

Sequencing of T-Superfamily Conotoxins from *Conus virgo*: Pyroglutamic Acid Identification and Disulfide Arrangement by MALDI Mass Spectrometry

Amit Kumar Mandal,^a Mani Ramakrishnan Santhana Ramasamy,^b Varatharajan Sabareesh,^a Matthew E. Openshaw,^c Kozhalmannom S. Krishnan,^{b,d} and Padmanabhan Balaram^a

^a Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

^b National Center for Biological Sciences, Bangalore, India

^c Shimadzu Biotech, Manchester, United Kingdom

^d Department of Biological Sciences, Tata Institute of Fundamental Research, Mumbai, India

De novo mass spectrometric sequencing of two *Conus* peptides, Vi1359 and Vi1361, from the vermicorous cone snail *Conus virgo*, found off the southern Indian coast, is presented. The peptides, whose masses differ only by 2 Da, possess two disulfide bonds and an amidated C-terminus. Simple chemical modifications and enzymatic cleavage coupled with matrix assisted laser desorption ionization (MALDI) mass spectrometric analysis aided in establishing the sequences of Vi1359, ZCCITIPECCRI-NH₂, and Vi1361, ZCCPTMPECCRI-NH₂, which differ only at residues 4 and 6 (Z = pyroglutamic acid). The presence of the pyroglutamyl residue at the N-terminus was unambiguously identified by chemical hydrolysis of the cyclic amide, followed by esterification. The presence of Ile residues in both the peptides was confirmed from high-energy collision induced dissociation (CID) studies, using the observation of w_n- and d_n-ions as a diagnostic. Differential cysteine labeling, in conjunction with MALDI-MS/MS, permitted establishment of disulfide connectivity in both peptides as Cys2-Cys9 and Cys3-Cys10. The cysteine pattern clearly reveals that the peptides belong to the class of T-superfamily conotoxins, in particular the T-1 superfamily. (J Am Soc Mass Spectrom 2007, 18, 1396–1404) © 2007 American Society for Mass Spectrometry

De novo sequencing by mass spectrometry is rapidly emerging as the method of choice for the characterization of the peptides from natural sources [1–4]. Contemporary MS techniques coupled with HPLC separation permit the rapid analysis of complex natural peptide libraries [5–7]. As part of a program to isolate peptides from the venom of cone snails found off the coast of India [1, 3, 8, 9], we have examined peptide mixtures obtained from the venom ducts of *Conus virgo*, a vermicorous snail. In a recent report, Quinton et al. have reported the use of Fourier transform mass spectrometry in the characterization of two major peptide components from *Conus virgo*, with molecular masses 1328.5 (ViVA) and 1358.5(ViVB) [2]. These authors have emphasized the importance of the high mass accuracy provided by FTMS for de novo sequencing. But, more importantly, they did not report the disulfide bridging arrangement or the distinction between the isobaric residues leucine (Leu), isoleucine (Ile), and hydroxyproline (Hyp).

We had independently been investigating the peptide

components of the venom from *Conus virgo* and had completed the sequence assignment of the peptide with mass 1358.5 Da (ViVB, following the nomenclature of Quinton et al., labeled as Vi1359 in this study) using MALDI mass spectrometry. In the present paper, we establish the disulfide connectivity and use collision induced dissociation (CID) to establish the presence of three Ile residues in Vi1359. Further, we describe a simple, selective hydrolysis procedure that establishes the presence of pyroglutamic acid in Vi1359. Additionally, we have also characterized a closely related peptide with mass 1360.6 Da (Vi1361), which shows variation at two positions compared with the sequence of Vi1359. Both peptides belong to the T-superfamily of conotoxins. Our approach in this study demonstrates that direct sequencing is possible even when the accuracy of mass determination is significantly lower.

Materials and Methods

Materials

Dithiothreitol (DTT), iodoacetamide (IAM), N-ethylmaleimide (NEM), tris (2-carboxyethyl) phosphine,

Address reprint requests to Dr. P. Balaram, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India. E-mail: pb@mbu.iisc.ernet.in

α -cyano-4-hydroxycinnamic acid, trypsin (TPCK treated), endoproteinase gluC, and Sephadex G-10 were purchased from Sigma Chemical (St. Louis, MO). All other chemicals used were of analytical grade.

Methods

Peptide purification. The venom duct of *Conus virgo*, collected off the Southeast coast of India, was dissected and extracted in methanol. The crude methanol extract was centrifuged at 6000 rpm for 15 mins at 25 °C. The clear supernatant was taken in micro centrifuge tubes and concentrated in a Speed Vac for 1 h. The concentrated crude venom extract was subjected to reverse phase HPLC purification using a Phenomenex C₁₈ (10 mm × 250 mm, 4 μ particle size, 90 Å pore size) column. A linear gradient of acetonitrile and water containing 0.1% trifluoroacetic acid was used. The flow rate was 1 mL/min and absorbance was monitored at 226 nm.

