Carmabins A and B, New Lipopeptides from the Caribbean Cyanobacterium *Lyngbya majuscula*

Gregory J. Hooper,[†] Jimmy Orjala,[†] Randall C. Schatzman,[‡] and William H. Gerwick*,[†]

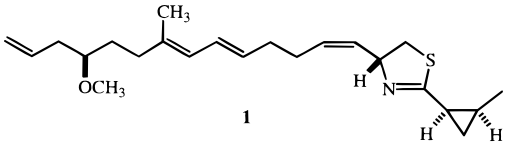
College of Pharmacy, Oregon State University, Corvallis, Oregon 97331, and Institute of Biochemistry and Cell Biology, Syntex Research, Palo Alto, California 94304

Received September 24, 1997

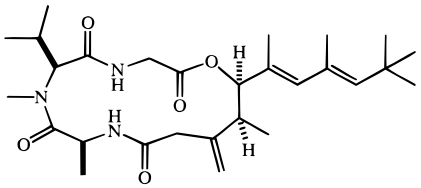
Carmabins A and B have been isolated as linear lipotetrapeptides from the BuOH extract of the marine cyanobacterium *Lyngbya majuscula*. The planar structures were elucidated by extensive 2D NMR analysis, including $^{1}H-^{15}N$ HMBC and HMQC−TOCSY experiments, together with MS measurements.

A Caribbean variety of the marine cyanobacterium *Lyngbya majuscula* Gomont (family Oscillatoriaceae) has provided our laboratory with a rich array of new and biologically active secondary metabolites. Bioassay-guided fractionation of the lipid extract of this cyanobacterium, collected in December 1991, gave rise to the potently active natural products curacin A[1] (**1**, antiproliferative), antillatoxin[2] (**2**, ichthyotoxic), malyngamide H[3] (**3**, ichthyotoxic), and barbamide[4] (**4**, molluscicidal). An investigation of the more polar fraction of the lipid extract, which showed antiproliferative activity,[5] revealed an interesting lipopeptide-like compound, trivially named carmabin A, in minute quantity (ca. 1 mg). Initial detection of this compound was by bioassay and ^{1}H NMR analysis, as it proved refractory to normal methods of TLC analysis. A subsequent recollection of the cyanobacterium in December 1993, and fractionation of its BuOH extract, using ^{1}H NMR to monitor fractions, resulted in reisolation of carmabin A (**5**, 0.016% dry wt), as well as an additional minor constituent, carmabin B (**6**, 0.002% dry wt). Considering the bioassay-guided nature of the initial isolation, it was surprising that these reisolated pure compounds proved inactive in the NCI's 60-cell-line bioassay (a panel that does not contain the MRC-5 cell line utilized for the original bioassay), a result that we are unable to explain at present. It is possible that either the "new" isolate is subtly different in some structural feature from the original isolate or that some unrecognized impurity in the original isolate gave rise to its potent antiproliferative activity. The major compound, carmabin A (**5**), was subsequently also isolated from a laboratory culture of the original collection. In the past few years a number of linear lipopeptide-type compounds, often with impressive biological activities, have been isolated from various collections of *Lyngbya majuscula*. Examples of these include the majusculamides A, B, and D;[6,7] microcolins A and B;[8] ypaomide;[9] barbamide;[4] and curacins A, B, and C.[1,10]

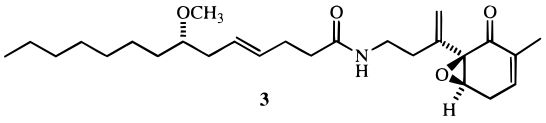
Carmabin A (**5**) was obtained from reversed-phase HPLC as a white solid. The molecular formula was established as $C_{40}H_{57}N_5O_6$ from positive HRFABMS. Various features of the ^{1}H and ^{13}C NMR data (Table 1)



