Cycloviolins A−D, Anti-HIV Macrocyclic Peptides from *Leonia cymosa*[1]

Yali F. Hallock,[†,⊥] Raymond C. Sowder II,[‡] Lewis K. Pannell,[§] Courtney B. Hughes,[‡] Donald G. Johnson,[‡] Robert Gulakowski,[†,∥] John H. Cardellina II,[†,▽] and Michael R. Boyd*[,†]

Laboratory of Drug Discovery Research and Development, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Frederick, Maryland 21702, Science Applications International Corporation, Frederick Cancer Research and Development Center, Frederick, Maryland 21702, and Laboratory of Analytical Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892

Received June 11, 1999

Four novel anti-HIV macrocyclic peptides containing 28−31 amino acid residues, named cycloviolins A−D, have been isolated from the hitherto unstudied tropical plant *Leonia cymosa*. Their primary structure, including amino acid composition and sequence, was determined by a combination of MALDI-TOF and FAB MS and by enzymatic digestion of reduced derivatives, followed by Edman degradation and mass analyses. All of the cycloviolins contain six cysteines, which are present as three intramolecular disulfide bridges. Intriguingly, cycloviolins A−D showed high degrees of sequence homology to the known cyclopsychotride A and circulins A and B from the Rubiaceae family but much less homology to the varv peptides from *Viola*, a member of the same family (Violaceae).

Leonia belongs to the Violaceae plant family, comprising some 23 genera and more than 900 species.[3] Except for *Viola*, all other genera in this family are tropical. There have been no previous reports on chemical investigations of any *Leonia* spp., and there is only limited work on the entire family Violaceae. In their study of *Viola japonica* and *V. lactiflora*, Moon et al isolated flavonol glycosides.[4] In separate reports, the chemical constituents of *Hybanthus enneaspermus*[5] and a partially characterized sulfated polysaccharide from the Chinese medicinal herb *V. yedoensis*[6] were described. Recently, two different groups have reported macrocyclic peptides, violapeptide A[7] and varv peptides A−H,[8,9] from *V. arvensis* Murray.

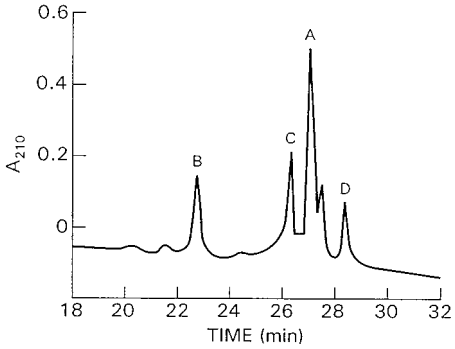

Figure 1. Reversed-phase HPLC chromatogram for the purification of cycloviolins A−D. Letters on the peaks correspond to names of the peptides. The trailing shoulder on the cycloviolin A peak analyzed by FAB MS had the same MW as A but was not further characterized.

We screened for anti-HIV activity[2] several extracts of *Leonia* sp. from different South American collections, including extracts from *L. cymosa* C. Martius and *L. glycycarpa*. Interestingly, only the *L. cymosa* extracts were active. The anti-HIV activity and the unknown chemistry of the genus *Leonia* spp. prompted us to select these extracts for bioassay-guided fractionation.

Results and Discussion

The crude aqueous extract was subjected to a solvent−solvent partitioning protocol, which concentrated the anti-HIV activity in the butanol-soluble fraction. This fraction was separated by gel permeation on Sephadex LH-20, followed by reverse-phase HPLC, to give four compounds, named cycloviolins A−D (Figure 1). All of these compounds gave positive ninhydrin tests.

Proton NMR spectra of these four compounds suggested that they were closely related and showed remarkable similarity to those of circulins A and B, cyclic peptides previously from *Chassalia parvifolia* K.Schum.-

* Correspondence address: Bldg. 1052, Rm. 121, NCI-FCRDC, Frederick, MD 21702-1201. Tel: 301-846-5391. FAX: 301-846-6919.
† Laboratory of Drug Discovery Research and Development.
‡ Science Applications International Corporation.
§ Laboratory of Analytical Chemistry.
⊥ Current address: Grants and Contracts Operations Branch, Developmental Therapeutics Program, NCI, 6130 Executive Boulevard, Suite 841, Bethesda, MD 20892-7456.
∥ Current address: National Institute of Allergy and Infectious Diseases, Solar Building, Room 3B08, Bethesda, MD 20892-7630.
▽ Current address: Council for Responsible Nutrition, 1875 Eye Street, NW Suite 400, Washington, DC20006-5409