Reduction and alkylation. Ten μ l of purified peptide was dissolved in 20 μ l 50 mM NH₄HCO₃ buffer pH 8. Reduction of the peptide was carried out by adding a DTT stock solution to a final concentration of 70 mM, in the presence of 50 mM NH₄HCO₃ buffer (pH 8) followed by incubation at 37 °C for 3 h. For alkylation, an IAM stock solution was added to the above reaction mixture to maintain a final concentration of 50 mM. The reaction was allowed to proceed for 1 h at 25 °C in the dark. The resulting reaction mixture was analyzed by MALDI-MS.

Acetylation. To 5 μ l of reduced and alkylated Vi1359, 2 μ l of 0.5 M NH₄HCO₃ pH 8.0, and 2 μ l acetic anhydride were added and the volume was made up to 20 μ l with distilled water. After incubating the reaction mixture for 1 h at 25 °C, the product was analyzed in MALDI-MS.

Esterification. Methanolic HCl was prepared by dropwise addition of 20 μ l acetyl chloride to 100 μ l of ice cold, anhydrous methanol. Four μ l of reduced and alkylated Vi1359 was added to 16 μ l of methanolic HCl. The esterification was allowed to proceed at 25 °C. The reaction mixture was analyzed by MALDI-MS at different incubation times of 1, 6, and 72 h.

Disulfide connectivity. Partial reduction and alkylation of disulfide bonds of Vi1359 was performed by following the kinetics of reduction using tris (2-carboxyethyl) phosphine (TCEP) solution as a reducing agent with a final concentration of 25 mM in 50 mM sodium acetate pH 3.0. All reduction reactions were carried out by chilling the solution in an ice bath. After 6 min of reduction, a solution of NEM was added to a final concentration of 90 mM to alkylate free cysteines. Alkylation was allowed to proceed for 40 min in the dark. The whole reaction mixture was passed through a Sephadex G-10 spin column to remove salts and excess

reagents. The pH was adjusted to 8.0 by adding a stock solution of NH₄HCO₃. DTT was added to a final concentration of 50 mM. Reduction of the remaining disulfide bond was allowed to proceed for 2.5 h at 37 °C. Alkylation of the free thiol of cysteine residues was achieved by adding a stock solution of IAM to a final concentration of 70 mM and allowing the reaction to proceed at 25 °C in the dark. The differentially reduced and alkylated peptide was analyzed by MALDI-MS.

Mass spectrometry. Electrospray ionization (ESI) mass spectra were obtained on a Hewlett Packard, HP 1100 MSD series spectrometer equipped with a single quadrupole analyzer (Palo Alto, CA). The data were acquired over a range of 200 to 3000 *m/z* in positive ion mode and analyzed using HP LC/MSD Chemstation software (Palo Alto, CA). Mass scale calibration was done using ES tuning mix supplied by the manufacturer covering a mass range of 118.09 to 2121.9 *m/z*.

MALDI-TOF-MS analyses were performed using an Ultra flex TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) in reflectron (positive ion) mode, using a 90 ns time delay, and a 25 kV accelerating voltage in the positive ion mode. Identical conditions were maintained while analyzing samples in the negative ion mode. The system utilizes a 50 Hz pulsed nitrogen laser, emitting at 337 nm. The ion source and the flight tube were kept at a pressure of about 7×10^{-7} mbar by turbo molecular pumps. The sample was prepared by mixing an equal volume of peptide solution and a saturated solution of matrix (α -cyano-4-hydroxycinnamic acid) in 1:1 (vol/vol) acetonitrile-water mixture. A standard peptide mixture was used for external calibration.

MS-MS spectra were acquired by selecting the precursor mass with a 10 Da window, and fragments were generated in post-source decay mode (PSD). A single acquisition was the sum of at least 50 series with 600 shots added shots to generate the MS-MS spectra. Mass spectra were analyzed using Flex-Analysis software (Bruker Daltonics).

The high-energy collision induced dissociation (CID) MALDI-MS/MS data were acquired using an AXIMA-TOF² MALDI-TOF mass spectrometer (Shimadzu Biotech/Kratos, Manchester, UK). Data were acquired in reflectron (positive ion) mode, with the pulsed extraction optimized for mass 1500. For MS/MS experiments, helium was used as collision gas. As the ion beam is not decelerated at any point following acceleration from the ion source, any collisions between the analyte ions and helium molecules occur with a collision energy of 20 keV (laboratory frame of reference).