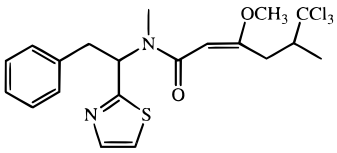
1



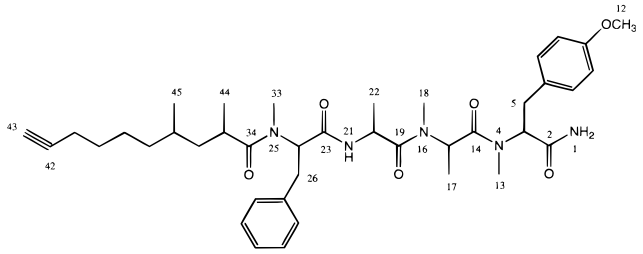
2



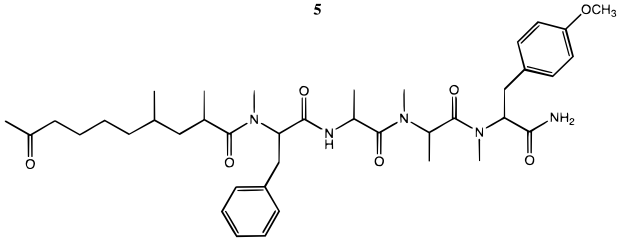
3



4



5



6

* To whom correspondence should be addressed. Tel.: (541)737-5801. Fax: (541)737-3999. E-mail, gerwickw@ccmail.orst.edu.
† Oregon State University.
‡ Syntex Research; present address: Progenitor, Inc., 4040 Campbell Ave., Menlo Park, CA 94025.

S0163-3864(97)00443-6 CCC: $15.00
Published on Web 04/03/1998

Table 1. NMR Data for Carmabin A at 150 MHz (^{13}C) and 600 MHz (^{1}H) in DMSO-d_6

unit	no.	^{13}C (mult.[a])	^{1}H (mult, J = Hz)	COSY to	HMBC[b] to
NH_2	1		7.33 (s)		C: 2
	2	171.9 (s)			
	3	57.3 (d)	5.31 (obscured)	H: 5	C: 2, 5, 6, 13
	5	33.0 (t)	3.25 (m)	H: 3	C: 2, 3, 6, 7/11, 8/10
			2.79 (m)		
N-Me*O*-MeTyr	6	129.9 (s)			
	7	129.7 (d)	7.13 (d, 7.0)		C: 3, 5, 6, 8/10, 9, 11
	8	113.8 (d)	6.84 (d, 7.7)		C: 6, 7/11, 9, 10
	9	158.0 (s)			
	10	113.8 (d)	6.84 (d, 7.7)		C: 6, 7/11, 8, 9
	11	129.7 (d)	7.13 (d, 7.0)		C: 3, 5, 6, 7, 8/10, 9
	12	55.1 (q)	3.72 (s)		C: 8/10, 9
	13	29.9 (q)	2.66 (s)		C: 3, 14[c]
N-MeAla	14	170.9 (s)			
	15	49.3 (d)	5.32 (obscured)	H: 17	C: 14[c], 17, 18
	17	13.9 (q)	0.97 (d, 5.6)	H: 15	C: 14[c], 15
	18	28.3 (q)	1.99 (s)		C: 15, 19[c]
Ala	19	170.9 (s)			
	20	44.9 (d)	4.44 (t, 6.9)	H: 21, 22	C: 19[c], 22, 23
	21		7.83 (d, 6.8)	H: 20	C: 20, 22, 23
	22	17.2 (q)	1.04 (d, 6.4)	H: 20	C: 19[c], 20
	23	169.7 (s)			
	24	56.1 (d)	5.39 (d, 11.3)	H: 26	C: 23, 26, 27, 33, 34
	26	34.2 (t)	3.16 (m)	H: 24	C: 23, 24, 27, 28/32
			2.87 (m)		
N-MePhe	27	137.7 (s)			
	28	128.0 (d)	7.21 (d)		
	29	128.7 (d)	7.27 (d)		
	30	126.2 (d)	7.14 (d)		C: 28/32, 29/31
	31	128.7 (d)	7.27 (d)		
	32	128.0 (d)	7.21 (d)		
	33	31.1 (q)	2.84 (s)		C: 24, 34
	34	176.7 (s)			
	35	32.6 (d)	2.59 (m)		C: 34, 36, 37, 44
	36	40.5 (t)	0.83 (obscured)		
	37	29.6 (d)	0.98 (obscured)		
	38	35.9 (t)	0.98 (obscured)		
			0.84 (obscured)		
2,4-diMe-9-decynoic acid	39	25.1 (t)	1.33 (obscured)		
	40	28.3 (t)	1.32 (obscured)		
	41	17.7 (t)	2.11 (br s)	H: 43[d]	C: 39, 40, 42, 43
	42	84.5 (s)			
	43	71.1 (d)	2.75 (s)	H: 41[d]	C: 40, 41, 42
	44	16.9 (q)	0.86 (d, 6.0)		
	45	19.5 (q)	0.61 (d, 4.8)		C: 36, 37, 38

