Available online at www.sciencedirect.com
ScienceDirect
Phytochemistry 69 (2008) 939–952
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

# The alpine violet, *Viola biflora*, is a rich source of cyclotides with potent cytotoxicity

Anders Herrmann [a], Robert Burman [a], Joshua S. Mylne [b], Gustav Karlsson [a,b], Joachim Gullbo [c], David J. Craik [b], Richard J. Clark [b], Ulf Göransson [a,*]

[a] *Division of Pharmacognosy, Department of Medicinal Chemistry, Uppsala University, Biomedical Centre, P.O. Box 574, SE-751 23 Uppsala, Sweden*
[b] *The Institute for Molecular Biosciences, The University of Queensland, Brisbane, Queensland, 4072, Australia*
[c] *Division of Clinical Pharmacology, Department of Medical Sciences, University Hospital, Uppsala University, SE-751 85 Uppsala, Sweden*

Received 26 July 2007; received in revised form 15 October 2007
Available online 14 January 2008

**Abstract**

The cyclotides are currently the largest known family of head-to-tail cyclic proteins. The complex structure of these small plant proteins, which consist of approximately 30 amino acid residues, contains both a circular peptide backbone and a cystine knot, the combination of which produces the cyclic cystine knot motif. To date, cyclotides have been found in plants from the Rubiaceae, Violaceace and Cucurbitaceae families, and are believed to be part of the host defence system. In addition to their insecticidal effect, cyclotides have also been shown to be cytotoxic, anti-HIV, antimicrobial and haemolytic agents. In this study, we show that the alpine violet *Viola biflora* (Violaceae) is a rich source of cyclotides. The sequences of 11 cyclotides, vibi A-K, were determined by isolation and MS/MS sequencing of proteins and screening of a cDNA library of *V. biflora* in parallel. For the cDNA screening, a degenerate primer against a conserved (AAFALPA) motif in the cyclotide precursor ER signal sequence yielded a series of predicted cyclotide sequences that were correlated to those of the isolated proteins. There was an apparent discrepancy between the results of the two strategies as only one of the isolated proteins could be identified as a cDNA clone. Finally, to correlate amino acid sequence to cytotoxic potency, vibi D, E, G and H were analysed using a fluorometric microculture cytotoxicity assay using a lymphoma cell line. The $IC_{50}$-values of the bracelet cyclotides vibi E, G and H ranged between 0.96 and 5.0 μM while the Möbius cyclotide vibi D was not cytotoxic at 30 μM.


*Keywords:* *Viola biflora L.*; Violaceae; Arctic yellow-violet; Two-flower violet; Cyclotides; Circular proteins; cDNA screening; Cytotoxicity; MS/MS; Anticancer

## 1. Introduction

The cyclotides are a family of plant-derived proteins containing 28–37 amino acids. These mini-proteins have an extraordinary structure: they are head-to-tail macrocyclic and they have three disulfides arranged in a cystine knot, in which two disulfide bonds and their connecting backbone segments form an embedded ring that is penetrated by the third disulfide bond (Craik et al., 2006; Göransson et al., 2003). Together these features define the cyclic cystine knot (CCK) motif (Craik et al., 1999), which is conserved within the cyclotide family and makes them extremely stable: they withstand both enzymatic and thermal degradation (Colgrave and Craik, 2004). The key features of the cyclotide structure are shown in Fig. 1A. In addition, the disulfide network fills the core of the protein, which forces the hydrophobic parts of the protein to be exposed at the molecular surface. These unique properties make cyclotides a prime target for protein engineering with possible applications in both medicine and agriculture; however our knowledge about their expression, biosynthesis, and distribution *in planta* is still limited.

* Corresponding author. Tel.: +46 18 4715031; fax: +46 18 509101.
*E-mail address:* ulf.goransson@fkog.uu.se (U. Göransson).

0031-9422/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2007.10.023

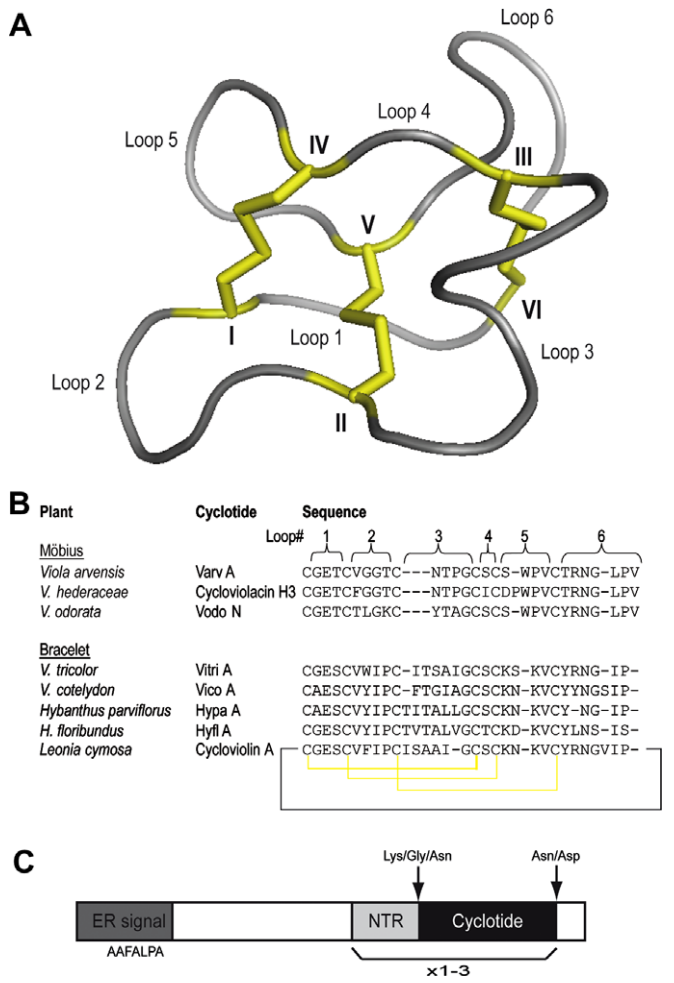


Fig. 1. Structure, sequences and arrangement of the genetic structure of cyclotides. (A) Backbone ribbon and the disulfide network of cyclotides with inter-cysteine loops 1–6 and cystine residues I-VI (cycloviolacin O1, PDB file 1NBJ) (Rosengren et al., 2003). (B) Sequences of representative cyclotides from all Violaceae plants where proteins have been characterized: *V. arvensis* (Claeson et al., 1998), *V. hederaceae* (Chen et al., 2005), *V. odorata* (Svangård et al., 2003), *V. tricolor* (Svangård et al., 2004), *V. cotyledon* (Göransson et al., 2003), *Hybanthus parviflorus* (Chen et al., 2005), *H. floribundus* (Simonsen et al., 2005) and *Leonia cymosa* (Hallock et al., 2000). The number of Möbius and bracelet cyclotides in the figure reflects the proportions between the two subfamilies of all the 88 cyclotides characterized so far. Loop numbers and disulfide bonds are shown at the top and bottom, respectively. (C) Schematic block diagram of cyclotide precursors. The precursor is arranged in the same way in all *Viola* species investigated: A conserved endoplasmic reticulum (ER) signal (dark gray), a pro-region (white) a N-terminal repeat (NTR) signal (grey) followed by the mature cyclotide sequence (black) and a poly-end tail (white). Cleavage points of the mature cyclotide are after the conserved Lys/Gly/Asn residue and Asn/Asp. The NTR- and cyclotide-region can be repeated up to three times, encoding different or identical cyclotides within one gene. The conserved ER-region has been exploited when designing a degenerative primer, i.e. AAFALPA, in order to find genes in plants encoding for cyclotides. Information regarding sequences of naturally occurring cyclotides, synthetic derivatives, grafted analogs and acyclic permutants is available at a recently created database for backbone-cyclized proteins, CyBase, at http://research1t.imb.uq.edu.au/cybase (Mulvenna et al., 2006).

The discovery of the cyclotides dates back to the early 1970s when it was reported that women in Congo used a decoction of the plant *Oldenlandia affinis* DC. (Rubiaceae) to accelerate childbirth (Gran, 1973) The uterotonic agent was isolated and shown to be a protein of 29 amino acid residues in size. This protein, named kalata B1 after the native name for *O. affinis,* ''Kalata-Kalata'', is the founding member of the cyclotide family (Craik et al., 1999). It was reported that the N- and C-termini were blocked but it was not until 1995 that the macrocyclic nature of kalata B1 was revealed (Saether et al., 1995). At that time three similar proteins had been reported from two additional Rubiaceae plants: the circulins A and B from *Chassalia parvifolia* Schum. (Gustafson et al., 1994), and cyclopsychotride A from *Psychotria longipes* Muell. Arg. (Witherup et al., 1994). In addition, a peptide, named violapeptide-I, had been isolated and partially characterised from *Viola arvensis Murr.*, which belongs to the plant family of Violaceae (Schöpke et al., 1993). Today plants from the Violaceae plant family contribute to the majority of known cyclotides. Of the 88 cyclotides reported to date 60 originate from Violaceae, and cyclotides have been identified in all family members hitherto examined. In these plants, cyclotides are expressed as complex cocktails; for example the number of reported cyclotides from the single species *Viola odorata* L. is 30 (Ireland et al., 2006a), but it has been suggested that the number of different cyclotides in one species may be over 100 and that the total number in the family Violaceae alone might exceed 9000 (Simonsen et al., 2005). If this is true, the cyclotides would be one of the largest protein families known (Simonsen et al., 2005). Fig. 1B lists all of the Violaceae species where cyclotides have been identified, together with a representative cyclotide from each species.