(1) HIV-Inhibitory Natural Products. 59. For part 58, see ref 17.
(2) Boyd, M. R. In *AIDS Etiology, Diagnosis, Treatment and Prevention*; Devita, V. T., Jr., Hellman, S., Rosenberg, S. A., Eds.; Lippincott: Philadelphia, PA, 1988; pp 305−319.
(3) Mabberley, D. J. *The Plant Book*; Cambridge University Press: Cambridge, 1987; p 608.
(4) (a) Moon, C.; Yook, C. *Saengyak Hakhoechi* **1981**, *12*, 146; *Chem. Abstr.* **1981**, *96*, 3569. (b) Moon, C.; Yook, C. *Saengyak Hakhoechi* **1981**, *12*, 147−148; *Chem. Abstr.* **1981**, *96*, 214335.
(5) Majumder, P. L.; Basu, A.; Mal, D. *Indian J. Chem., Sect. B* **1979**, *17*, 297−298.
(6) Ngan, F.; Chang, R. S.; Tabba, H. D.; Smith, K. M. *Antiviral Res.* **1988**, *10*, 107−115.
(7) Schopke, T.; Hasan Agha, M. I.; Kraft, R.; Otto, A.; Hiller, K. *Sci. Pharm.* **1993**, *61*, 145−153.
(8) Claeson, P.; Goransson, U.; Johansson, S.; Luijendijk, T.; Bohlin, L. *J. Nat. Prod.* **1998**, *61*, 77−81.
(9) Goransson, U.; Luijendijk, T.; Johansson, S.; Bohlin, L.; Claeson, P. *J. Nat. Prod.* **1999**, *62*, 283−286.

