A Novel Anti-HIV Macrocyclic Peptide from *Palicourea condensata*

Heidi R. Bokesch,[†] Lewis K. Pannell,[‡] Pamela K. Cochran,[†] Raymond C. Sowder II,[§] Tawnya C. McKee,[†] and Michael R. Boyd*,[⊥]

Laboratory of Drug Discovery Research and Development, Division of Basic Sciences, NCI-Frederick, Frederick, Maryland 21702-1201, SAIC Frederick, Frederick, Maryland 21702-1201, and Laboratory of Bioorganic Chemistry, NIDDK, NIH, Bethesda, Maryland 20892

Received July 27, 2000

A 37 amino acid cyclic polypeptide has been isolated from the organic extract of the tropical tree *Palicourea condensata*. Palicourein (**1**) is the largest of a growing family of plant peptides that contain a cyclized amino acid backbone cross-linked via three internal disulfide bridges. Palicourein inhibits the in vitro cytopathic effects of HIV-1_{RF} infection of CEM-SS cells with an EC_{50} value of 0.1 μM and an IC_{50} value of 1.5 μM.

As part of our interest in natural products with HIV-inhibitory activity,[1] we undertook the investigation of the organic extract of *Palicourea condensata* Standl. (Rubiaceae), which showed activity in the NCI's XTT-tetrazolium based anti-HIV primary screen.[2] Previous studies of this genus have yielded triterpenes,[3] alkaloids,[4] and fluoroacetate.[5] Although there have been no reports of peptides from this genus, macrocyclic peptides have been described from three genera of plants in the Rubiaceae family[6−9] as well as several genera in the Violaceae family[10−14] and, more recently, from the Cucurbitaceae family.[15]

Plant macrocyclic peptides from the Rubiaceae family include circulins A−F from *Chassalia parvifolia*,[6,7] cyclopsychotride A from *Psychotria longipes*,[8] and kalata B1 from *Oldenlandia affinis*.[9] Reported peptides isolated from the Violaceae family are cycloviolacins O1−O11 and H1 from *Viola* species,[10] cycloviolins A−D from *Leonia cymosa*,[11] and violapeptide I[12] and varv A-H1[3−14] peptides from *Viola arvensis*. These plant polypeptides share a primary structure consisting of 28−31 amino acid residues covalently cyclized by the amide backbone and three internal disulfide bridges. Craik et al.[10] have proposed that this novel family of macrocylic peptides be referred to as cyclotides. A recent report describes the first isolation of a 34-residue macrocyclic peptide from Cucurbitaceae[15] that also possesses a cyclized amino acid backbone and disulfide bridges. Although it shares a three-dimensional structure similar to that of the peptides isolated from the Rubiaceae and Violeaceae families, there is little sequence homology apart from the six Cys residues. To date, palicourein, with 37 amino acids, is the largest naturally occurring cyclotide.

The water soluble portion of the organic extract of *P. condensata* was permeated through Sephadex LH-20; the resulting active fraction was subjected to vacuum-liquid chromatography on wide pore C4 gel, followed by reversed-phase HPLC to give **1** (1.04% yield).

Electrospray ionization mass spectrometry (ESIMS) of palicourein provided a molecular weight of 3904.1 Da. The mass calculated for the linear peptide with all cysteines in the reduced form was 3928.5, approximately 24 Da higher than the observed molecular weight. This suggested that palicourein was a cyclic peptide and accounted for the loss of one water molecule (18 Da) due to cyclization and six hydrogens (6 Da) due to the formation of the three disulfide bridges. Reduction and alkylation of **1** with 4-vinylpyridine generated the (*S*)-(*β*-4-pyridylethyl)cysteine (PEC) derivative, which gave an ESIMS molecular weight of 4541.0 Da. This is consistent with the presence of six disulfide-linked cysteines (+636.4 Da due to six pyridylethyl groups). Attempts to sequence the PEC derivative of **1** were unsuccessful, providing additional support that the N-terminus was blocked due to the cyclic nature of palicourein.

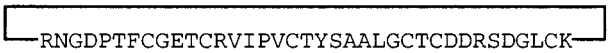

Figure 1. Cyclic amino acid sequence of palicourein established by N-terminal Edman degradation of the linear peptide generated by digestion with endoproteinase Lsy-C.

Amino acid analysis of the PEC derivative of palicourein indicated that it contained only one lysine; therefore it was subjected to enzymatic digestion with Lys-C to provide a derivatized, linear peptide that could be sequenced. Digestion of **1** with endoproteinase Lys-C gave a linear product from which the total sequence of amino acids could be deduced (Figure 1).

