Discovery, isolation and structural characterization of cyclotides from *Viola sumatrana* Miq.

Ploypat Niyomploy,[1,2] Lai Yue Chan,[1] Aaron G. Poth,[1] Michelle L. Colgrave,[3]

Polkit Sangvanich[4] and David J. Craik[1]

[1]Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia

[2]Program in Biotechnology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

[3]CSIRO Agriculture, 306 Carmody Rd, St Lucia, QLD 4067, Australia

[4]Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

AUTHOR INFORMATION

Please address correspondence to: David J. Craik, Institute for Molecular Bioscience, The University of Queensland, Brisbane 4072; E-mail: d.craik@imb.uq.edu.au

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/bip.22914

ABSTRACT

Cyclotides are bioactive cyclic peptides that have been discovered in plants from the Violaceae, Rubiaceae, Fabaceae, Cucurbitaceae and Solanaceae families. They are sparsely distributed in most of these families, but appear to be ubiquitous in the Violaceae, having been found in every plant so far screened from this family. However, not all geographic regions have been examined so far and this is the first study to report on the discovery of cyclotides from a *Viola* species from South-East Asia, specifically *Viola sumatrana* Miq. from Thailand. Two novel cyclotides (Visu 1 and 2) and two known cyclotides (kalata S and kalata B1) were identified in *V. sumatrana* after their isolation, purification and peptide sequencing. NMR studies revealed that kalata S and kalata B1 had similar secondary structures. The biological activities of kalata S and kalata B1 in *V. sumatrana* were determined in cytotoxicity assays; both had similar cytotoxic activity and both exhibited the highest cytotoxicity on U87 cells compared to other cell lines. Overall, the study strongly supports the ubiquity of cyclotides in the Violaceae and adds to our understanding of their distribution and cytotoxic activity.

Cyclotides[1,2] are the largest family of naturally occurring circular disulfide-rich peptides. They comprise 28–37 amino acid residues with a head-to-tail cyclized backbone and six conserved cysteine residues that form three interconnecting disulfide bonds (Cys^{I}-Cys^{IV}, Cys^{II}-Cys^{V} and Cys^{III}-Cys^{VI}). A small embedded ring is formed by two intra-cysteine backbone segments connected by two disulfide bonds (Cys^{I}-Cys^{IV}, Cys^{II}-Cys^{V}) and this ring is penetrated by the third disulfide bond (Cys^{III}-Cys^{VI}) to form a cystine knot. The combination of the cystine knot and cyclic backbone is known as a cyclic cystine knot (CCK) motif and renders cyclotides extremely resistant to thermal, chemical or enzymatic degradation.[1-3] Cyclotides are classified into two major subgroups, referred to as the Möbius and bracelet subfamilies, which are formally defined by the presence or absence of a X-Pro peptide bond in a *cis* configuration but also differ in the type and number of amino acid residues in the backbone loops between cysteine residues. Möbius subfamily cyclotides are characterized by a *cis*-proline residue in loop 5, which causes a 'twist' in the circular peptide backbone. Both subfamilies have high sequence homology, with certain amino acids showing strong conservation in loop 1 (glutamic acid, E) and loop 6 (aspartic acid, D; asparagine, N).[4, 5] A third cyclotide subfamily (the trypsin inhibitor subfamily) was based on the discovery of two trypsin inhibitor peptides, MCoTI-I and MCoTI-II, extracted from the seeds of *Momordica cochinchinensis*.[6, 7] Recently this subfamily has been expanded with the discovery of a series of related peptides in other *Momordica* species.[8, 9] MCoTI-II and related trypsin inhibitor cyclotides have been pivotal in understanding the biosynthetic processing of cyclic peptides in angiosperms.[9]

The prototypic cyclotide, kalata B1, first came to notice following a report by a Norwegian doctor who studied the bioactive component a Congolese traditional medicine "*kalata kalata*", which was derived from boiling leaves of *Oldenlandia affinis* (Roem & Schult. DC, Rubiaceae) to make a tea used for accelerating childbirth.[10, 11] The discovery of

the bioactivity of kalata B1 was followed many years later by the complete characterization of the peptide sequence, circular backbone, cystine knot and 3D structure by Saether *et al*.[12] A wide range of other cyclotides were later reported in *O. affinis*[13] and in other plants.[1,2]

The natural role of cyclotides in plants is postulated to be as host defense agents based on their insecticidal activities[14-16] but they have been reported to exhibit a wide range of other biological activities, including uterotonic activity,[9-12] anti-HIV activity,[17-20] hemolytic activity,[21] antimicrobial activity,[22] antifouling activity,[23] nematocidal activity,[24] molluscicidal activity,[25] neurotensin antagonism[26] as well as cytotoxic activity to mammalian cells,[27-29] green algae, and some soil bacteria.[30] Recently, cyclotides have also been reported as proteinase inhibitors[31] and immunosuppressive peptides.[32, 33] As cyclotides have a stable structure and exhibit a range of biological activities, they have found application in drug design and pharmaceutical development.[34, 35]

To date more than 400 cyclotides have been discovered in five plant families, including the Rubiaceae (50 cyclotides), Fabaceae (18 cyclotides), Cucurbitaceae (5 cyclotides), Solanaceae (3 cyclotides) and Violaceae (337 cyclotides), with the Violaceae containing the majority of known cyclotides.[36, 37] In an early study, Simonsen *et al*.[38] suggested that a single plant species could contain more than 100 cyclotides. Evidence for this level of complexity was provided in *Viola odorata* with 30 sequenced cyclotides, and more than 100 cyclotide-like masses detected.[39] Until recently, the discovery of cyclotides has probably been biased toward the detection of the most abundant peptides in a given extract, but combined transcriptomic and proteome mining approaches are beginning to be used, for example as recently utilised in the analysis of *Viola tricolor*, resulting in the characterization of 164 cyclotides.[40]

In the current study we focused specifically on the Violaceae (violets) family, which comprises 25 genera and approximately 900 species.[41] Cyclotides have been reported in many genera of the Violaceae, including *Gloeospermum*, *Hybanthus*, *Leonia, Melicytus*, *Rinorea* and *Viola sp.* and they are found most commonly in *Viola*, *Hybanthus and Rinorea*. *Viola* is the largest genus in the Violaceae, containing more than half of the total species.[41] Several studies have focused on the discovery and biological properties of cyclotides from plants in this genus, including *Viola abyssinica*,[42] *Viola arvensis*,[43] *Viola biflora*,[44] *Viola hederacea*,[20] *Viola odorata*,[22, 45, 46] *Viola tricolor*,[40, 47] *Viola yedoensis*[18] and *Viola philippica*,[27] but many more species have yet to be examined.

The previously studied *Viola* species were collected from far-reaching locations around the world, including Asia (China), Europe (Sweden), South America (Argentina), Africa (Ethiopia) and Australia. Although every *Viola* plant that has been studied so far contains cyclotides there have been no *Viola* species endemic to South-East Asia that have been examined for cyclotides. Here we examined cyclotide expression in a *Viola* species, *Viola sumatrana* Miq., also known as *Viola hossei* W. Becker or Hong-Ron (local Thai name), which is widely distributed in South-East Asian countries, including Indonesia, Malaysia, Myanmar, Vietnam, Thailand and China. Although some *Viola* species have been used as traditional medicines, a literature search (using SciFinder and NAPALERT) on *V. sumatrana* yielded only reports on its morphological characterization and biological diversity.[48]

We report the characterization of two novel cyclotides (Visu 1 and 2) and the presence of two known cyclotides (kalata S and kalata B1) in *V. sumatrana*. Additionally, the structure of kalata S was evaluated by NMR and found to be homologous to the prototypic cyclotide kalata B1. Kalata S and kalata B1 were tested in cytotoxicity assays and both exhibited toxicity against cancer cells at low micromolar concentrations.