Results

The mass detected HPLC profile (ESI-LC/MS) of a crude venom extract from *Conus virgo* reveals the presence of several peptide components corresponding to

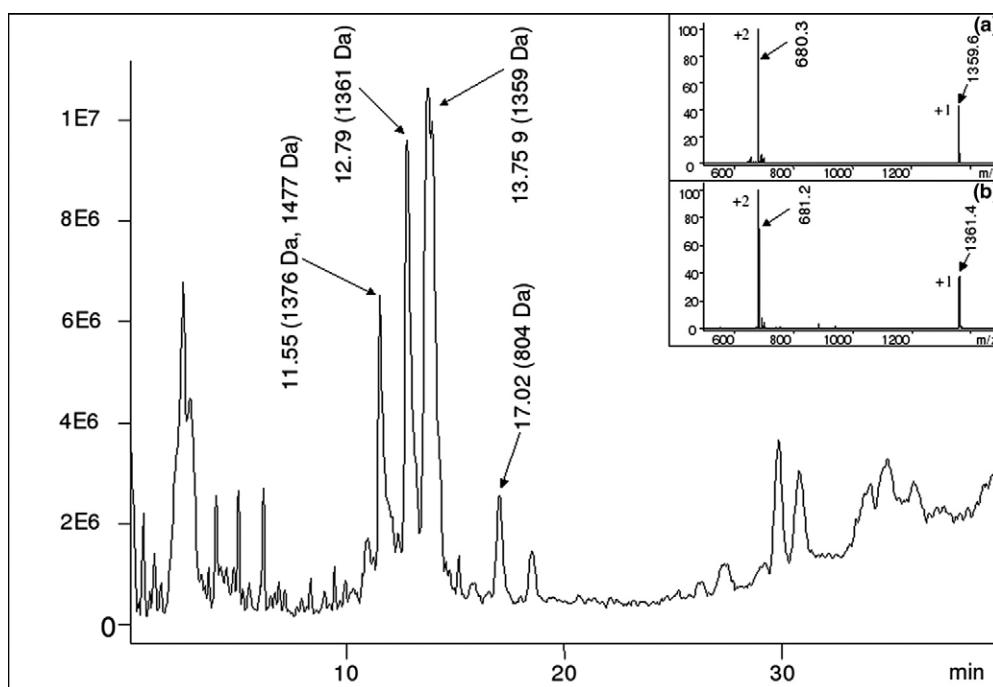


Figure 1. Total ion chromatogram obtained from LC-ESI/MS of methanol extract of crude venom from *C. virgo*. ESI mass spectra of two major peaks in HPLC profile has been shown as insets: (a) Vi1359, (b) Vi1361.

diverse *Conus* peptides (Figure 1). Two major peaks yielding masses of 1359.6 Da $[M + H]^+$ and 1361.6 Da $[M + H]^+$, shown as inset in Figure 1, were collected for further analysis. Reduction of Vi1359 with DTT followed by alkylation with IAM resulted in a product with a mass of 1591.6 Da. The mass change ($\Delta M = 232$ Da) is consistent with the presence of two disulfide bonds formed by four Cys residues in the sequence. Since the objective of the study was to establish the complete constitution of *Conus* peptides using mass spectrometry, we undertook preliminary studies employing functional group modification to facilitate the subsequent mass spectrometric sequencing.

Chemical Modification

Chemical modification studies were undertaken to identify the presence of specific functional groups. Acetylation did not yield a modified product, indicating the absence of primary amino groups, pointing to the absence of lysine (Lys) residues and the presence of an N-terminal blocking group. Figure 2 shows the MALDI mass spectrum, obtained following esterification, using a solution of acetyl chloride in dry methanol. Esterification at a single site proceeds rapidly with the almost complete disappearance of the parent ion at 1359.6 Da and the appearance of modified peptide ion at 1373.7 Da. The observed ΔM of 14 Da is consistent with the addition of a single methyl group, supporting the presence of one $-COOH$ group. After prolonged incubation under esterification conditions, two prominent species are observed at 1405.8 and 1407.8 Da. The

distribution of intensities within the isotopic cluster argues for the presence of the distinct chemical entities differing in mass by 2 Da. In addition, a low intensity cluster is observed centered at ~ 1390 Da. Inspection of this group of peaks (Figure 2b) suggests that multiple species contribute to this mass spectral feature. Tentatively, three distinct species, 1388.6, 1389.6, and 1391.6 Da, may be assigned based on the distribution of intensities. Prolonged incubation (72 h) in acetyl chloride and methanol results in the spectrum shown in Figure 2c. It is evident that the major species observed are at 1405.8 and 1420.8 Da. The peak at highest m/z value of 1420.8 exhibits the anticipated distribution of intensities for the isotopic peaks, suggesting that it arises from a single homogeneous molecular species. Under conditions of the esterification reaction, hydrolysis of an amide bond releasing a free carboxylic acid group is possible [10]. The final mass of 1420.8 Da corresponds to hydrolysis (+1 Da) and subsequent esterification at two additional sites (+28 Da). The observed mass difference between the single esterified species 1373.7 Da and the final product is 47 Da. This is consistent with the presence of a single amide group (hydrolysis + esterification) and a pyroglutamyl residue (hydrolysis + esterification). The intensity distribution in the isotopic clusters of intermediate species are may be explained by the presence of free carboxylic acid groups as a consequence of amide hydrolysis (+1 Da), which are subsequently esterified. Primary amide hydrolysis proceeds more rapidly than peptide backbone secondary amide bonds. Notably, hydrolysis of the cyclic secondary amide of the pyroglutamyl residue is

