Lecythidaceae family tree were originally collected as part of the joint U.S. National Cancer Institute/U.S. Department of Agriculture research under the direction of Drs. John D. Douros and Matthew I. Suffness in the NCI, and Robert E. Purdue and James A. Duke in the USDA. A voucher specimen of *G. hexapetala* has been deposited in the Medicinal Plant Resources Laboratory of the USDA, Beltsville, MD. A scale-up re-collection of the wood, bark, and twigs was again made in Peru in 1979, and they were received in our Institute in May 1980. The scale-up re-collection weighed 72.2 kg (dry weight) and was again attained through the NCI/USDA collaboration.

Extraction and Initial Separation of *G. hexapetala*. The 1979 re-collection of *G. hexapetala* was chipped and divided between two 200 L containers. To the plant material in each container was added 120 L of CH_2Cl_2−CH_3OH (1:1). After 10 days at ambient temperatures, the solvent was removed and 120 L of H_2O was added to cause phase separation. The CH_2Cl_2 fraction gave 393 g (P388 ED_{50} 66 μg/mL) and the CH_3OH−H_2O fraction 760 g (P388 ED_{50} >100 μg/mL). The ambient extraction with 1:1 CH_2Cl_2−CH_3OH was repeated a second time to yield 197 g of the CH_2Cl_2 fraction and 408 g of the CH_3OH−H_2O fraction. The total yield was 590 g from the CH_2Cl_2 and 1.168 kg from the CH_3OH−H_2O fractions. The CH_2Cl_2 fractions were combined and partitioned[14] between solvent mixtures comprising CH_3OH−H_2O (9:1 → 3:2) and hexane (for the 9:1 mix) followed by CH_2Cl_2 (for the 3:2 mix) to provide 178.5 g (P388 ED_{50} 27 μg/mL) of a CH_2Cl_2 fraction. The hexane fraction weighed 320.5 g (P388 ED_{50} > 100 μg/mL), and the remaining CH_3OH−H_2O fraction amounted to 54.3 g (P388 ED_{50} > 100 μg/mL).

Isolation of Gustastatin (1) and Constituents 2−**5.** The following isolation procedure was guided by bioassay using human cancer cell lines (see Table 1) in conjunction with P388 murine lymphocytic leukemia. A CH_3OH solution of the CH_2Cl_2 fraction (178.5 g) was loaded onto a Sephadex LH-20 column and eluted with CH_3OH. Cell line (both P388 and human cancer) active fractions a, b, c, and d were obtained. Fractions b and c were combined and chromatographed on a Sephadex LH-20 column, using CH_3OH−CH_2Cl_2 (3:2) as eluent, to provide three active fractions (e, f, and g). Fraction e was successively rechromatographed with hexane−CH_3OH−2-propanol (3:1:1), hexane−ethyl acetate−CH_3OH (70:30:15), and toluene−2-propanol−1-butanol (6:1:1) and afforded the new compound gustastatin (**1**, 11 mg, 1.5×10^{-5} % yield) and an active fraction (h). Fraction f was rechromatographed on a Sephadex LH-20 column, using hexane−CH_3OH−2-propanol

(3:1:1) as eluent, and two active fractions (i and j) were obtained. Fraction i was rechromatographed on a Sephadex LH-20 column, using hexane−ethyl acetate−CH_3OH (70:30:15) as eluent, from which β-amyrin (16 mg, 2.2×10^{-5} % yield) and an active fraction (k) were obtained. Fraction k was rechromatographed on a Sephadex LH-20 column, using hexane−toluene−CH_2Cl_2−EtOH (17:1:1:1) as eluent, and betulinic acid (**2**, 512 mg, 7.1×10^{-4} % yield) and an active mixture of olenolic/ursolic acids (90 mg, 1.3×10^{-4} % yield) were obtained. Fraction j was chromatographed on a Sephadex LH-20 column, using hexane−toluene−CH_2Cl_2−EtOH (17:1:1:1) as eluent, which led to two active fractions (l and m). Fraction l was further separated by HPLC, using a Zonbrax SB C18 column and an isocratic mobile phase (55% CH_3CN in H_2O for 20 min). Peaks were monitored at 210 nm; flow rate was 1.5 mL/min. Pure portentol (**5**, 3.8 mg, 5.3×10^{-6} % yield) was obtained. Fraction d was rechromatographed on columns of Sephadex LH-20 and eluted successively with (a) CH_2Cl_2−CH_3OH (3:2), (b) hexane−2-propanol−CH_3OH (3:1:1), (c) hexane−CH_2Cl_2−2-propanol−CH_3OH (8:1:2:2), and (d) hexane−toluene−CH_2Cl_2−EtOH (17:1:1:1). That led to isolation of xanthyletin (**3**, 60 mg, 8.3×10^{-5} % yield) and lichexanthone (**4**, 2.6 mg, 3.6×10^{-6} % yield).