10.1021/jo990952r CCC: $19.00
Published on Web 12/16/1999

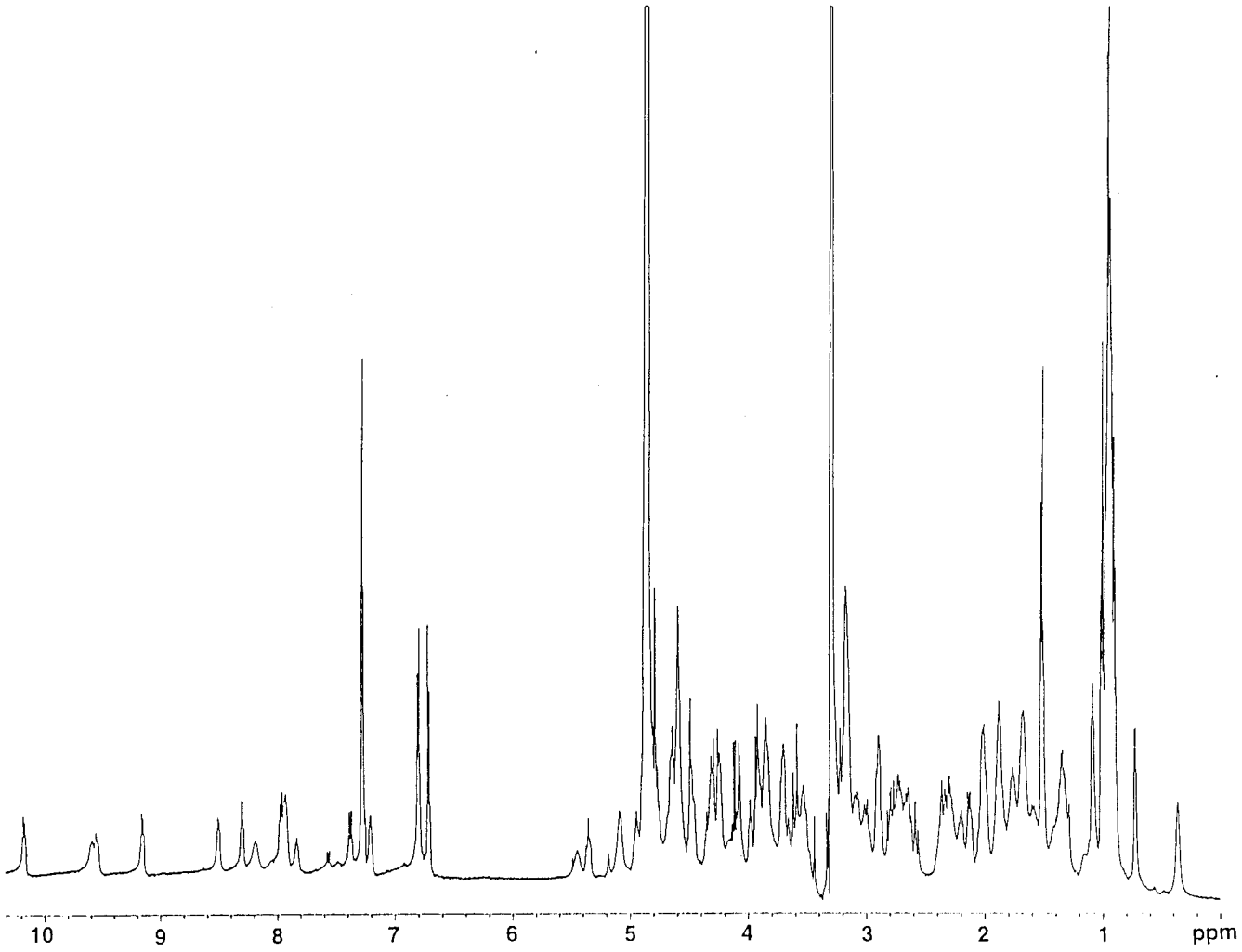

Figure 2. The 500 MHz ^{1}H NMR spectrum of cycloviolin A in CD_3OD.

(Rubiaceae).[10] The proton NMR spectrum of cycloviolin A taken in CD_3OD (Figure 2) showed a complex pattern of proton signals from 0.3 to 5 ppm. Most interesting were a broad singlet and a doublet at 0.38 and 0.75 ppm, respectively.

FAB mass spectral analysis of cycloviolin A gave a pseudomolecular ion (MH^+) at *m*/*z* 3213.4. Acid hydrolysis, followed by amino acid analysis, indicated that cycloviolin A contained 31 amino acids (Table 1), with over 65% homology to circulins A and B, as well as to cyclopsychotride A (Figure 3), a peptide from *Psychotria longipes* Muell. Arg. (Rubiaceae).[11] The calculated molecular weight of a linear peptide, based on the amino acid residues, was approximately 24 mass units higher than the observed molecular weight. This suggested that cycloviolin A, like circulins and cyclopsychotride A, was a cyclic peptide ($-H_2O$, 18 Da) with three intramolecular disulfide bridges (−6 Da).

Cycloviolins A−D were reduced by mercaptoethanol and then derivatized with 4-vinylpyridine to give the corresponding *S*-(4-pyridylethyl)cysteine (PEC) derivatives. After purification, all PEC derivatives were analyzed by FAB MS. The observed mass was consistent with six cysteine residues in each cycloviolin, in good agreement with the expected mass. In the case of cycloviolin A, the PEC derivative gave a FAB MS molecular ion at 3850.7, which coincided with the calculated theoretical mass for the proposed cyclic peptide containing three internal disulfide bonds [3212.8 + 636.6 (PEC) = 3849.4]. Because amino acid analysis of the parent cycloviolins indicated the presence of only one glutamic acid residue in each peptide, the corresponding PEC derivatives were subjected to enzymatic digestion with Glu-C to provide linear peptide PEC derivatives, which were amenable to Edman sequencing. These linear PEC derivatives were also analyzed by FAB MS. For cycloviolin A, a MH^+ ion at *m*/*z* 3869.3 was observed (calculated: 3212.8 + 18 for H_2O + 636.6 for PEC = 3867.4), further confirming a cyclic primary structure for the peptide. Similar comparisons of the observed with the calculated molecular weights of the other linear cycloviolin PEC derivatives were consistent with analogous cyclic structures containing three disulfide bridges.

Cycloviolins A−D tested against HIV-1 in an XTT-tetrazolium based metabolic screen[2,12] showed comparable antiviral activities (EC_{50} values ≈ 130 nM). The

(10) Gustafson, K. R.; Sowder, R. C., II; Henderson, L. E.; Parsons, I. C.; Kashman, Y.; Cardellina, J. H., II; McMahon, J. B.; Buckheit, R. W., Jr.; Pannell, L. K.; Boyd, M. R. *J. Am. Chem. Soc.* **1994**, *116*, 9337−9338.