RESULTS AND DISCUSSION

Isolation, purification and mass spectrometric sequencing of cyclotides from *V. sumatrana*. Fresh leaves of *V. sumatrana* were extracted using 50% (v/v) acetonitrile in 1% formic acid and the resultant extract was freeze-dried, yielding 2 g of crude extract. The extract was subjected to nanoelectrospray LC-MS, as shown in Figure 1. The three most abundant peaks contained signals corresponding to four masses in the molecular weight range typical of cyclotides (2800–3500 Da). Kalata S and kalata B1 co-eluted at the same retention time (40.03 min), while Visu 1 and Visu 2 eluted slightly later (41.46 min and 41.73 min, respectively). The crude extract was partially purified using solid-phase extraction (SPE) in which cyclotides were eluted using a stepwise acetonitrile (10% steps) gradient. SPE fractions corresponding to elution with 30–50% acetonitrile were found to contain cyclotides and were further purified using RP-HPLC. The HPLC fractions were subjected to peptide sequencing and cytotoxicity assays.

All cyclotides were first reduced with dithiothreitol followed by alkylation of the cysteines with iodoacetamide. An increase in mass of 348 Da following reduction and alkylation indicated the presence of six cysteines, and hence three disulfide bonds, in the native peptides. After reduction and alkylation, each cyclotide was subjected to enzymatic digestion with two proteolytic enzymes (endo-GluC and trypsin, individually and in combination) before tandem MS sequencing. More than ten cyclotide-like masses were detected in the crude extract, most of which were of low abundance. From these, we fully characterized two new cyclotides (Visu 1 and Visu 2) and two known cyclotides (kalata S and kalata B1), which were the four most abundant cyclotides in the extract. The isolated cyclotide sequences and their expected and theoretical masses are shown in Table 1.

Sequence homology of the new peptides to those previously reported was used to guide assignment of the positions of leucine and isoleucine in the new cyclotides. The new sequences contain Glu7 in loop 1 and Asn29 in loop 6, with the latter known to be critical for excision from the precursor protein and required for enzymatic cyclization. Visu 2 is categorized as a bracelet cyclotide based on the lack of a Pro residue in loop 5, whereas kalata S and kalata B1 belong to the Möbius subfamily. Loop 5 of Visu 1 (TWPV) has a single conservative amino acid substitution compared to 'SWPV' commonly observed in the corresponding position among Möbius cyclotides, which suggests that a *cis*-proline conformation is more likely to be maintained in this loop than not, hence the classification of Visu 1 here as a member of the Möbius subfamily. Due to an insufficient amount of purified Visu 1, further characterization of the *cis*-proline could not be made by NMR. Visu 1 has the same molecular weight as cycloviolacin O22, but has two conservative amino acid changes in loops 5 and 6 (Val4 to Ile4 and Thr22 to Ser22, respectively). Similarly, Visu 2 has the same molecular weight as the known cyclotide, cycloviolacin O1. These isobaric cyclotides contain different amino acid sequences in loops 3 (Ile15 to Val15) and 5 (Lys24 to Ser24, Ser25 to Asn25 and Lys26 to Arg26).

MS/MS spectra used for the characterization of Visu 1 and Visu 2, with assigned b- and y- ion series, after reduction, alkylation and enzymatic digestion are presented in Figures 2 and 3, respectively. Interestingly, we observed deamidation of Asn29 of Visu 1 and Asn30 of Visu 2 as artifacts of sample preparation. The monoisotopic mass $[M+H]^+$ of Visu 2 shifted from 3115.4 to 3116.4 Da after adjusting the pH of the sample prior to enzymatic digestion (Figure 4). Deamidation of Asn residues has been reported previously as a direct result of sample handling, with Asn residues immediately before Gly residues known to be particularly susceptible.[49] The conditions employed in proteomic analyses such as temperature, ionic strength of the buffer and pH all directly influence the rate of Asn

deamidation. Scotchier and co-workers[50] reported the maximum stability of Asn in peptides is at pH 6. Although previous analyses reported artefactual deamidation for cyclotides following reduction and alkylation reactions involving prolonged incubation at elevated temperatures,[40, 42] in the current study we found that simply dissolving them in ammonium bicarbonate buffer (pH 8.0) resulted in rapid and extensive (approximately 90%) deamidation of Asn residues within 30 minutes, consistent with results reported by Scotchier *et al*.[50]

Structural analysis of kalata S by NMR. Kalata S (alternatively named varv A) was the first cyclotide discovered from *V. arvensis* (Violaceae) and is also found in several other members of the Violaceae family. To date, only amino acid sequencing data and biological activities have been reported for cyclotides from this plant, and therefore it was of interest to determine the structure of kalata S as no NMR data have been reported so far. A NOESY sequential assignment based on αH_i-$NH_{(i+1)}$ connectivities is shown in Figure 5, with a complete list of chemical shifts shown in Table 2. To determine the secondary structure of kalata S, its αH secondary chemical shifts were compared with those of kalata B1, which has been well characterized structurally. Kalata S showed similar secondary shifts to kalata B1, indicating that both peptides indeed have similar secondary structure (Figure 6).

The presence of a proline in loop 5 of kalata S classifies it as a member of the Möbius subfamily, assuming that the Pro residue is in a cis conformation. This was confirmed by noting that Pro24 in kalata S has similar αH and *β*H chemical shifts to kalata B1, and $\alpha H_{(i-1)}$-$\alpha H_{(i)}$ NOE, indicating a conserved *cis*-peptide bond conformation. Overall, the similarity observed for the NMR data and amino acid sequencing from mass spectrometry revealed that kalata S contains the unique CCK motif and is structurally similar to kalata B1. On the one hand, this is not unexpected since the sequences of kalata S and kalata B1 differ by only one

amino acid in loop 4, i.e. Ser20 to Thr20, but on the other hand, this single residue in loop 4 is thought to be involved in crucial stabilizing hydrogen bonding interactions via the hydroxyl sidechain[5] and the structural data reported here are the first experimental validation of that suggestion.

Cytotoxic activity. The cytotoxicity of kalata B1 against HT29 cells has been reported previously using the FMCA assay.[51] Therefore it was of interest to assess the cytotoxicity of three well-known cyclotides (kalata S, kalata B1 and cycloviolacin O2) against a range of different human cell lines under standardized conditions with an MTT assay. Kalata S and kalata B1 were synthesized due to the limited amounts that could be isolated from the plant extract. Both cyclotides were evaluated using a MTT assay against HUVEC cells (a non-cancer cell line) and four different human cancer cell lines (U87, U251, HT29 and MCF7) to determine the specificity of any toxicity for cell type. Cycloviolacin O2 was selected as a positive control based on previous reports[47, 51, 52] that describe it as the most cytotoxic cyclotide tested to date. Previous publications have reported cytotoxicity screening of cycloviolacin O2 against U251, HT29 and MCF7 using SRB,[47] FMCA[51] and MTT assays[52], respectively.