Gustastatin (1): amorphous solid; mp 118−121 °C; UV ($CHCl_3$) λ_{max} (log ϵ) 240 (3.88), 269 (4.13), 302 (3.85) nm; IR ν_{max} 2952, 2931, 2854, 1714, 1652, 1597, 1458, 1431, 1331, 1254, 1161, 1146, 1073, 1043, 952, 737 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$) δ 11.28 (1 H, s, 2-OH), 6.59 (1 H, d, J = 1.5 Hz, H-3′), 6.57 (1 H, d, J = 1.5 Hz, H-5′), 6.46 (1 H, d, J = 2.5 Hz, H-3), 6.29 (1 H, d, J = 2.5 Hz, H-5), 4.07 (2 H, s, H-1″), 3.88 (3 H, s, CO_2CH_3), 3.85 (3 H, s, 4-OCH_3), 3.83 (3 H, s, 2′-OCH_3), 3.70 (2 H, s, H-1‴), 2.44 (1 H, t, J = 7.5 Hz, H-3″), 2.41 (1 H, t, J = 7.5 Hz, H-3‴), 1.56 (1 H, pent, J = 7.5 Hz, H-4″), 1.53 (1 H, pent, J = 7.5 Hz, H-4‴), 1.27 (2 H, m, H-6″, H-6‴), 1.21 (2 H, m, H-5″, H-5‴); ^{13}C NMR (125 MHz, $CDCl_3$) δ 207.4 (C-2″), 206.5 (C-2‴), 169.1 (C-7), 167.5 (C-7′), 166.5 (C-2), 164.9 (C-4), 158.4 (C-2′), 151.3 (C-4′), 138.9 (C-6), 135.4 (C-6′), 121.8 (C-1′), 116.2 (C-5′), 113.4 (C-5), 104.5 (C-3′), 104.3 (C-1), 100.1 (C-3), 56.3 (2′-OCH_3), 55.5 (4-OCH_3), 52.3 (CO_2CH_3), 51.2 (C-1″), 47.5 (C-1‴), 42.5 (C-3″), 42.2 (C-3‴), 31.3 (C-5″, C-5‴), 23.4 (C-4″), 23.3 (C-4‴), 22.4 (C-6″, C-6‴), 13.9 (C-7″), 13.8 (C-7‴); HRFABMS m/z 557.2750 $[M + H]^+$ (calcd for $C_{31}H_{41}O_9$, 557.2750).

Betulinic acid (2):[7] colorless needles from CH_3OH; mp 282 °C; HRAPCIMS m/z 439.35364 $[M + H - H_2O]^+$ (calcd for $C_{30}H_{47}O_2$, 439.35761).

Xanthyletin (3):[8] colorless plates from CH_3OH; mp 110−112 °C; HRAPCIMS m/z 229.08860 $[M + H]^+$ (calcd for $C_{14}H_{13}O_3$, 229.08648).

Lichexanthone (4):[9] colorless needles from CH_3OH; mp 175−176 °C; HRAPCIMS m/z 287.08813 $[M + H]^+$ (calcd for $C_{16}H_{15}O_5$, 287.09195).

Oleanolic acid and ursolic acid isomeric mixture:[10,11] amorphous powder from CH_3OH; HRAPCIMS m/z 439.35621 $[M + H - H_2O]^+$ (calcd for $C_{30}H_{47}O_2$, 439.35761).

Portentol (5):[12] colorless needles from CH_3OH; mp 225 °C; HRAPCIMS m/z 293.17577 $[M + H - H_2O]^+$ (calcd for $C_{17}H_{25}O_4$, 293.17529); ^1H NMR (500 MHz, $CDCl_3$) δ 0.89 (3H, d, J = 7.0 Hz, 15-H_3), 0.90 (3H, d, J = 7.5 Hz, 16-H_3), 1.10 (3H, d, J = 6.0 Hz, 17-H_3), 1.23 (3H, d, J = 7.5 Hz, 13-H_3), 1.29 (1H, qd, J = 7.0, 2.0 Hz, 9-H), 1.33 (3H, s, 12-H_3), 1.48 (3H, d, J = 7.0 Hz, 14-H_3), 1.89 (1H, qd, J = 7.0, 2.0 Hz, 7-H), 2.35 (1H, qd, J = 7.0, 1.0 Hz, 3-H), 2.55 (1H, q, J = 7.0 Hz, 5-H), 3.54 (1H, t, J = 2.0 Hz, 8-H), 3.85 (1H, hex, J = 6.0 Hz, 10-H), 4.40 (1H, s, 4-H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 8.6 (C12), 14.3 (C15), 14.3 (C13), 15.1 (C14), 15.4 (C16), 19.6 (C17), 38.5 (C5), 38.6 (C7), 41.6 (C9), 46.0 (C3), 67.9 (C1), 70.2 (C10), 74.6 (C8), 80.7 (C6), 84.2 (C4), 168.9 (C11), 208.2 (C2). The structure was confirmed by single-crystal X-ray analysis (see Supporting Information).