(11) Witherup, K. M.; Bogusky, M. J.; Anderson, P. S.; Ramjit, H.; Ransom, R. W.; Wood, T.; Sardana, M. *J. Nat. Prod.* **1994**, *57*, 1619−1625.

(12) Gulakowski, R. J.; McMahon, J. B.; Stanley, P. G.; Moran, R. A.; Boyd, M. R. *J. Virol. Methods* **1991**, *22*, 87−100.

Table 1. Amino Acid Compositions and Molecular Properties of Cycloviolins A−D

amino acid		A		B		C		D	
		seq	AAA	seq	AAA	seq	AAA	seq	AAA
Asp + Asn	(D + N)	2	2.0	1	1.0	2	2.0	2	2.0
Thr	(T)	0	0.0	4	3.9	2	2.0	0	0.0
Ser	(S)	3	3.0	3	3.0	2	2.0	3	3.1
Glu + Gln	(E + Q)	1	1.0	2	2.1	1	1.0	1	1.0
Pro	(P)	2	2.0	1	1.0	2	2.0	2	2.0
Gly	(G)	3	3.0	3	3.1	3	3.0	3	3.0
Ala	(A)	2	2.0	1	1.0	1	1.0	2	1.9
$Cys^{(PEC)}$	($C^{(PEC)}$)	6[a]	5.5[a]	6[a]	5.5[a]	6[a]	5.7[a]	6[a]	5.6[a]
Val	(V)	3	2.7	2	1.9	3	2.9	2	2.0
Met	(M)	0	0.0	0	0.0	0	0.0	0	0.0
Ile	(I)	4	3.6	0	0.0	2	2.0	3	3.0
Leu	(L)	0	0.0	1	1.1	1	1.0	0	0.0
Tyr	(Y)	1	1.0	1	1.0	1	1.0	1	1.0
Phe	(F)	1	1.0	2	2.0	1	1.0	2	1.9
His	(H)	0	0.0	0	0.0	0	0.0	0	0.0
Lys	(K)	2	2.0	1	1.0	2	1.9	2	1.9
Arg	(R)	1	1.0	0	0.0	1	1.1	1	1.1
Trp	(W)	0	ND[b]	0	ND[b]	0	ND[b]	0	ND[b]
ϵ_{280}[c]		1640		1640		1640		1640	
I. E. point[d]		7.93		5.90		7.93		7.93	
residues		31		28		30		30	
charge[e]		+2		0		+2		+2	
$MS^{(linear-PEC)}$ [f]		3849.7		3524.5		3781.2		3785.4	
linear-PEC MW[g]		3850.1		3524.1		3782.4		3786.0	
$MS^{(linear-red.)}$ [h]		3236.4		2910.1		3168.4		3172.0	
linear−red. MW[i]		3235.8		2910.3		3166.7		3170.7	
$MS^{(circular)}$ [j]		3213.4		2887.1		3145.4		3149.0	
circular MW[k]		3211.8		2886.3		3142.7		3146.7	

[a] Cysteine was determined as the *S*-*β*-(4-pyridylethyl)-L-cysteine (PEC) derivative, following reduction and alkylation with 4-vinylpyridine. [b] As a result of destruction during hydrolysis, tryptophan was not determined (ND) by amino acid analysis. [c,d,e,i] Values were calculated using GCG software from the Wisconsin Package (Version 10.0, Genetics Computer Group, Madison, WI), which assumes a reduced linear form of the peptide. [f,g] Values are for peptides with PEC-derivitized cysteines. [g,i,k] Values were calculated on the basis of the sequence using average masses. [j,k] Values for peptides in their native macrocyclic form were calculated as −18 Da due to loss of water upon circularization and −6 Da due to the formation of three disulfide bonds when compared with the reduced linearized form. [f,h,j] Masses were determined by FAB mass spectrometry.

direct cytotoxicity of the compounds to the host cells (CEM-SS) was also comparable to that of circulins A and B (IC_{50} values ≈ 560 nM). Thus, the variations in the amino acid sequences did not appear to have a measurable impact on the overall activity profile.