The results of the cytotoxicity assays are shown in Figure 7 and a summary of IC_{50} values is given in Table 3. Among the tested cyclotides, cycloviolacin O2 has the lowest (most potent) IC_{50} value. The IC_{50} values of cycloviolacin O2 for HUVEC, U87, U251, HT29 and MCF7 cells were 0.35, 0.45, 0.79, 0.87 and 0.29 μM, respectively. Cycloviolacin O2 was found to be the most cytotoxic against the MCF7 cell line, consistent with previous reports.[47, 51, 52] The higher activity of cycloviolacin O2 might be due to its different spatial distribution of hydrophobic amino acids on the cyclotide surface compared to other cyclotides and net charge,[51] since both bracelet and Möbius families have similar ratios of hydrophobic residues but different spatial distributions.[53] Although the cytotoxicity of cycloviolacin O2 against

MCF7 has been reported using an MTT assay by Gerlach *et al*.,[17] different experimental conditions, including incubation periods and dyes for assessing cell viability were used in the current study but the results were comparable.

It is interesting to compare the IC_{50} values of kalata S and kalata B1 as they differ in only one amino acid residue but that residue forms part of the embedded ring of the cystine knot motif and, as noted earlier, is involved in crucial hydrogen bonding interactions. Both cyclotides were found to have similar IC_{50} values for all tested cell lines, suggesting that both maintain a similar cyclotide-membrane interaction and have similar structures, with the difference in loop 4 (i.e. Ser or Thr) not playing a crucial role in activity. Among the cell lines tested, kalata B1 and kalata S exhibited the highest cytotoxicity on U87 with an IC_{50} of 3-5 μM, which is 2-3 fold lower than for other cell lines (Table 3).

EXPERIMENTAL SECTION

Plant material. Fresh leaves of *V. sumatrana* were collected from Chanthaburi province, Thailand in May 2013. A voucher specimen (SN065812) has been deposited at the Queen Sirikit Botanical Garden Herbarium: QBG, Mae Rim, Chiang Mai, Thailand.

Isolation and extraction of cyclotides from *V. sumatrana*. Fresh leaves of *V. sumatrana* were cooled with liquid nitrogen and ground with a mortar and pestle, giving 2 g of total crude powder. The crude powder was dissolved in 100 mL of 50% (v/v) acetonitrile in 1% formic acid and stirred at room temperature (25°C) for 4 h during peptide extraction. The crude solution was centrifuged for 45 min at 8,000 rpm and the supernatant was retained. The supernatant was lyophilized using a freeze drier (CHRIST Alpha 2-4 LD Freeze dryer).

Solid-phase extraction and RP-HPLC purification. The crude extract was partially purified by solid-phase extraction (SPE) and followed by peptide isolation using reversed-

phase high performance liquid chromatography (RP-HPLC). Crude extract (2 g) was dissolved in 20 mL of 1% (v/v) formic acid before SPE. C18 SPE cartridges (Waters, 500 mg) were activated and equilibrated with 10 mL of methanol and 1% (v/v) formic acid, respectively. The crude extract was then loaded to the cartridges and eluted sequentially with 10 mL volumes of 20–80% (v/v) acetonitrile in 1% (v/v) formic acid for cyclotide separation. Cyclotide-containing fractions were freeze-dried after fractionation and re-dissolved in 0.05% (v/v) trifluoroacetic acid/water (buffer A). This solution was then loaded to a Phenomenex C18 semi-preparative column (250 mm x 10 mm, 10 µm, 300 Å, 3 mL min^{-1} flow rate) on a RP-HPLC (LC10, Shimadzu) and run with a 0.5% min^{-1} gradient from 20–80% solvent B (0.45% trifluoroacetic acid/90% acetonitrile) to yield pure cyclotides. All eluting cyclotides were detected at 215 nm and 280 nm. This purification step was repeated several times to obtain high purity cyclotides for use in biological assays and further peptide characterization.

Reduction and alkylation of cyclotides. Each isolated cyclotide was dissolved in 100 µL 10 mM NH_4HCO_3 (pH 8). To reduce disulfide bonds, 10 µL of 100 mM dithiothreitol (DTT) was added and the solution was incubated for 30 min (60ºC, under nitrogen). To block cysteines and prevent sample reoxidation, 10 µL of 250 mM iodoacetamide was added and the solutions were incubated for 30 min at room temperature. Finally, the reduced and alkylated sample was desalted with C18 Ziptips (Millipore) and eluted in 10 µL of 80% (v/v) acetonitrile with 1% formic acid. An aliquot of the desalted cyclotide was mixed in a 1:1 ratio with 7 mg/mL of MALDI matrix (α-cyano-4-hydroxycinnamic acid (CHCA) in 50% (v/v) acetonitrile/ 1% (v/v) formic acid) and analyzed with MALDI-TOF MS to determine the molecular weight of each cyclotide.

Enzymatic digestion coupled with nanospray and MALDI-TOF MS/MS sequencing. Reduced and alkylated samples prepared as above were also used for enzymatic digestion. Enzymes used in this study are endoproteinase Glu-C, trypsin or a mixture of both.

For single enzyme digestions, 2 µL of 1 µg/µL endoproteinase Glu-C or trypsin enzyme was added into a 20-µL sample. Meanwhile 20 µL of sample was mixed with 2 µL of both enzymes for double enzyme digestion. All digested peptides were then incubated at 37°C for 6 h and desalted using C18 ziptips before MS/MS peptide sequencing. All digested peptides were mixed with CHCA matrix (7 mg/mL in 50% ACN and 2% formic acid) in a ratio 1:1 before spotting 1 µL of the mixture onto the MALDI target plates. The peptides were analyzed using MALDI-TOF mass spectrometry (Bruker UltrafleXtreme TOF-TOF MS and SCIEX 4700 TOF-TOF MS). For nanoelectrospray-MS/MS, 10 µL of digested peptides was loaded into Thermo Fisher Scientific ES380 nanoES metal-coated spray capillaries before analysis by static nanospray on a QSTAR Pulsar mass spectrometer (SCIEX) equipped with a Proxeon II nano ion source. The nanospray voltage was set to 900 V with a declustering potential of 70 V. TOF-MS data was acquired in the mass range 400–2000 Da and MS/MS data was acquired by incrementally increasing the collision energy (from 15 to 50 V) to obtain fragment ion coverage and enhanced signal. FlexControl and FlexAnalysis (Bruker) and Analyst QS (SCIEX) software packages were used to acquire and process the MS/MS data.

NMR structure analysis. We focused on kalata S for NMR studies. This peptide was dissolved in 90% H_2O/10% D_2O (pH 3.29) and added 1% of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as a chemical shift reference for spectral calibration. ^1H spectra and two-dimensional spectra (TOCSY and NOESY) were acquired. TOCSY and NOESY spectra were acquired with a mixing time of 80 ms and 200 ms, respectively. All spectra were recorded on a Bruker Avance 600 MHz spectrometer at 298 K and were processed using TOPSIN 2.1 (Bruker) program. NMR data were analyzed using CCPNMR spectra assignment program version win32 2.3.1.[54] and all sequential assignments were done according to the sequential assignment procedure of Wuthrich *et al*.[55]