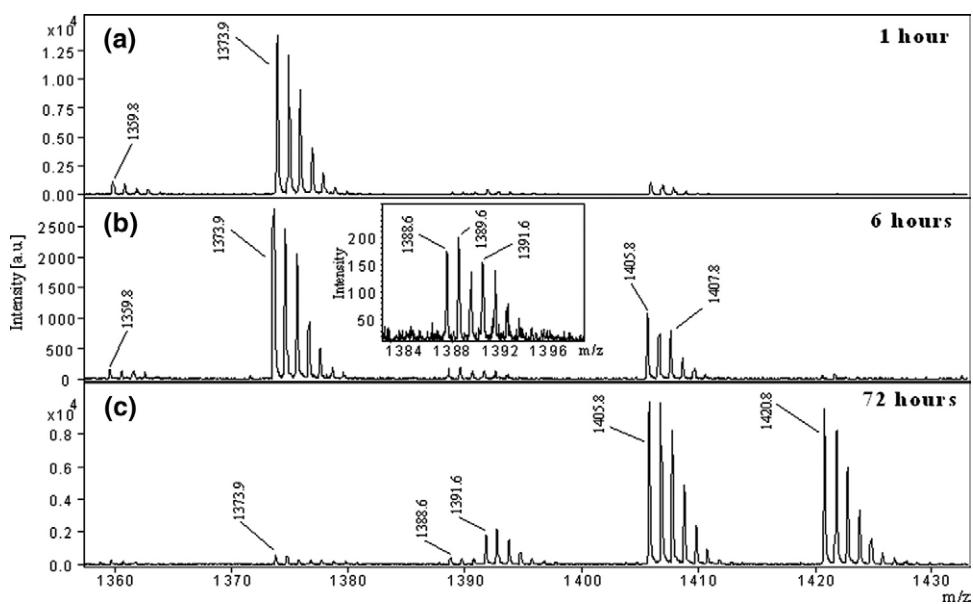


Figure 2. MALDI mass spectra obtained as a function of time during esterification of Vi1359 by incubation in a solution of acetyl chloride and dry methanol. (b) inset) Cluster of peaks centered at ~ 1390 Da. For all ions adjacent peaks are separated by 1 Da.

even more rapid. The chemical modification studies thus establish the absence of primary amino groups, the presence of a pyroglutamic acid residue at the N-terminus and a primary amide group in the sequence, which may arise from the presence of C-terminal amidation, a common post-translation modification observed in *Conus* peptides [11].

Enzymatic Cleavage

Tryptic cleavage of reduced and alkylated Vi1359 ($[M + H]^+ = 1591.6$ Da) yields a peptide of mass 1479.7 Da, establishing the presence of a single Arg residue in the sequence. The mass difference of 112 Da suggests that the C-terminus residue is Leu/Ile with the C-terminus being amidated. The isobaric residue hydroxyproline (Hyp), 113 Da, is eliminated since the Arg/Lys-Pro is normally resistant to cleavage by trypsin [12].

Digestion with the enzyme gluC yielded a peptide fragment with mass 607.4 Da, establishing the presence of a single glutamic acid (Glu) residue in the sequence. This enzyme cleaves peptides at the C-terminus of Glu residues at pH 8.0 [12].

MS/MS Sequencing

Sequencing of Vi1359 was undertaken using MALDI-MS-MS spectra obtained on both the intact peptide and a tryptic fragment, which lacked the C-terminus L/I residue. In both cases, rich fragment ion spectra were obtained with intense peaks being distributed over the m/z region 200 to 900. Figure 3 illustrates the observed fragmentation pattern obtained for the intact peptide Vi1359 after reduction and alkylation ($[M + H]^+ = 1591.5$ Da). Four pairs of "b" and "y" ions (272.3/1319.8,

432.2/1160.2, 546.4/1047.4, 647.4/946.6) could be readily recognized. Further extension of the "y" series could also be achieved by inspection of mass differences. The "b₂" ion with m/z 272.3 was assigned tentatively to the dipeptide sequence ZC* (Z = pyroglutamic acid, C* = carboxyamidomethyl cysteine). Esterification data suggested the absence of a C-terminus carboxylic acid residue, implying the presence of a C-terminus amide, which is a common post-translation modification in *Conus* peptides. The presence of the "y₁" ion at m/z 175.9 in the MS/MS spectrum of the tryptic fragment of reduced and alkylated Vi1359 (data not shown), losses of ammonia from "y₆" and "y₄" ions (Figure 3) are consistent with the presence of an arginine residue at the C-terminus. The intense "b₅-H₂O" peak at m/z 629.4 confirms the presence of Thr at position 5. The immonium ions corresponding to Pro, Leu/Ile, and C* are also readily identified. The sequence Z C* C* L₁ T₁ P EC* C* R₁ -NH₂ may thus be derived directly from the observed MS/MS spectrum. Acetylation experiments using pyridine/acetic anhydride did not yield additional acetylation sites, suggesting the absence of Hyp at positions 4 and 6 in Vi1359.

