Kahalalides: Bioactive Peptides from a Marine Mollusk *Elysia rufescens* and Its Algal Diet *Bryopsis* sp.[1]

Mark T. Hamann, Clifton S. Otto, and Paul J. Scheuer*

Department of Chemistry, University of Hawaii of Manoa, Honolulu, Hawaii 96822

D. Chuck Dunbar

Department of Pharmacognosy, School of Pharmacy, The University of Mississippi, University, Mississippi 38677

Received May 14, 1996[⊗]

In addition to the previously reported bioactive kahalalide F six new peptides are described. Six of these, including kahalalide F, are cyclic depsipeptides, ranging from a C_{31} tripeptide to a C_{75} tridecapeptide isolated from a sacoglossan mollusk, *Elysia rufescens*. The mollusk feeds on a green alga, *Bryopsis* sp., which has also been shown to elaborate some of these peptides in smaller yields, in addition to an acyclic analog of F, kahalalide G. The bioassay results of antitumor, antiviral, antimalarial, and OI (activity against AIDS opportunistic infections) tests are reported.

Six cyclic depsipeptides, kahalalides A−F, ranging from a C_{31} tripeptide to a C_{75} tridecapeptide, were isolated from a sacoglossan mollusk, *Elysia rufescens*. The mollusk feeds on a green alga, *Bryopsis* sp., which also elaborates an identical C_{75} peptide and its acyclic analog, kahalalide G. Except for the rare dehydroaminobutyric acid in kahalalide F and G, all constituent amino acids are of common occurrence. Each peptide also contains a fatty acid. Structures were determined by spectral methods. Absolute configuration of most amino acids was determined by chiral GC−MS, HPLC, and TLC. Kahalalide A shows modest antimalarial activity against *Plasmodium falciparum*; kahalalide E exhibits selective activity against *Herpes simplex* II virus (HSV II). Kahalalide F exhibits selective activity against solid tumors and some of the AIDS opportunistic infections (OI).

Among opisthobranch marine mollusks (a subclass of Gastropods), members of two orders, the sluglike sea hares and nudibranchs have become frequent targets for chemical research, since most of them are conspicuous and accumulate secondary metabolites from their algal[2] or invertebrate animal[3] diets. Another opisthobranch order, the sacoglossans, some of which have shells, generally feed on green algae and ingest functioning chloroplasts,[4] which may participate in the biosynthesis of secondary metabolites, frequently polypropionates. This was demonstrated for the sacoglossan *Placobranchus ocellatus*.[5] A related genus *Elysia*, also in the family Plakobranchidae, is represented in Hawaii by a number of species.[6] *Elysia degeneri* is reported to feed on calcareous green algae, *Udotea* spp., while no food preference is mentioned for *E. rufescens*. *Udotea* spp. have been investigated by Paul, who isolated diterpene aldehydes possessing feeding deterrent properties.[7] While collecting *E. rufescens*, we observed the animals feeding on a green alga, *Bryopsis* sp. Our research revealed that the constituents of *E. rufescens* and of *Bryopsis* sp. are of neither diterpenoid nor polypropionate origin, but are amino and fatty acid-derived depsipeptides. We have reported the structure and some of the properties of the major constituent, kahalalide F, in a preliminary communication.[8] In this paper we describe the entire spectrum of peptides isolated from *E. rufescens* and from *Bryopsis* sp.

E. rufescens (Gould, 1852) is an orange-fringed blue-green, small (1−4 cm), soft-bodied sacoglossan. During April and May 1991, approximately 200 animals (300 g wet) were collected by snorkeling near Black Point, O'ahu; they were extracted with ethanol and subjected to flash chromatography on silica followed by HPLC on a reversed-phase C_{18} column. Gradient elution with acetonitrile/water/TFA ranging from 30:70:0.1 to 55:45:0.1 furnished six peptides which were eluted in the following order: kahalalide[9] D (**4**, 12 mg, 0.004%), F (**6**, 40 mg, 0.01%), C (**3**, 4 mg, 0.001%), A (**1**, 25 mg, 0.008%), B (**2**, 25 mg, 0.008%), and E (**5**, 10 mg, 0.003%).[10]

In May 1993, it became necessary to recollect the animals for *in vivo* assays and to maximize the yield of kahalalide F (**6**). The animals (293 specimen, 216 g) were drained of excess sea water and extracted with ethanol. After removal of ethanol, the aqueous suspension was partitioned against hexane. To the organic phase were added equal volumes of methanol and hexane. The methanolic phase produced a residual greenish oil (3.6 g). This oil was subjected in 600 mg batches to countercurrent chromatography (ethyl acetate/heptane/methanol/water, 7:4:4:3), yielding 2.1 g (1%) of kahalalide F (**6**), representing a significant improvement in yield.

[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1996.

(1) From the Ph.D. Dissertation of M.T.H., University of Hawaii, Honolulu, HI, 1992.

(2) Scheuer, P. J. *Isr. J. Chem.* **1977**, *16*, 52−56.

(3) Karuso, P. In *Bioorganic Marine Chemistry*; Scheuer, P. J., Ed.; Springer: Berlin, 1987; Vol. 1, pp 31−60.

(4) Trench, R. K.; Trench, M. E.; Muscatine, L. *Biol. Bull.* **1972**, *142*, 335−349.

(5) Ireland, C.; Scheuer, P. J. *Science* **1979**, *205*, 922−923.

(6) Kay, E. A. *Hawaiian Marine Shells*; Bishop Museum Press: Honolulu, 1979; pp 451−454.

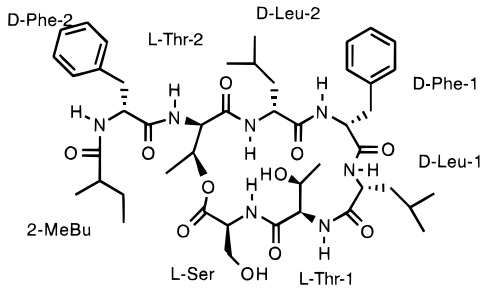
(7) Paul, V. J. In *Ecological Roles of Marine Natural Products*; Paul, V. J., Ed.; Cornell University Press: Ithaca, NY, 1992; pp 24−50.

(8) Hamann, M. T.; Scheuer, P. J. *J. Am. Chem. Soc.* **1993**, *115*, 5825−5826.

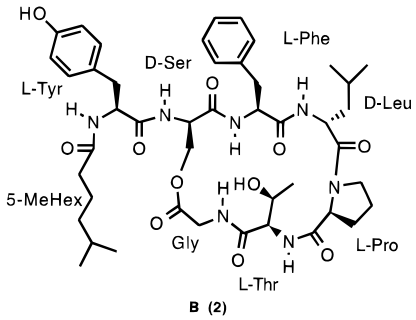
(9) Black Point, the collection site, is in the Kahala district of the island of O'ahu.