***β*-Amyrin:**[13] colorless needles from CH_3OH; mp 194 °C; HRAPCIMS m/z 409.38518 $[M + H - H_2O]^+$ (calcd for $C_{30}H_{49}$, 409.38343).

X-ray Crystal Structure Analysis of Portentol (5). In 1970, the X-ray determination of portentol was performed on a heavy-atom derivative, the *p*-bromobenzoate.[12h]

Acknowledgment. We are pleased to thank for financial assistance the support of Outstanding Investigator Grant CA44344-01A1-12, and Grant RO1 CA-90441-01-03 from the Division of Cancer Treatment Diagnosis, NCI, DHHS; the Arizona Disease Control Research Commission; Dr. Alec D. Keith; J. W. Kieckhefer Foundation; Margaret T. Morris Foundation; the Robert B. Dalton Endowment Fund; the Caitlin Robb Foundation; Gary L. and Diane Tooker; Polly J. Trautman; Sally Schloegel; the Eagles Art Ehrmann Cancer Fund; and the Ladies Auxiliary to the Veterans of Foreign Wars. In addition, we thank for technical assistance Drs. G. M. Cragg, J.-C. Chapuis, F. Hogan, V. Mukku, and J. C. Knight as well as F. Craciunescu and L. Williams.

Supporting Information Available: Tables of X-ray crystallographic data for compound **5** are available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Contribution 528 of Antineoplastic Agents. For preceding part, see: Pettit, G. R.; Melody, N. *J. Nat. Prod.*, accepted.
(2) Hartwell, J. L. *Plants Used Against Cancer*; Quarterman Publications: Lawrence, MA, 1982; p 293.
(3) De Souza, A. D. L.; Da Rocha, A. F. I.; Pinheiro, M. L. B.; Andrade, C. H. de S.; Galotta, A. L. de A. Q.; dos Santos, M. do P. S. S. *Quim. Nova* **2001**, *24*, 439−442.
(4) El-Seedi, H. R.; Zayed, M. F.; Morono, F. G.; Torssell, K. B. G. *Rev. Lat. Am. Quim.* **1999**, *27*, 56−57.
(5) Friese, F. W. *Apoth. Ztg.* **1929**, *44*, 1480−1482.
(6) Berkov, A.; Meurer-Grimes, B.; Purzycki, K. L. *Biotropica* **2000**, *32*, 440−451.
(7) Peng, C.; Bodenhausen, G.; Qiu, S.; Fong, H. S.; Farnsworth, N. R.; Yuan, S.; Zheng, C. *Magn. Reson. Chem.* **1998**, *36*, 267−278.
(8) Manzini, P.; Rodighiero, P.; Pastorini, G.; Guiotto, A. *Magn. Reson. Chem.* **1987**, *25*, 740−742.
(9) Jiménez, C.; Marcos, M.; Villaverde, M. C.; Riguera, R.; Castedo, L.; Stermitz, F. *Phytochemistry* **1989**, *28*, 1992−1993.
(10) Lin, G.; Xu, X.; Liu, D.; Ju, J.; Yang, J. *Zhongguo Yaoxue Zazhi* **2000**, *35*, 298−300.
(11) Pettit, G. R.; Klinger, H.; Jorgensen, N.-O. N.; Occolowitz, J. *Phytochemistry* **1966**, *5*, 301−309.
(12) (a) Huneck, S.; Follmann, G. *Z. Naturforsch., B: Chem. Sci.* **1967**, *22*, 1185−1188. (b) Aberhart, D. J.; Overton, K. H. *J. Chem. Soc., Chem. Commun.* **1969**, 162−163. (c) Huneck, S.; Mathey, A.; Trotet, G. *Z. Naturforsch., B: Chem. Sci.* **1967**, *22*, 1367−1368. (d) Huneck, S.; Trotet, G. *Z. Naturforsch., B: Chem. Sci.* **1967**, *22*, 363. (e) Huneck, S.; Roenneberg, H.; Liaaen-Jensen, S. *Pharmazie* **1982**, *37*, 866. (f) Quilhot, W.; Garbarino, J. A.; Gambaro, V. *J. Nat. Prod.* **1983**, *46*, 593. (g) Aberhart, D. J.; Overton, K. H.; Huneck, S. *J. Chem. Soc. (C)* **1970**, 1612−1623. (h) Ferguson, G.; Mackay, I. R. *J. Chem. Soc., Chem. Commun.* **1970**, 665−666.
(13) Knight, S. A. *Org. Magn. Reson.* **1974**, *6*, 603−611.
(14) Bligh, E. G.; Dyer, W. J. *Can. J. Biochem. Physiol.* **1959**, *37*, 911−917.
(15) Pettit, G. R.; Meng, Y.; Stevenson, C. A.; Doubek, D. L.; Knight, J. C.; Cichacz, Z.; Pettit, R. K.; Chapuis, J.-C.; Schmidt, J. M. *J. Nat. Prod.* **2003**, *66* (2), 259−262.

NP030509I