Palicourein shows a high degree of homology to the previously isolated cyclotides. A report by Craik et al.[10] has classified these 31 cyclotides into two subfamilies based on the backbone structure and charge of the peptides. Palicourein does not fit distinctly into either of these, containing a mixing of partial sequences between the two subfamilies that thus far has not been observed (Figure 2). For loop 1, the third residue is S in subfamily 1 and T in subfamily 2. For loop 6, the P is adjacent to C in subfamily 1 and separated by an intervening hydrophobic residue in subfamily 2. Using these criteria, palicourein appears to fit in subfamily 2. This however changes when comparing loops 2 and 3. For loop 2, in subfamily 1 the fourth residue is always P, and in subfamily 2, the two cental residues are GG. For loop 3, subfamily 1 contains up to seven residues, while subfamily 2 contains only four. Palicourein appears to fit in subfamily 1 with these constraints. In terms of net charge, peptides from subfamily 1 have a net charge of +2 and those in subfamily 2 are neutral or negative, so palicourein with a net charge of −1 would fit into subfamily 2.

The three-dimensional structure of kalata B1, circulin A, and cycloviolacin O1 has been determined and shows a

* Corresponding author. Tel: (301) 846-5393. Fax: (301) 846-6919. E-mail: boyd@dtpax2.ncifcrf.gov, http://dtp.nci.nih.gov/epnrd/lddrd.html.
[†] SAIC Frederick.
[‡] Laboratory of Bioorganic Chemistry.
[§] AIDS Vaccine Program, SAIC.
[⊥] Laboratory of Drug Discovery Research and Development.

10.1021/np000372l CCC: $20.00
Published on Web 01/20/2001

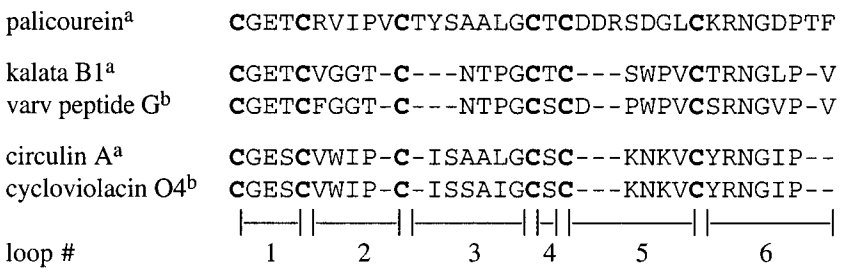

Figure 2. Sequence comparison of select cyclotides from the [a]Rubiaceae and [b]Violeaceae families. The six conserved cysteine residues are highlighted in bold. The four previously identified cyclotides are grouped into subfamilies with the backbone segments indicated by loops 1−6 as per Craik et al.[10]

consensus of a cyclic backbone with conserved cysteine residues and a cyclic knot motif. Although the disulfide bonding pattern of palicourein has not been determined, the Cys I−Cys IV, Cys II−Cys V, and Cys III−Cys VI cyclic cysteine knot motif is conserved in all of the examples isolated to date.

Palicourein inhibited the cytopathic effects of HIV-1_{RF} infection in a human T-lymphoblastoid cell line (CEM-SS),[2] with an EC_{50} value of 0.10 μM and an IC_{50} value of 1.5 μM. This compares well with the anti-HIV activity of the circulins A−F and the cycloviolins A−D with EC_{50}'s ranging from 0.05 to 0.28 μM.

Experimental Section

General Experimental Procedures. ESIMS were acquired on a Hewlett-Packard HP1 100 integrated LC-MS system. Amino acid analysis was performed with a Beckman System 6300 amino acid analyzer according to the protocols of the manufacturer. Amino acid sequences were determined by Edman degradation using an Applied Biosystems Model 494 sequencer. All HPLC mobile phases included 0.05% (v/v) TFA.

Collection and Extraction. Bark from the tree *Palicourea condensata* Standley (Rubiaceae) was collected under NCI contract in the province of Pastaza in Ecuador, March 1988, by B. Boom. Voucher specimens (Q65T1001) are maintained at the New York Botanical Garden and the Smithsonian Institution. The dried plant material (362 g) was ground and sequentially extracted with 1:1 MeOH−CH_2Cl_2 followed by 100% MeOH. The combined organic extracts were evaporated in vacuo to give 17.44 g of extract.

Isolation. A portion of the organic extract (5.30 g) was subjected to a solvent−solvent partitioning scheme[16] that concentrated the anti-HIV activity in the water soluble fraction (3.53 g). Size exclusion chromatography of the active material on Sephadex LH-20 (5.5 × 68 cm) with MeOH−H_2O (7:3) provided an early eluting fraction (0.19 g), which was further separated by vacuum-liquid chromatography (VLC) on wide pore C4 media with a 0−100% MeOH in H_2O gradient. The 100% MeOH fraction containing the active component was purified by reversed-phase HPLC (Rainin Dynamax, C_{18}, 300 Å, 1 × 25 cm) eluting with a 25−60% CH_3CN in H_2O gradient to give 55.9 mg (1.04% yield) of palicourein (**1**). **1**: white powder; $[\alpha]_D = -68.1°$ (*c* 0.79, MeOH); UV (MeOH) λ_{max} (log ϵ) 223 (4.72), 203 (5.21).