Evaluation of cyclotides in mammalian cell cytotoxicity assays. Cyclotides kalata S and kalata B1 were synthesized using Fmoc-based synthesis of disulfide-rich cyclic peptides.[56] The cytotoxic effects of cyclotides on human umbilical vein endothelial cells (HUVEC), two human brain cell lines (U87 and U251), a human colon adenocarcinoma cell line (HT29) and a breast cancer cell line (MCF7) were evaluated using an MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; Sigma) assay. HUVECs and two human gliomablastoma cell lines (U87 and U251) were plated in 96-well plates at 3×10^3 cells $well^{-1}$ (100 µL) in 10% FBS/EBM-2 media supplemented with SingleQuots (which includes growth factors, cytokines, antibiotics) (Lonza) and 10% FBS/DMEM (Dulbecco's Modified Eagle Medium) (Gibco), respectively. On the other hand, HT29 and MCF7 were plated at 2×10^4 cells $well^{-1}$ (100 µL) in 10% FBS/DMEM. All cells were incubated at 37ºC in 5% CO_2 for 24 h. Prior to the addition of the test compounds, media were removed and replaced with fresh serum-free EBM-2 and DMEM media (100 µL $well^{-1}$). All cyclotides were tested in triplicate with final peptide concentrations ranging from 0.05 to 100 µM, and with incubation with cells for 2 h. Controls used in this assay included the vehicle control (negative control) and 1% (v/v) Triton (positive control). After 2 h of incubation, 10 µL of MTT (5 mg mL^{-1} in phosphate buffered saline) was added to each well and further incubated for 3 h before removing supernatant. The MTT formazan crystals observed in each well were dissolved in 100 µL of dimethylsulfoxide and absorbance was measured using BioTek PowerWave XS spectrophotometer at 600 nm. Data were analyzed using GraphPad Prism® software and IC_{50} values were obtained from the sigmoidal dose-response curve.

ACKNOWLEDGEMENTS

We thank the Thailand Research Fund through the Royal Golden Jubilee PhD Program (Grant No. PHD/0013/2552) and Bruker for financial support. We also thank A. Jones, P. Harvey and O. Cheneval from The University of Queensland, Australia for their help and suggestions regarding mass spectrometry, NMR and peptide synthesis, respectively and also thank Bob Harwood for plant material. We are grateful to the Proteomics Mass Spectrometry Facility at the Institute for Molecular Bioscience at The University of Queensland for access to the mass spectrometers used in this study, and acknowledge access to the facilities of the Queensland NMR Network. This work was supported by grants from the National Health and Medical Research Council (APP1028509) and the Australian Research Council (ARC; DP150100443). DJC is an ARC Australian Laureate Fellow (FL150100146).

REFERENCES

1. Burman, R.; Gunasekera, S.; Stromstedt, A. A.; Göransson, U. J. Nat. Prod. 2014, 77, 724-736.
2. Craik, D. J.; Daly, N. L.; Bond, T.; Waine, C. J. Mol. Biol. 1999, 294, 1327-1336.
3. Colgrave, M. L.; Craik, D. J. Biochemistry 2004, 43, 5965-5975.
4. Ireland, D. C.; Clark, R. J.; Daly, N. L.; Craik, D. J. J. Nat. Prod. 2010, 73, 1610-1622.
5. Rosengren, K. J.; Daly, N. L.; Plan, M. R.; Waine, C.; Craik, D. J. J. Biol. Chem. 2003, 278, 8606-8616.
6. Hernandez, J. F.; Gagnon, J.; Chiche, L.; Nguyen, T. M.; Andrieu, J. P.; Heitz, A.; Trinh Hong, T.; Pham, T. T.; Le Nguyen, D. Biochemistry 2000, 39, 5722-5730.
7. Chan, L. Y.; He, W.; Tan, N.; Zeng, G.; Craik, D. J.; Daly, N. L. Peptides 2013, 39, 29-35.
8. Mahatmanto, T.; Mylne, J. S.; Poth, A. G.; Swedberg, J. E.; Kaas, Q.; Schaefer, H.; Craik, D. J. Mol. Biol. Evol. 2015, 32, 392-405.
9. Mylne, J. S.; Chan, L. Y.; Chanson, A. H.; Daly, N. L.; Schaefer, H.; Bailey, T. L.; Nguyencong, P.; Cascales, L.; Craik, D. J. Plant Cell 2012, 24, 2765-2778.
10. Gran, L. Medd. Nor. Farm. Selsk. 1970, 12, 173-180.
11. Gran, L. Acta Pharmacol. Toxicol. 1973, 33, 400-408.
12. Saether, O.; Craik, D. J.; Campbell, I. D.; Sletten, K.; Juul, J.; Norman, D. G. Biochemistry 1995, 34, 4147-4158.
13. Plan, M. R. R.; Göransson, U.; Clark, R. J.; Daly, N. L.; Colgrave, M. L.; Craik, D. J. ChemBioChem 2007, 8, 1001-1011.
14. Craik, D. J. Toxins 2012, 4, 139-156.

15. Jennings, C.; West, J.; Waine, C.; Craik, D.; Anderson, M. Proc. Natl. Acad. Sci. U. S. A. 2001, 98, 10614-10619.

16. 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. Biochemistry 2005, 44, 851-860.

17. Gerlach, S. L.; Yeshak, M.; Göransson, U.; Roy, U.; Izadpanah, R.; Mondal, D. Biopolymers 2013, 100, 471-479.

18. Wang, C. K.; Colgrave, M. L.; Gustafson, K. R.; Ireland, D. C.; Göransson, U.; Craik, D. J. J. Nat. Prod. 2008, 71, 47-52.

19. Daly, N. L.; Koltay, A.; Gustafson, K. R.; Boyd, M. R.; Casas-Finet, J. R.; Craik, D. J. J. Mol. Biol. 1999, 285, 333-345.

20. Chen, B.; Colgrave, M. L.; Daly, N. L.; Rosengren, K. J.; Gustafson, K. R.; Craik, D. J. J. Biol. Chem. 2005, 280, 22395-22405.

21. Chen, B.; Colgrave, M. L.; Wang, C.; Craik, D. J. J. Nat. Prod. 2006, 69, 23-28.

22. Pranting, M.; Loov, C.; Burman, R.; Göransson, U.; Andersson, D. I. J. Antimicrob. Chemother. 2010, 65, 1964-1971.

23. Göransson, U.; Sjögren, M.; Svangård, E.; Claeson, P.; Bohlin, L. J. Nat. Prod. 2004, 67, 1287-1290.

24. Colgrave, M. L.; Kotze, A. C.; Huang, Y. H.; O'Grady, J.; Simonsen, S. M.; Craik, D. J. Biochemistry 2008, 47, 5581-5589.

25. Plan, M. R. R.; Saska, I.; Cagauan, A. G.; Craik, D. J. J. Agric. Food Chem. 2008, 56, 5237-5241.

26. Witherup, K. M.; Bogusky, M. J.; Anderson, P. S.; Ramjit, H.; Ransom, R. W.; Wood, T.; Sardana, M. J. Nat. Prod. 1994, 57, 1619-1625.

27. He, W.; Chan, L. Y.; Zeng, G.; Daly, N. L.; Craik, D. J.; Tan, N. Peptides 2011, 32, 1719-1723.

28. Lindholm, P.; Göransson, U.; Johansson, S.; Claeson, P.; Gullbo, J.; Larsson, R.; Bohlin, L.; Backlund, A. Mol. Cancer Ther. 2002, 1, 365-369.

29. Svangård, E.; Göransson, U.; Hocaoglu, Z.; Gullbo, J.; Larsson, R.; Claeson, P.; Bohlin, L. J. Nat. Prod. 2004, 67, 144-147.

30. Ovesen, R. G.; Brandt, K. K.; Göransson, U.; Nielsen, J.; Hansen, H. C.; Cedergreen, N. Environ. Toxicol. Chem. 2011, 30, 1190-1196.

