THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 45, Issue of November 5, pp. 46858–46867, 2004
Printed in U.S.A.

Conserved Structural and Sequence Elements Implicated in the Processing of Gene-encoded Circular Proteins*

Received for publication, July 2, 2004, and in revised form, August 18, 2004
Published, JBC Papers in Press, August 24, 2004, DOI 10.1074/jbc.M407421200

Julie L. Dutton‡§, Rosemary F. Renda¶, Clement Waine‡‖, Richard J. Clark‡, Norelle L. Daly‡, Cameron V. Jennings§¶, Marilyn A. Anderson¶, and David J. Craik‡

From the ‡Institute for Molecular Bioscience, University of Queensland, Brisbane, Quuensland 4072 and the ¶Department of Biochemistry, La Trobe University, Melbourne, Victoria 3086, Australia

The cyclotides are the largest family of naturally occurring circular proteins. The mechanism by which the termini of these gene-encoded proteins are linked seamlessly with a peptide bond to form a circular backbone is unknown. Here we report cyclotide-encoding cDNA sequences from the plant *Viola odorata* and compare them with those from an evolutionarily distinct species, *Oldenlandia affinis*. Individual members of this multigene family encode one to three mature cyclotide domains. These domains are preceded by N-terminal repeat regions (NTRs) that are conserved within a plant species but not between species. We have structurally characterized peptides corresponding to these NTRs and show that, despite them having no sequence homology, they form a structurally conserved α-helical motif. This structural conservation suggests a vital role for the NTR in the *in vivo* folding, processing, or detoxification of cyclotide domains from the precursor protein.

The discovery of naturally occurring circular proteins has in recent years introduced a new topological paradigm into the field of protein structure. The cyclotides, backbone cyclic proteins of 28–35 amino acid residues isolated from plants, are the largest family of naturally occurring circular proteins (1, 2) (see Table I). Interest in these molecules has been driven by their topological novelty and associated resistance to thermal and enzymatic degradation. They were initially discovered as active components in a tea used in parts of Africa to accelerate childbirth. The uterotonic activity of the prototypic cyclotide kalata B1 was maintained after boiling of the plant *Oldenlandia affinis* to make the medicinal tea (3). The remarkable stability of these proteins has been attributed to their novel structure, the cyclic cystine knot motif (4, 5) (Fig. 1), in which two disulfide bonds and the surrounding peptide sequence form an embedded ring through which a third disulfide bond is threaded. Although the cystine knot motif is found in a wide range of other peptides from sources as diverse as fungi, insects, plants, and animals (6), it is only in the cyclotides that this motif is coupled with a cyclic backbone to form the cyclic cystine knot.

The number of cyclotide sequences has grown to over 50 in recent years (7). It is thought that their natural function is in plant defense (8), but there is still relatively little known about the process by which these peptides are synthesized in plants. Our previous isolation of cDNA clones encoding cyclotides from *O. affinis*, from the Rubiaceae family, has shown that in contrast to well known microbially derived cyclic peptides, which typically contain fewer than 12 residues and are non-ribosomally synthesized by enzyme complexes, the cyclotides are directly encoded by genes (8). Four clones, named Oak1 through Oak4 (*O. affinis* kalata-encoding clones 1–4), were isolated in that study, which until now was the only study to report nucleic acid sequences of cyclotides. The precursor proteins predicted from the Oak cDNA clones feature a signal peptide sequence and one to three cyclotide repeats corresponding to mature cyclotide domains. The predicted precursors also exhibit a highly conserved sequence near the N terminus of each of the cyclotide sequences that has been termed the N-terminal repeat (NTR).[1]

Until now all that was known about the genes that encode cyclotides came from a single Rubiaceae family species, and all of these clones encoded cyclotides from the so-called Moebius sub-class. The cyclotides have been classified into two subclasses depending on whether they have a proline residue in the fifth loop as shown in Fig. 1. Moebius cyclotides have a *cis* proline that induces a local backbone twist, whereas bracelet cyclotides do not. In the current study we have isolated cyclotide-encoding cDNA clones from the other major plant family known to express cyclotides, the Violaceae. In particular, we have isolated four cDNA clones from *Viola odorata* and have found that the precursor proteins predicted from these clones feature the same general arrangement, having an endoplasmic reticulum signal peptide, an N-terminal prodomain, and one to three cyclotide domains each preceded by a conserved sequence of 25 residues at their N termini.

As well as reporting the first cDNA sequences encoding cyclotides from the Violaceae family, and isolation of the first cDNA clones encoding bracelet cyclotides, this report examines the structure and function of conserved elements of sequence in the processing of mature cyclotides from their precursor proteins. Many peptides and proteins are produced through processing of larger precursor molecules. Sequences within the

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY630563–AY630566.

The atomic coordinates and structure factors (codes 1WN4 and 1WN8) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Supported by Australian Postgraduate Awards.
‖ Supported by an AusAID scholarship.

** An Australian Research Council Professorial Fellow. To whom correspondence should be addressed. Tel.: 61-7-3346-2019; Fax: 61-7-3346-2029; E-mail: d.craik@imb.uq.edu.au.

[1] The abbreviations used are: NTR, N-terminal repeat; *Oa*NTR, *Oldenlandia affinis* NTR; *Vo*NTR, *Viola odorata* NTR; TFE, trifluoroethanol; NOE, nuclear Overhauser Effect; NOESY, nuclear Overhauser effect spectroscopy; r.m.s.d., root mean square deviation; DQF-COSY, double quantum filtered correlation spectroscopy; ECOSY, exclusive correlation spectroscopy.

precursor that are not part of the mature protein may modulate protein folding and/or structure (9, 10). Indeed, a diverse range of actions has been found for such propeptides, including modulation of protein folding or protein function, detoxification, and targeting. A recent study of a family of conotoxins, which are disulfide-rich proteins similar in size to the cyclotides, suggests that propeptides may play a role in the protein disulfide isomerase-catalyzed folding of precursor proteins (11).

In this report, the sequences of the precursor proteins predicted from cDNA clones have been determined and are used to identify residues that may be important for cyclotide processing. The solution structures of synthetic peptides based on the NTRs from *O. affinis* and *V. odorata* are presented, and the significance of the NTR structure to cyclotide processing is discussed.

EXPERIMENTAL PROCEDURES

RNA Isolation and Production of Partial Clones—Total RNA was isolated from *V. odorata* leaves using TRIzol® reagent (Invitrogen). The partial clones were obtained by reverse transcriptase-PCR using total RNA and the Gene Amp® Gold RNA PCR kit from PE Biosystems. The Vok1 cDNA was amplified with Kal2 and oligo(dT) primers as described by Jennings *et al.* (8) (Kal2: GGGGATCCGTITGYGGIGARACITG (BamHI-VCGET)) for amplification of Moebius cyclotide sequences from *O. affinis* RNA. The Voc1, Voc 2, and Voc3 partial cDNAs were amplified using oligo(dT) with a degenerate forward primer, cyc1, which is complementary to a conserved sequence in the bracelet family of cyclotides (cyc1: TGTGTTTGGATACCTTGC (CVWIPC)). The PCR conditions comprised 30 cycles with the following regime: 94 °C, 1 min; 47 °C, 1 min; and 72 °C, 1 min. The PCR fragments were gel-purified and cloned into the pCR2.1 vector (Invitrogen) for sequencing.

Preparation of the V. odorata cDNA Library and Isolation of Clones—Poly(A)$^+$ RNA was isolated from total RNA (1 mg) using the PolyATract® mRNA isolation system (Promega). The mRNA (5 μg) was used to construct a cDNA library with the Lambda ZAP-cDNA synthesis kit and packaging extracts from Stratagene. Full-length cDNAs were obtained by screening the library with the [α-^{32}P]dCTP-labeled PCR fragments. Secondary structure of the predicted precursor proteins was predicted in MacVector (12) using Chou-Fasman and Robson-Garnier algorithms.

