Full Papers

New Circulin Macrocyclic Polypeptides from *Chassalia parvifolia*[1]

Kirk R. Gustafson,[†] Linda K. Walton,[†] Raymond C. Sowder, II,[‡] Donald G. Johnson,[‡] Lewis K. Pannell,[§] 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 Cancer Research and Development Center, Building 1052, Room 102, Frederick, Maryland 21702-1201, AIDS Vaccine Program, SAIC, Inc., NCI−FCRDC, Frederick, Maryland 21702-1201, and Laboratory of Bioorganic Chemistry, NIDDK, NIH, Bethesda, Maryland 20892

Received September 3, 1999

Four new macrocyclic polypeptides were isolated and identified from an extract of the tropical tree *Chassalia parvifolia*. Circulins C−F are 29−30 amino acid cyclic peptides in which the entire primary amino acid chain is covalently cyclized via peptide bonds. Their structures were deduced from a combination of FABMS analyses, N-terminal Edman degradation, endoproteinase digestion, and amino acid analyses. All the peptides share a high degree of sequence homology and contain six cysteine residues forming three intramolecular disulfide bridges. Circulins C−F inhibited the cytopathic effects of in vitro HIV-1 infection with EC_{50} values of 50−275 nM.

A novel class of biologically active macrocyclic peptides has recently been described from three different genera of plants in the Rubiaceae family. The HIV-inhibitory cyclic polypeptides, circulins A and B, were isolated from a Tanzanian collection of *Chassalia parvifolia* (Schumann).[2] Cyclopsychotride A, found in extracts of *Psychotria longipes* from Brazil, reportedly blocked neurotensin binding and increased cytosolic Ca^{2+} levels.[3] The kalata B1 peptide, isolated from African collections of *Oldenlandia affinis*, has uterotonic activity.[4] Additional representatives of this class of macrocyclic peptide have also been found recently in several *Viola* and *Leonia* (Violaceae) species.[5−8] These metabolites all contain 28−31 amino acid residues, forming a macrocycle due to peptide bond cyclization of the entire amide backbone. They share a high degree of sequence homology and contain six cysteine residues that form three intramolecular disulfide bonds. The disulfide linkage pattern in circulins A and B was previously elucidated by mass spectral analysis of peptide fragments generated after mild acid hydrolysis.[9] Three-dimensional (3D) NMR solution structures of both circulin A[10] and kalata B1[4] have been determined. The peptides form compact structures with closely related 3D folds that are dominated by β-turns and β-strands. Their disulfide connectivities form an "inhibitor-type", cysteine-knot motif,[11] similar to that found in ω-conotoxin.[12] The cyclic amide backbone and high disulfide content combine to produce rigid, highly ordered polypeptides that are resistant to proteolytic enzymes. The tertiary structure of these peptides appears critical to their biological function, because reductive cleavage of the disulfides in circulins A and B resulted in complete loss of anti-HIV activity.[2]

An evaluation of the constituents of *C. parvifolia* was originally based upon HIV-inhibitory activity[13] of both the organic and aqueous extracts. Continued bioassay-guided separations of *C. parvifolia* metabolites has provided four additional minor peptides, in addition to circulins A and B. We report herein the isolation, amino acid sequence determination, and anti-HIV activity of the macrocyclic polypeptides, circulins C−F.

Results and Discussion

Fractionation of the organic extract (8.2 g) of stems from *C. parvifolia* by a combination of solvent−solvent partitioning, gel permeation on Sephadex LH-20, and centrifugal partition chromatography concentrated its anti-HIV constituents into a peptide-enriched fraction. Final purification of this material by repeated C_{18} HPLC, eluted with a linear gradient of CH_3CN in H_2O (0.05% TFA), provided samples of circulin C (2.2 mg), circulin D (1.8 mg), circulin E (1.4 mg), and circulin F (1.4 mg).

Circulin C was isolated as an amorphous white solid that provided a FABMS $[MH]^+$ molecular ion at *m*/*z* 3102.3.[14] The peptidic nature of circulin C was readily apparent from characteristic resonances in its ^1H NMR spectrum. However, native circulin C was resistant to both direct N-terminal Edman degradation and enzymatic digestion with a variety of proteases, suggesting it was a macrocyclic analogue of circulins A and B. Reductive alkylation of the peptide by treatment with β-mercaptoethanol and 4-vinylpyridine provided the *S*-(β-4-pyridylethyl)cysteine (PEC) derivative of circulin C. FABMS analysis of the PEC derivative revealed a $[MH]^+$ molecular ion at *m*/*z* 3739.9, which was consistent with the presence of six cysteine residues that formed three disulfide bonds in native circulin C.[15] Treatment of PEC−circulin C with the endoproteinase Arg C provided a single peptide product with a FABMS molecular ion at *m*/*z* 3758.3. The increase in mass of approximately 18 Da was indicative of enzymatic cleavage of a cyclic peptide to a linear form. This was confirmed when N-terminal sequencing of the digestion product established the complete amino acid sequence of a linear 30 amino acid residue peptide. The difference between the

* To whom correspondence should be addressed. Tel.: (301) 846-5391. Fax: (301) 846-6177. E-mail: boyd@dtpax2.ncifcrf.gov.
† Laboratory of Drug Discovery Research and Development, NCI.
‡ AIDS Vaccine Program, SAIC.
§ Laboratory of Bioorganic Chemistry, NIDDK.