31. Hellinger, R.; Koehbach, J.; Puigpinós, A.; Clark, R. J.; Tarragó, T.; Giralt, E.; Gruber, C. W. J. Nat. Prod. 2015, 78, 1073-1082.

32. Grundemann, C.; Koehbach, J.; Huber, R.; Gruber, C. W. J. Nat. Prod. 2012, 75, 167-174.

33. Thell, K.; Hellinger, R.; Sahin, E.; Schabbauer, G.; Gruber, C. W. J. Pept. Sci. 2014, 20, S100-S101.

34. Northfield, S. E.; Wang, C. K.; Schroeder, C. I.; Durek, T.; Kan, M. W.; Swedberg, J. E.; Craik, D. J. Eur. J. Med. Chem. 2014, 77, 248-257.

35. Henriques, S. T.; Craik, D. J. Drug Discov. Today 2010, 15, 57-64.

36. Craik, D. J.; Malik, U. Curr. Opin. Chem. Biol. 2013, 17, 546-554.

37. Kaas, Q.; Craik, D. J. Biopolymers 2010, 94, 584-591.

38. Simonsen, S. M.; Sando, L.; Ireland, D. C.; Colgrave, M. L.; Bharathi, R.; Göransson, U.; Craik, D. J. Plant Cell 2005, 17, 3176-3189.

39. Colgrave, M. L.; Poth, A. G.; Kaas, Q.; Craik, D. J. Biopolymers 2010, 94, 592-601.

40. Hellinger, R.; Koehbach, J.; Soltis, D. E.; Carpenter, E. J.; Wong, G. K.-S.; Gruber, C. W. J. Proteome Res. 2015, 14, 4851-4862.

41. Ballard, H. E.; Sytsma, K. J.; Kowal, R. R. Syst. Bot. 1999, 23, 439-458.

42. Yeshak, M. Y.; Burman, R.; Asres, K.; Göransson, U. J. Nat. Prod. 2011, 74, 727-731.

43. Göransson, U.; Luijendijk, T.; Johansson, S.; Bohlin, L.; Claeson, P. J. Nat. Prod. 1999, 62, 283-286.

44. Herrmann, A.; Burman, R.; Mylne, J. S.; Karlsson, G.; Gullbo, J.; Craik, D. J.; Clark, R. J.; Goransson, U. Phytochemistry 2008, 69, 939-952.

45. Ireland, D. C.; Colgrave, M. L.; Craik, D. J. Biochem. J. 2006, 400, 1-12.

46. Ireland, D. C.; Colgrave, M. L.; Nguyencong, P.; Daly, N. L.; Craik, D. J. J. Mol. Biol. 2006, 357, 1522-1535.

47. Tang, J.; Wang, C. K.; Pan, X.; Yan, H.; Zeng, G.; Xu, W.; He, W.; Daly, N. L.; Craik, D. J.; Tan, N. Peptides 2010, 31, 1434-1440.

48. Chen, Y.; Yang, Q.; Ohba, H.; Nikitin, V. V. Flora of China 2007, 13, 72-111.

49. Mckerrow, J. H.; Robinson, A. B. Anal. Biochem. 1971, 42, 565-568.

50. Scotchler, J. W.; Robinson, A. B. Anal. Biochem. 1974, 59, 319-322.

51. Burman, R.; Stromstedt, A. A.; Malmsten, M.; Göransson, U. Biochim. Biophys. Acta 2011, 1808, 2665-2673.

52. Gerlach, S. L.; Rathinakumar, R.; Chakravarty, G.; Göransson, U.; Wimley, W. C.; Darwin, S. P.; Mondal, D. Biopolymers 2010, 94, 617-625.

53. Simonsen, S. M.; Sando, L.; Rosengren, K. J.; Wang, C. K.; Colgrave, M. L.; Daly, N. L.; Craik, D. J. J. Biol. Chem. 2008, 283, 9805-9813.

54. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. Proteins 2005, 59, 687-696.

55. Wüthrich, K. NMR of Proteins and Nucleic Acids; New York: Wiley-Interscience, 1986.

56. Cheneval, O.; Schroeder, C. I.; Durek, T.; Walsh, P.; Huang, Y. H.; Liras, S.; Price, D. A.; Craik, D. J. J. Org. Chem. 2014, 79, 5538-5544.

57. Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. J. Biomol. NMR 1995, 5, 67-81.

Table 1. Amino acid sequences of cyclotides from *V. sumatrana*.

Cyclotide	Retention time (min)	Experimental *m/z* [a]	Experimental mass (Da) [b]	Theoretical mass (Da)	Amino acid sequence
kalata S	40.03	959.72	2876.16	2876.18	GLPV**C**GET**C**VGGT**C**NTPG---**C**S**C**SWPV**C**TRN
kalata B1	40.03	964.39	2890.17	2890.20	GLPV**C**GET**C**VGGT**C**NTPG---**C**T**C**SWPV**C**TRN
Visu 1	41.46	969.06	2904.18	2904.22	GLPV**C**GET**C**VGGT**C**NTPG---**C**T**C**TWPV**C**TRN
Visu 2	41.73	1039.46	3114.44	3114.43	GIP—**C**AES**C**VYIP**C**TITALLG**C**S**C**KSKV**C**YN

[a] *m/z* (+3) experimental masses
[b] All experimental masses were within 10 ppm of theoretical values

Table 2. NMR Chemical shifts (ppm) of kalata S at 298K, pH 3.29

Position	Residue	HN	Hα	Hβ	Others
1	Gly	8.63	4.25, 3.61		
2	Leu	7.77	5.07	1.94, 1.51	$_{\gamma}$H 1.71; $_{\delta}CH_3$ 0.99, 0.93
3	Pro		5.07	2.46, 1.75	CH_2 2.18, 2.05; CH_2 3.81, 3.81
4	Val	8.16	4.67	2.6	$_{\delta}CH_3$ 0.86
5	Cys	7.99	4.47	3.36, 3.00	
6	Gly	8.52	3.85, 3.75		
7	Glu	7.17	4.81	2.01, 1.89	CH_2 2.53
8	Thr	8.44	4.58	4.46	$_{\gamma}CH_3$ 1.17
9	Cys	8.33	4.93	3.18, 2.95	
10	Val	8.51	3.9	2.05	$_{\gamma}CH_3$ 1.00
11	Gly	8.73	4.24, 3.86		
12	Gly	8.24	4.41, 4.05		
13	Thr	7.85	4.71	4.11	$_{\gamma}CH_3$ 1.16
14	Cys	8.67	4.72	3.06, 2.78	
15	Asn	10.37	4.77		
16	Thr	8.47	4.49	4.22	$_{\gamma}CH_3$ 1.34
17	Pro		4.24	2.32. 1.91	CH_2 2.01, 2.15; CH_2 4.16, 3.72
18	Gly	8.78	4.19, 3.71		
19	Cys	7.7	5.32	3.83, 2.63	
20	Ser	9.54	4.69	3.75	
21	Cys	9.08	4.58	3.12, 2.82	
22	Ser	8.94	4.79	3.86	
23	Trp	8.02	4.09	3.27, 3.27	2H 7.32; 4H 7.56; 5H 7.16; 6H 7.26; 7H 7.46; NH 10.39
24	Pro		3.46	1.73, -0.13	CH_2 1.31, 1.42; CH_2 3.25, 3.25
25	Val	8.25	4.22	1.95	
26	Cys	7.75	5.12	3.22, 2.75	
27	Thr	9.84	5.08	3.72	$_{\gamma}CH_3$ 0.89
28	Arg	8.75	4.76	1.67	$_{\gamma}CH_2$ 1.44; $_{\delta}CH_2$ 3.24
29	Asn	9.55	4.41	3.11, 2.82	

*chemical shifts are referenced to DSS

Table 3. Cytotoxic activity of cyclotides against cancerous and non-cancerous cells.