Cyclotides are divided into two main subfamilies. The division is based upon the presence of a *cis*-Pro peptide bond in loop 5, causing a 'twist' in the protein backbone. Cyclotides with this feature form the Möbius subfamily, while those lacking the *cis*-Pro form the bracelet subfamily (Craik et al., 1999). As illustrated in Fig. 1B, the subfamilies also differ in the size and amino acid content of the inter-cysteine loops, i.e. the number and types of amino acids between adjacent cysteines. For example, bracelet cyclotides generally contain a higher number of cationic residues (in loops 5 and 6), and a hydrophobic α-helix that is absent in the Möbius cyclotides (in loop 3). Loops 1 and 4 are, on the other hand, relatively conserved between the subfamilies. Structurally important residues are well conserved: this includes the glutamate residue in loop 1 and the strictly conserved cysteines. The former is found in all cyclotides characterised to date except one, in which it is replaced by aspartate (Plan et al., 2007), and is involved in defining and stabilizing the structure through multiple intra-molecular hydrogen bonds (Rosengren et al., 2003).

The role of cyclotides in plants is likely connected to host defence because of their potent insecticidal properties (Jennings et al., 2001). The fact that cyclotide-expression varies over the growth season (Trabi et al., 2004) supports this suggestion together with the differential expression that is seen between different plant parts (Trabi and Craik,

2004). The mechanism of action for the cytotoxic properties of the cyclotides has been shown to involve disruption of cell membranes (Kamimori et al., 2005; Shenkarev et al., 2006; Svangård et al., 2007). This effect on membranes could also explain the broad range of biological activities exhibited by the cyclotides: in addition to being cytotoxic, uterotonic and insecticidal, cyclotides have also been shown to be haemolytic (Schöpke et al., 1993) antifouling (Göransson et al., 2004) and anti-HIV active (Gustafson et al., 2004). Cytotoxicity measurements, in particular, can be utilised as an important tool for elucidating structure/activity relationships of the cyclotides in an effort to reveal the reasons for the diversity of cyclotide sequences observed in nature.

Cyclotides are gene products (Dutton et al., 2004; Jennings et al., 2001; Mulvenna et al., 2005; Simonsen et al., 2005) in contrast to many small circular peptide derivatives found in nature. For example peptides such as cyclosporine and bacitracin are both assembled and cyclized by non-ribosomal peptide synthetases (Walsh, 2004), and the majority of the cyclic plant peptides known today other than cyclotides are likely not gene products (Tan and Zhou, 2006). The first cyclotide-encoding cDNA clones were isolated from *O. affinis* (Jennings et al., 2001), and it was reported that cyclotides are expressed as precursor proteins. As shown in Fig. 1C, these precursor proteins start with an endoplasmic reticulum (ER) signal peptide domain followed by a pro-region, a conserved N-terminal repeat (NTR), the cyclotide domain, and end with a short hydrophobic tail. The cyclotide domain may be found in 1, 2 or 3 repeats, and a single cDNA clone may thus encode for one cyclotide sequence in up to three repeats (Jennings et al., 2001). Clones have also been found to contain two different cyclotide sequences (Jennings et al., 2001; Mulvenna et al., 2005). The post-translational processing that ultimately leads to the macrocyclic and cystine knotted protein is not yet fully understood. However the mature cyclotide sequence is most likely obtained by cleavage of precursor after a conserved Lys/Gly/Asn residue and a Asn/Asp (located in loop 6) as illustrated in Fig. 1C (Dutton et al., 2004; Jennings et al., 2001).

To date, cyclotide-encoding mRNA has been isolated from 10 species from the Violaceae plant family. This includes two species from the genus *Viola*, namely *V. tricolor* (Mulvenna et al., 2005) and *V. odorata* (Dutton et al., 2004; Ireland et al., 2006b) and eight species from the genus *Hybanthus*: *H. floribundus, H. monopetalus, H. stellarioides, H. vernonii, H. calycinus, H. debilissiumus, H. ennaeaspermus* and *H. epacroides* (Simonsen et al., 2005). For the two *Viola* species, partial cDNA clones were obtained using primers based on known cyclotide sequences. After sequencing of the partial clones, full length clones were obtained by screening specifically designed primers towards either a cDNA library (Dutton et al., 2004), as in the case for the Rubiaceae plant *O. affinis* (Jennings et al., 2001), or by rapid amplification of cDNA ends (RACE) (Mulvenna et al., 2005). These studies showed that the overall arrangement of Violaceae cDNA isolates was similar to the one from *O. affinis*. However, a conserved sequence element was identified in the ER-signal peptide of the Violaceae species (Simonsen et al., 2005). By using primers targeted to this conserved sequence (AAFALPA), cDNA clones revealing nearly full length precursors (missing only the upstream region of the ER-signal) could then be obtained from several *Hybanthus* species (Simonsen et al., 2005). Hence, the AAFALPA sequence seems conserved and widespread in the family Violaceae and could possibly provide a tool to monitor total cyclotide mRNA expression and a facile route to discovery of novel cyclotide sequences.

In this work, we examine the cyclotide expression of *V. biflora* L. (Arctic Yellow-violet, Two-flower violet), a species growing in the alpine regions of Europe and the northern parts of Asia and America. A previous study showed that this particular species displays a chemically diverse set of cyclotide sequences, as judged by liquid chromatography/mass spectrometry (LC/MS) (Göransson et al., 2003). Here we utilise both a traditional peptide isolation based approach and the AAFALPA primer sequence cDNA screening technique as methods for characterising the cyclotide content of *V. biflora*. In addition to the discovery of a number of novel and biologically active cyclotides, this study also provides valuable insights into the development of efficient screening methods for the discovery of novel cyclotide sequences in the future.

## 2. Results

We analysed the cyclotide content in *V. biflora* on both protein and mRNA level, as outlined in Fig. 2. Eleven cyclotides were isolated from an aqueous plant extract, after a first fingerprint of the protein expression was obtained by LC–MS, and their sequences were determined by tandem mass spectrometry. Eight of the isolated proteins were previously unknown; they were named vibi A–H. In parallel, we profiled the cyclotide mRNA expression using a degenerate primer targeting a conserved part in the ER region of the cyclotide precursor. In total six different cDNA clones were identified, here termed *V. biflora* cyclotide clones (*Vbc*) 1–6. The two most abundant clones encoded the previously known cyclotides cycloviolacin O9 and vitri A. One of the clones encoded vibi E and three of them coded for sequences that were not found as proteins: we named these vibi I–K. The sequences of all cyclotides identified in *V. biflora* are summarized in Table 1. The four most abundant cyclotides, vibi D, E, G and H, were tested for their cytotoxic activity in the fluorometric microculture cytotoxicity assay (FMCA) using the human lymphoma cell line U-937 GTB.

### 2.1. Protein profiling and isolation

Fresh plant material was subjected to extraction with $MeOH/H_2O$, followed by removal of chlorophyll and other