(10) The letters A, B... refer to the sequence in which the structures were elucidated.

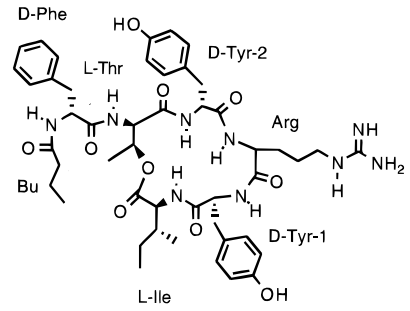
S0022-3263(96)00877-8 CCC: $12.00 © 1996 American Chemical Society



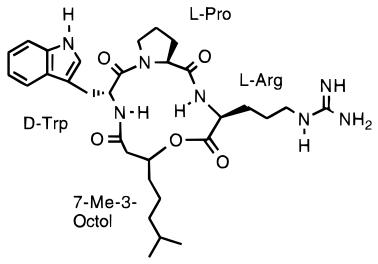
A (1)



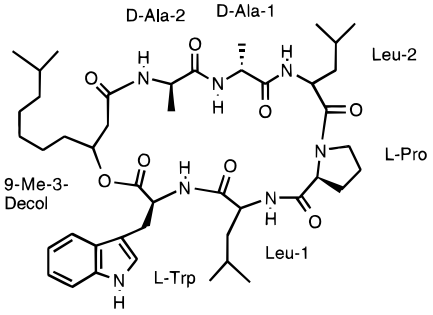
B (2)



C (3)

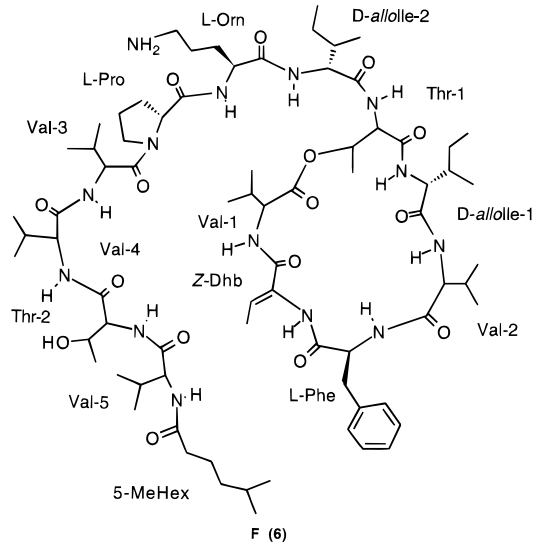


D (4)



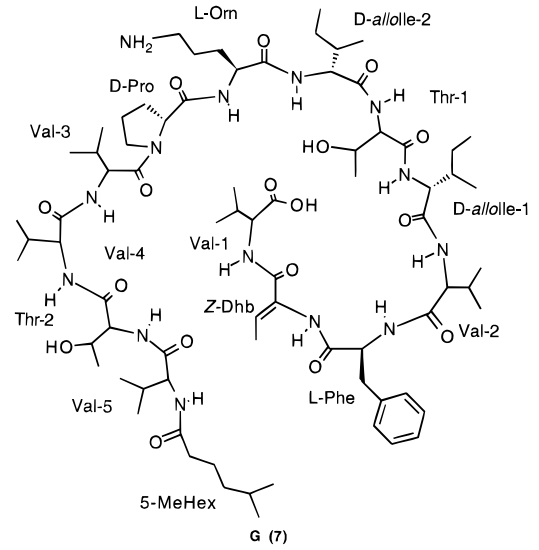
E (5)

The green alga, *Bryopsis* sp., on which *E. rufescens* grazes, also was collected at Black Point, O'ahu, in June 1992. Ethanol extraction of the alga (3 kg) was followed by flash chromatography on silica (ethyl acetate/methanol, 1:1) and HPLC, first on a reversed phase C_{18} column (water, isopropyl alcohol, TFA, 55:45:0.1), which yielded kahalalide A (**1**, 15 mg, 0.0005%) and B (**2**, 50 mg, 0.002%); further chromatography on an amino-bonded column (methanol/ethyl acetate, 65:35) yielded kahalalide



F (6)

G (**7**) as a white amorphous powder (15 mg, 0.0005%) and kahalalide F (**6**, 5 mg, 0.0002%).



G (7)

Kahalalide D (**4**) is the smallest (see Table 1) of the peptides and serves as a model for elucidation of all the other peptide structures. High-resolution FAB MS data furnished a molecular formula of $C_{31}H_{44}N_7O_5$ for kahalalide D (**4**). This peptide is made up of three amino acids, L-arginine, L-proline, and D-tryptophan. The β-hydroxy group of its fatty acid component, 3-hydroxy-7-methyloctanoic acid (7-Me-3-Octol), furnishes the ester linkage of the depsipeptide cycle. It is an isoacid, as are most of the fatty acid components of the kahalalides. The structures of all amino acids were determined by 2D-NMR including HMQC,[11] HMBC,[12] COSY, and TOCSY[13] experiments. Chiral GCMS,[14] HPLC,[15] and TLC experiments were used to determine the absolute configuration. Because of the modest size of kahalalide D the ^{1}H NMR signals were well resolved so that the amino acid

(11) Sklenar, V.; Bax, A. *J. Magn. Reson.* **1987**, *71*, 379−383.
(12) Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2094−2096.
(13) Inagaki, F.; Shimada, I.; Kohda, D.; Suzuki, A.; Bax, A. *J. Magn. Reson.* **1989**, *81*, 186−190.
(14) Nicholson, G. J.; Hartmut, F.; Bayer, E. *HRC & CC, J. High Resol. Chromatogr. Chromatogr. Commun.* **1979**, *Volume 2*, 411−415.
(15) Gubitz, G.; Jellenz, W. *J. Chromatogr.* **1981**, *203*, 377−384.