Chemically and structurally, cycloviolins A−D are new plant-derived members of an emerging class of cyclic peptides with 28−31 amino acid residues, six of which are cysteines that are present as internal disulfide bridges. Of the related cyclopeptides, violapeptide I[7] and the most recently reported varv peptides A−H[9] are from the family Violaceae, and kalata B1,[13,14] cyclopsychotride A, and circulins A and B are all from the family Rubiaceae (Figure 3). Interestingly, the amino acid sequences of cycloviolins A−D showed a higher degree of homology with Rubiaceae cyclic peptides, such as circulins A and B and cyclopsychotride A, than with the varv peptides or violapeptide I from *Violaceae*. Intriguingly, the smallest member of this class of peptides, cycloviolin B, seems not to fit in either group shown in Figure 3. This peptide lacks a proline residue and has a unique TSSQ sequence expanding from residue 17 to 20. Further, the new cycloviolins A−D are composed of 28−31 amino acid residues, whereas the varv peptides and violapeptide I have a slightly more restricted length of 29−30 amino acids. In addition, the varv peptides are more homologous in sequence to kalata B1 of the Rubiaceae family. These high interfamily and low intrafamily homologies are of potential chemotaxonomic interest. Since plants in the genus *Viola* are temperate, while the *Leonia* and Rubiaceae peptides are all from tropical or neotropical regions, the production of these cyclopeptides may be determined more by environmental factors than by taxonomy. Given the limited number of plant cyclopeptides known to date, this kind of comparison may become more meaningful as additional related metabolites are discovered that may shed more light on the function of these peptides in the plants. Also, a more thorough understanding and comparison of the genetic machinery regulating the expression of these peptides would provide more definitive answers to their intriguing interrelationships.

The disulfide bonds may contribute to chemical stability[14] and potentially render the peptides resistant to proteolysis.[15] A 3D structure analysis of kalata B1 and circulin A indicated a similar stabilized global fold from a rigid structure arising from an embedded loop, formed by two of the three disulfide bonds and from a third disulfide bond that is referred to as a "cysteine-knot", threading through the center of the loop.[14,1616−17] For kalata B1 and circulins A and B, the disulfides linked Cys(1) with Cys(4), Cys(2) with Cys(5), and Cys(3) with

(13) Gran, L. *Lloydia* **1973**, *36*, 174−178.
(14) Saether, O.; Craik, D. J.; Campbell, I. D.; Sletten, K.; Juul, J.; Norman, D. G. *Biochemistry* **1995**, *34*, 4147−4158.
(15) Derua, R.; Gustafson, K. R.; Pannell, L. K. *Biochem. Biophys. Res. Comm.* **1996**, *228*, 632−638.
(16) Daly, N. L.; Koltay, A.; Gustafson, K. R.; Boyd, M. R.; Casas-Finet, J. R.; Craik, D. J. *J. Mol. Biol.* **1999**, *285*, 333−345.
(17) Ford, P. W.; Gustafson, K. R.; McKee, T. C.; Shigematsu, N.; Maurizi, L. R.; Pannell, L. K.; Williams, D. E.; de Silva, E. D.; Lassota, P.; Allen, T. M.; Van Soest, R.; Anderson, R. J.; Boyd, M. R. *J. Am. Chem. Soc.*, in press.