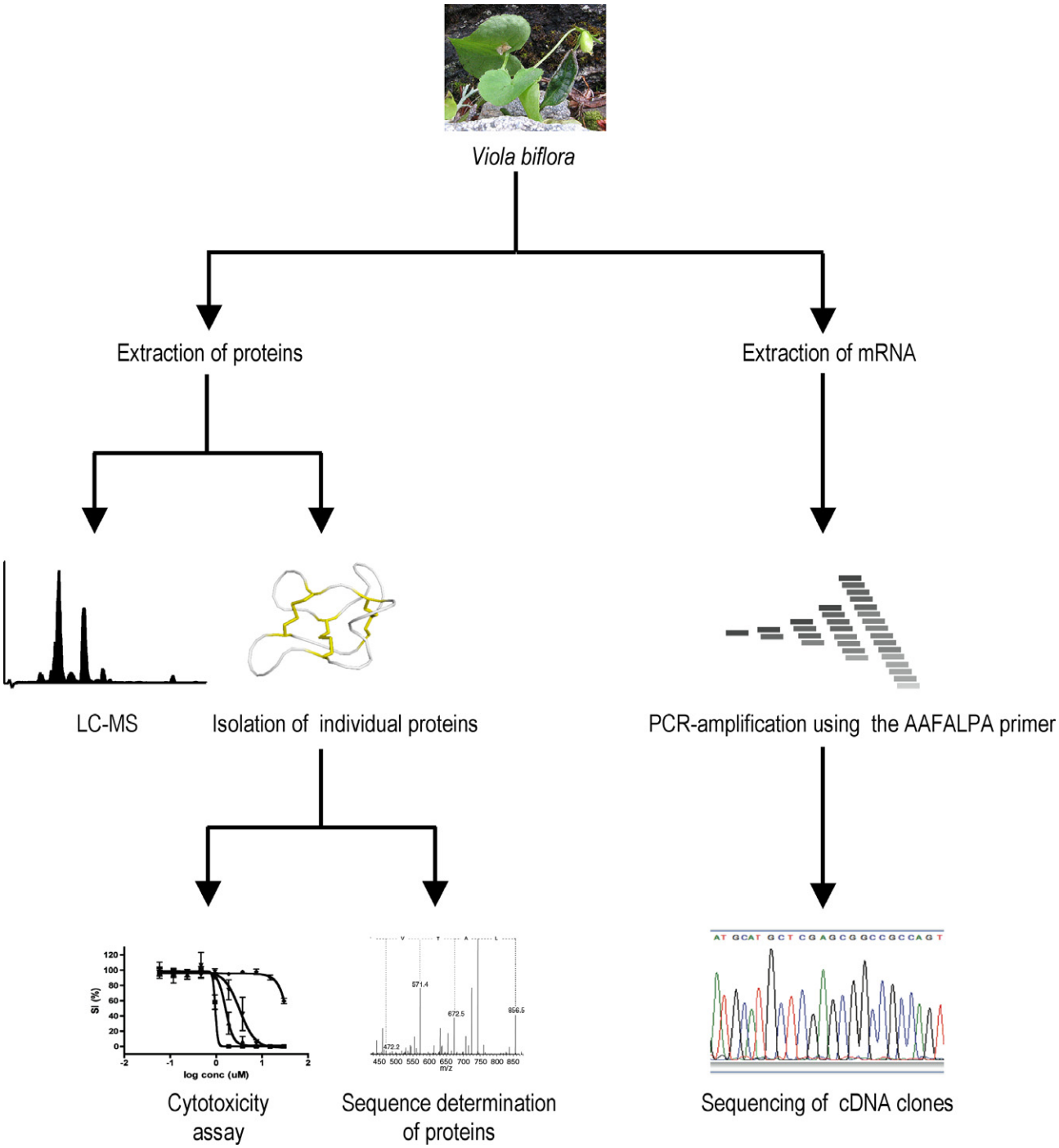


Fig. 2. Strategy for the mapping of cyclotide expression in *Viola biflora*. Proteins and mRNA were extracted from the fresh plant material as described in the Experimental section. Cyclotide content in the protein extract was revealed by LC–MS followed by isolation of individual cyclotides. Sequences were then determined using MS/MS (vibi A-H, cycloviolacin O9 and varv A), and highly abundant proteins were tested for their cytotoxic activity (vibi D, E, G and H). In parallel, cyclotide mRNA expression was probed by amplification of a cDNA library using the AAFALPA primer, which targets a conserved region in the ER-region. Six cDNA clones (*Vbc1-6*), were subsequently sequenced, which revealed the identity of the encoded proteins vibi E, I, J, K, cycloviolacin O9 and vitri A.

hydrophobic substances by liquid–liquid extraction using $CH_2Cl_2$. The aqueous, cyclotide-containing, phase was then diluted with $H_2O$ to increase the polarity of the extract before cyclotides were captured using solid phase extraction with $C_{18}$ material. This has shown to be an effective extraction method as the cyclotides are strongly retained on RP-material due to the exposed hydrophobic faces of the protein. To get an immediate overview of the cyclotide content in *V. biflora*, the purified extract was then analysed by LC–MS, as shown in Fig. 3A. Cyclotides were identified in this fingerprint by their late retention time and a molecular mass of approximately 3 kDa: in total 53 peaks with

Table 1
Alignment of cyclotide sequences from isolated and from cDNA clones predicted proteins in *Viola biflora*

| Protein | cDNA clone | Subtype | MW (Da)[a] | Sequence | Source | $IC_{50}$ (μM) |
|---|---|---|---|---|---|---|
| Vibi A | N/A | Möbius | 2916 | G-LPVCGETCFGGTCNT---PGCSCSYPICTRN | Protein | N/A |
| Vibi B[b] | N/A | Möbius | 2930 | G-LPVCGETCFGGTCNT---PGCTCSYPICTRN | Protein | N/A |
| Vibi C | N/A | Möbius | 2974 | G-LPVCGETCAFGSCYT---PGCSCSWPVCTRN | Protein | N/A |
| Vibi D | N/A | Möbius | 2985 | G-LPVCGETCFGGRCNT---PGCTCSYPICTRN | Protein | > 30 |
| Vibi E | *Vbc*1 | Bracelet | 3081 | G-IP-CAESCVWIPCTVTALIGCGCSNKVCY-N | Protein and cDNA | 3.2±0.8 |
| Vibi F | N/A | Bracelet | 3188 | GTIP-CGESCVFIPC-LTSALGCSCKSKVCYKN | Protein | N/A |
| Vibi G | N/A | Bracelet | 3222 | GTFP-CGESCVFIPC-LTSAIGCSCKSKVCYKN | Protein | 0.96±0.1 |
| Vibi H | N/A | Bracelet | 3272 | GLLP-CAESCVYIPC-LTTVIGCSCKSKVCYKN | Protein | 1.6±0.2 |
| Vibi I | *Vbc*2 | Bracelet | 3170 | G-IP-CGESCVWIPC-LTSTVGCSCKSKVCYRN | cDNA | N/A |
| Vibi J | *Vbc*3 | Bracelet | 3203 | GTFP-CGESCVWIPC-ISKVIGCACKSKVCYKN | cDNA | N/A |
| Vibi K | *Vbc*4 | Bracelet | 3424 | G-IP-CGESCVWIPC-LTSAVGCPCKSKVCYRN | cDNA | N/A |
| CyO9[c] | *Vbc*5 | Bracelet | 3140 | G-IP-CGESCVWIPC-LTSAVGCSCKSKVCYRN | Protein and cDNA | N/A |
| Vitri A | *Vbc*6 | Bracelet | 3154 | G-IP-CGESCVWIPC-ITSAIGCSCKSKVCYRN | cDNA | 0.60[e] |
| VarvA | N/A | Möbius | 2877 | G-LPVCGETCVGGTCNT---PGCSCSWPVCTRN | Protein | 5.8[f] |
| CyO2[d] | N/A | Bracelet | 3140 | G-IP-CGESCVWIPC-ISSAIGCSCKSKVCYRN | Protein | 0.75±0.1[g] |

[a]Average isotopic molecular weights. [b]The identity of residue 2 is Leu or Ile. [c]CyO9 = cycloviolacin O9. [d]CyO2 = cycloviolacin O2. [e]From (Svangård et al., 2004). [f]From (Lindholm et al., 2002). [g]From (Herrmann et al., 2006).

unique retention times and/or masses, ranging from 2828 to 3296 Da, corresponding to cyclotides were found. However, as illustrated in Fig. 3A, a smaller number of cyclotides dominated the extract and occurred in amounts large enough for their isolation and sequencing. The total content of cyclotides was calculated to be 50 μg per g of fresh plant material.

The cyclotide extract was subjected to preparative reverse phase HPLC, and collected fractions were analysed by MS. Individual cyclotides were then isolated by repetitive RP-HPLC. For that isolation, we exploited the fact that cyclotide retention times and elution order changed when shifting between $HCO_2H$ and TFA as ion pairing agents. As exemplified in Fig. 3B, some cyclotides in particular were more sensitive to the choice of acid, and that phenomena could effectively be used for the final purification. After determination of the sequences vibi A–H, we found that bracelet cyclotides were more affected by the choice of acid than the Möbius ones. In retrospect, this is not surprising as those cyclotides contain a higher number of positively charged residues, which serve the ion pair counterparts of $HCO_2H$ or TFA.

### 2.2. Protein characterization

Amino acid sequences of the isolated cyclotides were determined primarily by MS. First, the number of anionic residues in the novel cyclotides was established by esterification of carboxylic acids by treatment with acetyl chloride in MeOH. The molecular mass increased by 14 Da for all proteins, which showed that a single Glu or Asp residue was present in all sequences. Not surprisingly, this increase was later shown to be due to the conserved Glu residue in loop 1.

Cyclotides were then subjected to a combination of chemical and enzymatic reactions prior to MS/MS analysis, first to break disulfide bonds and then to cleave the circular backbone, as shown in Fig. 4. In short, cystines were reduced and *S*-carbamidomethylated. Modified proteins were purified by RP-HPLC and analysed by MS before enzymatic cleavage. The molecular masses of each cyclotide increased by 348 Da on alkylation, which is consistent with reduction of three disulfide bonds and addition of six carbamidomethyl groups. Modified proteins were digested in two separate experiments, using either trypsin or endo-

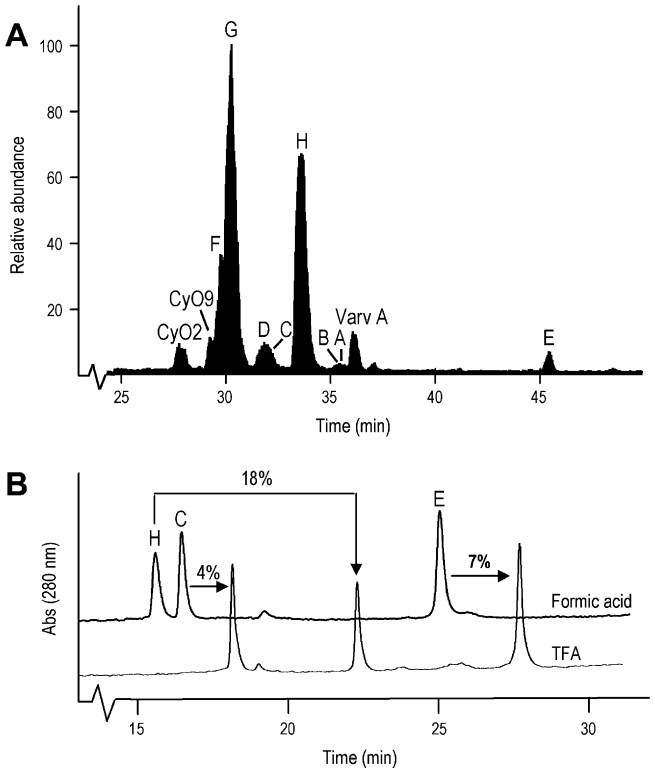


Fig. 3. Cyclotide proteins in *Viola biflora*. (A) Base peak chromatogram (*m/z* 800–1800) showing vibi A–H and three previously known cyclotides [cycloviolacin O2 (cyO2), and O9 (cyO9) and varv A]. (B) Choice of ion pairing acid changes retention times and selectivity for cyclotides in RP-HPLC. UV traces show an example for a mixture of vibi C, E and H that was separated at identical experimental conditions, but with $HCO_2H$ (above) and TFA (below). The bracelet cyclotide with net charge +2 (vibi H) is most affected, while the Möbius cyclotide with zero in net charge (vibi C) is least affected by the choice of acid. This observation was used in HPLC experiments for the final purifying of the vibi cyclotides. The percentage indicates the difference in % of eluent B when the cyclotides eluted.

proteinase GluC. Cleavage with endoproteinase GluC resulted in one single linear fragment that had a mass increase of 18 Da, as all vibi proteins contained one Glu residue. After cleavage with trypsin, one to three tryptic fragments were obtained depending on the number of cationic residues. To achieve full sequence coverage, several cyclotides had to be digested with a combination of the two enzymes, being first cleaved with trypsin and subsequently cleaved with endoproteinase GluC.