Table 1. Comparative Composition of *Elysia* and *Bryopsis* Peptides

			amino acids													
	mol formula	mol wt	Ala	Arg	Dhb	Gly	Ile	Leu	Orn	Phe	Pro	Ser	Thr	Trp	Tyr	Val	fatty acid
kahalalide																	
A	$C_{46}H_{67}N_7O_{11}$	894.1						2		2		1	2				2-MeBu
B	$C_{45}H_{63}N_7O_{11}$	878.0				1		1		1	1	1	1		1		5-MeHex
C	$C_{47}H_{63}N_9O_{10}$	914.1		1			1			1			1		2		Bu
D	$C_{31}H_{45}N_7O_5$	595.7		1							1			1			7-Me3Octol
E	$C_{45}H_{69}N_7O_8$	836.1	2					2			1			1			9-Me3Decol
F	$C_{75}H_{124}N_{14}O_{16}$	1477.9			1		2		1	1	1		2			5	5-MeHex
G	$C_{75}H_{126}N_{14}O_{17}$	1495.9			1		2		1	1	1		2			5	5-MeHex

sequence could be determined independently by ROESY[16] or HMBC experiments. The HMBC experiment provided a correlation between the NH of tryptophan (8.80 ppm) and the carbonyl of 7-Me-3-Octol (171.64 ppm). There is also an HMBC correlation between the NH of arginine (6.78 ppm) and the carbonyl of proline (171.08 ppm). The ester linkage was established by an HMBC correlation between the β-proton (5.03 ppm) of the hydroxy acid and the carbonyl of arginine (170.26 ppm), leaving one connection to be made between tryptophan and proline. A ROESY experiment provided the Pro-Trp connection by correlating the α-proton (4.55 ppm) of tryptophan with the δ-protons (3.70, 2.70 ppm) of proline. The ROESY experiment also correlated the NH of tryptophan (8.80 ppm) to the α-protons (2.52, 2.59 ppm) of the hydroxy acid. The NH of arginine (6.78 ppm) was correlated by a ROESY experiment to the α-proton (4.07 ppm) of proline, leaving only the ester bond between arginine and the hydroxy acid to complete the cycle. The absolute configuration of the proline residue was determined to be L-Pro by chiral GCMS. We were unable to resolve arginine and tryptophan by chiral GCMS. Arginine was determined to be L-Arg by chiral HPLC (ligand exchange), and tryptophan proved to be D-Trp by chiral HPLC and TLC. The stereochemistry of the 3-hydroxy-7-methyloctanoic acid is still under investigation and will require isolation of additional material. Kahalalide D has only been isolated from *E. rufescens* in 0.005% yield (wet).

Next in size are kahalalide A (**1**), B (**2**), C (**3**), and E (**5**), with either six (C and E) or seven (A and B) amino acid residues. Kahalalide A (**1**), $C_{46}H_{67}N_7O_{11}$, contains three hydroxy amino acid residues (two threonines and one serine) and only two other amino acids: two residues each of leucine and phenylalanine. Its ester linkage arises from the carbonyl of serine and the hydroxyl of threonine-2. 2-Methylbutyric acid is the fatty acid, which forms an amide with phenylalanine-2. Kahalalide A and B are soluble in MeCN and provide sharp and detailed proton-detected 2D NMR data. The amino acid sequence of kahalalide A (**1**) was assembled by HMBC correlations between NH protons and vicinal carbonyl carbons. In MeCN the NH protons also gave distinct HMBC correlations to the α- and β-carbons of the corresponding amino acid. The ester bond was assigned by an HMBC correlation between the β-proton (5.45 ppm) of threonine-2 and the carbonyl (170.1 ppm) of serine. The absolute configuration of the amino acids was determined by chiral GCMS. Both threonines and the serine residue have the L configuration, while phenylalanine and leucine residues have the D configuration. The absolute configuration of 2-methylbutyric acid remains to be determined.

Kahalalide B (**2**), $C_{45}H_{63}N_7O_{11}$, while of approximately the same size as A, is made up of seven different amino acids (Gly, Thr, Pro, Leu, Phe, Ser, Tyr). 5-Methylhexanoic acid, which is also present in F(**6**) and G(**7**), is the fatty acid that forms an amide with Tyr and is part of the side chain. The ester linkage of the cyclic depsipeptide is formed from the serine hydroxyl and the carbonyl of Gly. The two β-protons of Ser are part of the rigid 19-carbon ring system and resonate at 3.38 and 4.39 ppm. The latter proton provides an HMBC correlation to the carbonyl carbon of Gly (169.81 ppm), while the proton at 3.38 ppm shows an HMBC correlation to the Ser carbonyl at 169.26 ppm. The three-residue side chain of kahalalide B includes 5-methylhexanoic acid. Because of the restricted rotation around the two amide bonds the molecule exists in two conformations, which seem unaffected by changes in temperature or solvent. These conformational differences can be seen in the ^{1}H and ^{13}C NMR spectra. The NH protons at 6.92, 7.91, and 7.73 show distinct shoulders indicating the existence of two conformers. This can be seen most clearly in the carbon spectrum, where a subset of carbons resonate at approximately 20% of the intensity of the rest of the carbons. These signals can be assembled into 5-methylhexanoic acid by HMQC, HMBC, and COSY data. Both conformers lead to the same tyrosine residue in an HMBC experiment. Chiral GCMS allowed the complete assignment of absolute configuration and showed that the amino acid residues Tyr, Phe, Pro, and Thr have the L configuration, while Leu and Ser have the D configuration.

Kahalalide C (**3**), $C_{47}H_{63}N_9O_{10}$, is the only other arginine-containing peptide besides D. The carboxyl group of isoleucine is esterified with the hydroxyl of threonine. This connection could be made with an HMBC correlation between the β-proton of Thr (5.36 ppm) and the carbonyl of Ile (169.9 ppm). The Ile-Tyr-1 sequence was established with an HMBC correlation between the NH proton of Ile (7.78 ppm) and the carbonyl of Tyr-1 (170.3 ppm). Similarly, it was possible to establish the Tyr-1-Arg-Tyr-2-Thr sequence by an HMBC experiment utilizing the NH protons (8.16, 8.40, 7.67 ppm) and their connection to the carbonyl of the vicinal amino acid (171.4, 171.2, 167.8 ppm), respectively. Kahalalide C contains the simplest fatty acid, butyric, which forms an amide with the amino group of phenylalanine as part of the side chain. The Thr-Phe sequence was established with an HMBC correlation between the α-proton of Thr and the carbonyl of Phe. The final connection between Phe-But could then be made without the use of spectroscopy. Under our experimental conditions we were unable to see any correlations between the NH or α-proton of Phe and the carboxylic acid of the butyric acid residue. Chiral GCMS showed that the amino acid residues Ile and Thr have the L configuration, while Phe and both Tyr residues have the D configuration. It will be necessary to isolate

(16) Bothner-By, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.