Peptide Synthesis—Boc-L-amino acids were obtained from Novabiochem (Laufelfingen, Switzerland) or the Peptide Institute (Osaka, Japan); *t*-Boc-Pro-OCH$_2$-PAM resin was obtained from PerkinElmer Life Sciences (Brisbane, Australia). 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate was obtained from Richelieu Biotechnologies (Quebec, Canada). The following reagents were of peptide synthesis grade and obtained from Auspep (Melbourne, Australia): trifluoroacetic acid, *N*,*N*-diisopropylethylene, and *N*,*N*-dimethylformamide. Acetonitrile (high-performance liquid chromatography grade) was purchased from BDH (Poole, England).

A peptide based on the NTR from the *O. affinis* clones and another based on the NTR of the *V. odorata* clones were assembled manually by stepwise solid-phase peptide synthesis using the *in situ* neutralization protocol of Boc chemistry (13) starting from *t*-Boc-Pro-OCH$_2$-PAM resin on a 0.25-mmol scale. The following side-chain protected amino acids were used: Leu-OH·H$_2$O, Gly-OH, Lys(Clz)-OH, Gln(Xan-OH, Met-OH, Glu-(OcHex)-OH, Phe-OH, Thr(Bzl)-OH, Ser(Bzl)-OH, and Val-OH. The crude peptides were purified by preparative reverse-phase high-performance liquid chromatography (Vydac C$_{18}$) on a Waters high-performance liquid chromatography system using a linear gradient of 0–80% acetonitrile in water and 0.1% trifluoroacetic acid over 80 min. Mass spectrometry data were obtained using an atmospheric pressure ionization electrospray mass spectrometer (PerkinElmer Life Sciences, PE-Sciex, Canada).

NMR Spectroscopy—Samples for various ^{1}H NMR measurements contained 0.2–3.5 mM peptide in 90% ^{1}H$_2$O/10% ^{2}H$_2$O at pH 3.6–5.3. A sample of the *Vo*NTR peptide was also prepared in 20% deuterated trifluoroethanol (TFE)/70% ^{1}H$_2$O/10% ^{2}H$_2$O. For structure determination samples at 3.5 mM and pH 3.6 were used. Spectra were acquired on Bruker ARX 500- and 750-MHz spectrometers, mainly at 298 K, but for clarification of some assignments additional spectra were acquired at 280 K and 310 K. NOESY spectra were acquired with mixing times of 250 and 300 ms for *Oa*NTR and 200, 150, and 100 ms for *Vo*NTR. Solvent suppression in DQF-COSY and ECOSY experiments was achieved using selective low power irradiation of the water resonance during a relaxation delay of 1.8 s. In NOESY and TOCSY experiments, a modified WATERGATE sequence was used to achieve solvent suppression. Chemical shifts were internally referenced to sodium 2,2-dimethyl-2-silapentane-5-sulponate.

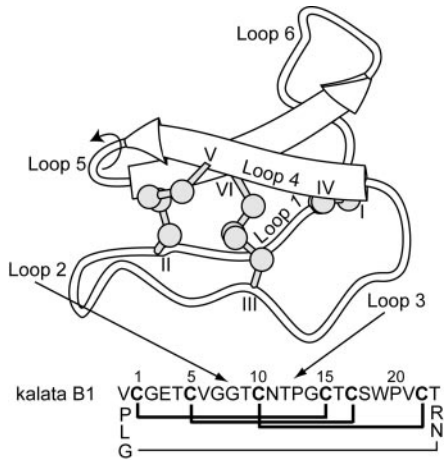


FIG. 1. **The cyclotide framework.** The structure is of the prototypic cyclotide kalata B1 (PDB code 1NB1) and illustrates the region of β-sheet (*broad arrows*) and disulfide bonds (in *ball-and-stick mode*). The Cys residues are labeled with *Roman numerals*, and the backbone loops between them are numbered *loops 1–6*. The *curved arrow* indicates the point in the sequence where a *cis*-Pro residue defines the Moebius sub-class of cyclotides. The sequence of kalata B1 is given *below the structure*.

Slow exchange experiments were performed following dissolution of the protonated peptides in ^{2}H$_2$O or in 20% deuterated TFE/80% ^{2}H$_2$O. TOCSY and one-dimensional experiments were conducted to monitor the exchange of the amide protons with the solvent. Amide proton signals that still appeared 60 min after dissolution in ^{2}H$_2$O were regarded as slow exchanging.

Spectra were processed using XWINNMR software (Bruker) on an SGI Octane2 workstation. Distance restraints were derived from the intensities of cross-peaks in NOESY spectra recorded with mixing times of 100–250 ms. Inter-proton distance restraints and pseudo-atom corrections were applied where necessary, and empirical corrections were added to intensities associated with methyl protons (14). Dihedral angle assignments were made from a combination of coupling constants and NOE (nuclear Overhauser effect) intensities: ϕ angle restraints were assigned based on $^3J_{HN\text{-}H\alpha}$ couplings measured from one-dimensional and DQF-COSY spectra, whereas χ_1 angles were assigned based on $^3J_{H\alpha\text{-}H\beta}$ couplings measured from ECOSY spectra and observed NOE intensities.

Preliminary three-dimensional structures of the peptides were determined using X-PLOR version 3.851 (15). Starting from a template structure with randomized ϕ and ψ angles and extended side chains, the program generated an ensemble of 50 structures using an *ab initio* simulated annealing protocol. These structures were subjected to high temperature dynamics (1000 K), with low initial weighting on the force constant and NOE restraints, prior to cooling to 100 K. Each structure was energy-minimized for 1000 cycles using the Powell algorithm and a refined CHARMM force-field. Final structures were calculated in CNS (16) with energy minimization in a water shell (17). The structures were displayed using Insight II (Biosym Technologies, San Diego, CA) or MolMol (18) and analyzed using PROMOTIF (19) and PROCHECK (20).

RESULTS

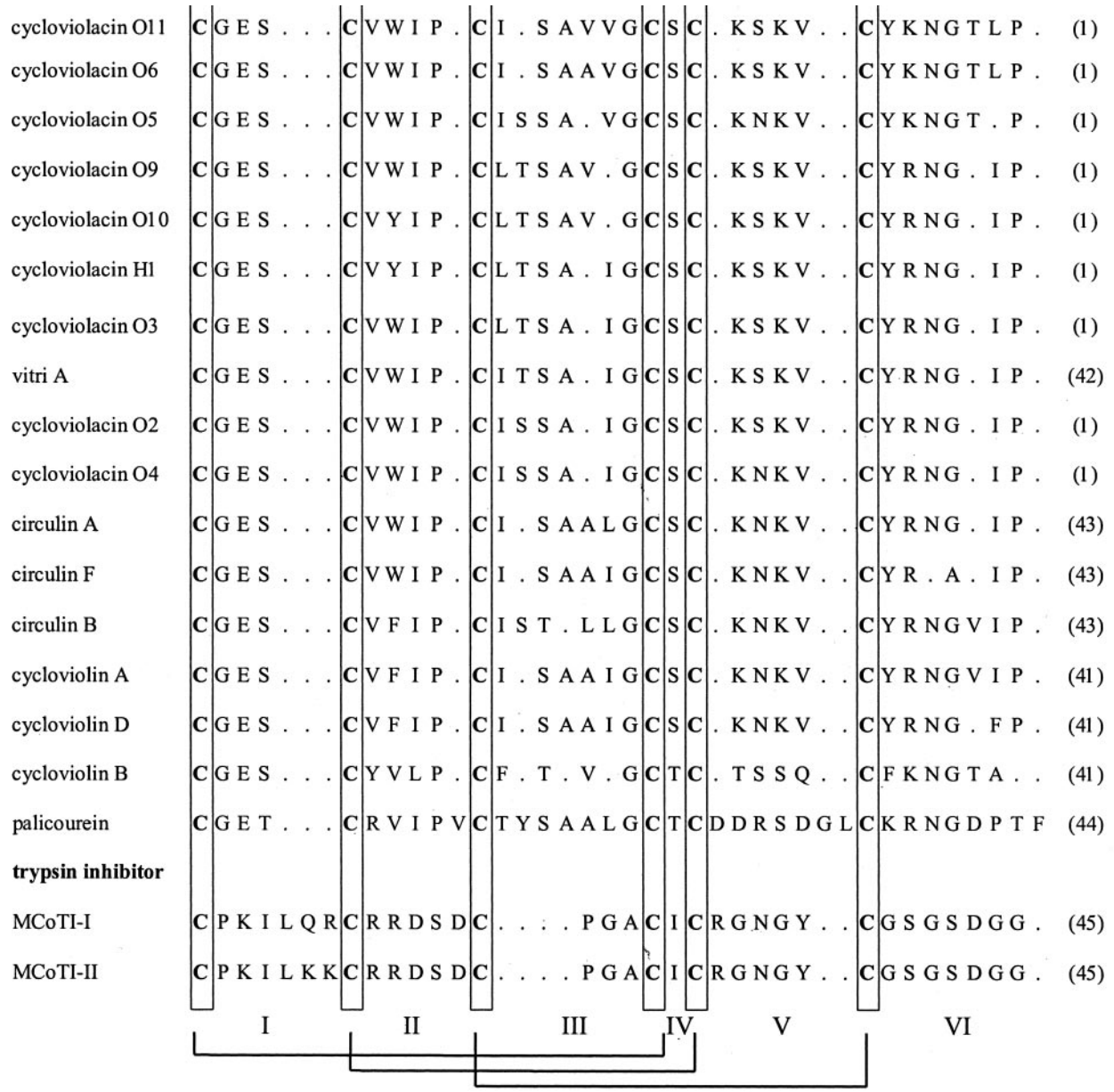
Previously we reported the isolation of four cDNA clones encoding cyclotide precursors from *O. affinis*, from the Rubiaceae, and prior to the current report these were the only known cyclotide-encoding clones. The *O. affinis* clones only encode cyclotides from the Moebius sub-class. In the current study it was of interest to isolate cDNAs from a plant species that is evolutionarily distant from *O. affinis* and also to isolate cDNAs encoding cyclotides from the other sub-class, namely bracelet cyclotides. *V. odorata* was screened, because it is both evolutionarily distant from *O. affinis* and a rich source of bracelet cyclotides (1, 21).