Distinguishing between amino acid residues that are isobaric, i.e., Ile/Leu and Lys/Gln, or nearly isobaric, i.e., Asn/Asp and Gln/Glu, are intrinsic problems for MS/MS sequencing. Here, we used chymotrypsin digestion to distinguish Leu from Ile; that enzyme cleaves the peptide bond C-terminal of hydrophobic residues, including Leu, but not Ile. In combination with the number of Ile and Leu as revealed by quantitative amino acid analysis the position of these residues could then be assigned by the cleavage pattern alone. As shown by the example of vibi E in Fig. 5, the results were also confirmed by MS/MS. In the case of vibi B however, sufficient amount of pure protein could not be isolated for amino acid analysis; the assignment of the Leu residue in loop 6 in that case is based solely on sequence homology.

By analogy, tryptic cleavage distinguishes Lys from Gln/Glu, as Lys is the only one of these amino acids to be cleaved by that enzyme. As described above we could then capitalize on the esterification reaction to identify acidic residues from their respective amide, i.e., Glu from Gln and Asp from Asn. Notably, Glu/Gln and Asp/Asn, respectively, cannot be told apart from the strategy used for quantitative amino acid analysis, i.e., ion exchange chromatography and post column ninhydrin detection. However, combined the experiments confirmed that there are no Gln or Asp residues present in any of the cyclotides vibi A–H.

Together, enzymatic digestions, MS/MS analyses and quantitative amino acid analyses gave complete sequences for vibi A–H [UniProt accession numbers P85239 (vibi A), P85240 (vibi B), P85241 (vibi C), P85242 (vibi D), P85243 (vibi E), P85244 (vibi F), P85245 (vibi G), and P85246 (vibi H)], which are shown in Table 1. The peaks corresponding to the respective cyclotides are marked in the LC–MS trace in Fig. 3A. As shown by that trace, *V. biflora* also contains the previously reported cyclotides varv A (Claeson et al., 1998) and cycloviolacins O2 and O9 (Craik et al., 1999). These cyclotides were identified by MS and, for varv A and cycloviolacin O2, also confirmed by co-injection of the isolated proteins together with reference material of the respective cyclotide. The results from the MS/MS experiments and the amino acid analyses are available as Supplementary data.

### 2.3. cDNA clones from *V. biflora*

Using a primer that targets a conserved region in the ER signal (AAFALPA) upstream of the mature cyclotide in 3′ RACE PCR gave products of around 600 bp. Forty-one sequences in total were analyzed and resulted in the identification of six unique clones, *Vbc1-6* [GenBank accession numbers EU046618 (*Vbc1*), EU046619 (*Vbc2*), EU046620 (*Vbc3*), EU046621 (*Vbc4*), EU046622 (*Vbc5*), and EU046623 (*Vbc6*)], which all code for bracelet cyclotide sequences. Of these 41 sequences, the majority (27) corresponded to a precursor protein (*Vbc5*) encoding the previously described peptide cycloviolacin O9 (Craik et al., 1999), nine to a precursor (*Vbc1*) of the novel peptide vibi E, two to a precursor (*Vbc6*) of vitri A, which was originally discovered in *V. tricolor* (Svangård et al., 2004) and finally three novel peptides, vibi I–K, were predicted from one single clone each, *Vbc2–4*, respectively. Notably, of the six sequenced clones, only cycloviolacin O9 and vibi E were found on the protein level, and then in relatively small amounts as shown in Fig. 3A.

The predicted amino acid sequences of *Vbc1-6* are shown in Fig. 6. They show a high degree of homology, in particular *Vbc2* and *Vbc4-6*, which have almost identical sequences both in the mature cyclotide region and in the NTR/pro-region. Compared to other Violaceae precursors

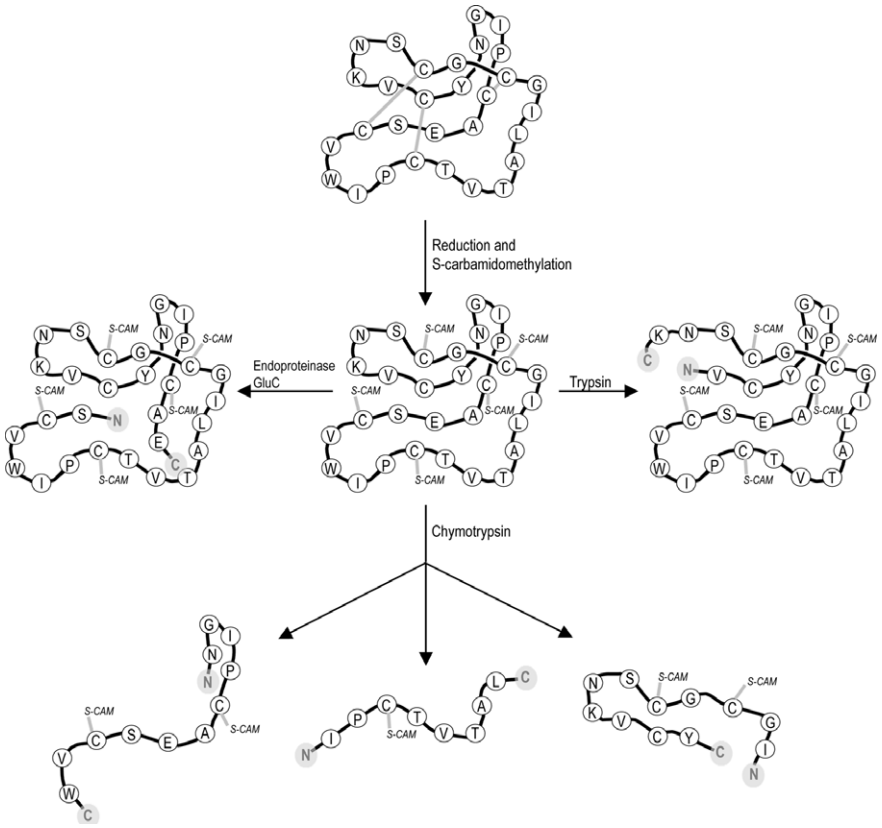


Fig. 4. Reduction, alkylation and enzymatic cleavage of vibi E. Prior to enzymatic digestion, the native protein (top) was reduced and *S*-carbamidomethylated (*S*-CAM). The reduced and Cys-protected protein (middle) was then digested using trypsin, endoproteinase GluC and chymotrypsin, in separate experiments. Since vibi E contains one Glu and one Lys residue, one single fragment was obtained after cleavage with either trypsin or endoproteinase GluC (right and left, respectively). Chymotryptic digestion yielded three fragments after cleavage at the C-terminal of Trp, Leu and Tyr residues (below). The N- and C-termini of the enzymatic digestion products are highlighted.

the overall arrangement seems identical, and the structurally important NTR region and the first parts of the pro-region show high sequence similarities. An alignment of all known Violaceae precursor NTRs and pro-regions is included as Supplementary data (sequences were obtained from (Dutton et al., 2004; Ireland et al., 2006b; Mulvenna et al., 2005)).

### 2.4. Novel sequences from *V. biflora*

Table 1 summarizes the sequences for all cyclotides that were identified on the protein and/or mRNA level. Four of the isolated cyclotide proteins, vibi A-D, belong to the Möbius subfamily: besides the six conserved Cys residues and the single Glu residue in loop 1, they all have one Pro residue in loop 5 and one Arg residue in loop 6, which all are typical sequence features of Möbius cyclotides. Two unusual sequence segments were nevertheless observed: first the Trp residue, which is conserved in all Möbius cyclotides known to date and precedes the *cis*-Pro in loop 5, is substituted with a Tyr in vibi A, B and D. Secondly, loop 2 of vibi D contains one additional Arg residue to the conserved one in loop 5, which gives this cyclotide a net charge of +1. Although this is the first cyclotide with an Arg residue at this position, five examples of cyclotides containing a positively charged Lys residue at this position are known (vodo M and N (Svangård et al., 2003) and cycloviolacin O14–O16 (Ireland et al., 2006a), all isolated from *V. odorata*).

Cyclotides vibi E–K all belong to the bracelet subfamily, and they show some noteworthy sequence segments. First vibi E, which was found expressed both on protein and mRNA level, contains only one cationic residue, a Lys located in loop 5. This makes the net charge of that cyclotide neutral, which is contrary to most bracelets. In addition vibi E has a Gly residue as the single amino acid residue in loop 4 in contrast to the vast majority of the cyclotides, which have a Ser or Thr residue at this position. Second, two of the sequences obtained from cDNA clones are conspicuous: vibi J (cDNA clone *Vbc3*) has a Lys residue located in the normally hydrophobic loop 3, and Ala instead of Ser/Thr in loop 4. Lastly, vibi K (cDNA clone *Vbc4*) is almost identical to cycloviolacin O9 (cDNA clone *Vbc5*), only differing in loop 4 where the Ser residue is replaced with a Pro which is highly unusual at this position.