10.1021/np990432r CCC: $19.00
Published on Web 01/13/2000

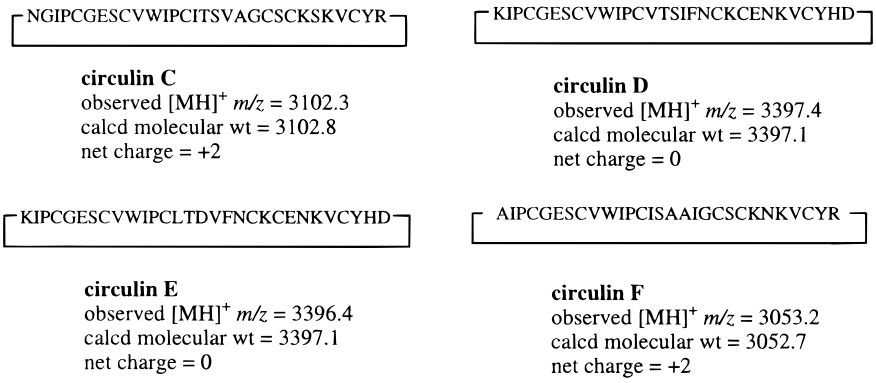

Figure 1. Cyclic amino acid sequences of circulins C−F. Solid lines connecting the two ends of each sequence represent cyclization of the peptide backbone. Sequences were established by N-terminal Edman degradation of linear peptides generated by digestion of the cyclic starting material with endoproteinase Arg C. Calculated molecular weights are those for the protonated molecular ion of the native peptide.

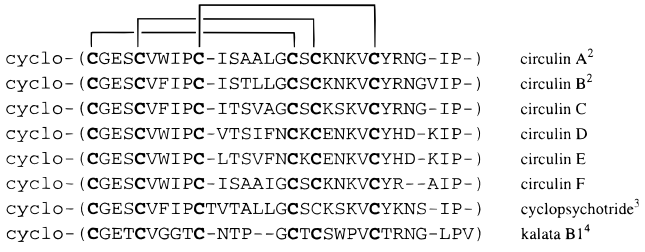

Figure 2. Sequence comparison of macrocyclic peptides isolated from plants in the family Rubiaceae. The six highly conserved cysteine residues in each peptide are highlighted in bold print. The disulfide bonding pattern between Cys I−Cys IV, Cys II−Cys V, and Cys III−Cys VI, which was established previously for this class of polypeptides,[4,9,10] is illustrated by the solid lines above circulin A.

theoretical molecular weight calculated for this linear sequence (3125.5 Da) and the molecular weight of circulin C measured experimentally by MS, was accountable to the three disulfide bonds and cyclic peptide backbone in native circulin C. The sequence of the linear peptide was also consistent with the amino acid composition of circulin C as determined by quantitative amino acid analysis. Thus, circulin C was established as a macrocyclic polypeptide with the amino acid sequence illustrated in Figure 1.

The cyclic structures and amino acid sequences of circulins D−F were established in a manner analogous to that described above. Native circulin D, circulin E, and circulin F provided FABMS $[MH]^+$ molecular ions at *m/z* 3397.4, 3396.4, and 3053.2, respectively. The PEC derivatives of each peptide were treated with Arg C, and the digestion products were purified by HPLC. FABMS analysis of the enzymatically cleaved PEC derivatives of circulins D−F revealed $[MH]^+$ molecular ions at *m/z* 4052.4, 4051.9, and 3707.7, respectively. Thus, circulins D−F were also macrocyclic, and each had six cysteine residues that formed three disulfide bonds. Direct N-terminal Edman sequencing of the enzymatically cleaved peptides established their complete amino acid sequences (Figure 1). These sequences were all consistent with the mass spectral measurements and subsequent quantitative amino acid analyses of the parent peptides. It is interesting that even though circulins D and E do not contain arginine, digestion with the enzyme Arg C resulted in cleavage of the amide link between aspartic acid (D) and lysine (K), albeit in low yield.