TABLE I
The cyclotide family
The cysteine residues are labeled with Roman numbers; the disulfide connectivity is indicated at the bottom of the table.

	I	II	III	IV	V	VI		
Moebius								
kalata B1	C	GET . . .	C	VGGT .	C	NT . . . PG	CTC	. SWPV . . C TRNGL . PV (4)
kalata B4	C	GET . . .	C	VGGT .	C	NT . . . PG	CTC	. SWPV . . C TRDGL . PV (1)
kalata S/ varv A[1]	C	GET . . .	C	VGGT .	C	NT . . . PG	CSC	. SWPV . . C TRNGL . PV (34)
kalata B2	C	GET . . .	C	FGGT .	C	NT . . . PG	CSC	. TWPI . . C TRDGL . PV (1)
kalata B3	C	GET . . .	C	FGGT .	C	NT . . . PG	CTC	DPWPI . . C TRDGL . PT (1)
kalata B6	C	GET . . .	C	FGGT .	C	NT . . . PG	CSC	SSWPI . . C TRNGL . PT (8)
varv B	C	GET . . .	C	FGGT .	C	NT . . . PG	CSC	DPWPM . . C SRNGL . PV (35)
varv G	C	GET . . .	C	FGGT .	C	NT . . . PG	CSC	DPWPV . . C SRNGV . PV (35)
varv H	C	GET . . .	C	FGGT .	C	NT . . . PG	CSC	ETWPV . . C SRNGL . PV (35)
violapeptide I	C	GET . . .	C	VGGT .	C	NT . . . PG	CSC	. SRPV . . C TXNGL . PV (36)
varv D	C	GET . . .	C	VGGS .	C	NT . . . PG	CSC	. SWPV . . C TRNGL . PI (35)
varv E[1]	C	GET . . .	C	VGGT .	C	NT . . . PG	CSC	. SWPV . . C TRNGL . PI (35)
varv C	C	GET . . .	C	VGGT .	C	NT . . . PG	CSC	. SWPV . . C TRNGV . PI (35)
kalata B7	C	GET . . .	C	TLGT .	C	YT . . . QG	CTC	. SWPI . . C KRNGL . PV (8)
vodo N	C	GET . . .	C	TLGK .	C	YT . . . AG	CSC	. SWPV . . C YRNGL . PV (21)
varv F	C	GET . . .	C	TLGT .	C	YT . . . AG	CSC	. SWPV . . C TRNGV . PI (35)
vodo M	C	GES . . .	C	FTGK .	C	YT . . . VQ	CSC	. SWPV . . C TRNGA . PI (21)
bracelet								
kalata B5	C	GES . . .	C	VYIP .	C	IS . GVIG	CSC	. TDKV . . C YLNGT . P . (1)
cyclopsychotride A	C	GES . . .	C	VFIP .	C	TVTALLG	CSC	. KSKV . . C YKNS . IP . (37)
cycloviolacin O7	C	GES . . .	C	VWIP .	C	TITALAG	CKC	. KSKV . . C Y . NS . IP . (1)
hypa A	C	AES . . .	C	VYIP .	C	TITALLG	CSC	. KNKV . . C Y . NG . IP . (38)
cycloviolacin O1	C	AES . . .	C	VYIP .	C	TVTALLG	CSC	. SNRV . . C Y . NG . IP . (1)
vico A	C	AES . . .	C	VYIP .	C	F . TGIAG	CSC	. KNKV . . C YYNGS IP . (39)
vico B	C	AES . . .	C	VYIP .	C	I . TGIAG	CSC	. KNKV . . C YYNGS IP . (39)
circulin D	C	GES . . .	C	VWIP .	C	VTS . IFN	CKC	. ENKV . . C YHDK . IP . (40)
circulin E	C	GES . . .	C	VWIP .	C	LTS . VFN	CKC	. ENKV . . C YHDK . IP . (40)
circulin C	C	GES . . .	C	VFIP .	C	ITSVA . G	CSC	. KSKV . . C YRNG . IP . (40)
cycloviolin C	C	GES . . .	C	VFIP .	C	LTTVA . G	CSC	. KNKV . . C YRNG . IP . (41)
cycloviolacin O8	C	GES . . .	C	VWIP .	C	ISS . VVG	CSC	. KSKV . . C YKNGTLP . (1)

[1] Both varv A and varv E were also isolated from other plant species and their sequences published under the names of kalata S and cycloviolacin O12.

TABLE I—*continued*



Isolation of V. odorata cDNA Clones and Comparison with the Oak Clones—Forward primers based on cyclotide sequences were used with oligo(dT) in reverse transcription-PCR to amplify cDNA-encoding cyclotides from *V. odorata* total leaf RNA. The partial cDNA clones were sequenced and used individually to screen a leaf cDNA library for full-length clones. Using the same forward primer that was used to isolate the Oak clones from *O. affinis*, which is based on loop 1 of the Moebius cyclotide kalata B1 (Fig. 1 and Table I; VCGET), we cloned a Moebius sub-class member from *V. odorata*. Conservation of the general structure of the genes, discussed in more detail below and shown in Fig. 2, indicates that they have evolved from an ancestral gene present in flowering plants before diversification of the asterid and rosid lineages in the eudicots. To isolate clones encoding members of the bracelet class of cyclotides, we designed a new primer complementary to a sequence that is conserved in loop 2 (Fig. 1), CVWIPC (Table I), of the bracelet cyclotides. This approach led to the identification of three cDNAs encoding bracelet cyclotides. In total, four full-length cDNA clones were isolated. The predicted protein sequences are shown in the *upper panel* of Fig. 2. For comparison, the earlier clones from *O. affinis* are shown in the *lower panel* of Fig. 2.

Nomenclature for the *V. odorata* clones follows that used for the *O. affinis* clones, where kalata-encoding cDNA clones were named OaKx, with x related to the order of discovery. The Moebius clone encodes a precursor with a single copy of kalata B1 and two copies of kalata S and was therefore named Vok1 (*V. odorata* kalata encoding clone 1). The Vok1 precursor is remarkably similar in structure to the Oak4 precursor, shown in Fig. 2, with three cyclotide repeats separated by conserved NTRs, an N-terminal signal sequence, and a small C-terminal tail. The bracelet cyclotide clones, Voc1, Voc2, and Voc3 (*V. odorata* cycloviolacin encoding clones), encode cycloviolacin O8, cycloviolacin O11, and cycloviolacin O13, a cycloviolacin/circulin-like peptide that has not been reported previously (numbering is based on the reported peptide sequences in Table I). These clones all resemble Oak1, encoding precursors with just one cyclotide repeat, with an endoplasmic reticulum signal sequence, an N-terminal pro-domain, NTR, and small C-terminal tail.