### 2.5. Cytotoxicity of vibi D, E, G and H

The cytotoxic effect of the four most abundant cyclotides, vibi D, E, G and H, was tested in the FMCA using

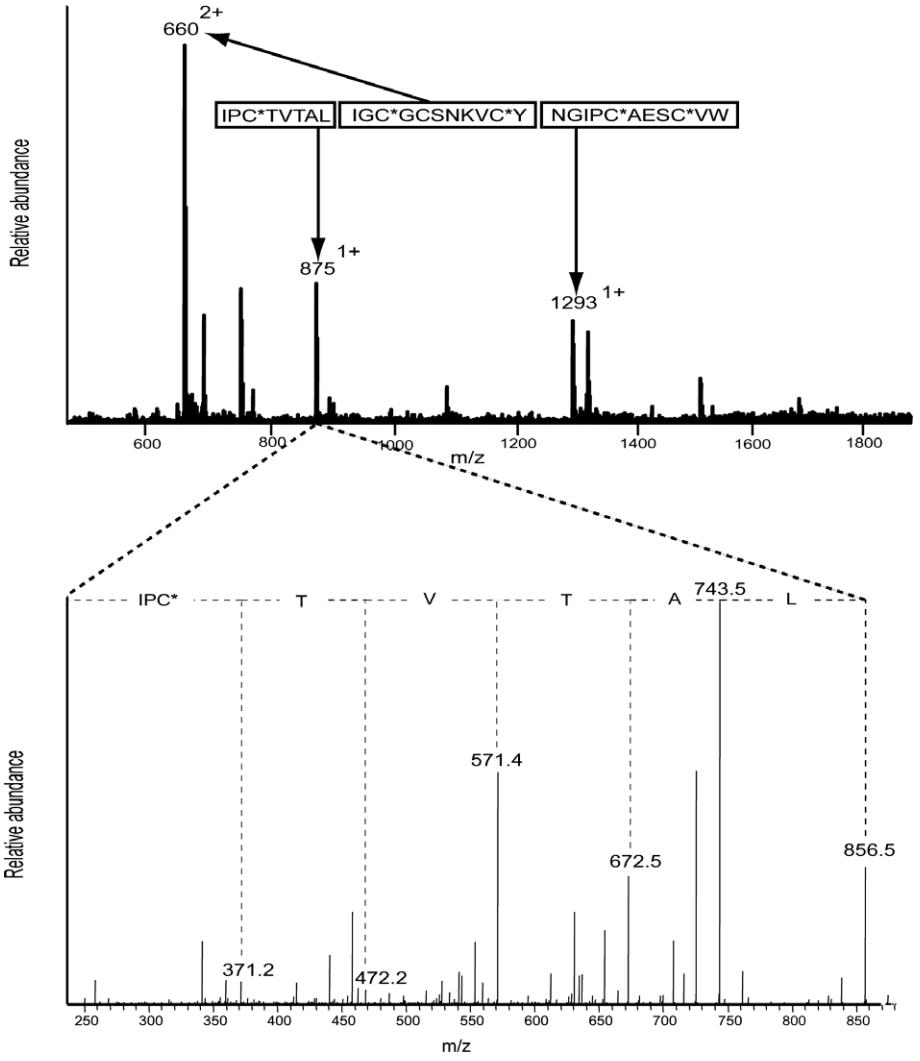


Fig. 5. Determination of Ile/Leu residues in vibi E. Nanospray MS of the chymotryptic digest was used to determine Ile/Leu positions. Major fragments were sequenced, as shown at the top, by MS/MS. The latter is exemplified by the identification of the b-series of the fragment IPCTVTAL in the spectrum at the bottom. Ile and Leu content in the sequence was confirmed by quantitative amino acid analyses and, when possible, by sequencing of cDNA clones. C* represents *S*-carbamidomethylated cysteine residues.

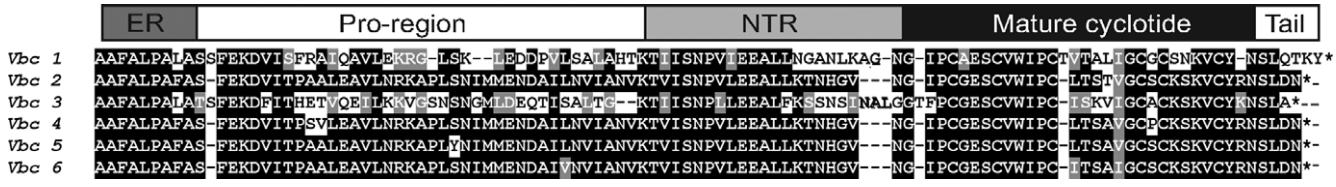


Fig. 6. Deduced amino acid sequences of cyclotide precursors identified in *Viola biflora*. *Vbc1-6* show a high degree of sequence homology and have the overall arrangement in common with previously known precursors; an ER signal region followed by a pro-region, an N-terminal repeat (NTR), the mature cyclotide and a short C-terminal, as shown at the top.

the human lymphoma cell line U-937 GTB. The dose–response curves and $IC_{50}$ values of the tested cyclotides are shown in Fig. 7.

The two most potent cyclotides, vibi G and H, are both bracelet cyclotides, and although they differ in 6 out of their 31 amino acids they have similar $IC_{50}$-values, 0.96 and 1.6 μM, respectively. These values are essentially comparable to those for the prototypic bracelet cyclotide cycloviolacin O2, and vitri A, as shown in Table 1. The degree of homology between all these cyclotides is similar to that between vibi G and H and a close examination of the sequences reveals that substitutions have not significantly changed the amino acid polarity. For example vibi G and H differ from cycloviolacin O2 and most other bracelets by having Phe and Tyr, respectively, instead of Trp in loop 2, and Lys instead of Arg in loop 6. Accord-

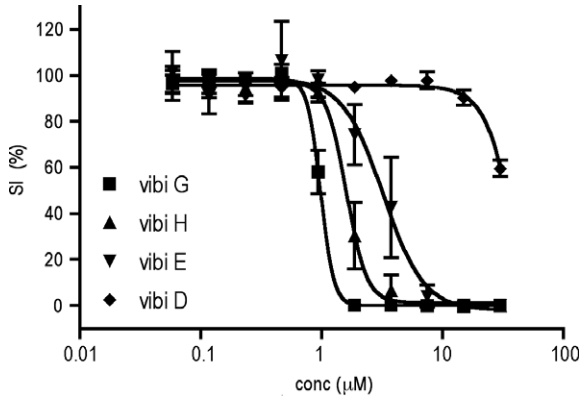


Fig. 7. Effect on the human lymphoma cell line U-937 GTB. Dose-response curves for, from left to right, vibi G (squares, $IC_{50}$-value 0.96 μM), vibi H (upward-pointing triangles, 1.6 μM), vibi E (downward-pointing triangles, 3.2 μM) and vibi D (diamonds, >30 μM). The error bars represent standard mean error.

ingly one would expect the surface properties of these cyclotides to be similar; hence it might not be surprising that they have similar cytotoxic potency.

Vibi E belongs to the same subfamily, but shows some anomalies compared to the typical bracelet cyclotide. For example it contains only one cationic residue, a Lys in loop 5, which gives vibi E a neutral net charge instead of +2 as for the bracelet cyclotides described above. This difference in charge is reflected in the activity of vibi E as its $IC_{50}$ is threefold less than that of vibi G.

Vibi D was the only Möbius cyclotide that could be isolated in sufficient amount for testing. With an $IC_{50}$ exceeding 30 μM it is the least potent cyclotide tested in the assay to date; compared to other Möbius cyclotides it is 5–8 times less potent (Lindholm et al., 2002). Vibi D is distinguished from the other Möbius cyclotides previously tested by two main features, it contains an additional Arg residue in loop 2, and the Trp residue normally found in loop 5 in Möbius cyclotides is replaced by Tyr. Plausibly, the Trp/Tyr substitution plays a minor part for the decreased activity (by analogy with the results obtained for the bracelets as described above) suggesting that it is the presence of the Arg in loop 2 that is the major contributing factor to the lower cytotoxicity of vibi D.

## 3. Discussion

In this study, we used a strategy involving protein profiling by LC–MS, sequencing of isolated proteins by tandem mass spectrometry and screening for cyclotide mRNA to discover novel cyclotides in *V. biflora*. This resulted in the identification of 15 cyclotide sequences, 11 as proteins and six as expressed mRNA, but only two of them were found at both mRNA and protein levels. Furthermore 11 of the sequences were previously unknown, vibi A to K, while four had been previously characterized on the protein level from other *Viola* species, namely varv A from *V. arvensis* (Claeson et al., 1998), vitri A from *V. tricolor* (Svangård et al., 2004), and cycloviolacin O2 and O9 from *V. odorata* (Craik et al., 1999). Thus like previously examined cyclotide-containing plants, *V. biflora* expresses a cocktail of different cyclotides of which some are common to other plants while others seem to represent a signature set specific for that species. Highly abundant cyclotides were then tested for cytotoxic activity to reveal if structural diversity is reflected in biological diversity.