Further, the sequence of Vi1361 was deduced from the MALDI-MS/MS spectrum of its reduced and alkylated product ($[M + H]^+ = 1593.6$ Da) (Figure 4). The sequence of Vi1361 differs from that of Vi1359 at only two positions. The residues at positions 4 and 6 in Vi1359 are replaced by Pro and Met, respectively, in Vi1361.

Determination of Disulfide Connectivity

Vi1359 possesses four Cys residues and two disulfide bridges resulting in three possible disulfide pairing

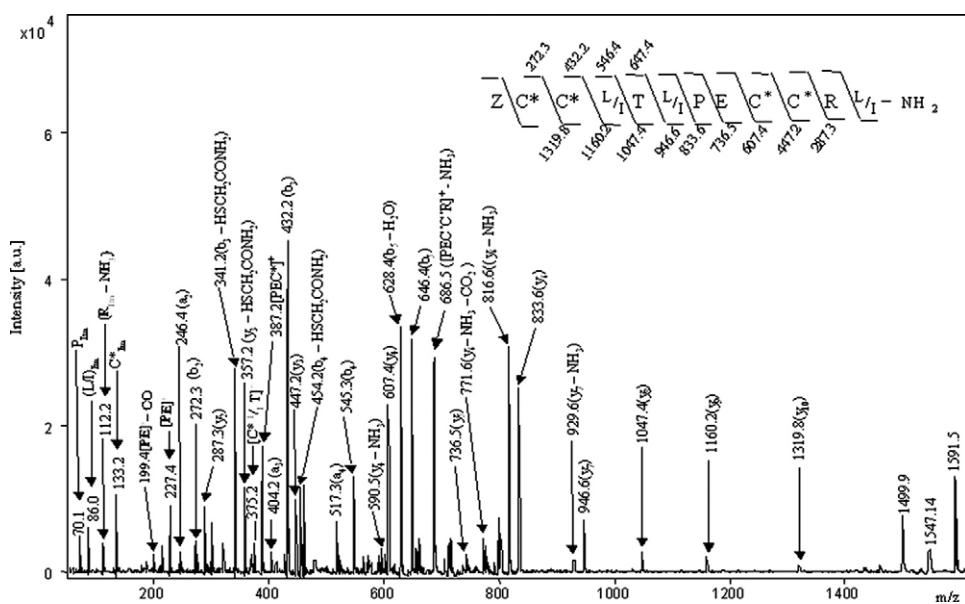


Figure 3. MALDI-MS/MS spectrum of reduced and alkylated Vi1359 ($[M + H]^+ = 1591.5$ Da). Series of "y" and "b" ions and molecular ions obtained by neutral losses are labeled.

schemes, C_2/C_3 , C_9/C_{10} ; C_2/C_9 , C_3/C_{10} ; C_2/C_{10} , C_3/C_9 . Selective reduction conditions were standardized to yield the species, which contain only one disulfide bond (see in the Methods section). At this point, alkylation was effected with N-ethylmaleimide (NEM) resulting in a mass increase of 126 Da per sulfhydryl group. Subsequent reduction with DTT, followed by alkylation with IAM, yielded a peptide containing four modified Cys residues, two of which were carboxyamidomethylated (+58 Da), while the other two were modified with

NEM. Scheme 1 summarizes the strategy adopted and highlights the key fragment, which permits the unambiguous determination of the connectivity pattern. It may be noted that the third disulfide paring arrangement, which requires disulfide formation between adjacent Cys residues is relatively rare. Figure 5 shows the MALDI-MS/MS spectrum of the modified Vi1359 ($[M + H]^+ = 1727.4$ Da). Once again the observation of an almost complete series of y-ions, together with a few b-ions, permitted identification of the nature of chemi-