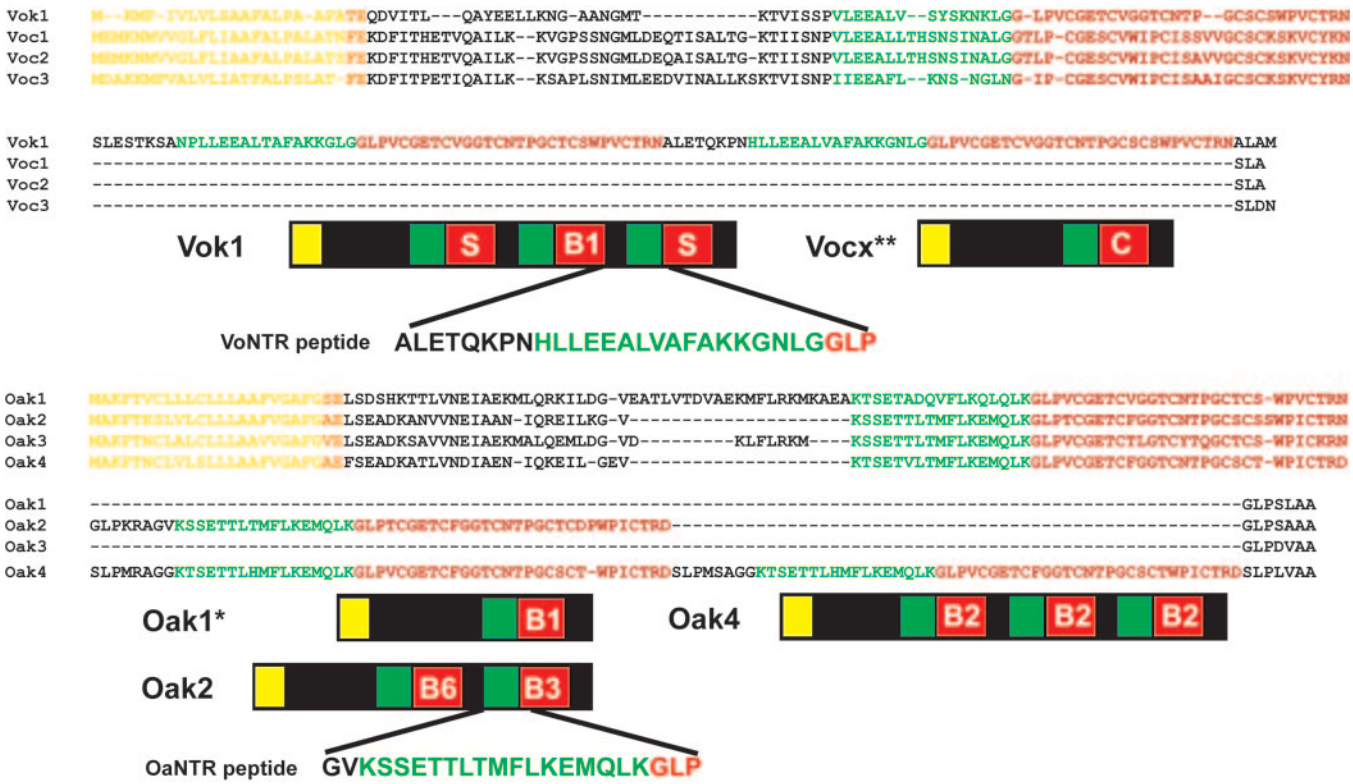


FIG. 2. **Predicted cyclotide precursor proteins from *V. odorata* and *O. affinis* cDNA clones (*upper* and *lower panels*, respectively).** Schematic representations of the general organization of the predicted protein precursor encoded by each type of cDNA clone are given along with alignments of the amino acid. The endoplasmic reticulum signals, cyclotide repeats, and most highly conserved part of the NTR are indicated in *yellow* (and *orange*), *red*, and *green*, respectively. The sequences corresponding to the synthetic NTR peptides are given, and their position in the predicted precursor proteins is indicated. Cleavage of the endoplasmic reticulum signal from the precursor is predicted to occur between the *yellow* and *orange* residues. *, Oak3 has the same structure as Oak1 but encodes kalata B7 instead of kalata B1. **, for Voc1, Voc2, and Voc3 the cyclotide designated by *C* is cycloviolacin O8, cycloviolacin O11, and [S13A] cycloviolacin O2, respectively.

The amino acid identities of the cyclotide and NTR domains of the *O. affinis* and *V. odorata* clones relative to Oak1 and Voc1 are summarized in Table II. Cyclotide domains in the Moebius group share greater than 79% identity at the amino acid level irrespective of whether they are sequences from the Rubiaceae or Violaceae. When the Moebius and bracelet mature cyclotide domains are compared they share only 40% identity. Interestingly, the NTR sequences are not as tightly conserved, and the sequences cluster according to the plant of origin rather than the Moebius or bracelet groupings. Conservation of the N-terminal signal sequences and NTRs between Vok1 and the Voc clones is not as great as in the Oak clones, but then this is to be expected, because the Oak clones all encode kalata variants, whereas the *V. odorata* clones encode cyclotides from different sub-classes, *i.e.* kalata and cycloviolacin variants.

Synthesis and Three-dimensional Structure Characterization of NTR Peptides—Because the NTR sequences are conserved within an individual plant species, it was hypothesized that these regions may play a role in cyclotide processing. Surprisingly, however, the NTRs are not conserved in sequence between species, so it was of interest to investigate possible structural conservation. Structure prediction suggested that the NTRs are helical. To experimentally verify this, two peptides were synthesized; one based on the NTR from the *O. affinis* clones and another based on the NTR of the *V. odorata* clones (named *Oa*NTR and *Vo*NTR, respectively). Both peptides were designed from the NTR sequences preceding a C-terminal cyclotide repeat domain and were synthesized using solid phase peptide chemistry. The sequences of these peptides are in Fig. 2. The *Oa*NTR peptide was designed based on the conserved 19 amino acids just upstream of the cyclotide repeat in Oak2. The NTR sequence from Oak2 was used, because it has a higher homology with the other repeat fragments than the NTR of Oak1. The first three amino acids of the mature cyclotide domain, Gly-Leu-Pro, were also included in the peptide. The *Vo*NTR peptide incorporated all of the sequence between the second and third cyclotide domains in Vok1, that is, the conserved 19 residues and the 6 remaining N-terminal residues. As with the *Oa*NTR peptide, the first 3 residues of the adjacent cyclotide domain (Gly, Leu, and Pro) were included.

TABLE II

Percentage amino acid identity of cyclotide and NTR domains with the cyclotide and NTR domains of the Oak1 and Voc1 precursors

NTR (17) includes the 17 relatively highly conserved amino acids N-terminal to the GLP of the cyclotide domain, and NTR (25) includes the 25 amino acids N-terminal to the GLP of the cyclotide domain.