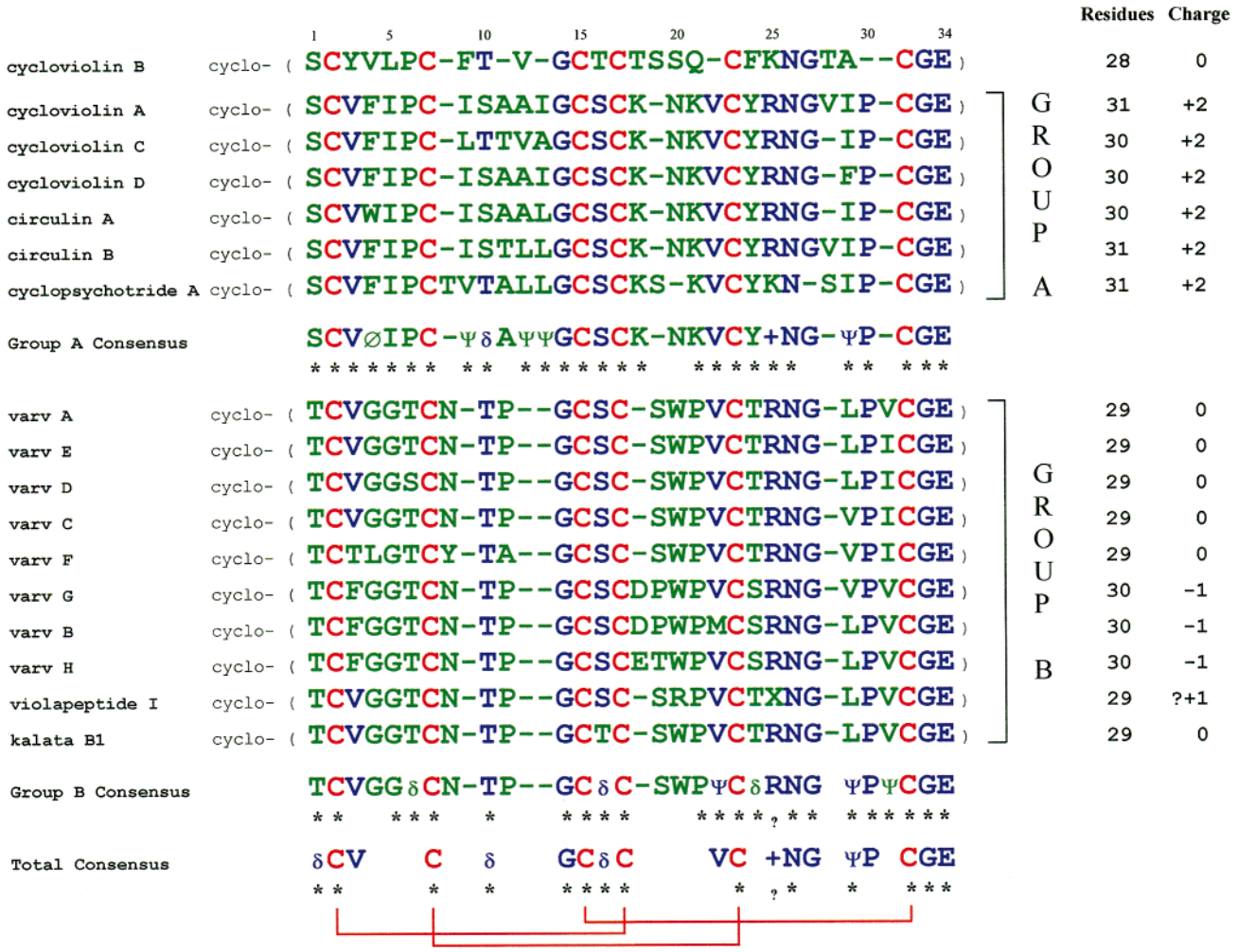

Figure 3. Sequences of macrocyclic peptides aligned at the endoproteinase Glu-C cleavage site. Vertical alignment indicates a relative order of homology defining two distinct subgroups: Group A includes cycloviolins (with the exception of cycloviolin B, having characteristics of both groups), circulins, and cyclopsychotride A, and Group B includes the varv isolates, violapeptide I, and kalata B1. Gaps inserted for maximum alignment are indicated by −, conserved residues or functionalities by *, conserved hydroxy residues by δ, conserved hydrophobic residues by Ψ, conserved aromatic residues by Ø, conserved basic residues by +, and disulfide bonds by |____|. Blue letters indicate residues that are highly conserved within the context of the entire alignment, green indicates poorly conserved positions, and red highlights cysteine residues involved in disulfide linkages. Group A has 67% and Group B has 53% (57% if the unidentified residue (X) in violapeptide I is an arginine) identities. However, if functionally similar residues are considered, the number of conserved positions for Group A is 90% and for Group B 73−77% (based on an averaged length of 30 residues). For macrocyclic peptides overall, the conservation is 30−37% identities and 47−50% invariantly conserved features.

Cys(6), where the numbers designate their respective order in the sequence lineup (Figure 3). Starting with the cysteine at position 2, the spacing of residues for the known plant peptides is $C-X_4-C-X_{(4-6)}-C-X_1-C-X_{4-5}-C-X_{(5-7)}-C-X_3$, where X is any residue other than cysteine. With alignment gaps introduced for maximum homology, all cysteine residues in both groups A and B (Figure 3) could be considered totally conserved. The 3D folds of both kalata B1 and circulins allow the presentation of an exposed hydrophobic surface, flanked by an aromatic residue on one side and the charged residues on the opposite side. NMR analysis of the 3D structure also revealed several regions of β-turns and β-sheet in kalata B1 as well as in the circulins. It is conceivable that the diverse biological activities displayed by these structurally related peptides could, in part, be due to a common topology of the 3D structure, which presents a hydrophobic surface for docking to receptor and the charged residues for initial recognition and binding.