Table 2. *In Vitro* Activity of Kahalalide F Against AIDS OI Pathogens

assay	kahalalide F	amphotericin B	rifampin
OI pathogen (MIC μg/mL)			
Candida albicans	1.56	0.39	
Cryptococcus neoformans	1.56	0.39	
Aspergillus flavus	25.00	1.56	
Aspergillus fumigatus	6.25	1.56	
*M*ycobacterium intracellulare	25.00		0.20
antimicrobial (MIC μg/mL)			
Trychophyton mentagrophytes	1.56		
OI pathogen (MFC μg/mL)			
C. albicans	25.00		
C. neoformans	12.50		
A. flavus	100.00		
A. fumigatus	100.00		
M. intracellulare	50.00		
antimicrobial (MFC μg/mL)			
T. mentagrophytes	6.25		

Table 3. ^{1}H NMR and ^{13}C NMR Data for Kahalalide A (1) in CD_3CN

amino acid	carbon[a]	^{13}C NMR, ppm	mult	^{1}H NMR,[b] ppm	multiplicity
serine	1	170.1	s	(NH) 6.91	d, J = 5.8
	2	57.3	d	3.99	m
	3	61.9	t	3.52	d, J = 4.8
threonine-1	1	171.0	s	NH, 7.92	d, J = 7.2
	2	61.2	d	3.99	m
	3	66.6	d	4.48	dq, J = 6.4, 1.9
	4	20.9	q	1.31	d, J = 6.4
leucine-1	1	175.4	s	NH 7.21	m
	2	52.5	d	4.80	m
	3	43.5	t	1.47, 1.40	m
	4	25.6	d	1.54	m
	5	23.0	q	0.91	d, J = 6.0
	6	23.0	q	0.90	d, J = 6.0
phenylalanine-1	1	171.9	s	NH 7.62	d, J = 9.8
	2	55.6	d	4.80	m
	3	39.6	t	3.25, 2.81	dd, J = 14.3, 5.0; dd, J = 14.3, 10.3
	4	138.1	s		m
	5, 5′	129.8	d	7.21	m
	6, 6′	129.4	d	7.27	m
	7	127.7	d	7.22	
leucine-2	1	172.3	s	NH 7.45	d, J=9.3
	2	54.3	d	4.29	q, J = 9.5
	3	42.4	t	1.21	t, J = 7.7
	4	25.3	d	1.05	m
	5	22.7	q	0.69	d, J = 6.4
	6	22.1	q	0.75	d, J = 6.6
threonine-2	1	169.4	s	NH 8.22	J = 9.5
	2	56.7	d	4.40	dd, J = 9.5, 2.1
	3	70.0	d	5.45	dq, J = 6.4, 2.1
	4	16.1	q	0.63	d, J = 6.4
phenylalanine-2	1	174.4	s	NH 7.34	d, J = 4.8
	2	56.9	d	5.06	dt, J = 7.7, 5.0
	3	37.9	t	3.01	dd, J = 7.4, 1.3
	4	137.5	s		
	5, 5′	130.3	d	7.20	m
	6, 6′	129.4	d	7.27	m
	7	127.9	d	7.22	m
2-methylbutyric acid	1	180.7	s		
	2	43.0	d	2.36	m
	3	28.4	t	1.65, 1.50	m, m
	4	12.4	q	0.89	t, J = 7.5
	5	18.0	q	1.13	d, J = 6.9

[a] At 125 MHz. [b] At 500 MHz.

additional material to determine the absolute configuration of the Arg residue by chiral HPLC.

The last of these four kahalalides of intermediate size, E (**5**), $C_{45}H_{69}N_7O_8$, is made up of only four amino acids, two each of alanine and leucine and one each of proline and tryptophan. It embraces the largest (22-membered) ring and the largest fatty acid, 3-hydroxy-9-methyldecanoic (9-Me-3-Decol). Its hydroxyl forms the ester linkage with the carbonyl of tryptophan (170.0 ppm) as established by an HMBC correlation with the β-proton of the 9-Me-3-Decol residue (5.11 ppm). The sequence of the amino acids could be established by both HMBC and ROESY experiments. The NH proton (8.07 ppm) on Trp shows both an HMBC correlation to the carbonyl (171.2 ppm) and a ROESY correlation to the α-proton (4.26 ppm) on Leu-1. The NH proton (7.95 ppm) of Leu-1 also shows an HMBC correlation to the carbonyl (171.1 ppm), and a ROESY correlation to the α-proton (4.28 ppm) of Pro. The Pro-Leu-2 linkage was established by ROESY correlations between the δ-protons of Pro (3.67 & 3.46 ppm) and the α-proton (4.45 ppm) and β-protons (1.30 and 1.55 ppm) of Leu-2. The Leu-1-Ala-1-Ala-2 linkage was established in the same way, utilizing the NH (7.03, 7.70 ppm) correlations to the carbonyls (171.1, 172.0 ppm) in the HMBC and the α-protons (4.19, 4.06

Table 4. ^{1}H NMR and ^{13}C NMR Data for Kahalalide B (2) in CD_3CN

amino acid	carbon[a]	^{13}C NMR, ppm	mult	^{1}H NMR,[b] ppm	multiplicity
glycine	1	169.8	s	NH 7.18	m
	2	42.5	t	4.03, 3.48	dd, J = 17.4, 5.7; dd, J = 17.4, 5.7
threonine	1	172.2	s	NH 7.21	m
	2	60.8	d	4.24	dd, J = 8.7, 2.7
	3	68.3	d	4.52	m
				OH 5.15	br, s
	4	20.2	q	1.27	d, J = 6.5
proline	1	173.0	s		
	2	62.0	d	4.41	m
	3	30.0	t	2.13	m
	4	25.0	t	2.01,1.9	m,m
	5	48.7	t	3.98; 3.52	dt, J = 7.9, 2.4; m
leucine	1	173.4	s	NH 7.26	d, J = 7.6
	2	52.1	d	4.08	dt, J = 7.1, 3.25
	3	39.4	t	1.47; 1.37	m, m
	4	25.2	d	1.36	m
	5	22.0	q	0.80	d, J = 5.7
	6	23.4	q	0.86	d, J = 6.0
phenylalanine	1	174.7	s	NH 7.73	d, J = 9.0
	2	54.9	d	4.73	q, J = 8.4
	3	39.5	t	3.00	dd, J = 7.3, 3.0
	4	138.1	s		
	5, 5′	130.2	d	7.17	d, J = 7.1
	6, 6′	129.3	d	7.26	t, J = 7.6
	7	127.7	d	7.20	t, J = 7.6
serine	1	169.3	s	NH 7.91	d, J = 9.8
	2	51.1	d	4.50	dt, J = 9.8, 2.2
	3	66.1	t	4.39, 3.38	dd, J = 6.8, 1.9 dd, J = 10.9, 2.7
tyrosine	1	172.4	s	NH 6.92	d, J = 5.2
	2	57.6	d	4.28	q, J = 5.2
	3	36.8	t	2.90, 2.81	dd, J = 13.6, 7.6; dd, J = 13.6, 8.4
	4	128.6	s		
	5, 5′	131.0	d	7.02	d, J = 8.4
	6, 6′	116.3	d	6.75	d, J = 8.7
	7	156.8	s		
5-MeHex	1	175.0	s		
	2	36.3	t	2.15	t, J = 7.3
	3	24.2	t	1.50	m
	4	39.0	t	1.12	m
	5	28.5	d	1.45	m
	6	22.7	q	0.78	d, J = 4.1
	7	22.7	q	0.77	d, J = 4.0
5-MeHex (second conformn)	1	175.3	s		
	2	33.9	t	2.15	m
	3	33.0	t	1.50; 1.30	m; m
	4	29.8	t	1.08	m
	5	34.7	d	1.30	m
	6	19.1	q	0.80	m
	7	11.5	q	0.80	m