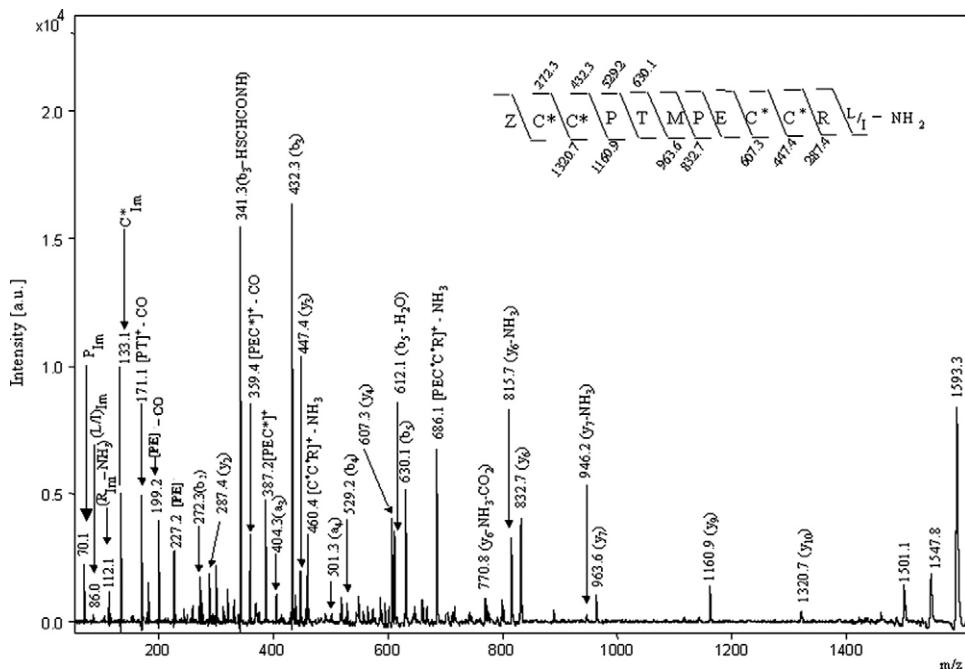
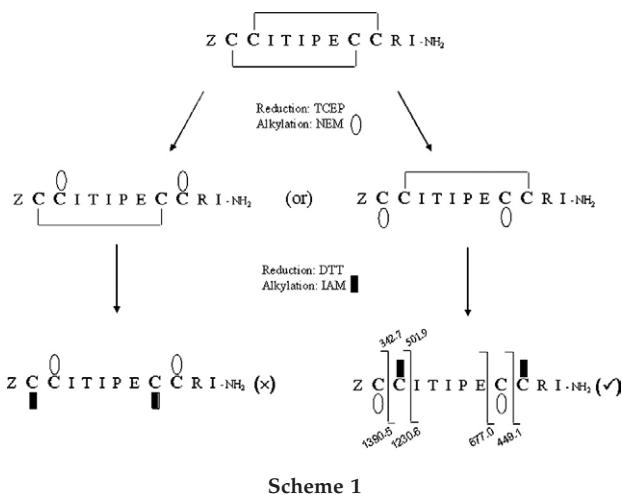


Figure 4. MALDI-MS/MS spectrum of reduced and alkylated Vi1361 ($[M + H]^+ = 1593.5$ Da). Series of "y" and "b" ions and molecular ions obtained by neutral losses are labeled.



cal modification of the different Cys residues in the sequence. Cys (2) and Cys (9) were modified with NEM and Cys (3) and Cys (10) were carboxyamidomethylated. The disulfide bonding pattern is thus established as Cys (2) to Cys (9) and Cys (3) to Cys (10). Likewise the disulfide connectivity of Vi1361 was also determined to be same as that of Vi1359, Cys (2) to Cys (9) and Cys (3) to Cys (10) (data not shown).

Leu/Ile Differentiation Using High-Energy Collision Induced Dissociation Study

The primary sequence determination of Vi1359 by MALDI-MS/MS procedures described above still contains ambiguities at positions 4, 6, 12, which correspond to residue masses of 113 Da and are assigned to Leu/Ile. The Leu/Ile differentiation was

achieved by high-energy collision induced dissociation (CID) (vide methodology). Figure 6 shows the high-energy CID spectra for carboxyamidomethylated Vi1359 (precursor ion m/z 1591.6). Figure 7 illustrates an expanded view of specific ions obtained by side-chain fragmentation under high-energy CID condition. The observation of v_n -type ions (v_7^- and v_9^- in this case) confirms the presence of Leu/Ile at positions 4 and 6 in Vi1359. However, v_n -ions that arise by the $C^\alpha-C^\beta$ cleavage do not aid in distinguishing between Leu and Ile. The d_n - and w_n -type ions, arises from side-chain cleavage at the $C^\beta-C^\gamma$ bond, facilitate in distinguishing Leu from Ile unambiguously [13]. The observation of w_n - and d_n -ions, at m/z 901.34 (w_7^a), 1115.50 (w_9^a), and 1518.66 (d_{12}^a) in the high-energy CID spectrum of carboxyamidomethylated Vi1359 (Figure 6) are all supportive of the presence of Ile at positions 4, 6, and 12. Conversely, Leu at all these three positions should have resulted in peaks at m/z 887.45 (w_7^a), 1101.59 (w_9^a), and 1504.68 (d_{12}^a) resulting from the side-chain fragmentation of Leu, which are, in fact, absent in the high-energy CID spectrum, thereby ruling out the possibility of any Leu residues in Vi1359. Furthermore, the presence of "Ile" at positions 4, 6, and 12 in Vi1359 is also confirmed by 700 MHz 1 H NMR spectroscopy, which reveals the presence of three sets of triplet resonance corresponding to the $C^{\beta}H_3$ groups of Ile (data not shown). Similarly, high-energy CID spectra of carboxyamidomethylated Vi1361 ($[M + H]^+ = 1593.6$ Da) aided in the identification of Ile at the C-terminus as evidenced by the presence of d_{12}^a and d_{-12}^b fragment ions (data not shown).