	Cyclotide		NTR (25)		NTR (17)	
	Oak1	Voc1	Oak1	Voc1	Oak1	Voc1
Moebius subclass						
Oak1 (kalata B1)	100	44	100	4	100	5
Oak2 (kalata B6)	86	36	45	18	58	17
Oak2 (kalata B3)	82	30	50	13	58	17
Oak3 (kalata B7)	79	37	45	13	58	17
Oak4 (kalata B2 repeat 1)	82	37	57	14	64	17
Oak4 (kalata B2 repeat 2)	82	37	59	13	64	17
Oak4 (kalata B2 repeat 3)	82	37	59	13	64	17
Vok1 (kalata S repeat 1)	96	48	4	54	5	41
Vok1 (kalata B1)	100	44	9	31	23	29
Vok1 (kalata S repeat 2)	96	48	4	27	29	29
Bracelet subclass						
Voc1 (cycloviolacin O8)	44	100	4	100	5	100
Voc2 (cycloviolacin O11)	44	96	4	100	5	100
Voc3 (cycloviolacin O13)	44	80	14	61	17	41

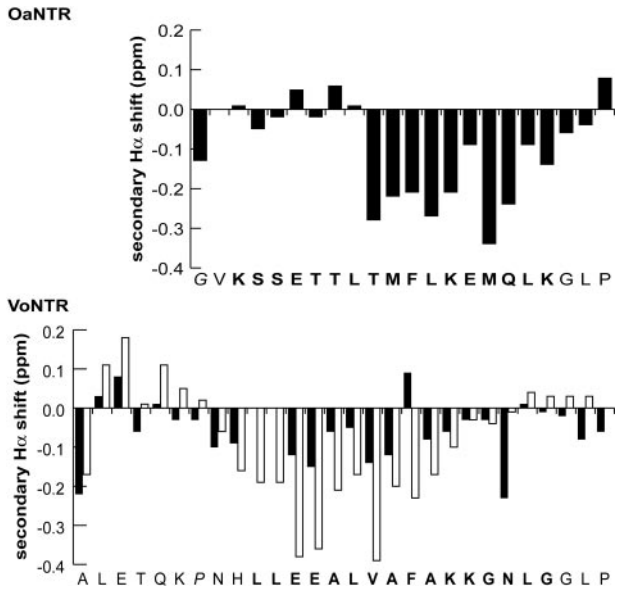


FIG. 3. **Secondary *α* proton shifts of the NTR peptides from *O. affinis* (*top*) and *V. odorata* (*bottom*).** Secondary α proton shifts more negative than -0.1 ppm are generally accepted to be indicative of helix. The α proton shifts for the *Oa*NTR and *Vo*NTR in 90% ^1H_2O/10%^2H_2O (*black bars*) and the *Vo*NTR in 20%TFE/70%^1H_2O/10% ^2H_2O (*white bars*) are plotted. The amino acids that are in the highly conserved part of the NTR are in *bold*. The first amino acid in the *Oa*NTR peptide and the equivalent position in the *Vo*NTR peptide are in *italics*.

NMR spectra were recorded, and the amino acid spin systems of the *Oa*NTR and *Vo*NTR peptides were identified using a combination of TOCSY and DQF-COSY spectra at 298 K. Sequential assignments were completed using TOCSY and NOESY spectra. Ambiguities due to overlap were resolved from additional spectra acquired at 280 K or 310 K. Secondary αH shifts, that is, the differences between observed αH shifts in the peptide of interest and those for the corresponding residues in random coil peptides, are indicative of secondary structure. In aqueous solution the C-terminal half of *Oa*NTR exhibits negative secondary shifts typical of helix, as may be seen in Fig. 3. The sample conditions of 298 K, 3.5 mM, and pH 3.6 were chosen, respectively, to provide good dispersion, good sensitivity, and to minimize exchange broadening of amide signals. A series of one-dimensional and TOCSY spectra recorded with concentrations ranging from 0.2 to 3.5 mM and pH from 3.6 to 5.3 showed no substantial changes in αH chemical shifts, suggesting that the helical tendency is not critically dependent on solution concentration or pH over these ranges. The secondary shifts of the *Vo*NTR peptide in water are also indicative of nascent helix in the conserved region highlighted in Fig. 3. To evaluate this helical tendency, TFE, a solvent that is known to stabilize helices in peptides that have an intrinsic propensity for helicity (22), was added to the solution. Comparison of the secondary shifts of the *V. odorata* repeat fragment dissolved in water and 20% TFE, shown in Fig. 3, reveals that the addition of TFE clearly increases the helicity of the peptide.

The secondary structures of the *Oa*NTR molecule were further inferred from a qualitative analysis of sequential and medium-range NOE intensities, amide proton slow exchange information and coupling constants, which are summarized in Fig. 4. Many strong to medium strength $d_{NN(i,\,i+1)}$ and medium strength $d_{\alpha N(i,\,i+1)}$ and some medium range NOEs were present between residues in the region of Ser-5 to Leu-21. Slow exchanging amide protons were detected from a series of one-dimensional spectra recorded up to 60 min after dissolution in ^2H_2O. These slow exchange protons correspond to residues Phe-12, Leu-13, and Glu-15 to Leu-21. All amide protons had exchanged by 60 min, and although this is faster than what is commonly seen for larger well structured proteins, the existence of an extensive set of observable amide protons in ^2H_2O indicates a high degree of structural ordering in this small peptide. Indeed the NOE pattern and slowly exchanging amide protons are strongly indicative of a helix in the region Phe-12 to Leu-21. Similarly, slow exchange and NOE data confirmed that the *Vo*NTR peptide forms a helix over residues His-9 to Lys-20. Dissolution of the fully protonated *Vo*NTR peptide in ^2H_2O identified four slow exchanging amide protons. Slow exchange experiments in the presence of 20% deuterated TFE identified the amide protons of residues Leu-10, Leu-11, and Glu-13 to Lys-21 as slow exchanging. This is consistent with the *Vo*NTR peptide featuring a nascent helical structure in water that is stabilized by the addition of 20% TFE.

The structures of the *Oa*NTR peptide were generated using 172 inter-proton distances derived from 110 sequential, 59 medium range, and four long range NOEs, and 12 backbone (ϕ) and two side-chain (χ_1) dihedral angle restraints. The residues Leu-9 to Leu-13, Met-16, Leu-18, Lys-19, and Leu-21 were assigned ϕ angles of $-65 \pm 30°$ and ϕ restraints of $-100 \pm 80°$ were applied for Ser-4, Glu-6, and Thr-7 based on additional coupling and NOE data. Stereospecific assignments of methylene protons and χ_1 dihedral angle restraints ($60 \pm 30°$) were derived for Ser-4 and Ser-5. Two NOEs were observed for $H\alpha_{i-1}$-$H\delta_i$ for Leu-21 and Pro-22, indicating that Pro-22 was in the *trans* conformation.

The structures of the *Vo*NTR peptide were calculated using 143 inter-proton distances, including 82 sequential and 61 medium range NOEs derived from NOESY spectra acquired with the peptide dissolved in 20% TFE/80% ^2H_2O. Although the COSY couplings measured for the *Vo*NTR were inconclusive due to large linewidths, some ϕ angles could be inferred due to the distinct differences in cross-peak intensities observed in the spectrum. In particular the cross-peaks for residues His-9 to Lys-20 were weak relative to those for all other residues so the ϕ angles for these residues could confidently be assigned as $-65 \pm 15°$. The β-methylene protons of residue Glu-3 were stereospecifically assigned, and the corresponding χ^1 angle was restrained to $-60 \pm 30°$. Sequential NOEs were observed for $H\alpha_{i-1}$-$H\delta_i$ for Lys-6 and Pro-7 and for Leu-27 and Pro-28, indicating that Pro-7 and Pro-28 were both in the *trans* conformation. The restraints used to calculate both structures are summarized in Fig. 4, from which it is apparent that there are a greater number of NOEs that are typical of helices found in *Oa*NTR than in *Vo*NTR. However, in the *Vo*NTR spectra there is some overlap of Hα shifts, and as a result potential $d_{\alpha N(i,\,i+4)}$ and $d_{\alpha N(i,\,i+3)}$ NOEs were overlapped with sequential NOEs.

The final set of *Oa*NTR structures, generated from energy minimization in a water shell, was analyzed, and the 20 structures with the lowest overall energies were selected. The ensemble of structures is shown in Fig. 5, superimposed over the backbone atoms of residues Thr-7 to Gly-20. A ribbon depiction of a representative structure is also included in Fig. 5 and shows that the C-terminal part of the peptide, between residues Thr-7 to Gly-20 adopts an α-helical conformation. The N-terminal region is less ordered. The view down the helical axis in Fig. 5 shows that the *Oa*NTR peptide forms an amphipathic helix, with hydrophobic residues clustered on the opposite side of the helix to the charged residues.