[a] At 125 MHz. [b] At 500 MHz.

Table 5. ^{1}H NMR and ^{13}C NMR Data for Kahalalide C (3) in DMSO-d_6

amino acid	carbon[a]	^{13}C NMR, ppm	mult	^{1}H NMR,[b] ppm	multiplicity
isoleucine	1	169.9	s	NH 7.78	d, J = 7.1
	2	56.3	d	4.19	t, J = 7.1
	3	36.3	d	1.75	m
	4	14.7	q	0.69	d, J = 6.8
	5	24.7	t	1.18	m
	6	11.1	d	0.79	d, J = 7.1
tyrosine-1	1	170.3	s	NH 8.16	J = 7.8
	2	53.9	d	4.37	m
	3	35.5	t	2.79, 2.67	m,m
	4	127.6	s		
	5, 5′	129.9	d	6.98	d, J = 8.4
	6, 6′	114.9	d	6.61	d, J = 8.4
	7	156.0	s		
arginine	1	171.4	s	NH 8.40	d, J = 6.4
	2	52.6	d	3.87	q, J = 6.4
	3	26.4	t	1.45; 1.35	m; m
	4	24.6	t	1.08	m
	5	49.6	t	2.90; NH 9.79	m; br s
	6	157.6	s	NH 9.14	br s
tyrosine-2	1	171.2	s	NH, 7.67	d, J = 7.4
	2	54.1	d	4.40	m
	3	37.4	t	2.74; 2.64	m; m
	4	127.1	s		
	5, 5′	129.9	d	6.86	d, J = 8.4
	6, 6′	114.8	d	6.57	d, J = 8.4
	7	155.9	s		
threonine	1	167.8	s	NH, 8.45	d, J = 7.8
	2	56.1	d	4.23	d, J = 10.5
	3	69.3	d	5.26	q, J = 6.7
	4	16.7	q	0.77	d. J = 7.7
phenylalanine	1	172.4	s	NH 8.45	d, J = 7.8
	2	54.7	d	4.73	q, J = 7.1
	3	37.3	t	2.94	m
	4	137.4	s		
	5, 5′	128.0	d	7.31	d, J = 7.4
	6, 6′	129.1	d	7.25	t, J = 7.4
	7	126.2	d	7.17	t, J = 7.1
Bu	1	172.6	s		
	2	36.8	t	2.08	m
	3	18.4	t	1.45	sextet, 6.41
	4	13.4	q	0.76	t, 7.4

[a] At 125 MHz. [b] At 500 MHz.

Table 6. ^{1}H NMR and ^{13}C NMR Data for Kahalalide D (4) in DMSO-d_6

amino acid	carbon[a]	^{13}C NMR, ppm	mult	^{1}H NMR,[b] ppm	multiplicity
arginine	1	170.3	s	NH, 7.48	q, J = 6.2
	2	51.5	d	4.1	q, J = 5.5
	3	27.5	t	1.68; 1.58	m; m
	4	24.5	t	1.35	m
	5	40.3	t	3.03; NH 8.23	t, J = 5.5; br s
	6	157.0	s	NH 7.48	br s
proline	1	171.1	s		
	2	60.6	d	4.07	dd, J = 8.6, 4.1
	3	29.0	t	1.86, 1.64	m, m
	4	23.9	t	1.67	m
	5	46.0	t	3.70; 2.7	q, J = 8.9; m
tryptophan	1	171.6	s	NH 8.80	d, J = 6.5
	2	52.8	d	4.55	q, J = 8.2
	3	25.3	t	3.14; 2.98	dd, J = 14.0, 9.2; dd, J = 14.1, 6.5
	4	109.2	s		
	5	123.9	d	7.19	s
	6	136.0	s		
	7	126.7	s		
	8	111.3	d	7.30	d, J = 7.9
	9	120.9	d	7.05	t, J = 6.9
	10	118.3	d	6.97	t, J = 7.2
	11	118.1	d	7.48	d, J = 7.9
7-Me-3-Octol	1	171.6	s		
	2	38.4	t	2.59, 2.52	dd, J = 15.8, 5.2; dd, J = 15.8, 9.3 d, 2.4
	3	71.9	d	5.03	m; m
	4	33.6	t	1.67; 1.57	m
	5	22.3	t	1.25	m
	6	38.0	d	1.15	m
	7	27.2	q	1.50	d, J = 6.5
	8, 9	22.3	q	0.84	

[a] At 125 MHz. [b] At 500 MHz.

ppm) in the ROESY experiment. The Ala-2 could be linked to the 9-Me-3-Decol residue by an HMBC correlation between the NH (7.92 ppm) of Ala and the carbonyl (169.7 ppm) of 9-Me-3-Decol. Chiral GCMS showed that proline has the L configuration and both Ala units have the D configuration. GCMS also showed that one of the Leu residues is D and one is L. The configuration of the Trp residue was shown to be L by removing the Trp residue from the hydrolysis mixture using RP C18 and determining its configuration by chiral TLC. Kahalalide E is the only peptide other than F that has demonstrated notable biological activity; it is active against HSV II at 5 μg/mL. Determination of the absolute configuration of the hydroxy acid and position of the L and D leucine residues will require isolation of additional material for partial hydrolysis and preparation of Mosher's ester as modified by Ohtani *et al.*[17]

The largest of the *Elysia* peptides, kahalalide F (**6**), was described in the preliminary communication.[8] It is worth noting that its rare amino acid, *Z*-dehydroaminobutyric acid, has recently been reported from a peptide isolated from a terrestrial *Streptomyces sp*. The peptide cypemycin contains four-Dhb residues.[18]