[a] Multiplicity was determined using the DEPT sequence. ^{1}H−^{13}C connectivities assigned by HMQC. [b] HMBC optimized for 6 Hz coupling. [c] Overlapping ^{13}C NMR resonances. [d] Long-range COSY correlation.

suggested a lipopeptide structure. The NMR spectral profiles were exceptionally complex due to the presence of three *N*-methyl substituents, which resulted in five discernible conformers by NMR in an approximate ratio of 23:6:5:1:1. As variable temperature NMR analysis did not appreciably improve this feature, structure elucidation was based on NMR analysis of the major conformer at 298 K in DMSO-d_6.

Standard amino acid analysis revealed the presence of one residue of Ala as the only non-*N*-methylated amino acid; its secondary amide proton resonated at δ 7.83 (d, J = 7 Hz). The spin systems for three other amino acids (*N*-Me,*O*-MeTyr, *N*-MePhe, and *N*-MeAla) were deduced from pulsed-field gradient ^{1}H−^{1}H COSY, HMQC, HMBC, TOCSY, and HMQC−TOCSY NMR experiments.[11] The presence of five carbonyls (δ 176.7, 171.9, 170.9, 170.9, 169.7); six aromatic double bonds (δ 158.0, 137.7, 129.9, 129.7, 129.7, 128.7, 128.7, 128.0, 128.0, 126.2, 113.8, 113.8) associated with the *N*-Me,*O*-MeTyr, and *N*-MePhe amino acids; and a terminal acetylene recognized from a combination of ^{1}H, ^{13}C, DEPT, and HMQC NMR experiments, accounted for all 15 degrees of unsaturation required by the molecular formula, and hence, implied a linear lipopeptide structure.

The furthest downfield carbonyl (δ 176.7, C-34) was attributed to a fatty acid amide linkage as a result of the HMBC correlations from H-33, H-35, and H-44 to the C-34 carbonyl. The fatty acid chain was defined as 2,4-dimethyl-9-decynoic acid from COSY, TOCSY, HMQC−TOCSY (mixing times of 28 and 55 ms), and HMBC experiments. Due to overlap in the methylene region, the proton resonances of H-39 and H-40 could not be unambiguously assigned. Several lipopeptides with a 2,4-dimethyloctanoic acid chain have been isolated from *Lyngbya majuscula*,[7,8] but this appears to be the first instance that includes a terminal acetylene functionality in the fatty acid chain.

A ^{1}H−^{15}N HMBC NMR experiment[12] clearly showed the presence of five distinct nitrogen resonances, four

Table 2. NMR Data for Carmabin B at 150 MHz (^{13}C) and 600 MHz (^{1}H) in DMSO-d_6

unit	no.	^{13}C (mult.[a])	^{1}H (mult, J = Hz)	COSY to	HMBC[b] to
NH_2	1		7.33 (br s)		C: 2
	2	171.9 (s)			
	3	57.3 (d)	5.30 (obscured)	H: 5	C: 2, 5, 13
	5	33.0 (t)	3.25 (m)	H: 3	C: 3, 7/11
			2.79 (m)		
N-Me*O*-MeTyr	6	129.9 (s)			
	7	129.7 (d)	7.13 (d, 8.4)	H: 8	C: 5, 6, 8/10, 11
	8	113.8 (d)	6.84 (d, 8.4)	H: 7	C: 7/11, 9, 10
	9	158.0 (s)			
	10	113.8 (d)	6.84 (d, 8.4)	H: 11	C: 8, 9, 7/11
	11	129.7 (d)	7.13 (d, 8.4)	H: 10	C: 5, 6, 7, 8/10
	12	55.1 (q)	3.72 (s)		C: 9
	13	30.3 (q)	2.66 (s)		C: 3, 14[c]
N-MeAla	14	170.8 (s)			
	15	49.3 (d)	5.32 (obscured)	H: 17	C: 14[c], 17, 18
	17	13.9 (q)	0.97 (d, 6.5)	H: 15	C: 14[c], 15
	18	28.3 (q)	1.99 (s)		C: 15, 19[c]
Ala	19	170.8 (s)			
	20	44.9 (d)	4.45 (t, 6.7)	H: 21, 22	C: 19[c], 22
	21		7.83 (d, 7.3)	H: 20	C: 23
	22	17.2 (q)	1.04 (d, 6.8)	H: 20	C: 19[c], 20
	23	169.7 (s)			
	24	56.0 (d)	5.38 (dd, 11.6, 4.9)	H: 26	C: 23, 26, 33
	26	34.1 (t)	3.16 (m)	H: 24	C: 24, 27
			2.88 (m)		
N-MePhe	27	137.7 (s)			
	28	128.0 (d)	7.20 (d)		
	29	128.7 (d)	7.20 (d)		
	30	126.1 (d)	7.13 (d)		C: 28/32, 29/31
	31	128.7 (d)	7.20 (d)		
	32	128.0 (d)	7.20 (d)		
	33	31.1 (q)	2.84 (s)		C: 24, 34
	34	176.7 (s)			
	35	32.5 (d)	2.59 (m)		C: 34, 44
	36	40.5 (t)	0.83 (obscured)		
	37	29.7 (d)	0.98 (obscured)		
	38	36.3 (t)	0.97 (obscured)		
			0.84 (obscured)		
2,4-diMe-9-oxo-decanoic acid	39	25.6 (t)	1.28 (obscured)		
	40	23.6 (t)	1.34 (t, 7.2)		C: 41, 42
	41	42.8 (t)	2.37 (t, 7.2)		C: 39, 40, 42
	42	208.0 (s)			
	43	29.5 (q)	2.07 (s)		C: 41, 42
	44	16.9 (q)	0.86 (d, 6.5)		
	45	19.4 (q)	0.60 (d, 6.2)		C: 36, 37, 38