The geometric and energetic statistics that define the family of low energy *Oa*NTR structures are given in Table III. No significant deviations from idealized covalent geometry were observed. The superimposition of the ensemble of structures is excellent over residues Thr-7 to Gly-20, with a pairwise backbone r.m.s.d. of 0.49 $\pm$ 0.16 Å. There is slight fraying at the

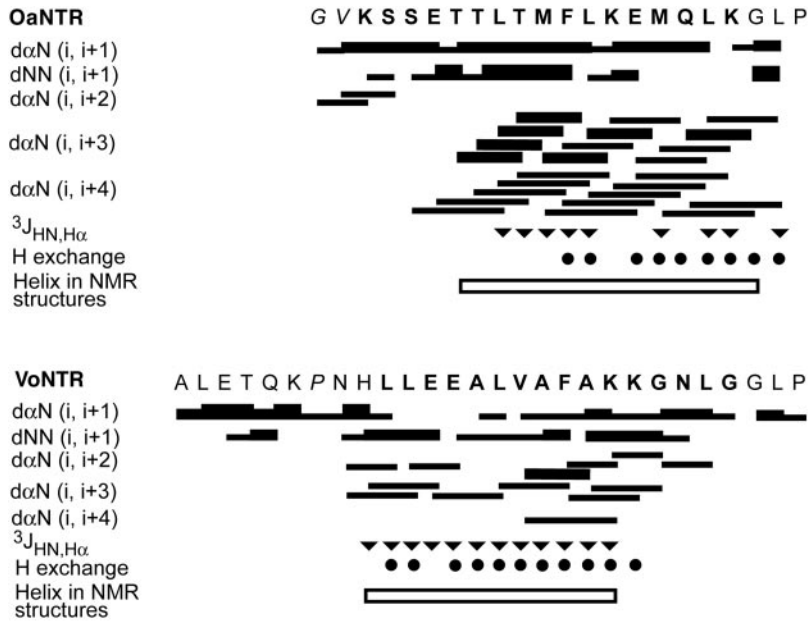


FIG. 4. **Restraint summary for *O. affinis* in 90% $^{1}H_2O$/10% $^{2}H_2O$ (*top*) and *V. odorata* in 20% TFE/70% $^{1}H_2O$/10% $^{2}H_2O$ (*bottom*).** The NOE intensities, weak, medium, and strong are shown as *thin*, *medium*, and *thick black bars*. Residues with small $^{3}J_{HN\text{-}H\alpha}$ couplings are indicated with an *arrowhead*. Slow exchanging amide protons are indicated with *circles*.

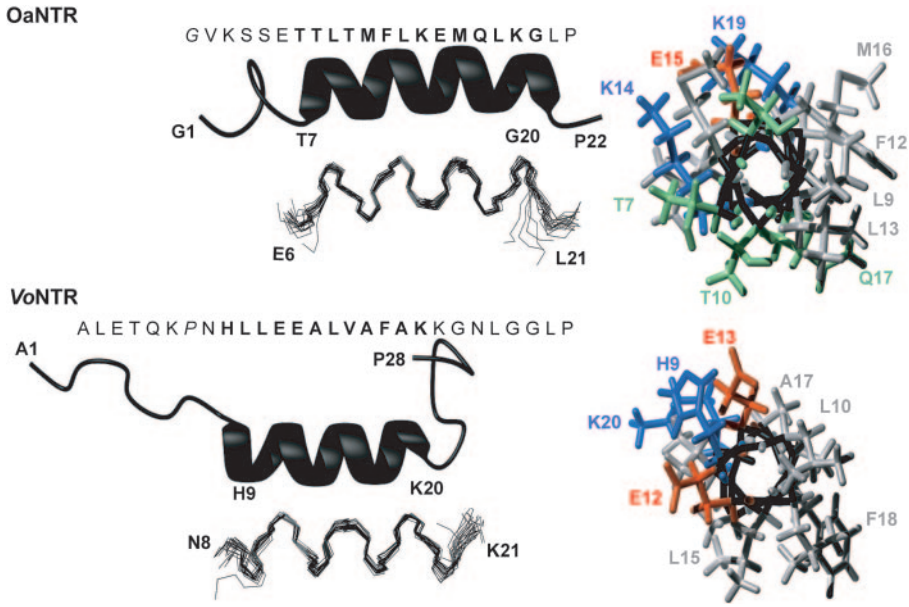


FIG. 5. **Structures of the *O. affinis* NTR peptide and the *V. odorata* NTR peptide.** Overlays of the 20 lowest energy structures are given and side chains are omitted. The *Oa*NTR structures were overlaid over the backbone atoms of Thr-7 to Gly-20, whereas the *Vo*NTR structures were overlaid over His-9 to Lys-20. Residues Gly-1 to Ser-5 and Pro-22 of *Oa*NTR peptide and residues Ala-1 to Pro-7 and Gly-22 to Pro-28 of the *Vo*NTR peptide have been omitted from the overlays for clarity. *Ribbon depictions* of representative structures are shown. The sequences of the peptides are given with the helical region in *bold*. The two residues in *italics* indicate equivalent positions in the two peptides (with regard to the precursor sequence). In the *right panel* representative structures of *Oa*NTR and *Vo*NTR show the amphipathic nature of the helices. The polar side chains, positively charged side chains, and negatively charged side chains are *green*, *blue*, and *red*, respectively. Note that His-9 in *Vo*NTR is *colored blue*, because it would be charged under the conditions used for NMR structure determination.

C-terminal and a degree of disorder near the N-terminal, with an overall pairwise backbone r.m.s.d. of 1.75 ± 0.64 Å. This reflects the general structure of the molecule, where the C-terminal half adopts a well defined α-helix and the N-terminal region is relatively disordered. Hydrogen bond restraints were not included in the structure calculations but were identified to be Phe-12 HN–Thr-8 O, Leu-13 HN–Leu-9 O, Glu-15 HN–Met-11 O, Met-16 HN–Phe-12 O, Asn-17 HN–Leu-13 O, Leu-18 HN–Lys-14 O, and Lys-19 HN–Glu-15 O. These hydrogen bonds were identified based on their detection in more than 50% of the calculated structures, and all of the relevant amide protons were detected as slow exchanging. The quality of the structures is excellent, and the small proportion of residues found in disallowed regions of the Ramachandran plot are from the unstructured region of the peptide, that is, residues Val-2 to Glu-6.

A superposition of the 20 lowest energy structures of *Vo*NTR in 20% TFE is shown in Fig. 5. Residues Ala-1 to Pro-7 and residues Gly-22 to Pro-28 have been omitted for clarity, because they are in unstructured regions. Residues His-9 to Lys-20 form an α-helix. A helix over residues His-9 to Lys-20 is consistent with the slow exchange data that identified the amide protons of residues 10 and 14 to 20 as taking more than 12 h to exchange with the solvent. The *Vo*NTR helix, like the *Oa*NTR helix, is amphipathic, with the hydrophobic residues clustered on one face and the charged residues on the other.

The geometric and energetic statistics that define the family of low energy *Vo*NTR structures are given in Table III and

TABLE III
Structural and energetic statistics for the O. affinis and V. odorata N-terminal repeat fragment peptides

	*Vo*NTR	*Oa*NTR
Mean r.m.s.d. from experimental restraints		
NOE (Å)	0.02 ± 0.00	0.03 ± 0.00
cdih (°)	0.15 ± 0.16	0.07 ± 0.14
Mean r.m.s.d. from idealized covalent geometry		
Bonds (Å)	0.00 ± 0.01	0.00 ± 0.00
Angles (°)	0.52 ± 0.03	0.61 ± 0.06
Impropers (°)	0.38 ± 0.03	0.42 ± 0.06
Mean total energy (kJ mol^{-1})	−1129 ± 53	−850 ± 33
Pairwise r.m.s.d.		
Backbone atoms (N, Cα, C) (Å)	(9–20) 0.41 ± 0.17	(1–22) 1.75 ± 0.64
	(6–21) 1.10 ± 0.33	(7–20) 0.49 ± 0.16
		(3–21) 1.09 ± 0.36
Heavy atoms (Å)	(9–20) 1.68 ± 0.34	(1–22) 2.42 ± 0.55
	(6–21) 2.32 ± 0.40	(7–20) 1.53 ± 0.25
		(3–21) 1.98 ± 0.34
Ramachandran plot		
Most favored regions (%)	72.0	78.7
Additional allowed regions (%)	22.5	13.4
Generously allowed regions (%)	3.4	6.1
Disallowed regions (%)	2.0	1.8

indicate no significant deviations from idealized covalent geometry. Over the helical region of the peptide, His-9 to Lys-20, the pairwise backbone r.m.s.d. is 0.41 ± 0.17 Å. Hydrogen bond restraints were not included in the structure calculations but were identified to be Glu-13 HN–His-9 O, Leu-15 HN–Glu-11 O, Val-16 HN–Glu-12 O, Ala-17 HN–Glu-13 O, Phe-18 HN–Ala-14 O, Ala-19 HN–Leu-15 O, and Lys-20 HN–Val-16 O. As with *Oa*NTR, these hydrogen bonds were detected in more than 50% of the calculated structures, and all of the corresponding amide protons were detected as slow exchanging. Hydrogen bonds were predicted in a minority of the ensemble for residues at either end of the helix, including Leu-10, Leu-11, and Lys-21, which were all found to be slow exchanging. The Ramachandran statistics are excellent with the small percentage of residues in the disallowed regions corresponding to disordered parts of the peptide.