Kahalalide G (**7**), the algal peptide, has a molecular formula of $C_{75}H_{126}N_{14}O_{17}$, which differs from F (**6**) by one additional mole of water. Indeed, it is the only acyclic peptide. The ester linkage between the carboxyl of valine-1 and the hydroxyl of threonine-1 is lacking. The alga also contained the corresponding cyclized kahalalide F and small quantities of A and B. Table 1 summarizes the composition and makeup of the seven compounds. Kahalalide G is also the only amphoteric peptide in the series and will form a gel in most solvents. Obtaining suitable NMR data required that the compound be kept from congealing by acquiring data at 40 °C and with dilute solutions (10 mg/mL). The sequence of the amino acids was established utilizing both HMBC and ROESY experiments. The NH proton (7.30 ppm) of Val-1 shows an HMBC correlation to the carbonyl (164.8 ppm) of Dhb. A correlation between the methyl (1.60 ppm) and NH signal of Dhb (9.36 ppm) in the ROESY experiment indicated *Z* stereochemistry of this uncommon amino acid. The Dhb-Phe linkage was established by an HMBC correlation between the Dhb NH resonance and the Phe carbonyl (172.1 ppm). The Phe-Val-2 connection was made in the same way with the NH resonance (7.68 ppm) of Phe and carbonyl of Val-2 (172.3 ppm). The NH signal (8.14 ppm) of Val-2 shows an HMBC correlation to the carbonyl of Ileu-1 that then shows an NH resonance (7.71 ppm) with a correlation to the Thr-1 carbonyl (171.0 ppm). The β-proton signal of Thr-1 (4.23 ppm) is upfield from the same signal of kahalalide F by 0.84 ppm, providing additional evidence for the lack of ester functionality at this position; this is in addition to absence of the HMBC correlation to the Val-1 carbonyl (173.4 ppm).

(17) Ohtani, I.; Kusumi, T.; Kashman, Y.; Kakisawa, H. *J. Am. Chem. Soc.* **1991**, *113*, 4092−4096.

(18) Minami, Y.; Yoshida, K. I.; Azuma, R.; Krakawa, A.; Rawaudi, T.; Otari, T.; Roniiyama, K.; Omura, S. *Tetrahedron Lett.* **1994**, *35*, 8001−8004.

Table 7. ^{1}H NMR and ^{13}C NMR Data for Kahalalide E (5) in DMSO-d_6

amino acid	carbon[a]	^{13}C NMR, ppm	mult	^{1}H NMR,[b] ppm	multiplicity
tryptophan	1	170.3	s	NH, 8.07	d, J = 7.7
	2	53.6	d	4.42	q, J = 7.4
	3	26.8	t	3.13; 3.05	m; dd, J = 14.7, 8.1
	4	109.4	s		
	5	123.5	d	7.0; NH, 10.66	s; s
	6	136.9	s		
	7	126.9	s		
	8	111.0	d	7.30	d, J = 8.1
	9	120.4	d	7.03	dt, J = 7.0, 1.1
	10	117.9	d	6.95	dt, J = 7.0, 0.7
	11	117.6	d	7.49	d, J = 8.1
leucine-1	1	171.2	s	NH, 7.95	d, J = 8.0
	2	50.5	d	4.26	m
	3	39.5	t	1.50; 1.39	m; m
	4	24.0	d	1.50	m
	5	21.3	q	0.85	m
	6	22.4	q	0.79	d, J = 5.9
proline	1	171.1	s		
	2	69.2	d	4.28	m
	3	28.1	t	2.10; 1.82	m; m
	4	24.6	t	2.10; 1.88	m; m
	5	46.4	t	3.67; 3.46	m; m
leucine-2	1	170.2	s	NH, 7.03	br d
	2	58.9	d	4.45	m
	3	39.4	t	1.55, 1.30	m, m
	4	23.7	d	1.60	m
	5	22.7	q	0.85	m
	6	20.9	q	0.82	m
alanine-1	1	171.1	s	NH, 7.70	d, J = 7.7
	2	47.7	d	4.19	p, J = 7.3
	3	16.3	q	1.19	d, J = 7.0
alanine-2	1	172.0	s	NH 7.92	d, J = 5.6
	2	50.1	d	4.06	p, J = 6.3
	3	17.0	q	1.24	d, J = 7.4
9-Me-3-Decol	1	169.7	s		
	2	39.5	t	2.64; 2.46	dd, J = 14.7, 5.6; dd, J = 14.7, 7.0
	3	70.8	d	5.11	p, 6.3
	4	32.8	t	1.42	m
	5	24.0	t	1.15; 1.05	m; m
	6	28.7	t	1.14	m
	7	26.2	t	1.16	m
	8	38.1	t	1.10	m
	9	27.0	d	1.50	m
	10, 11	22.1	q	0.84	d, J = 6.7

[a] At 125 MHz. [b] At 500 MHz.

The Thr1-Ileu2-Orn-Pro connections could be made with the HMBC experiment and the correlations from the NH signals (7.78, 8.00, 8.30 ppm, respectively) to the vicinal carbonyl (172.4, 171.8, and 172.8 ppm, respectively). As in the other proline-containing peptides the Pro-Val-3 linkage is established using the ROESY correlation between the Pro δ-proton signals (3.70, 3.88 ppm) and the Val-3 α-proton resonance (4.42 ppm). The final sequence of Val3-Val4-Thr2-Val5-5-MeHex was established by the NH resonances (7.77, 8.05, 7.82, 7.92 ppm, respectively) showing HMBC correlations to the vicinal carbonyls (172.3, 172.8, 172.7, and 173.8 ppm, respectively). As in kahalalides B and F, there exist two conformations of 5-MeHex. These conformational differences can be seen in the ^{1}H and ^{13}C NMR spectra. Some of the NH protons have shoulders indicating the existence of two conformers. These differences can again be seen most clearly in the carbon spectrum, where a subset of carbons resonate at approximately 20% of the intensity of the rest of the carbons. These signals can be assembled into 5-methylhexanoic acid by HMQC, HMBC, ROESY, and COSY data. Amino acid analysis by chiral GC revealed 12 amino acids: D-alloisoleucine (2), L-