[a] Multiplicity was determined using the DEPT sequence. ^{1}H−^{13}C connectivities assigned by HMQC. [b] HMBC optimized for 6 Hz coupling. [c] Overlapping ^{13}C NMR resonances.

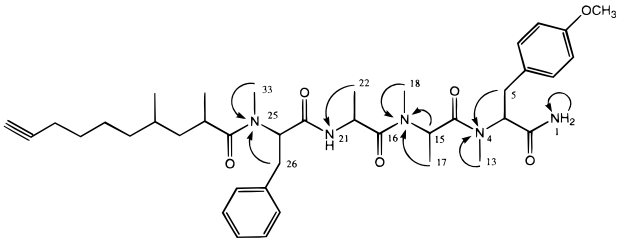


Figure 1. Correlations observed in the ^{1}H−^{15}N HMBC NMR experiment of carmabin A (**5**).

of which were readily assignable to the nitrogens of the *N*-Me,*O*-MeTyr, *N*-MeAla, Ala, and *N*-MePhe amino acids (Figure 1). The remaining nitrogen showed a strong correlation to a broad singlet immediately adjacent to the aromatic proton resonances in the ^{1}H NMR experiment. Initially this was erroneously attributed to one of the conformers; however, a substantial reduction in the signal after the addition of D_2O revealed its exchangeable nature. This, together with an HMBC correlation to C-2 of *N*-Me,*O*MeTyr, implied a terminal primary amino group. This was corroborated by HR-FABMS data, which showed an intense $[M - NH_2]^+$ ion (20%) at m/z 687.4120 (Δ −0.1 mmu).

The above substructures accounted for all of the atoms in **5**. Their sequence was dictated by FABMS and HMBC NMR analysis. A sequential loss of NH_2 (−16), *N*-Me-*O*-MeTyr (−191), *N*-MeAla (−85), and Ala (−71) (Figure 2), together with HMBC correlations from the N-21 proton of Ala to the C-23 carbonyl carbon of *N*-MePhe and from the C-33 *N*-Me protons of *N*-MePhe to the C-34 carbonyl carbon of 2,4-dimethyl-9-decynoic acid, defined the structure of carmabin A (**5**).

Carmabin B (**6**), also isolated as a white solid, had a molecular formula of $C_{40}H_{59}N_5O_7$ based on its positive HRFABMS. Hence, the molecular formula of **6** differed by 18 mass units compared to **5**, a consequence of an additional oxygen and two hydrogen atoms. Once again, the spectral profiles were complicated by the presence of three *N*-methyl substituents resulting in five discern-