Circulins C−E are each composed of 30 amino acid residues, while circulin F contains 29 residues. The compounds all share considerable sequence homology with the other *Chassalia* macrocyclic peptides (Figure 2), and with related peptides isolated from various *Viola* and *Leonia* species.[5−8] The position of the six cysteine residues is highly conserved. This strongly suggests the peptides all possess the same Cys I−Cys IV, Cys II−Cys V, and Cys III−Cys VI cysteine-knot[11] pattern of disulfide bond formation that has already been established for circulins A and B[9,10] and kalata B1.[4]

Table 1. Anti-HIV Data for Circulin C[a]

HIV-1 strain	target cell line	EC_{50} (nM)
III_B	CEM-SS	73
RF	CEM-SS	165
214	CEM-SS	213
205	CEM-SS	155
G	CEM-SS	111
SK1	CEM-SS	51
N119	CEM-SS	241
A17	MT-2	247
G9106	MT-2	67
H1122	MT-2	48
III_B	MT-2	200

[a] Descriptions and sources of the virus strains and target cell lines,[17,18] and experimental details of the in vitro XTT tetrazolium-based assay[16] have been previously published.

The new circulin polypeptides were evaluated for their HIV-inhibitory properties in a cell-based in vitro assay, details of which have been described previously.[16] All exhibited anti-HIV activity comparable to that of circulins A and B. Circulins C−F inhibited the cytopathic effects of productive HIV-1 infection with EC_{50} values (concentration that is 50% effective) that ranged from 50 to 275 nM, depending upon the particular virus strain and host cell line used in the assay. Assay results for circulin C are provided in Table 1.

Experimental Section

Collection and Extraction. Stems from the tree *Chassalia parvifolia* Schumann (Rubiaceae) were collected in the Iringa Region of Tanzania in December 1988, by J. Lovett. The plant material was air-dried, ground into fine chips, and sequentially extracted with 1:1 MeOH−CH_2Cl_2 followed by 100% MeOH. The combined organic extracts were evaporated in vacuo and then separated by a solvent−solvent partitioning protocol.[19] The water-soluble fraction was partitioned between H_2O and *n*-BuOH, and the resulting *n*-BuOH-soluble material was further fractionated by gel permeation on Sephadex LH-20 eluted with 100% MeOH. Early eluting fractions from the LH-20 column were subjected to centrifugal partition chromatography (Sanki Instruments, model NMS) using H_2O−*n*-BuOH−HOAc−EtOH, 10:8:1:1 in the descending mode. Final purification of circulins C, D, E, and F was achieved by HPLC on a Rainin, Dynamax C_{18} column eluted with a 30-min linear gradient from 20% to 50% CH_3CN in H_2O (0.05% v/v of trifluoroacetic acid in the mobile phase). The purity of the isolated peptides was confirmed by FABMS analysis.

Mass Spectral Analysis. Samples were analyzed by FABMS, using glycerol as a matrix, on a JEOL SX102 mass spectrometer. Mass spectral data were acquired by a JEOL XMS data system. The experimental masses quoted and the calculated theoretical values represent average isotopic masses: circulin C $[MH]^+$ *m/z* 3102.3 (calcd 3102.8); PEC−circulin C $[MH]^+$ *m/z* 3739.9 (calcd 3739.6); linear PEC−circulin C $[MH]^+$ *m/z* 3758.3 (calcd 3757.6); circulin D $[MH]^+$ *m/z* 3397.4 (calcd 3397.1); linear PEC−circulin D $[MH]^+$ *m/z* 4052.4 (calcd 4051.9); circulin E $[MH]^+$ *m/z* 3396.4 (calcd 3397.1); linear PEC−circulin E $[MH]^+$ *m/z* 4051.9 (calcd 4051.9); circulin F $[MH]^+$ *m/z* 3053.2 (calcd 3052.7); linear PEC−circulin F $[MH]^+$ *m/z* 3707.7 (calcd 3707.6).

Amino Acid Analysis. Peptide samples (5−10 μg) were placed in a glass tube and hydrolyzed with the vapor phase of 6 N HCl containing 0.1% phenol, within an evacuated glass vial that was heated at 110 °C for 18 h. Residual condensed solvent was removed in vacuo and the residue analyzed with a Beckman model 6300 automated amino acid analyzer.

Reduction and Alkylation of Cysteine Residues. HPLC-purified peptide samples (750 μg) were dissolved in 800 μL of 8 N guanidine HCl and treated with 40 μL of 4 M Tris HCl (pH 8.5), followed by 40 μL of β-mercaptoethanol. The reaction mixture was kept in the dark, at room temperature, under an atmosphere of N_2 for 2 h. Excess reagent was removed under a stream of N_2, and then 120 μL of 4-vinyl pyridine was added, and the reaction proceeded for 2 h at room temperature, in the dark, under N_2. Derivatized peptides were purified by C_{18} HPLC using a linear gradient of CH_3CN in H_2O (from 20% to 50% CH_3CN over 30 min with 0.1% TFA v/v).