Table 8. ^{1}H NMR and ^{13}C NMR Data for Kahalalide G (7) in DMSO-d_6

amino acid	carbon[a]	^{13}C NMR, ppm	mult	^{1}H NMR,[b] ppm	multiplicity
valine-1	1	173.4	s	(NH) 7.30	d, J = 8.0
	2	61.3	d	4.23	m
	3	31.1	d	2.18	m
	4	19.3	q	0.95	m
	5	18.0	q	0.91	m
Dhb	1	164.8	s	(NH) 9.36	s
	2	131.1	s		
	3	130.5	d	6.59	q, J = 7.0
	4	13.2	q	1.60	d, J = 7.0
phenylalanine	1	172.1	s	(NH) 7.68	d, J = 9.0
	2	56.4	d	4.70	m
	3	36.6	t	2.06	m
	4	138.7	s		
	5, 5′	130.3	d	7.44	d, J = 7.0
	6, 6′	128.8	d	7.29	m
	7	127.0	d	7.21	t, J = 7.5
valine-2	1	172.3	s	(NH) 8.14	d, J = 7.5
	2	58.6	d	4.37	m
	3	32.9	d	2.30	m
	4	19.6	q	0.95	m
	5	18.0	q	0.75	d, J = 6.5
isoleucine-1	1	172.5	s	(NH) 7.71	d, J = 8.0
	2	57.7	d	4.23	m
	3	37.8	d	1.94	m
	4	14.9	q	0.90	m
	5	26.6	t	1.20, 1.42	m, m
	6	11.9	q	0.83	m
threonine-1	1	171.0	s	(NH) 7.78	d, J = 8.5
	2	60.0	d	4.49	m
	3	67.3	d	4.23	m
	4	20.8	q	1.20	d, J = 6.4
isoleucine-2	1	172.4	s	(NH) 8.00	s
	2	57.3	d	4.40	m
	3	38.0	d	3.10	m
	4	14.9	q	0.90	m
	5	26.5	t	1.20, 1.42	m,m
	6	11.8	q		
ornithine	1	171.8	s	(NH) 8.30	d, J = 8.5
	2	52.9	d	4.47	m
	3	29.6	t	1.28	m
	4	24.6	t	1.85	m
	5	40.0	t	3.90	m
proline	1	172.8	s		
	2	57.4	d	4.44	m
	3	29.5	t	1.80	m
	4	25.5	t	1.88, 2.03	m
	5	48.2	t	3.70, 3.88	q,q J = 5.5,5.5
valine-3	1	171.2	s	(NH) 7.77	m
	2	57.8	d	4.42	m
	3	30.7	d	2.17	m
	4	19.7	q	0.95	m
	5	18.5	q	0.77	d, J = 7.0
valine-4	1	172.3	s	(NH) 8.05	d, J = 8.0
	2	59.1	d	4.24	m
	3	31.5	d	2.21	m
	4	19.6	q	0.95	m
	5	18.5	q	0.93	d, J = 5.2
threonine-2	1	172.8	s	(NH) 7.82	m
	2	58.9	d	4.43	m
	3	68.2	d	4.08	p, J = 6.8
	4	19.8	q	1.12	d, J = 6.4
valine-5	1	172.7	s	(NH) 7.92	d, J = 8.5
	2	59.6	d	4.33	m
	3	30.8	d	2.17	m
	4	19.7	q	0.95	m
	5	18.9	q	0.85	m
5-MeHex	1	173.8	s		
	2	36.3	t	2.28	m
	3	24.1	t	1.60	m
	4	39.0	t	1.20	m
	5	28.2	d	1.52	m
	6	22.7	q	0.85	d, J = 6.5
	7	22.6	q	0.85	d, J = 6.5
5-MeHex (second conformn)	1	174.0	s		
	2	33.9	t	2.29	m
	3	31.4	t	2.15	m
	4	29.9	t	1.28	m
	5	34.6	d	1.58	m
	6	19.0	q	0.76	m
	7	11.4	q	0.88	m

[a] At 125 MHz. [b] At 500 MHz.

ornithine, D-proline, L-threonine, D-allothreonine, D-valine (3), L-valine (2), and L-phenylalanine. This is identical to the amino acid composition for kahalalide F, and the absolute stereochemistry of the individual valines, and of threonine, is still under investigation.

Remarkably, only kahalalide F (**6**) exhibits noteworthy biological activity (Table 2) and has been the subject of a patent application.[8] The biological activity is lost when the ester bond is hydrolyzed and kahalalide F is transformed to G (**7**).

Experimental Section

General Experimental Procedures. Optical rotations were measured on a Jasco DIP-370 digital polarimeter. Infrared spectra were recorded on a Nicolet MX-5 FTIR spectrometer. Gas chromatography was performed using a Hewlett-Packard Model 5890 instrument. Mass spectra were measured on a VG-70SE magnetic sector mass spectrometer. NMR spectra were measured on a General Electric QE-300 or a GN OMEGA 500 instrument. ^{1}H NMR chemical shifts are reported in ppm with the chemical shift of the residual protons of the solvent used as internal standard. ^{13}C NMR chemical shifts are reported in ppm by using the natural abundance ^{13}C of the solvent as an internal standard. Ultraviolet spectra were recorded on a Hewlett-Packard Model 8452A diode array spectrophotometer.

Collection and Extraction. Two hundred sacoglossans (*E. rufescens*) were collected at Black Point, O'ahu, during April and May 1991 and extracted three times with EtOH. The animal was identified by Drs. Marilyn Dunlap and Allison Kay. A voucher specimen is deposited at the Bernice P. Bishop Museum, BPBM 247679. May−June appears to be the time of year when *E. rufescens* is in greatest abundance at Black Point. The combined extracts were then chromatographed using silica gel flash chromatography (hexane, hexane/EtOAc (1:1), EtOAc, EtOAc/MeOH (1:1), MeOH, and MeOH/HOAc (98:2)). The depsipeptides were found in the EtOAc/MeOH (1:1) fraction. HPLC using an RP 18 column and a gradient from H_2O/MeCN/TFA (70/30/1)−H_2O/MeCN/TFA (45/55/1) yielded six depsipeptides. Final purification was done isocratically and is described for each individual compound.

Kahalalide A (1). Final purification was accomplished by HPLC on RP 18 MeCN/H_2O/TFA (55/45/1). Physical data: $[\alpha]_D$ −19° (*c* 0.98, MeOH); ^{1}H NMR (Table 3, 500 MHz, CD_3CN); ^{13}C NMR (Table 3, 125 MHz, CD_3CN); IR neat (NaCl) 3305 (s, br), 2961 (s, br), 1742 (s), 1650 (s), 1550 (s), 1248 (s), 699 (m) cm^{-1}; mass spectrum HRFAB *m/z* (fragment) 894.4980 (M^+ + 1, 35) (calcd for $C_{46}H_{68}N_7O_{11}$ 894.4977); UV (MeOH) λ_{max} 204 (33 640) nm.