Cyclotide	IC_{50} (μM) ± S.D.				
	HUVEC	U87	U251	HT29	MCF7
kalata B1	6.43±0.04	3.21±0.07	10.88±0.03	11.43±0.08	5.76±0.05
kalata S	9.73±0.06	4.63±0.07	8.35±0.04	10.69±0.04	5.46±0.02
cycloviolacin O2	0.35±0.06	0.45±0.02	0.79±0.06	0.87±0.08	0.29±0.04

FIGURE CAPTIONS

Figure 1. LC-MS total ion chromatogram of crude protein from *V. sumatrana.*

Figure 2. MS-MS sequencing of Visu 1 after combined trypsin and endoproteinase Glu-C digest. (A) MS-MS of precursor *m/z* 1091.33^{3+} (3271.0 Da). (B) MS-MS of precursor *m/z* 1222.99^{2+} (2444.0 Da). (C) MS-MS of precursor *m/z* 423.68^{2+} (845.4 Da).

Figure 3. MS-MS sequencing of Visu 2 after combined trypsin and endoproteinase Glu-C digest. (A) MS-MS of precursor *m/z* 1161.52^{3+} (3481.6 Da). (B) MS-MS of precursor *m/z* 773.37^{3+} (2317.1 Da). (C) MS-MS of precursor *m/z* 592.24^{2+} (1182.5 Da).

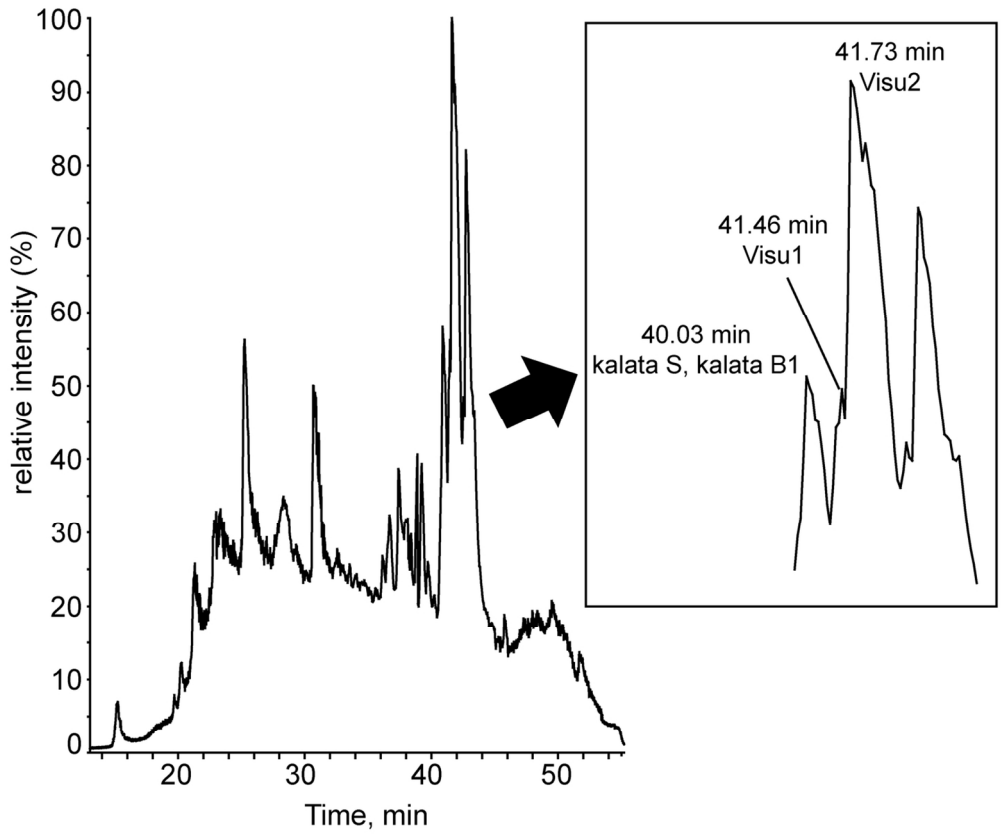
Figure 4. Data showing deamidation of Visu 2. (A) Monoisotopic mass of Visu 2 after being dissolved in water and (B) adjusted with ammonium bicarbonate (pH = 8).

Figure 5. (A) NOESY spectrum and (B) TOCSY spectrum of kalata S in 90% H_2O and 10% D_2O at 298 K, pH 3.29. Peaks are marked using the one-letter amino acid code. R28 (SC) refers to the side-chain.

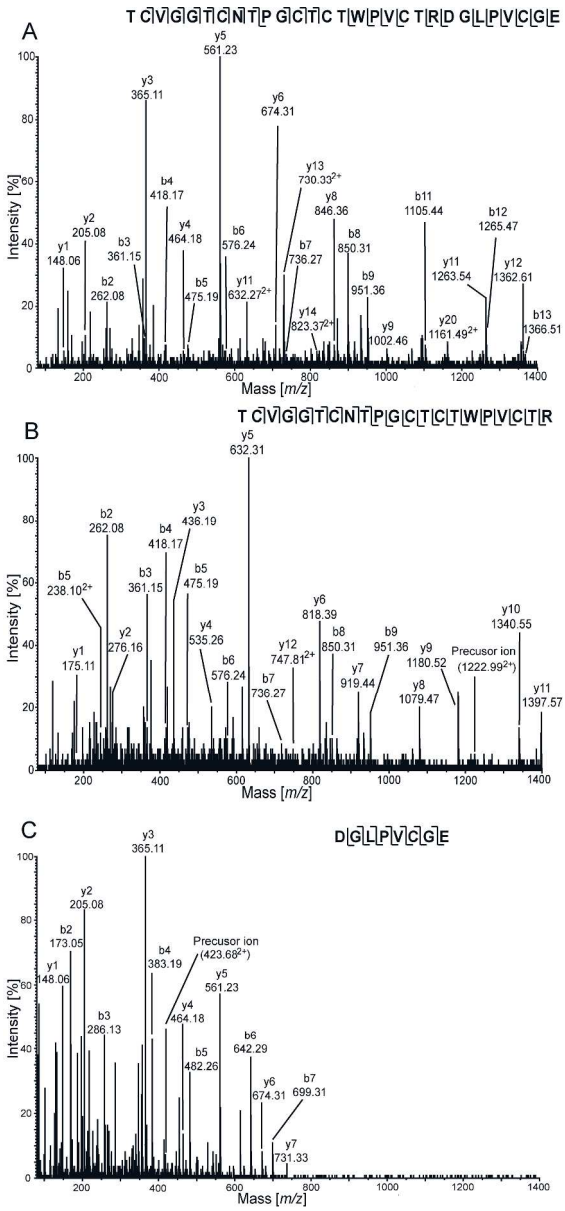
Figure 6. αH secondary chemical shift comparison of kalata S and kalata B1. Both ^1H NMR spectra were recorded at 298 K and the αH secondary shifts were calculated by subtracting

the random coil ^{1}H NMR chemical shifts of Wishart *et al*.[57] from the experimental αH chemical shifts. The average of the two αH proton shifts was used for glycine (G) residues.