Enzymatic Cleavage. Reduced, alkylated peptide samples were dissolved in 400 μL of 0.1 M Tris-HCl (pH 8.5) and treated with 5.0 μg of endoproteinase Arg C (Sigma). After incubating at 37 °C for 18 h, the cleaved peptide products were isolated by C_{18} HPLC using a linear gradient of CH_3CN in H_2O (from 20% to 40% CH_3CN over 30 min with 0.1% TFA v/v).

Sequence Determination. Amino acid sequences were determined by sequential Edman degradation using a pulsed, liquid-phase protein/peptide sequencer (Applied Biosystems, Inc., model 477A) equipped with an on-line phenylthiohydantoin analyzer (model 120A).

Anti-HIV Assay. The crude extract, chromatographic fractions, and purified peptides were dissolved in either 100% DMSO or H_2O−DMSO (3:1), diluted to the desired concentration, and tested in an XTT-based in vitro anti-HIV assay, the experimental details of which have been reported previously.[16]

Acknowledgment. We thank J. Lovett and G. Cragg (NPB) for collections, T. McCloud for extractions, D. Clanton for primary screening data, and J. McMahon and R. Buckheit for the 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. N01-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 63 in the series HIV-inhibitory Natural Products. For part 62, see Rashid et al.[20]
(2) 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**, *113*, 9337−9338.
(3) 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.
(4) Saether, O.; Craik, D. J.; Campbell, I. D.; Sletten, K.; Juul, J.; Norman, D. G. *Biochemistry* **1995**, *34*, 414−7−4158.
(5) Goransson, U.; Luijendijk, T.; Johansson, S.; Bohlin, L.; Claeson, P. *J. Nat. Prod.* **1999**, *62*, 283−286.
(6) Claeson, P.; Goransson, U.; Johansson, S.; Luijendijk, T.; Bohlin, L. *J. Nat. Prod.* **1998**, *61*, 77−81.
(7) Schopke, I.; Hasan Ahha, M. I.; Kraft, R.; Otto, A.; Hiller, K. *Sci. Pharm.* **1993**, *61*, 145−153.
(8) Hallock, Y. F.; Sowder, R. C., II; Pannell, L. K.; Hughes, C. B.; Johnson, D. G.; Gulakowski, R. J.; Cardellina, J. H., II; Boyd, M. R. *J. Org. Chem.* (in press).
(9) Derua, R.; Gustafson, K. R.; Pannell, L. K. *Biochem. Biophys. Res. Commun.* **1996**, *228*, 632−638.
(10) 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.
(11) Isaacs, N. W. *Curr. Opin. Struct. Biol.* **1995**, *5*, 391−395.
(12) Pallaghy, P. K.; Nielsen, K. J.; Craik, D. J.; Norton, R. S. *Protein Sci.* **1994**, *3*, 1833−1839.
(13) Boyd, M. R. In *AIDS Etiology, Diagnosis, Treatment and Prevention*; De Vita, V. T., Jr., Hellman, S., Rosenberg, S. A., Eds.; Lippincott: Philadelphia, 1988; pp 305−319.
(14) Mass values quoted are average molecular weights.
(15) Addition of each pyridylethyl group results in an increase of approximately 106.1 Da in the molecular weight of a PEC derivative, relative to the molecular weight of a native peptide in which all cysteines are disulfide linked.
(16) Gulakowski, R. J.; McMahon, J. B.; Staley, P. G.; Moran, R. A.; Boyd, M. R. *J. Virol. Methods* **1991**, *33*, 87−100.
(17) Currens, M. J.; Gulakowski, R. J.; Mariner, J. M.; Moran, R. A.; Buckheit, R. W., Jr.; Gustafson, K. R.; McMahon, J. B.; Boyd, M. R. *J. Pharmacol. Exp. Ther.* **1996**, *279*, 645−651.
(18) Buckheit, R. W., Jr.; Fliakas-Boltz, V.; Decker, W. D.; Robertson, J. L.; Pyle, C. A.; White, E. L.; Bowden, B. J.; McMahon, J. B.; Boyd, M. R.; Bader, J. P.; Nickell, D. G.; Barth, H.; Antonucci, T. K. *Antiviral Res.* **1994**, *25*, 43−56.
(19) Grode, S. H.; James, T. R.; Cardellina, J. H., II; Onan, K. D. *J. Org. Chem.* **1983**, *48*, 5203−5207.
(20) Rashid, M. A.; Gustafson, K. R.; Boyd, M. R. *J. Nat. Prod.* (in press).

NP990432R