Gly-1 of the *Oa*NTR peptide is equivalent in position to Pro-7 of the *Vo*NTR, indicated in Fig. 5 in *italics*. Considering the peptides in this way it is clear that the helix in the *Oa*NTR peptide is closer to the C-terminal and hence closer to the beginning of the cyclotide domain repeat than is the helix in *Vo*NTR. The *Oa*NTR, which included only conserved sequence, showed a tendency to form helix at the N-terminal end.

DISCUSSION

The mechanism by which mature cyclotides are processed from their precursor is unknown. In this report we describe the isolation of four *V. odorata* cDNA clones. This finding, taken with our previous work in *O. affinis*, shows that cyclotide genes are present and expressed, and that their overall gene structure is conserved, in two plant families from different lineages of the phylogenetic tree. In addition we have identified the first clones encoding bracelet cyclotides and have shown that, unlike the conotoxins where the prodomain is highly conserved and the mature conotoxin domains are hypervariable, for the cyclotides the prodomain sequence varies while the mature cyclotide domain is highly conserved (23). To investigate the significance of the observation that these NTR sequences are conserved within but not between species, we characterized their structures and here propose a role for the NTR in cyclotide processing/detoxification. Comparison of the predicted *V. odorata* precursor sequences to those from *O. affinis*, in combination with comparison of cyclotide sequences from various species and our structural analysis, has allowed the identification of residues and structural motifs that may be important in the cyclization process.

Structure and Significance of the NTR Domain—Conservation of the NTR within cyclotide precursors from the one species suggests that the NTR plays an important role in processing of cyclotides from the precursor. Our finding that the sequence of the NTR is not highly conserved between plant families is therefore surprising. To investigate the significance of this observation, and to gain some insights into the role of the NTR in cyclotide processing, the solution structures of synthetic peptides based on the NTR sequences from *O. affinis* and *V. odorata* cyclotide precursor proteins were determined. Although the propensity to form helix in aqueous solution is less pronounced in *Vo*NTR than in *Oa*NTR, the helix can be stabilized with the addition of 20% TFE. Given the large difference in NTR sequences from the two species, the conservation of structural characteristics suggests that formation of helices directly N-terminal to the cyclotide repeats is likely to be functionally important *in vivo*.

Studies of a wide range of propeptides indicate that these regions can be involved in the regulation of protein folding and modulate protein function. The main role of propeptides appears to be regulation of protein folding, catalyzing the folding reaction and in some cases inhibiting folding. Other less common functions of propeptides include protein transport, targeting, polymerization, and detoxification (9, 10). The exact role of the NTR region in cyclotide processing is uncertain, but as the position of the helix in the NTR appears to be variable between species, it seems more likely that the NTR is involved in the folding process rather than directing the exact site of cleavage from the precursor protein, which occurs at a very specific site. This is supported by an analysis of the surface properties of the helices in the two NTR peptides studied. Both are amphipathic helices (Fig. 5). Folded cyclotides, rather than burying hydrophobic residues as is common in most proteins, display their hydrophobic residues on their surface. *In vitro*, the hydrophobic solvent isopropanol aids the folding of synthetic linear peptides having cyclotide sequences (24). It is thus likely that the amphipathic helix in the NTR interacts with the hydrophobic residues of the cyclotide domain to aid the folding process. Alternatively, some cyclotides have been reported to be cytotoxic, and accumulation could potentially be deleterious to the plant, so it is possible that the role of the NTR is to detoxify the cyclotide domains during their biosynthesis inside the cell.

Studies of the cystine knot protein nerve growth factor (NGF) show that its prosequence facilitates folding of the protein, although, unlike the NTR described here, it is unstruc-

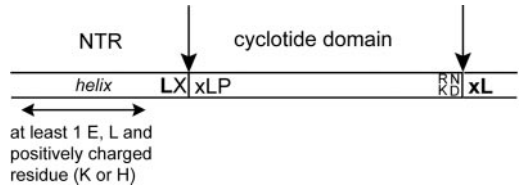


FIG. 6. **Summary of conserved sequence and structural elements implicated in precursor processing.** The sites at which the cyclotides are predicted to be processed from the precursor protein are indicated by *arrows*. Conserved residues are specified, with absolutely conserved residues in *bold*. *Uppercase X* indicates any amino acid; *lowercase x* indicates amino acids with small side chains.

tured unless covalently linked to NGF (25). The helical nature of the NTR has been noted for other propeptides, including, for example the N-terminal prosequence of mammalian prosomatostatin, the C-terminal prosequence of *Nicotiana alata* proteinase inhibitor, microcin B17 propeptide, and the prosequence of caricain (26–29). The first two of these are sorting signals, the third forms hydrophobic interactions with a processing synthetase, and the last acts to prevent access to the substrate-binding cleft.

The NTR domain could in principle play other roles, including for example host defense. It could be implicated in such a role because it forms a weak amphipathic helix that is typical of other host-defense peptides, including those isolated from bees and amphibians (30–32). Those peptides act against microbial pathogens by permeabilizing the membrane, leading to the breakdown of the transmembrane potential and ion gradients that cause leakage of cell contents and ultimately results in cell death. However, at this stage there is no evidence to suggest that NTR regions are released from the precursor, or that they would be active while attached to the precursor.

Sequence Analysis of Precursor and Mature Cyclotide Sequences—Initial analysis of the Oak clones (8) identified a potential cleavage site for removal of the cyclotide domains from the precursor proteins. Examination of the protein precursor sequence of the *V. odorata* cDNA clones confirmed that the cyclotides are cleaved from the precursor protein prior to the Gly residue in the Gly-Leu-Pro sequence of the cyclotide domain and following the Asn residue, as shown in Fig. 6. Comparison of the sequences surrounding the cleavage sites in the precursor proteins predicted from the Vok, Voc, and Oak clones reveals that there are a number of conserved residues that may be important in cyclotide processing (Fig. 6). All the precursor proteins have in the NTR region a conserved Leu residue two residues N-terminal to the N-terminal cleavage site. Still in the NTR region but further toward the N terminus, all clones have at least one Glu and one Leu residue. Conserved positively charged residues, usually Lys but in Voc1 and Voc2 His residues, are also present in the NTR region.

The residue on the C-terminal side of the N-terminal cleavage site is not absolutely conserved in all mature cyclotides. Although it is usually a Gly, in some cyclotides the position is occupied by a Lys, Ser, or Ala residue. Most cyclotides feature a hydrophobic residue immediately downstream of this, although again this is not absolutely conserved, with polar and charged residues sometimes occurring in this position. The Pro residue following the Gly and Leu in the precursors is conserved in all but one mature cyclotide sequence. There is also a highly conserved sequence, Cys(Ala/Gly)Glu(Thr/Ser)Cys, just downstream of that cleavage site (Fig. 2). The conservation of this sequence and the high conservation of the Pro residue may indicate that they are involved in processing, but alternatively may reflect important structural or functional roles for these residues in mature cyclotides.

The position preceding the C-terminal cleavage site, with the apparent exception of circulin F is occupied in cyclotide sequences by an Asn or Asp residue. In the precursor sequences derived from the cDNA sequence the Asn/Asp is followed by a small uncharged residue and a Leu residue. Upstream of the Asn residue in mature cyclotide sequences there is in most cases a positively charged residue. The positions five or seven residues upstream of the conserved Asn/Asp are also often occupied by positively charged residues. In all cases where there is not a positively charged residue immediately upstream of the Asn/Asp, one of the other two positively charged residues are present. A summary of the conserved residues is given in Fig. 6.