The cyclotide protein expression of *V. biflora* was first profiled by LC–MS, which offers a quick and easily interpreted overview of the qualitative and quantitative cyclotide content in the plant using characteristic molecular masses and retention times. The profile showed a small set of major cyclotide peaks, representing the most abundant cyclotides, accompanied by a large number of minor peaks. This is a typical pattern for cyclotide-containing plant species from the Violaceae family (Chen et al., 2005; Göransson et al., 2003, 1999; Ireland et al., 2006a; Simonsen et al., 2005; Trabi and Craik, 2004). Whether all of these peaks represents native cyclotide structures, or oxidised or linear cyclotide products as recently discovered (Chen et al., 2005; Ireland et al., 2006a; Plan et al., 2007; Simonsen et al., 2005) or even non-cyclotide proteins remains to be shown.

In parallel, we probed *V. biflora* cyclotide mRNA expression by using a forward primer directed to a conserved sequence element in the ER-signal of the cyclotide precursor, as outlined by the work by Simonsen et al. (2005) and highlighted in Fig. 1C. In 3′ RACE PCR, this primer yields the full cyclotide sequence as well the majority of the precursor protein. Using this strategy, we hoped to see mRNA expression that more or less matched the protein profile. On the contrary, no clones encoding highly abundant proteins were found with only two of the six clones matching any of the characterised proteins. The most readily amplified cDNA clone, *Vbc5*, encoded the cycloviolacin O9 precursor and Vbc1 encoded the vibi E precursor. Both of these cyclotides were found in relatively low abundance in the plant. This is a surprising fact, in particular for cycloviolacin O9, which was the predominant clone isolated on the mRNA level, and reveals a deficiency in utilising this cDNA screening approach as the sole experimental strategy for the discovery of novel cyclotide sequences.

A number of possibilities could possibly bias the experimental results, of which we consider plant collection, RNA stability and structure, and choice of primers the major ones. The RNA originates from the same population and time of collection as was used for protein profiling. In addition RNA samples were also taken from a number of different individual plants, which all were immediately placed in the RNA conserving solution and then immediately frozen on dry ice. Possibly mRNA stability might also vary for different cyclotide precursors, or some precursors might also have structures that are hard to amplify (e.g. hairpin structures). Regarding the choice of primers, we

screened *V. biflora* cDNA using the degenerative AAFALPA-primer, which has been shown to target the ER region in several Violaceae species. In those studies however no attempt was made to match the full mRNA profile to the one at the protein level: hence ambiguities between the profiles would not be revealed. From the results in the current study we cannot confirm that all *V. biflora* cyclotide precursors contain codons that are targeted by the AAFALPA primer, or if the missing precursors encoding for cyclotide proteins are just not expressed at the time of collection. It is interesting to note that the majority of cyclotides found using the protein isolation and MS sequencing approach, including the most abundant ones, could not be found at all on a mRNA level using the AAFALPA primer. This may be due to fact that the mRNA that encodes the precursors for the major cyclotides found in the plant extract have a much lower affinity for the primer than those that were found, particularly *Vbc5* (cycloviolacin O9) which was encoded by 27 of the 40 clones sequenced. Therefore, although the AAFALPA is a useful tool for discovery of new cyclotide sequences from Violaceae species, as it is not biased towards known cyclotide sequences, it is clearly not sufficient for a comprehensive screen of potential cyclotides within a plant.

Each screening method, protein or mRNA, targets a unique subset of cyclotides produced by the plant. By using a RNA screening approach the cyclotides being expressed at the time of plant collection will be those that are discovered, while it is likely that cyclotide proteins can be stored for extensive time in the plant. A complicating factor is that cyclotide protein expression varies with season, geographical distribution and in different tissue types (Göransson et al., 2003; Simonsen et al., 2005; Trabi and Craik, 2004; Trabi et al., 2004). Therefore, even if an intensive RNA screening process using multiple primers is undertaken, the cyclotide producing potential of the plant will not be fully revealed. This will also be the case when using a protein screening method but the bias will be introduced for different reasons. Firstly, the protein screening approach is limited by sensitivity, as a certain amount of protein is required to fully elucidate the sequence of the cyclotide. Secondly, the cyclotide protein pool can be influenced by a large range of factors between translation of the mRNA and formation of the final peptide. These factors may include differences in translation of mRNA transcripts, differences in the stability of the different precursor proteins, more efficient processing of particular precursor proteins and stability of the mature cyclotide. Due to the limitations of both methods clearly an integrated approach that utilise both protein isolation and RNA screening is optimal.

Then the main questions are what is it that triggers the expression, and what is the role for the cocktail of cyclotides that is found in the plant. The answers to those questions are yet unknown. It is however tempting to speculate that various external events trigger a plant to express a specific cyclotide or a set of cyclotides, which enables an optimal effect on a specific target. In particular this theory seems to go hand in hand with their proposed function in host defence, as indicated by their potent insecticidal activity (Jennings et al., 2001, 2005) but also by their antimicrobial effects (Tam et al., 1999). Then the trigger could be bacterial infections or grazing insects. Indeed a recent study showed this to be the case in for a cyclotide-like sequence in maize (*Zea mays*); the amount of mRNA was significantly increased after infection of the plant with smut fungus (*Ustilago maydis*) (Basse, 2005).

In this conjecture it is interesting to note that the cyclotide content of *V. biflora* has been profiled before by LC–MS, but then on plant material collected in Switzerland (Göransson et al., 2003). Interestingly, the LC–MS profile of the Switzerland sample showed some differences to the profile from the sample from Sweden used in this study: some molecular masses observed from the plant material collected in Switzerland were not observed in the Swedish plant material and *vice versa*. In the current study, we examined plant material collected at the same location and same time of year (June) but both in 2005 and 2006. The cyclotide-profiles for each collection were practically identical with the most abundant cyclotides and the relative abundance between the cyclotides being the same and no changes seen for the minor molecular masses either. This demonstrates the stability of the protein expression as seen for these collections from different years.

It is clear that cyclotides exhibit a broad range of biological effects, as demonstrated by their activity in a number of different bioassays. Most of these effects appear to be connected with the same mechanism of action, membrane disruption. This includes their insecticidal (Jennings et al., 2001, 2005), haemolytic (Schöpke et al., 1993), and cytotoxic effects (Svangård et al., 2007) of which we used the latter to probe the structure activity relationship of high abundant cyclotides in the current study. The results show that cyclotides are tolerable to substitutions such as Trp/Tyr and Arg/Lys with sustained effect. In addition, the result obtained for vibi E supports our previous study that showed that a lower content of positively charged residues only marginally influences potency (Herrmann et al., 2006). From that study and others it is now becoming clear that the number and position of charged residues is not the sole key for explaining cyclotide activity, but rather the amphipatic properties of the protein (Kamimori et al., 2005; Nourse et al., 2004; Shenkarev et al., 2006). This hypothesis is also supported by the decreased activity of the Möbius cyclotide vibi D, which has an Arg residue in loop 2. The residues in loop 2 comprise part of the hydrophobic patch that is thought to play a critical role in the inherent biological activity of the Möbius cyclotides (Clark et al., 2006). Therefore, it is likely that the Arg residue in loop 2 of vibi D disrupts the hydrophobic surface of this cyclotide, which in turn leads to decreased amphipatic properties and lower potency.

The total number of known cyclotide sequences is now close to 100. This is a significant number but, as shown by the results in the current study, each Violaceae species

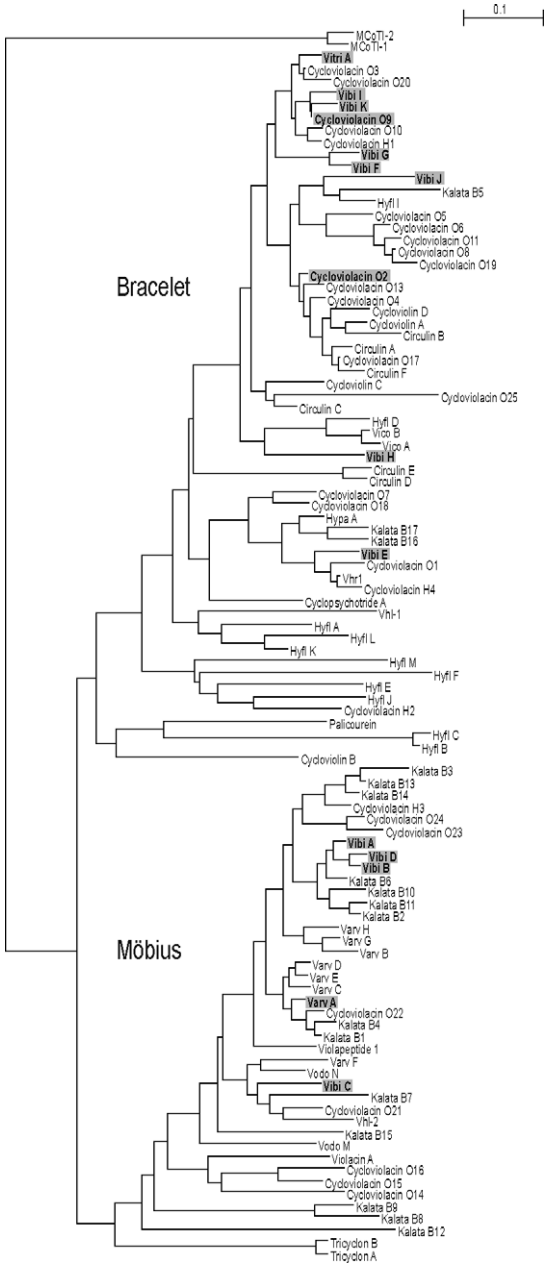


Fig. 8. Cluster analysis of the cyclotide protein family. The analysis comprises all known complete cyclotide sequences. Sequence alignment was made using ClustalW (Thompson et al., 1994) and the tree was made by neighbour joining. Cyclotides identified in *Viola biflora*, i.e. vibi A–K, vitri A, varv A and cycloviolacin O2 and O9, are highlighted. Note the separation into two branches corresponding to the bracelet and Möbius subfamilies. The hybrid cyclotides kalata B8–B9 and tricyclon A–B, with strong Möbius sequence similarity in loop 3, have been placed together with the Möbius cyclotides. MCoTI I–II are placed outside both subfamilies; these proteins have not much more in common with the bracelets or the Möbius cyclotides than a cyclic backbone and six cysteines.