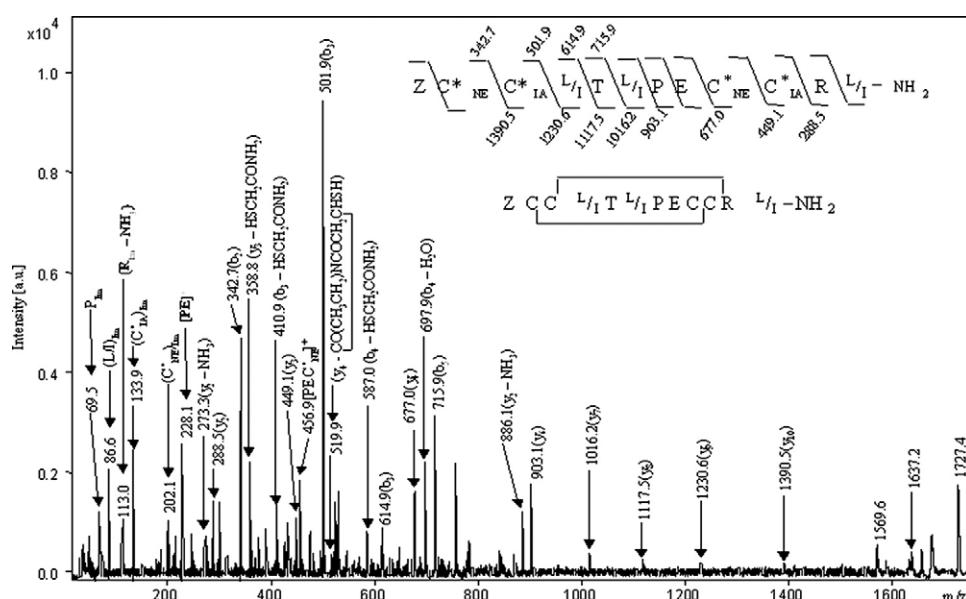


Figure 5. MALDI-MS/MS spectrum of differentially reduced and alkylated Vi1359 ($[M + H]^+ = 1727.4$ Da). Series of "y" and "b" ions and molecular ions obtained by neutral losses are labeled. Inset shows the sequence assignment of the peptide molecular ion of mass 1727.4 Da. The complete structure of Vi1359 is shown.

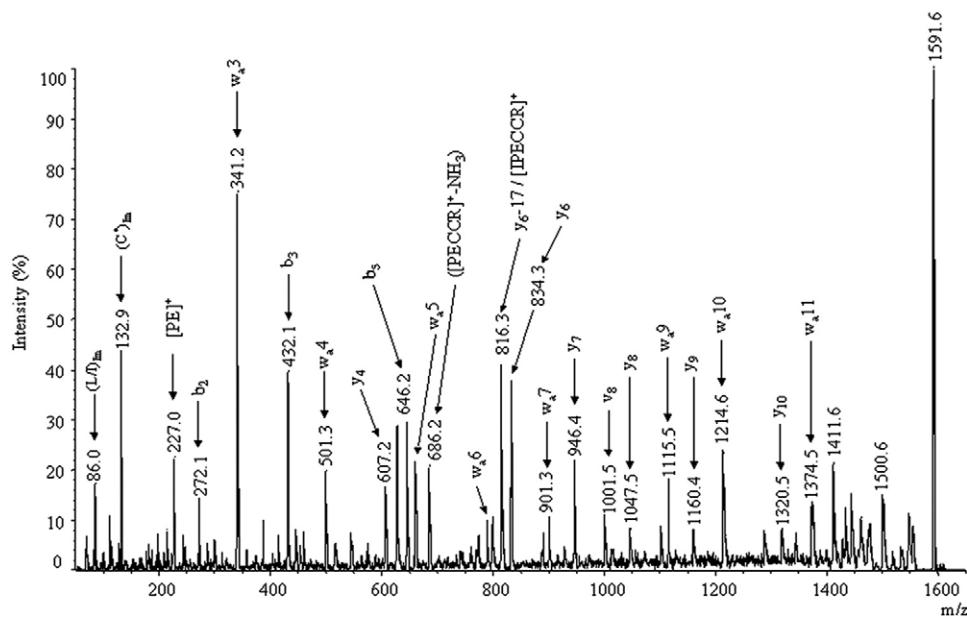


Figure 6. MALDI high-energy CID MS/MS spectrum of reduced/alkylated Vi1359 (precursor ion m/z 1591.7). Series of “y”, “b”, “v”, and “w” ions are labeled.