The significance of the conservation of the Asn/Asp residue at the C-terminal cleavage site derives from the fact that Asn and Asp residues are known to undergo a side-chain to backbone ring formation that results in cleavage of their C-terminal peptide bond. This reaction is observed in intein splicing, *i.e.* where specific protein sequences are spliced out in an autocatalytic process and the remaining proteins are ligated to form a mature protein. The post-translational processing of concanavalin A, a lectin from Jackbean cotyledons, also involves cleavages following Asn residues. This process, unlike intein splicing, is assisted by an enzyme, asparaginyl endopeptidase (33). It is possible then that the Asn (or Asp) residue that is conserved in the cyclotides forms a succinimide (or anhydride ring) resulting in cleavage at the C-terminal site. If the N-terminal site is cleaved by an enzyme, the succinimide (or anhydride in the case of Asp) may react with the free N-terminal to form the mature cyclotide.

In summary, isolation of cDNA clones from *V. odorata* has allowed us to gain a further insight into cyclotide biosynthesis. It is now clear that there is little sequence conservation of NTR sequences between plant species, but that the NTRs form a conserved structural motif in the form of an amphipathic helix. This structural conservation is suggestive of a role for the NTR sequence in precursor processing. Further studies are underway in our laboratories to clarify this role and more fully characterize cyclotide biosynthesis and processing.

Acknowledgment—C. W. thanks AusAID for a Ph.D. scholarship.

REFERENCES

1. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol.* **294,** 1327–1336
2. Trabi, M., and Craik, D. J. (2002) *Trends Biochem. Sci.* **27,** 132–138
3. Gran, L. (1973) *Acta Pharmacol. Toxicol.* **33,** 400–408
4. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) *Biochemistry* **34,** 4147–4158
5. Craik, D. J., Anderson, M. A., Barry, D. G., Clark, R. J., Daly, N. L., Jennings, C. V., and Mulvenna, J. (2002) *Lett. Peptide Sci.* **8,** 119–128
6. Craik, D. J. (2001) *Toxicon* **39,** 1809–1813
7. Craik, D. J., Daly, N. L., Mulvenna, J., Plan, M. R., and Trabi, M. (2004) *Curr. Protein Pept. Sci.* **5,** 297–315
8. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10614–10619
9. Braun, P., and Tommassen, J. (1998) *Trends Microbiol.* **6,** 6–8
10. Shinde, U., and Inouye, M. (2000) *Semin. Cell Dev. Biol.* **11,** 35–44
11. Buczek, O., Olivera, B. M., and Bulaj, G. (2004) *Biochemistry* **43,** 1093–1101
12. Olson, S. A. (1994) *Methods Mol. Biol.* **25,** 195–201
13. Schnolzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B. (1992) *Int. J. Pept. Protein Res.* **40,** 180–193
14. Clore, G. M., Brunger, A. T., Karplus, M., and Gronenborn, A. M. (1986) *J. Mol. Biol.* **191,** 523–551
15. Brünger, A. T. (1992) *X-PLOR Version 3.1 A System for X-ray Crystallography and NMR*, Yale University, New Haven
16. Brunger, A. T., Adams, P. D., and Rice, L. M. (1997) *Structure* **5,** 325–336
17. Rosengren, K. J., Clark, R. J., Daly, N. L., Goransson, U., Jones, A., and Craik, D. J. (2003) *J. Am. Chem. Soc.* **125,** 12464–12474
18. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph.* **14,** 51–55, 29–32
19. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci.* **5,** 212–220
20. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
21. Svangard, E., Goransson, U., Smith, D., Verma, C., Backlund, A., Bohlin, L., and Claeson, P. (2003) *Phytochemistry* **64,** 135–142
22. Buck, M. (1998) *Q. Rev. Biophys.* **31,** 297–355
23. Olivera, B. M., Walker, C., Cartier, G. E., Hooper, D., Santos, A. D., Schoenfeld, R., Shetty, R., Watkins, M., Bandyopadhyay, P., and Hillyard,

D. R. (1999) *Ann. N. Y. Acad. Sci.* **870,** 223–237
24. Daly, N. L., Love, S., Alewood, P. F., and Craik, D. J. (1999) *Biochemistry* **38,** 10606–10614
25. Kliemannel, M., Rattenholl, A., Golbik, R., Balbach, J., Lilie, H., Rudolph, R., and Schwarz, E. (2004) *FEBS Lett.* **566,** 207–212
26. Mouchantaf, R., Kumar, U., Sulea, T., and Patel, Y. C. (2001) *J. Biol. Chem.* **276,** 26308–26316
27. Nielsen, K. J., Hill, J. M., Anderson, M. A., and Craik, D. J. (1996) *Biochemistry* **35,** 369–378
28. Roy, R. S., Kim, S., Baleja, J. D., and Walsh, C. T. (1998) *Chem. Biol.* **5,** 217–228
29. Groves, M. R., Taylor, M. A., Scott, M., Cummings, N. J., Pickersgill, R. W., and Jenkins, J. A. (1996) *Structure* **4,** 1193–1203
30. Dathe, M., and Wieprecht, T. (1999) *Biochim. Biophys. Acta* **1462,** 71–87
31. Shai, Y. (1999) *Biochim. Biophys. Acta* **1462,** 55–70
32. Matsuzaki, K. (1999) *Biochim. Biophys. Acta* **1462,** 1–10
33. Sheldon, P. S., Keen, J. N., and Bowles, D. J. (1996) *Biochem. J.* **320,** 865–870
34. Claeson, P., Goransson, U., Johansson, S., Luijendijk, T., and Bohlin, L. (1998) *J. Nat. Prod.* **61,** 77–81
35. Goransson, U., Luijendijk, T., Johansson, S., Bohlin, L., and Claeson, P. (1999) *J. Nat. Prod.* **62,** 283–286
36. Schöpke, T., Hasan Agha, M. I., Kraft, R., Otto, A., and Hiller, K. (1993) *Sci. Pharm.* **61,** 145–153
37. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. (1994) *J. Nat. Prod.* **57,** 1619–1625
38. Broussalis, A. M., Goransson, U., Coussio, J. D., Ferraro, G., Martino, V., and Claeson, P. (2001) *Phytochemistry* **58,** 47–51
39. Goransson, U., Broussalis, A. M., and Claeson, P. (2003) *Anal. Biochem.* **318,** 107–117
40. Gustafson, K. R., Walton, L. K., Sowder, R. C. I., Johnson, D. G., Pannell, L. K., Cardellina, J. H. I., and Boyd, M. R. (2000) *J. Nat. Prod.* **63,** 176–178
41. Hallock, Y. F., Sowder, R. C., 2nd, Pannell, L. K., Hughes, C. B., Johnson, D. G., Gulakowski, R., Cardellina, J. H., 2nd, and Boyd, M. R. (2000) *J. Org. Chem.* **65,** 124–128
42. Svangard, E., Goransson, U., Hocaoglu, Z., Gulbo, J., Larsson, R., Claeson, P., and Bohlin, L. (2004) *J. Nat. Prod.* **167,** 144–147
43. Gustafson, K. R., Sowder, I. R. C., Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina, I. J. H., McMahon, J. B., Buckheit, J. R. W., Pannell, L. K., and Boyd, M. R. (1994) *J. Am. Chem. Soc.* **116,** 9337–9338
44. Bokesch, H. R., Pannell, L. K., Cochran, P. K., Sowder, R. C., 2nd, McKee, T. C., and Boyd, M. R. (2001) *J. Nat. Prod.* **64,** 249–250
45. Hernandez, J.-F., Gagnon, J., Chiche, L., Nguyen, T. M., Andrieu, J.-P., Heitz, A., Hong, T. T., Pham, T. T. C., and Nguyen, D. L. (2000) *Biochemistry* **39,** 5722–5730

Conserved Structural and Sequence Elements Implicated in the Processing of Gene-encoded Circular Proteins

Julie L. Dutton, Rosemary F. Renda, Clement Waine, Richard J. Clark, Norelle L. Daly, Cameron V. Jennings, Marilyn A. Anderson and David J. Craik

J. Biol. Chem. 2004, 279:46858-46867.
doi: 10.1074/jbc.M407421200 originally published online August 24, 2004