that is investigated has so far yielded novel sequences. As discussed earlier, some of the vibi-cyclotides show only small changes in sequence compared to previously known cyclotides, while some are set apart by atypical sequences. The un-rooted tree in Fig. 8 visualizes this varying degree of sequence similarity, and puts the vibi cyclotides in the context of the whole protein family. Notably, the Möbius vibi A, B and D cluster together because of the unique Tyr residue in loop 5, and vibi F and G form a cluster in the bracelets. As a group, however, there seems to be no common denominator characteristic for cyclotides from *V. biflora* as they are scattered throughout the entire tree. This is also the case for the other two species for which cyclotide content has been scrutinised: cycloviolacins from *V. odorata* (Craik et al., 1999; Ireland et al., 2006a) and kalata cyclotides (Plan et al., 2007) are also found throughout the tree. It is tempting to assume that there exists a set of common cyclotides (here represented by cycloviolacin O2 and varv A), which is complemented with species-specific sequences as a result of the evolution of individual species. Whether this is the case however, or if this current trend is merely a result of the limited set of yet investigated species remains to be shown.

To conclude, the results in the current study highlight the need for a wider screening of cyclotides and cyclotide-bearing plants to increase our knowledge about sequence diversity and plasticity of the cyclotide framework, and its structure activity relationships. Screening a plant with an upstream primer of the mature cyclotide domain, like the primer (AAFALPA) is in this way a good complementary method to normal extraction of cyclotides from a plant. In this way, complete sequences that are rare on protein level and probably not achievable by other means can be accessed. Our results indicate that if the overall aim is to establish all the cyclotide sequences that theoretically can be expressed from genes in a plant, both protein and mRNA isolation should be applied in combination.

## 4. Experimental

### 4.1. Plant material

Aerial parts of *V. biflora* L. (Violaceae) were collected in Jämtland, Sweden in June 2005 and June 2006. A voucher specimen is deposited at the Uppsala University Herbarium (UPS). Fresh plant material was stored at −20 °C until use. For the preservation of mRNA, a fraction of the plant material was stored in RNAlater (Ambion) at −20 °C until use.

### 4.2. RNA extraction and production of partial clones

Extraction of RNA from the leaves was done using the RNAqueous kit (Qiagen). From total RNA, cDNA was prepared using Superscript III reverse transcriptase (Invitrogen). *Vbc1–6* (*V. biflora* cyclotide encoding clones) were amplified using oligo(dT) with a degenerative forward primer (5′-GGG CHG CHT TYG CHC TTC CHG C-3′) that encodes the protein sequence AAFALPA, which was first described by (Simonsen et al., 2005). The resulting bands

were gel-purified and cloned by ligation into TOPO (Invitrogen) or pGEM-T Easy (Promega).

### 4.3. HPLC

An ÄKTA basic HPLC system (Amersham Biosciences, Uppsala, Sweden) was used for all HPLC experiments. Cyclotides were detected at 215, 254, and 280 nm, and peaks were integrated at 280 nm. A ReproSil-Pur C18-AQ column (250 × 20 mm i.d., 10 μm, 300 Å) was used for preparative HPLC, using a linear gradient from 10% MeCN in 0.05% TFA (buffer A) to 60% MeCN in 0.045% TFA (buffer B) over 45 min operated at a flow rate of 5 ml/min. For LC–MS experiments, a Grace Vydac Everest $C_{18}$ column (100 × 2.1 mm i.d., 5 μm, 300 Å) was used using a linear gradient from 10% MeCN in 0.05% $HCO_2H$ (buffer C) to 60% MeCN in 0.045% $HCO_2H$ (buffer D) over 80 min operated at a flow rate of 0.3 ml/min. A Grace Vydac Everest Analytical $C_{18}$ column (250 × 4.6 mm i.d., 5 μm, 300 Å) was used for purification and analysis of the cyclotides, using a linear gradient from buffer A to buffer B or buffer C to D over 40 min operated at a flow rate of 1 ml/min.

### 4.4. Mass spectrometry

For ESI-MS a Finnigan LCQ ion trap (Thermo Electron Co., Waltham, MA, USA) mass spectrometer was used in positive ion mode. The capillary temperature was set at 220 °C and the spray voltage at 4 kV. For nanospray MS, a Protana NanoES source (Proxeon Engineering, Odense, Denmark) was mounted on the same instrument. The capillary temperature was set at 150 °C, and the spray voltage at 0.5 kV. All samples were sprayed in 50% MeCN and 1% $HCO_2H$. For MS/MS sequencing, the CID was varied between 25% and 45%. Average isotopic masses were used for all calculated molecular weights.

### 4.5. Quantitative amino acid analysis

The amino acid content of vibi A and C–H was determined at the Amino acid Analyses Centre, Department of Biochemistry and Organic Chemistry, Uppsala University. The proteins were hydrolyzed for 24 h at 110 °C with 6 N HCl containing 2 mg/ml PhOH, and the hydrolysates were analyzed with an LKB model 4151 Alpha Plus amino acid analyzer using ninhydrin detection.

### 4.6. Isolation of cyclotides

For the isolation procedure, 300 g of plant material was subjected to three repeated extractions with 4 l of 60% aqueous MeOH. The extract was filtrated and subjected to liquid–liquid extraction with $CH_2Cl_2$ (2:1). The aqueous phase was diluted one time with $H_2O$ and the extract was loaded on silica gel ($C_{18}$), washed with 30% MeOH (aq) and finally eluted with MeOH. The MeOH phase was taken to dryness *in vacuo* and subjected to preparative HPLC. The cyclotides were purified with HPLC and obtained as a white powder after freeze-drying.

### 4.7. Reduction, alkylation and enzymatic cleavage of cyclotides

Prior to MS/MS, the proteins were reduced with DTE in 0.25 M Tris–HCl containing 1 mM EDTA and 6 M guanidine-HCl (pH 8.5) for 2 h at 37 °C under darkness and $N_2$ atmosphere. The free thiols of the protein were subsequently *S*-carbamidomethylated by adding IAM to the solution and reacting for 60 min at room temperature. The *S*-carbamidomethylated proteins were purified using HPLC and digested using trypsin, endoproteinase GluC or chymotrypsin in 50 mM $NH_4HCO_3$ for at least 4 h at 37 °C.

### 4.8. Esterification of Glu residues

For esterification of Glu residues, 0.1 ml (5 mmol) acetyl chloride was slowly added to 0.6 ml dry MeOH (Hunt et al., 1986). The mixture was stirred at room temperature for 5 min. Of this solution, 25 μl was added to 10 μg dry protein. The reaction was carried out at room temperature for 1 h. The reaction mixture was then diluted with 50 μl $H_2O$ and directly purified by HPLC.

### 4.9. Human cancer cell line assay

The human lymphoma cell line U-937 GTB (Sundström and Nilsson, 1976) was used in the fluorometric microculture cytotoxicity assay (FMCA) as described by Larsson and Nygren (1989). The cell line was procured and maintained as described earlier (Dhar et al., 1996). Vibi D, E, G and H were dissolved in 10% EtOH (aq) (yielding final concentrations of 1% EtOH in the assay) to a concentration of 300 μM. Dilution series with 10% EtOH (aq) were then made at a 1:1 ratio from these stock solutions. V-shaped, 96-well microtiter plates (Nunc, Roskilde, Denmark) were prepared using 20 μl of protein test solution per well in duplicates for each concentration. Also, six blank wells (200 μl per well of cell-growth medium), six negative-control wells (20 μl per well of PBS), and six solvent-control wells (20 μl per well of 10% EtOH) were prepared on each microtiter plate. Tumor cells suspended in cell-growth medium were dispensed on the microtiter plates (20,000 cells/180 μl per well) which then were incubated for 72 h at 37 °C and 5% $CO_2$, and fluorescence was measured in a 96-well scanning fluorometer at 538 nm, following excitation at 485 nm. The fluorescence is proportional to the number of living cells and cell survival is quantified as a survival index (SI) expressed in percent. The SI is defined as the fluorescence of the test wells relative to the average fluorescence of control wells (PBS), with the average for the blank wells subtracted. $IC_{50}$ values, which correspond to the concentration at an SI of 50%, were

calculated using non-linear regression in GraphPad Prism 4 (GraphPad Software, Inc., San Diego, CA, USA). The solvent used did not affect cell growth at final concentration (i.e. 1% EtOH). All experiments were performed three times.

## Acknowledgements

We thank Tala Jalilian for help with the isolation of cyclotide proteins. UG has been supported by a grant from the Disciplinary Domain of Medicine and Pharmacy, Uppsala University. DJC is an ARC Professorial Fellow and work in his lab on cyclotides is supported by Grants from the Australian Research Council and the National Health and Medical Research Council of Australia.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.phytochem.2007.10.023.

## References

Basse, C.W., 2005. Dissecting defense-related and developmental transcriptional responses of maize during *Ustilago maydis* infection and subsequent tumor formation. Plant Physiol. 138, 1774–1784.

Chen, B., Colgrave, M.L., Daly, N.L., Rosengren, K.J., Gustafson, K.R., Craik, D.J., 2005. Isolation and characterization of novel cyclotides from *Viola hederaceae*: solution structure and anti-HIV activity of vhl-1, a leaf-specific expressed cyclotide. J. Biol. Chem. 280, 22395–22405.