Kahalalide B (2). Final purification was carried out by HPLC on RP 18 MeCN/H_2O/TFA (55/45/1). Physical data: $[\alpha]_D$ +43° (*c* 2.53, MeOH); ^{1}H NMR (Table 4, 500 MHz, CD_3CN); ^{13}C NMR (Table 4, 125 MHz, CD_3CN); IR neat (NaCl) 3304 (s, br), 2958 (s, br), 1744 (s), 1647 (s), 1539 (s), 1451 (s), 1244 (s) cm^{-1}; mass spectrum HRFAB *m/z* (fragment) 878.4661 (M^+ + 1, 50) (calcd for $C_{45}H_{64}N_7O_{11}$ 878.4664); UV (MeOH) λ_{max} 204 (51 500), 226sh (21 320), 274 (17 580) nm.

Kahalalide C (3). Final purification was completed by HPLC on RP 18 MeCN/H_2O/TFA (55/45/1). Physical data: $[\alpha]_D$ +46° (*c* 0.41, MeOH); ^{1}H NMR (Table 5, 500 MHz, DMSO-d_6); ^{13}C NMR (Table 5, 125 MHz, DMSO-d_6); IR neat (NaCl) 3266 (s, br), 2965 (s, br), 1675 (s), 1559 (s), 1453 (s), 1406 (s), 1247 (s) cm^{-1}; mass spectrum HRFAB *m/z* (fragment) 914.4799 (M^+ + 1, 100) (calcd for $C_{47}H_{64}N_9O_{10}$ 914.4776); UV (MeOH) λ_{max} 202 (16 010), 224sh (1825) nm.

Kahalalide D (4). Final purification was performed by HPLC on RP 18 MeCN/H_2O/TFA (30/70/1). Physical data: $[\alpha]_D$ −51° (*c* 0.96, MeOH); ^{1}H NMR (Table 6, 500 MHz, DMSO-d_6); ^{13}C NMR (Table 6, 125 MHz, DMSO-d_6); IR neat (NaCl) 3304 (s, br), 2955 (s, br), 1741 (s), 1654 (s), 1532 (s), 1447 (s), 1182 (s), 1027 (s), 745 (s) cm^{-1}; mass spectrum HRFAB *m/z* (fragment) 596.3563 (M^+ + 1, 100) (calcd for $C_{31}H_{46}N_7O_5$ 596.3560); UV (MeOH) λ_{max} 206sh (16 895), 220 (18 710), 282 (3180) nm.

Kahalalide E (5). Final purification was accomplished by HPLC on RP 18 MeCN/H_2O/TFA (55/45/1). Physical data: $[\alpha]_D$ +5° (*c* 1.03, MeOH); ^{1}H NMR (Table 7, 500 MHz, DMSO-d_6); ^{13}C NMR (Table 7, 125 MHz, DMSO-d_6); IR neat (NaCl) 3301 (s, br), 2956 (s, br), 1744 (s), 1647 (s), 1539 (s), 1455 (s), 1244 (m), 1203 (m), 742 (m) cm^{-1}; mass spectrum HRFAB *m/z* (fragment) 836.5265 (M^+ + 1, 100) (calcd for $C_{45}H_{70}N_7O_8$ 836.5285); UV (MeOH) λ_{max} 204 (23 910), 222 (24 720), 272 (15 430) nm.

Kahalalide G (7). Final purification was carried out by HPLC on amino bonded silica MeOH/EtOAc (30/70). Physical data: $[\alpha]_D$ +22.5 (*c* 1.00, MeOH); ^{1}H NMR (Table 8, 500 MHz, DMSO-d_6); ^{13}C NMR (Table 8, 125 MHz, DMSO-d_6); IR neat (NaCl) 3276 (s, br), 3068 (w), 2964 (m), 1629 (s), 1543 (s), 1203 (w), 1138 (w); mass spectrum HRFAB *m/z* (fragment) 1495.9511 (M^+ + 1, 100) (calcd for $C_{75}H_{127}N_{14}O_{17}$ 1495.9504); UV (MeOH) λ_{max} 203 (58 300) nm.

Determination of Absolute Configuration. Each peptide was hydrolyzed by heating in a sealed vial at 100 °C for 22 h in 6 N HCl. The dried hydrolysate could then be used without derivatization for chiral HPLC and TLC analysis. Prior to GC−MS analysis the free amino acids were derivatized as the pentafluoropropyl amide and isopropyl ester (Alltech GC Derivatization Kit). Amino acid analysis was accomplished by GC−MS using a chiral GC column (Alltech, Chirasil-Val). All DL forms of amino acids could be resolved by GC with the exception of DL-arginine and DL-tryptophan. The enantiomers of DL-arginine and DL-tryptophan could be resolved using a chiral HPLC column based on ligand exchange chromatography (Macherey-Nagel·Duren, ET 250/8/4 Nucleosil Chiral-1), with a mobile phase of 0.5 mM $CuSO_4$, 60 °C, pH 5.6 and a flow rate of 1.0 mL/min. DL-Tryptophan was removed from the hydrolyzed peptide mixture using RP 18 H_2O/MeCN (85/15). DL-Tryptophan could also be separated using thin layer chromatography based on ligand exchange (Macherey-Nagel·Duren, Chiralplate). The R_f value for L-Trp is 0.6, and the R_f value for D-Trp is 0.5 in MeCN/MeOH/H_2O (4/1/1) solvent.

Acknowledgment. We thank Rhone K. Akee for his initial work on *E. rufescens*; Faith Caplan, Linda Kay Larsen, Alice Clark, Marsha Wright, and the biologists at Pharma Mar for bioassays; Walter Niemczura, Wesley Yoshida, Mike Burger, John Carney, Frank Wiggers, and Prof. Rinehart's group at the University of Illinois for spectral analyses; Marilyn Dunlap and Alison Kay for taxonomic identification; Toshio Ichiba, Jaroslaw Jurek, Bryan Yeung, Ellen Juagdan, Erwin Quebral, and Jay Corgiat for their help with collecting samples; and Kum Ming Woo for *Bryopsis* extraction. Financial support was provided by the National Science Foundation, the Sea Grant College Program, The American Association of Colleges of Pharmacy, Pharma Mar, S. A., and the American Chemical Society Project SEED for summer support of Kum Ming Woo.

Supporting Information Available: 125 MHz ^{13}C NMR, 500 MHz ^{1}H NMR, HMQC, HMBC, COSY, and ROESY spectra for **1** and **2** in MeCN-d_3 and **3**, **4**, **5**, and **7** in DMSO-d_6 and amino acid analysis by GC−MS on a Chirasil-Val column (50 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO960877+