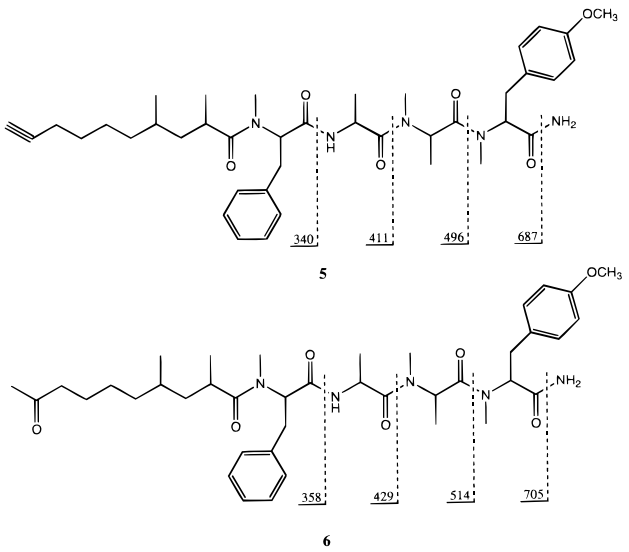


Figure 2. Mass spectral fragmentations for carmabins A (**5**) and B (**6**).

ible conformers in an approximate ratio of 18:5:4:1:1. Comparison of the ^{1}H and ^{13}C NMR data of carmabin B (Table 2) with that of carmabin A (**5**) indicated that the tetrapeptide portion with the terminal primary amide was the same in both compounds; thus, they differed only in the fatty acid portion. The absence of the acetylene ^{13}C NMR resonances (δ 84.5 and 71.1) and the presence of both an additional carbonyl and methyl resonance at δ 208.0 and 29.5, respectively, in the ^{13}C NMR spectrum of **6** compared with **5**, suggested that the acetylene functionality had been replaced with a methyl ketone. This proposal was consistent with mass spectral data (Figure 2) and was further supported by HMBC correlations from the methyl singlet at δ 2.07 (H-43) to the ketone carbonyl at δ 208.0 (C-42) and the methylene carbon at δ 42.8 (C-41), and from the methylene triplet at δ 1.34 (H-40) to the adjacent C-41 methylene carbon and the C-42 ketone carbonyl.

Although carmabins A and B were the inactive components of a fraction showing antiproliferative activity their biological evaluation is continuing. The carmabins are structurally unique in that they contain a primary amide and terminal functionalization of the lipid component.

Experimental Section

General Experimental Procedures. NMR spectra were recorded on a Bruker DRX600 spectrometer with the solvent as an internal standard (DMSO-d_6 at δ 2.50 for ^{1}H and 39.51 for ^{13}C). HRMS were recorded on a Kratos MS50TC mass spectrometer. Optical rotations were measured on a Perkin−Elmer 141 polarimeter. UV and IR spectra were recorded on Hewlett−Packard 8452A UV/vis and Nicolet 510 spectrophotometers, respectively. HPLC separations were performed on a Phenomenex Spherisorb ODS column (10 μm, 250 × 4.6 mm) with UV detection at 220 nm on a Waters LambdaMax 480 spectrophotometer. The standard amino acid analysis was performed by B. Herkenrath, Protein Structure Laboratory, University of California, Davis.

Collection. The marine cyanobacterium *L. majuscula* was collected by hand from shallow water on 15 December 1993, at Barbara Beach (Spanish Waters), Curaçao, Netherlands Antilles, and stored in 2-propanol at reduced temperature until workup. A voucher specimen is housed at Oregon State University and is available from WHG as NAC-15 Dec 93−1.

Extraction and Isolation. A total of 446 g (dry wt) of the alga was extracted with CH_2Cl_2−MeOH (1:1) five times to give a crude extract that was partitioned between *s*-BuOH and H_2O. A portion (800 mg) of the *sec*-BuOH extract (1.6 g) was subjected to reversed-phase vacuum liquid chromatography with a stepwise gradient from H_2O through MeOH. The fraction eluting with 80% MeOH−H_2O (91 mg) was further fractionated by reversed-phase HPLC (65% MeCN−H_2O) to yield compounds **5** (35.6 mg) and **6** (5.8 mg).

Carmabin A (5): white solid; $[\alpha]^{27}_D$ −109° (*c* 0.45, MeOH); UV λ_{max} (MeOH) 222 nm (ϵ 6900), 278 nm (ϵ 1400); IR ν_{max} (film) 3395, 3298, 2929, 2858, 1634, 1511, 1460, 1408, 1250, 1081 cm^{-1}; ^{1}H and ^{13}C NMR data, see Table 1; FABMS (3-NBA, positive) *m/z* (rel int) 726 (18), 704 (1), 687 (20), 496 (63), 411 (11), 340 (100); HR-FABMS (3-NBA, positive) *m/z* 726.4200 (M + Na)$^+$ calcd for $C_{40}H_{57}N_5O_6Na$ (Δ +0.3 mmu), 704.4390 (M + H)$^+$ calcd for $C_{40}H_{58}N_5O_6$ (Δ −0.7 mmu), 687.4120 (M − NH_2)$^+$ calcd for $C_{40}H_{55}N_4O_6$ (Δ −0.1 mmu).