Experimental Section

General. MALDI-TOF mass spectra of the cycloviolins and their PEC derivatives were obtained on samples dissolved in 0.1% TFA and 50% acetonitrile and applied to the target in a sinapinic acid matrix. For FAB MS the peptide samples were desorbed from a glycerol matrix using 6 keV xenon atoms, and mass measurements were performed at 10 000 resolution using electric field scans with the sample peak bracketed by two reference ions. For quantitative amino analysis, the peptides were hydrolyzed for 18 h at 110 °C with 6 N HCl and subjected to amino acid analysis using postcolumn ninhydrin detection.

Collection and Extraction. Bark material of the plant *Leonia cymosa* C. Martius was collected in the Loreto Department of the Maynas Province of Peru (73.75° west, 4.917° south) on February 15, 1988 by Douglas Daly of the New York Botanical Garden under NCI contract. A voucher specimen (Q65T-5650) was deposited at the New York Botanical Garden and the Department of Botany, Museum of Natural History, Smithsonian Institution. The dry bark material was ground by a hammer mill to a coarse powder (164 g). The entire mass

was first soaked in a percolator with 1:1 CH_2Cl_2/MeOH overnight and then washed with MeOH to give an organic extract (1.53 g). Subsequent extraction with 100% water overnight, followed by filtration and lyophilization, yielded the aqueous extract (7.68 g).

Isolation. Part of the aqueous extract (4.1 g) was partitioned between water and 1-butanol (1:1, 5 × 250 mL). The active 1-butanol fraction (639 mg) was then triturated with 50 mL of MeOH to give a soluble fraction, which was subjected to gel permeation on Sephadex LH-20 (8.5 cm × 46 cm) with MeOH as eluent. The peptide containing fraction B (208 mg) was further purified by C_{18} HPLC (5 μm, Rainin Microsorb, 1 cm × 25 cm) eluting with CH_3CN/H_2O-0.1% TFA (1:4). The peptide peaks eluted between 22 and 29 min and yielded pure cycloviolins A (2.9 mg), B (1.8 mg), C (1.1 mg), and D (1.3 mg), as determined by quantitative amino acid analysis.

Reduction and Conversion to Pyridylethyl (PEC) Derivatives. In a typical procedure, approximately 0.75 mg of a cycloviolin was dissolved in 0.85 μL of 8 M guanidine hydrochloride, 150 μL of deionized water, and 10 μL of Tris buffer (pH 8.5). After saturation with N_2 and addition of 20 μL of β-mercaptoethanol, the mixture was allowed to stand for 1.5 h at 37 °C in the dark. The reduced peptide was reacted with 30 μL of 4-vinylpyridine at room temperature in the dark for 2 h. Some of the excess volatile reagents were removed under a stream of N_2, and the resulting residue was purified by reversed-phase HPLC (Waters μBondapak C_{18}, 3.9 mm × 300 mm, 10 μm) under gradient conditions, starting with 100% solvent A (H_2O with 0.05% TFA, 2.0 mL/min). When the mobile phase reached 10% solvent B (CH_3CN) after 20 min, the flow rate was slowed to 1.0 mL/min and mobile phase was ramped to 40% solvent B in 60 min. Chromatography peaks containing peptide derivatives eluted from around 20% solvent B and ended at around 32% solvent B.

For sequence analysis, the PEC derivative of each of the cycloviolins was digested with endoproteinase Glu-C (1:25 w/w peptide/enzyme) in 50 mM ammonium acetate (pH 4, 2:1 w/v peptide/buffer) at 37 °C for 67 h. The digests were then dissolved with 1 mL of 8 M guanidine-HCl and acidified with 30 μL of 20% TFA before loading to a 10 μm Waters μBondapak C_{18} column (3.9 mm × 300 mm). Gradient elution from 10% to 35% solvent B (vide supra) over 130 min was used to purify the digested linear PEC peptide derivatives. For all digest mixtures, the peak corresponding to linear PEC derivative eluted between 20% and 24% solvent B. The amino acid sequences were then determined by automated Edman degradation.

Acknowledgment. We thank Dr. Gordon Cragg from Natural Products Branch and Mr. Douglas Daly from New York Botanical Gardens for collections, Mr. Thomas McCloud and his staff for extractions, Dr. David Clanton for screening assistance, and Mr. Bradley P. Kane for other technical assistance.

JO990952R