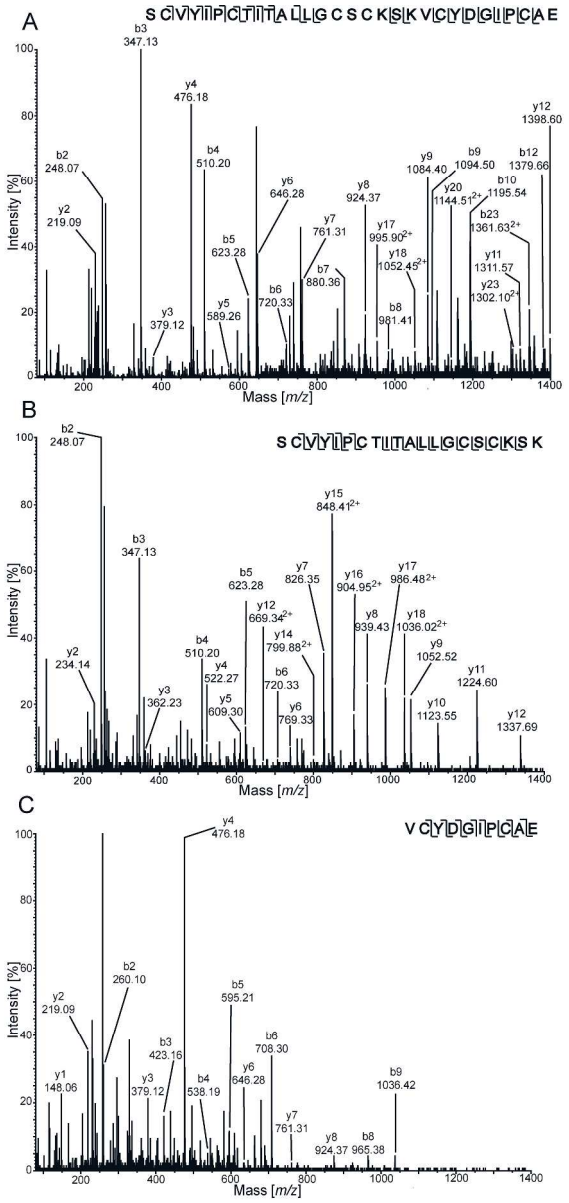
Figure 7. Cytotoxic activity of known cyclotides (kalata B1, kalata S, cycloviolacin O2 and cycloviolacin O12). (A) Human umbilical vein endothelial cell (HUVEC), a non-cancerous cell line; (B-C) human brain cancer cell lines, U87 and U251; (D) human colon adenocarcinoma cell line (HT29), and (E) breast cancer cell line (MCF7). IC_{50} values were obtained by plotting percentage of cell viability versus peptide concentration using GraphPad Prism®.



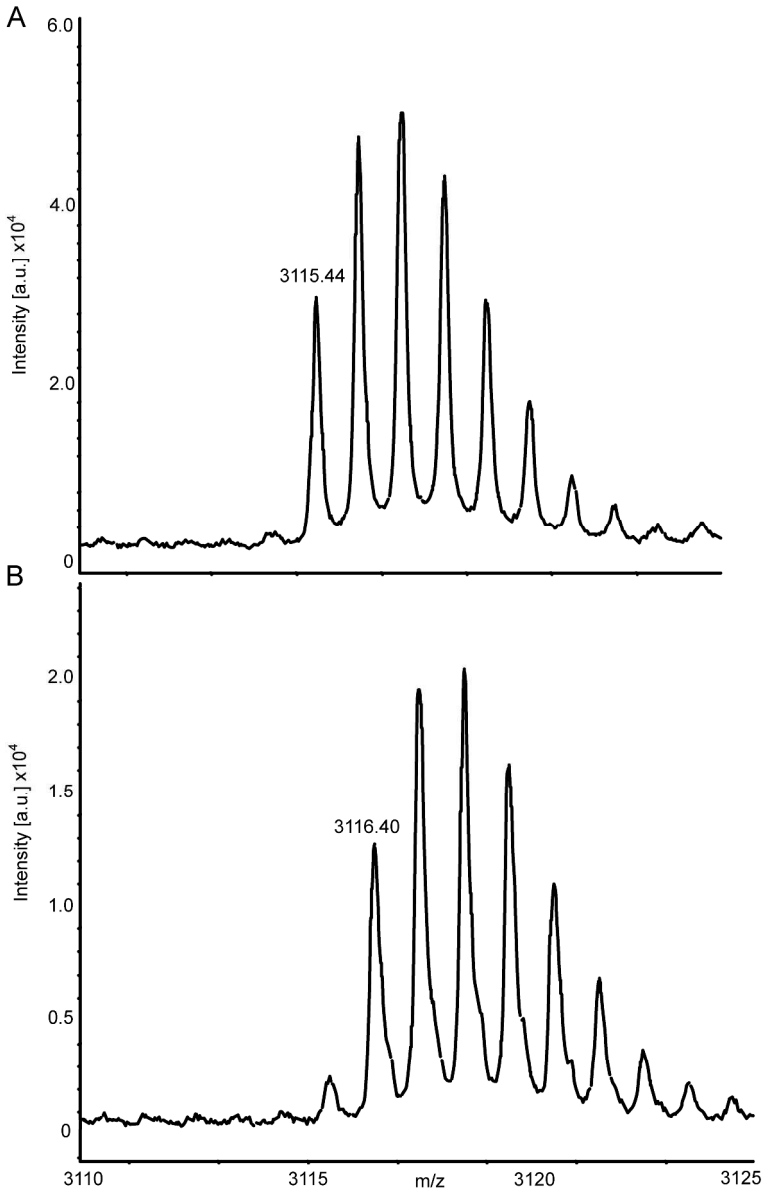
LC-MS total ion chromatogram of crude protein from *V. sumatrana*.
123x102mm (300 x 300 DPI)