Carmabin B (6): white solid; $[\alpha]^{27}_D$ −102° (*c* 0.19, MeOH); UV λ_{max} (MeOH) 222 nm (ϵ 7000), 278 nm (ϵ 1200); IR ν_{max} (film) 3395, 3313, 2924, 2863, 1634, 1511, 1460, 1403, 1250, 1081 cm^{-1}; ^{1}H and ^{13}C NMR data, see Table 2; FABMS (3-NBA, positive) *m/z* (rel int) 744 (22), 722 (4), 705 (22), 514 (54), 429 (12), 358 (100); HR-FABMS (3-NBA, positive) *m/z* 744.4309 (M + Na)$^+$ calcd for $C_{40}H_{59}N_5O_7Na$ (Δ −0.3 mmu), 722.4493 (M + H)$^+$ calcd for $C_{40}H_{60}N_5O_7$ (Δ −0.1 mmu), 705.42260 (M − NH_2)$^+$ calcd for $C_{40}H_{57}N_4O_7$ (Δ −0.1 mmu).

Acknowledgment. We are grateful to the CARMABI Tropical Research Center in Curaçao for their assistance in collection of the algal material and to M. A. Roberts for taxonomic identification and laboratory culture of the alga. We thank V. Hsu and M. Ivancic (Department of Biochemistry, OSU) and R. Kohnert and T. Williamson (Department of Chemistry, OSU) for assistance with NMR experiments and B. Arbogast (Department of Agricultural Chemistry, OSU) for help in obtaining mass spectral data. This work was supported by the National Cancer Institute (CA 52955).

References and Notes

(1) Gerwick, W. H.; Proteau, P. J.; Nagle, D. G.; Hamel, E.; Blokhin, A.; Slate, D. L. *J. Org. Chem.* **1994**, *59*, 1243−1245.
(2) Orjala, J.; Nagle, D. G.; Hsu, V. L.; Gerwick, W. H. *J. Am. Chem. Soc.* **1995**, *117*, 8281−8282.
(3) Orjala, J.; Nagle, D. G.; Gerwick, W. H. *J. Nat. Prod.* **1995**, 58, 764−768.
(4) Orjala, J.; Gerwick, W. H. *J. Nat. Prod.* **1996**, *59*, 427−430.
(5) Assay to the MRC-5 human embryonic lung cell line in confluent and proliferating states (cytotoxicity assessment assay, Syntex Discovery Research, unpublished); crude extract containing curacin A and carmabin A, IC_{50} = 6.58 ± 0.98 μg/mL confluent, $IC_{50} \leq 0.003$ μg/mL proliferating; carmabin A-containing fraction IC_{50} = 4.8 ± 0.6 μg/mL confluent, IC_{50} = 0.13 ± 0.06 μg/mL proliferating; curacin A-containing fraction IC_{50} = 50.2 ± 5.5 μg/mL confluent, IC_{50} = 0.01 ± 0.008 μg/mL proliferating.
(6) Marner, F.-J.; Moore, R. E.; Hirotsu, K.; Clardy, J. *J. Org. Chem.* **1977**, *42*, 2815−2819.
(7) Moore, R. E.; Entzeroth, M. *Phytochemistry*, **1988**, *27*, 3101−3103.

(8) Koehn, F. E.; Longley, R. E.; Reed, J. K. *J. Nat. Prod.* **1992**, *55*, 613−619.
(9) Nagle, D. G.; Paul, V. J.; Roberts, M. A. *Tetrahedron Lett.* **1996**, *37*, 6263−6266.
(10) Yoo, H.-D.; Gerwick, W. H. *J. Nat. Prod.* **1995**, *58*, 1961−1965.
(11) Lerner, L., Bax, A. *J. Magn. Reson.* **1986**, *69*, 375−380.
(12) Fukuzawa, S., Matsunaga, S., Fusetani, N. *Tetrahedron Lett.* **1996**, *37*, 1447−1448.

NP970443P