Claeson, P., Göransson, U., Johansson, S., Luijendijk, T., Bohlin, L., 1998. Fractionation protocol for the isolation of polypeptides from plant biomass. J. Nat. Prod. 61, 77–81.

Clark, R.J., Daly, N.L., Craik, D.J., 2006. Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design. Biochem. J. 394, 85–93.

Colgrave, M.L., Craik, D.J., 2004. Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. Biochemistry 43, 5965–5975.

Craik, D.J., Cemazar, M., Wang, C.K., Daly, N.L., 2006. The cyclotide family of circular miniproteins: nature's combinatorial peptide template. Biopolymers 84, 250–266.

Craik, D.J., Daly, N.L., Bond, T., Waine, C., 1999. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. J. Mol. Biol. 294, 1327–1336.

Dhar, S., Nygren, P., Csoka, K., Botling, J., Nilsson, K., Larsson, R., Huang, H.W., 1996. Anti-cancer drug characterisation using a human cell line panel representing defined types of drug resistance. Br. J. Cancer 74, 888–896.

Dutton, J.L., Renda, R.F., Waine, C., Clark, R.J., Daly, N.L., Jennings, C.V., Anderson, M.A., Craik, D.J., 2004. Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteins. J. Biol. Chem. 279, 46858–46867.

Gran, L., 1973. Oxytocic principles of *Oldenlandia affinis*. Lloydia 36, 174–178.

Gustafson, K.R., McKee, T.C., Bokesch, H.R., 2004. Anti-HIV cyclotides. Curr. Protein Pept. Sci. 5, 331–340.

Gustafson, K.R., Sowder II, R.C., Henderson, L.E., Parsons, I.C., Kashman, Y., Cardellina II, J.H., McMahon, J.B., Buckheit, R.W., Pannell, Boyd, M.R., 1994. Circulins A and B: novel HIV-inhibitory macrocyclic peptides from the tropical tree *Chassalia parvifolia*. J. Am. Chem. Soc. 116, 9337–9338.

Göransson, U., Broussalis, A.M., Claeson, P., 2003. Expression of *Viola* cyclotides by liquid chromatography-mass spectrometry and tandem mass spectrometry sequencing of intercysteine loops after introduction of charges and cleavage sites by aminoethylation. Anal. Biochem. 318, 107–117.

Göransson, U., Luijendijk, T., Johansson, S., Bohlin, L., Claeson, P., 1999. Seven novel macrocyclic polypeptides from *Viola arvensis*. J. Nat. Prod. 62, 283–286.

Göransson, U., Sjögren, M., Svangård, E., Claeson, P., Bohlin, L., 2004. Reversible antifouling effect of the cyclotide cycloviolacin O2 against barnacles. J. Nat. Prod. 67, 1287–1290.

Hallock, Y.F., Sowder, R.C., Pannell, L.K., Hughes, C.B., Johnson, D.G., Gulakowski, R., Cardellina, J.H., Boyd, M.R., 2000. Cycloviolins A–D, anti-HIV macrocyclic peptides from *Leonia cymosa*. J. Org. Chem. 65, 124–128.

Herrmann, A., Svangård, E., Claeson, P., Gullbo, J., Bohlin, L., Göransson, U., 2006. Key role of glutamic acid for the cytotoxic activity of the cyclotide cycloviolacin O2. Cell. Mol. Life. Sci. 63, 235–245.

Hunt, D.F., Yates 3rd, J.R., Shabanowitz, J., Winston, S., Hauer, C.R., 1986. Protein sequencing by tandem mass spectrometry. Proc. Natl. Acad. Sci. USA 83, 6233–6237.

Ireland, D.C., Colgrave, M.L., Craik, D.J., 2006a. A novel suite of cyclotides from *Viola odorata*: sequence variation and the implications for structure, function and stability. Biochem. J. 400, 1–12.

Ireland, D.C., Colgrave, M.L., Nguyencong, P., Daly, N.L., Craik, D.J., 2006b. Discovery and characterization of a linear cyclotide from *Viola odorata*: implications for the processing of circular proteins. J. Mol. Biol. 357, 1522–1535.

Jennings, C., West, J., Waine, C., Craik, D.J., Anderson, M., 2001. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. Proc. Natl. Acad. Sci. USA 98, 10614–10619.

Jennings, C.V., Rosengren, K.J., Daly, N.L., Plan, M., Stevens, J., Scanlon, M.J., Waine, C., Norman, D.G., Anderson, M.A., Craik, D.J., 2005. Isolation, solution structure, and insecticidal activity of kalata B2, a circular protein with a twist: do Möbius strips exist in nature? Biochemistry 44, 851–860.

Kamimori, H., Hall, K., Craik, D.J., Aguilar, M.I., 2005. Studies on the membrane interactions of the cyclotides kalata B1 and kalata B6 on model membrane systems by surface plasmon resonance. Anal. Biochem. 337, 149–153.

Larsson, R., Nygren, P., 1989. A rapid fluorometric method for semiautomated determination of cytotoxicity and cellular proliferation of human tumor cell lines in microculture. Anticancer Res. 9, 1111–1119.

Lindholm, P., Göransson, U., Johansson, S., Claeson, P., Gullbo, J., Larsson, R., Bohlin, L., Backlund, A., 2002. Cyclotides – a novel type of cytotoxic agents. Mol. Cancer Therap. 1, 365–369.

Mulvenna, J.P., Sando, L., Craik, D.J., 2005. Processing of a 22 kDa precursor protein to produce the circular protein tricyclon A. Structure 13, 691–701.

Mulvenna, J.P., Wang, C., Craik, D.J., 2006. CyBase: a database of cyclic protein sequence and structure. Nucleic Acids Res. 34, D192–D194.

Nourse, A., Trabi, M., Daly, N.L., Craik, D.J., 2004. A comparison of the self-association behavior of the plant cyclotides kalata B1 and kalata B2 via analytical ultracentrifugation. J. Biol. Chem. 279, 562–570.

Plan, M.R., Göransson, U., Clark, R.J., Daly, N.L., Colgrave, M.L., Craik, D.J., 2007. The cyclotide fingerprint in *Oldenlandia affinis*: elucidation of chemically modified, linear and novel macrocyclic peptides. Chem. Biol. Chem. 8, 1001–1011.

Rosengren, K.J., Daly, N.L., Plan, M.R., Waine, C., Craik, D.J., 2003. Twists, knots, and rings in proteins. Structural definition of the cyclotide framework. J. Biol. Chem. 278, 8606–8616.

Saether, O., Craik, D.J., Campbell, I.D., Sletten, K., Juul, J., Norman, D.G., 1995. Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. Biochemistry 34, 4147–4158.

Schöpke, T., Agha, M.I.H., Kraft, R., Otto, A., Hiller, K., 1993. Hämolytisch aktive komponenten aus *Viola tricolor* L. und *Viola arvensis* Murray. Sci. Pharm. 61, 145–153.

Shenkarev, Z.O., Nadezhdin, K.D., Sobol, V.A., Sobol, A.G., Skjeldal, L., Arseniev, A.S., 2006. Conformation and mode of membrane interaction in cyclotides. Spatial structure of kalata B1 bound to a dodecylphosphocholine micelle. FEBS J. 273, 2658–2672.

Simonsen, S.M., Sando, L., Ireland, D.C., Colgrave, M.L., Bharathi, R., Göransson, U., Craik, D.J., 2005. A continent of plant defense peptide diversity: cyclotides in Australian *Hybanthus* (Violaceae). Plant Cell 17, 3176–3189.

Sundström, C., Nilsson, K., 1976. Establishment and characterization of a human histiocytic lymphoma cell line (U-937). Int. J. Cancer 17, 565–577.

Svangård, E., Burman, R., Gunasekera, S., Lövborg, H., Gullbo, J., Göransson, U., 2007. Mechanism of action of cytotoxic cyclotides: cycloviolacin O2 disrupts lipid membranes. J. Nat. Prod. 70, 643–647.

Svangård, E., Göransson, U., Hocaoglu, Z., Gullbo, J., Larsson, R., Claeson, P., Bohlin, L., 2004. Cytotoxic cyclotides from *Viola tricolor*. J. Nat. Prod. 67, 144–147.

Svangård, E., Göransson, U., Smith, D., Verma, C., Backlund, A., Bohlin, L., Claeson, P., 2003. Primary and 3-D modelled structures of two cyclotides from *Viola odorata*. Phytochemistry 64, 135–142.

Tam, J.P., Lu, Y.A., Yang, J.L., Chiu, K.W., 1999. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. Proc. Natl. Acad. Sci. USA 96, 8913–8918.

Tan, N.H., Zhou, J., 2006. Plant cyclopeptides. Chem. Rev. 106, 840–895.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Trabi, M., Craik, D.J., 2004. Tissue-specific expression of head-to-tail cyclized miniproteins in Violaceae and structure determination of the root cyclotide *Viola hederacea* root cyclotide. Plant Cell 16, 2204–2216.

Trabi, M., Svangård, E., Herrmann, A., Göransson, U., Claeson, P., Craik, D.J., Bohlin, L., 2004. Variations in cyclotide expression in *Viola* species. J. Nat. Prod. 67, 806–810.

Walsh, C.T., 2004. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science 303, 1805–1810.

Witherup, K.M., Bogusky, M.J., Anderson, P.S., Ramjit, H., Ransom, R.W., Wood, T., Sardana, M., 1994. Cyclopsychotride A, a biologically active, 31-residue cyclic peptide isolated from *Psychotria longipes*. J. Nat. Prod. 57, 1619–1625.