Structure Determination. (a) Reduction and Alkylation of the Disulfide Bonds. Palicourein (1.0 mg) was dissolved in 400 μL of 8 M guanidine HCl. To this were added 20 μL of 3 M Tris HCl (pH 8.5), 15 μL of β-mercaptoethanol, and 100 μL of deionized water, and the reaction was kept in the dark, at room temperature, under an atmosphere of N_2. After 23 h, 45 μL of 4-vinylpyridine was added, and the mixture was again kept in the dark, at room temperature, under N_2 for 2 h. Excess reagent was removed under a stream of N_2, and the mixture was stored at −20 °C overnight. The derivatized peptide was purified by reversed-phase HPLC (Rainin Dynamax, C_{18}, 300 Å, 1 × 25 cm) using a gradient of 100% H_2O for 40 min, then increasing to 60% CH_3CN in H_2O over 50 min.

(b) Enzymatic Cleavage. Lys-C. The PEC derivative (250 μg) was dissolved in 250 μL of digestion buffer (25 mM Tris HCl, 1 mM EDTA, pH 8.5) and treated with 5 μg of endoproteinase Lys-C. The reaction was maintained at room temperature for 24 h, and the cleaved peptide product was purified by reversed-phase HPLC (Rainin Dynamax, C_{18}, 300 Å, 1 × 25 cm) using a gradient of 100% H_2O for 20 min, then increasing to 60% CH_3CN in H_2O over 100 min.

Acknowledgment. We thank G. Cragg (Natural Products Branch) for coordinating collections, T. McCloud for extractions, and R. Gardella and D. Rosser for anti-HIV evaluations. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-56000. The content of this article does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade name, commercial products, or organization imply endorsement by the U.S. Government.

References and Notes

(1) Part 70 in the series HIV-Inhibitory Natural Products; for part 69, see ref 17.
(2) Gulakowski, R. J.; McMahon, J. B.; Staley, P. G.; Moran, R. A.; Boyd, M. R. *J. Virol. Methods* **1991**, *33*, 87−100.
(3) Bolzani, V. Da S.; Trevisan, L. M. V.; Young, M. C. M. *Rev. Latinoam. Quim.* **1992**, *23*, 20−21.
(4) Morita, H.; Ichihara, Y.; Takeya, K.; Watanabe, K.; Itokawa, H.; Motidome, M. *Planta Med.* **1989**, *55*, 288−289.
(5) Kemmerling, W. *Z. Naturforsch.* **1996**, *51c*, 59−64.
(6) Gustafson, K. R.; Sowder, R. C., II; Henderson, L. E.; Parsons, I. C.; Kashman, Y.; Cardellina, J. C., II; McMahon, J. B.; Buckheit, R. W., Jr.; Pannell, L. K.; Boyd, M. R. *J. Am. Chem. Soc.* **1994**, *116*, 9337−9338.
(7) Gustafson, K. R.; Walton, L. K.; Sowder, R. C., II; Johnson, D. G.; Pannell, L. K.; Cardellina, J. C., II; Boyd, M. R. *J. Nat. Prod.* **2000**, *63*, 176−178.
(8) 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.
(9) Saether, O.; Craik, D. J.; Campbell, I. D.; Sletten, K.; Juul, J.; Norman, D. G. *Biochemistry* **1995**, *34*, 4147−4158.
(10) Craik, D. J.; Daly, N. L.; Bond, T.; Waine, C. *J. Mol. Biol.* **1999**, *294*, 1327−1336.
(11) Hallock, Y. F.; Sowder, R. C., II; Pannell, L. K.; Hughes, C. B.; Johnson, D. G.; Gulakowski, R.; Cardellina, J. C., II; Boyd. M. R. *J. Org. Chem.* **2000**, *65*, 124−128.
(12) Schopke, T.; Hasan Agha, M. I.; Kraft, R.; Otto, A.; Hiller, K. *Sci. Pharm.* **1993**, *61*, *145*−*153*.
(13) Claeson, P.; Goransson, U.; Johansson, S.; Luijendijk, T.; Bohlin, L. *J. Nat. Prod.* **1998**, *61*, 77−81.
(14) Goransson, U.; Luijendijk, T.; Johansson, S.; Bohlin, L.; Claeson, P. *J. Nat. Prod.* **1999**, *62*, 283−286.
(15) Hernandez, J.-F.; Gagnon, J.; Chiche, L.; Nguyen, T. M.; Andrieu, J.-P.; Heitz, A.; Hong, T. T.; Pham, T. T. C.; Le-Nguyen, D. *Biochemistry* **2000**, *39*, 5722−5730.
(16) Meragelman, K. M.; McKee, T. C.; Boyd, M. R. *J. Nat. Prod.* **2000**, *63*, 427−428.
(17) Bokesch, H. R.; Pannell, L. K.; McKee, T. C.; Boyd, M. R. *Tetrahedron Lett.* **2000**, *41*, 6305−6308.

NP000372L