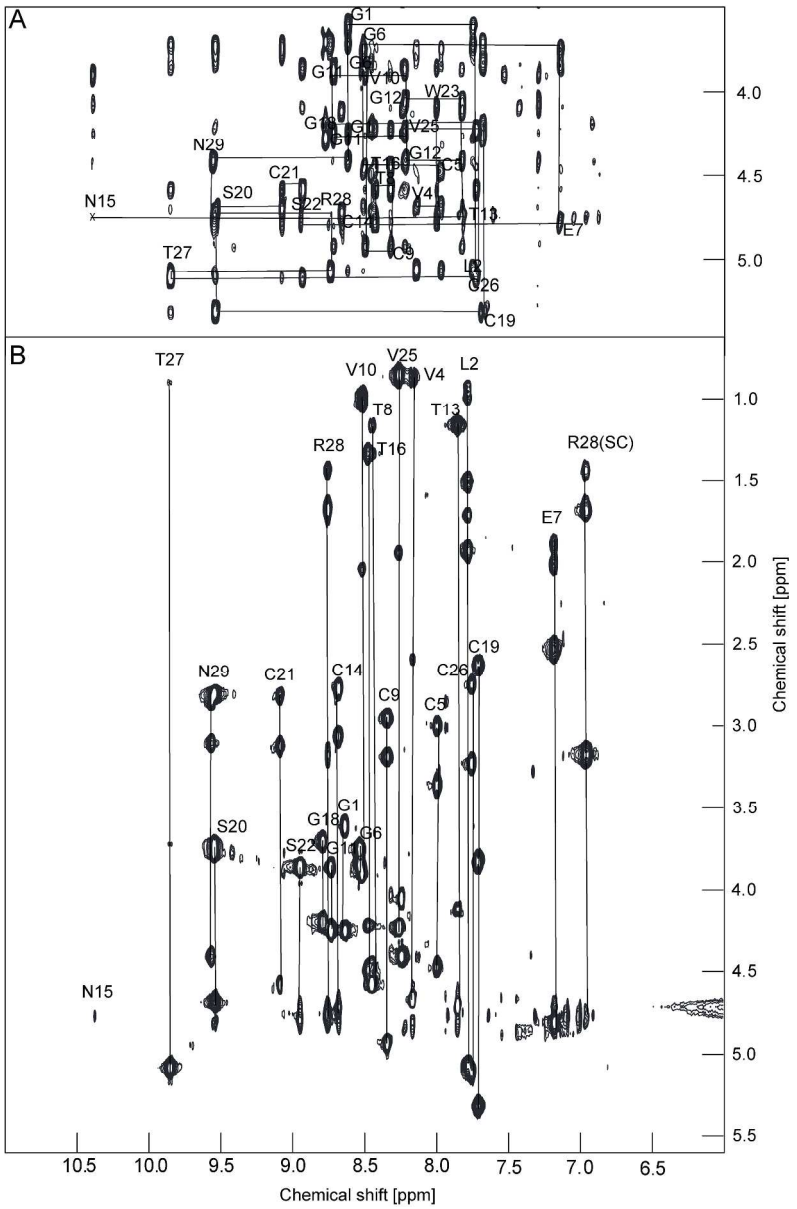
MS-MS sequencing of Visu 1 after combined trypsin and endoproteinase Glu-C digest. (A) MS-MS of precursor *m/z* 1091.33^{3+} (3271.0 Da). (B) MS-MS of precursor *m/z* 1222.99^{2+} (2444.0 Da). (C) MS-MS of precursor *m/z* 423.68^{2+} (845.4 Da).
289x613mm (300 x 300 DPI)



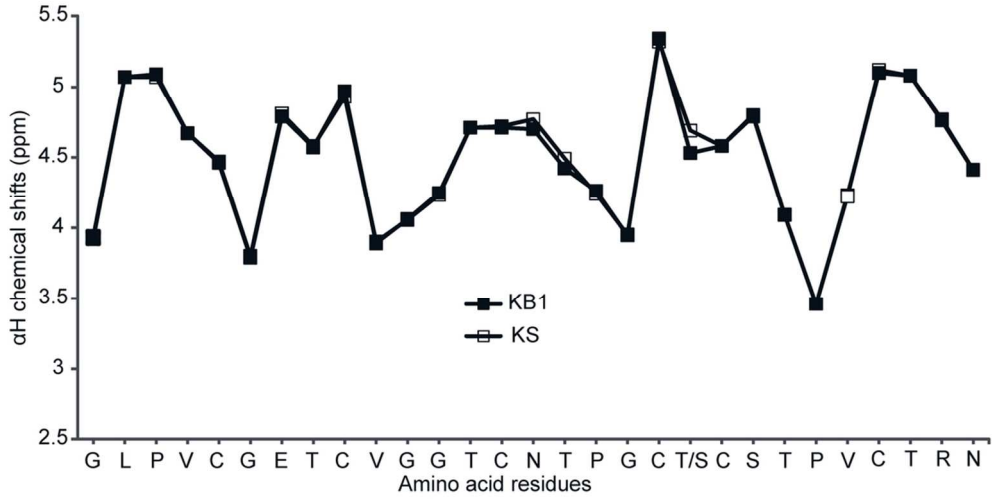
MS-MS sequencing of Visu 2 after combined trypsin and endoproteinase Glu-C digest. (A) MS-MS of precursor *m/z* 1161.52^{3+} (3481.6 Da). (B) MS-MS of precursor *m/z* 773.37^{3+} (2317.1 Da). (C) MS-MS of precursor *m/z* 592.24^{2+} (1182.5 Da).
285x595mm (300 x 300 DPI)



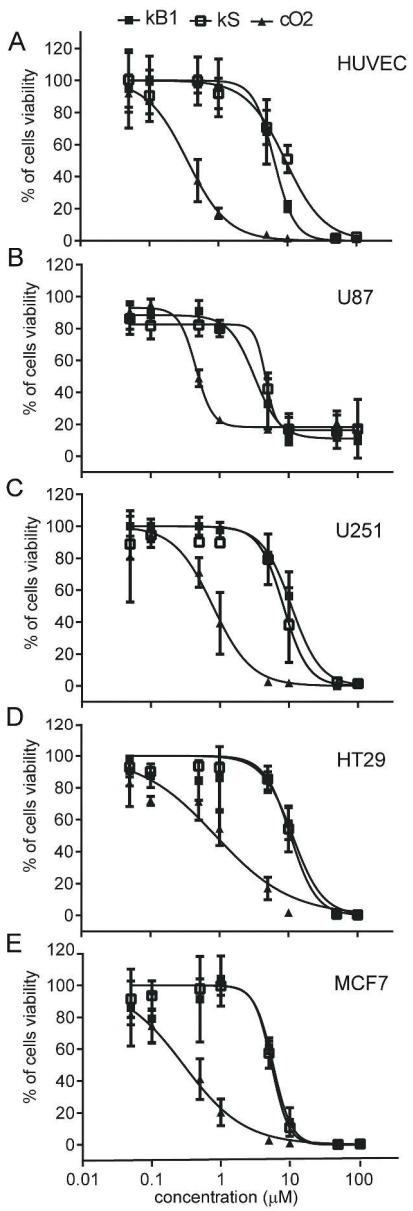
Data showing deamidation of Visu 2. (A) Monoisotopic mass of Visu 2 after being dissolved in water and (B) adjusted with ammonium bicarbonate (pH = 8).
272x431mm (300 x 300 DPI)



(A) NOESY spectrum and (B) TOCSY spectrum of kalata S in 90% H_2O and 10% D_2O at 298 K, pH 3.29. Peaks are marked using the one-letter amino acid code. R28 (SC) refers to the side-chain.
294x450mm (300 x 300 DPI)



αH secondary chemical shift comparison of kalata S and kalata B1. Both ^{1}H NMR spectra were recorded at 298 K and the αH secondary shifts were calculated by subtracting the random coil ^{1}H NMR chemical shifts of Wishart *et al.*[57] from the experimental αH chemical shifts. The average of the two αH proton shifts was used for glycine (G) residues.
99x50mm (300 x 300 DPI)



Cytotoxic activity of known cyclotides (kalata B1, kalata S, cycloviolacin O2 and cycloviolacin O12). (A) Human umbilical vein endothelial cell (HUVEC), a non-cancerous cell line; (B-C) human brain cancer cell lines, U87 and U251; (D) human colon adenocarcinoma cell line (HT29), and (E) breast cancer cell line (MCF7). IC_{50} values were obtained by plotting percentage of cell viability versus peptide concentration using GraphPad Prism®.
294x861mm (300 x 300 DPI)

Cyclotides are bioactive cyclic peptides from plants. Two novel cyclotides (Visu 1 and 2) and two known cyclotides were identified in *V. sumatrana* after their isolation, purification and peptide sequencing. Overall the study strongly supports the ubiquity of cyclotides in the Violaceae and adds to our understanding of their distribution and cytotoxic activity.
76x39mm (300 x 300 DPI)