Discussion

Comparison of the sequences of Vi1359 and Vi1361 with the reported sequences of *Conus* peptides reveal that the two peptides belong to the T-superfamily, more specifically, T-1-superfamily. T-1 conotoxins are characterized by a CCxxxxxCC motif with 1-3, 2-4 disulfide connectivity [2, 5, 14–24], while the cysteine framework in T-2 conotoxins is CCxxxxxCxxC, with 1-4, 2-3 disulfide connectivity [25–27]. Table 1 shows

the sequence alignment (based on the CC pair) of *Conus* peptides belonging to the T-1-superfamily, isolated from different species. The methodologies followed for the determination of sequences are also indicated. Two distinct groups can be identified in Table 1, based on the number of residues present between the CC pair. The first 6 sequences (ε -TxIX/tx5a/Tx5.2, vc5a/Vc5.1, mr5a and TeA31, TeAr193, Mo1274) possess the CC-x (4)-CC motif, while the

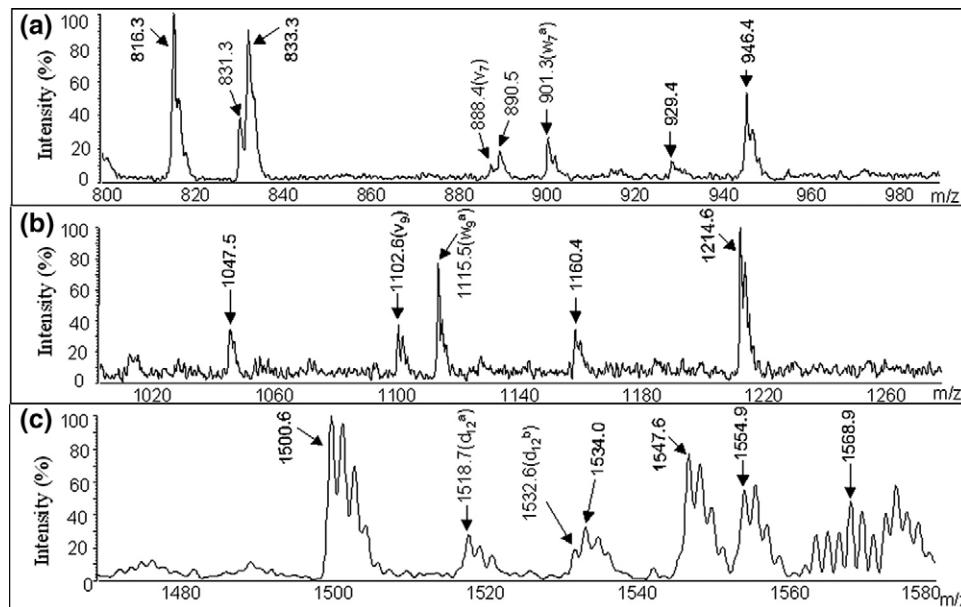


Figure 7. MALDI high-energy CID MS/MS spectrum of reduced/alkylated Vi1359 (precursor ion m/z 1591.7); (a) edited between m/z 800 and 990, showing peaks corresponding to v_7 and w_7^a ; (b) edited between m/z 1005 and 1280, showing peaks corresponding to v_9 and w_9^a ; (c) edited between m/z 1470 and 1580, showing peaks corresponding to d_{12}^a and d_{12}^b .

Table 1. Peptides belonging to T-1-superfamily, isolated from different *Conus* species

	Sequences	Name of the peptide	Species	Methodologies	References
1	<u>ECC</u> EDGX <u>CC</u> TAAO	<i>e</i> -TxIX, tx5a/Tx5.2	<i>C. textile</i>	Edman, MS, NMR, cDNA ^b	[22–24]
2	<u>CC</u> PGK <u>OC</u> CR <u>I</u> ^a	vc5a/Vc5.1	<i>C. victoriae</i>	cDNA, MS	[20]
3	N <u>AC</u> CI <u>VR</u> Q <u>CC</u>	mr5a	<i>C. marmoreus</i>	Edman, MS	[19]
4	N <u>CC</u> RRQ <u>I</u> <u>CC</u> GRTK	TeAr193	<i>C. textile</i>	cDNA	[17]
5	GNX <u>CC</u> SARV <u>CC</u>	Mo1274	<i>C. monile</i>	MS ^b	[16]
6	<u>GCC</u> PKQMR <u>CC</u> TL ^a	p5a/P5.1	<i>C. purpurascens</i>	MS, cDNA, cs ^b	[24]
7	<u>FCC</u> PFIRY <u>CC</u> W	au5a	<i>C. aulicus</i>	Edman, MS, cs ^b	[24]
8	<u>FCC</u> PVIRY <u>CC</u> W	au5b	<i>C. aulicus</i>	Edman, MS	[24]
9	<u>CC</u> QTFYW <u>CC</u> VQGK	Tx5.1	<i>C. textile</i>	cDNA	[24]
10	L <u>CC</u> VTEDW <u>CC</u> EW ^a W	Gm5.1	<i>C. gloriamaris</i>	cDNA	[24]
11	V <u>CC</u> RPVQD <u>CC</u> SGK	Gm5.2	<i>C. gloriamaris</i>	cDNA	[24]
12	DWN <u>CC</u> GKNPG <u>CC</u> PWGK	Im5.1	<i>C. imperialis</i>	cDNA	[24]
13	I <u>CC</u> YPN <u>EX</u> CCD	vc5c/Vc5.2	<i>C. victoriae</i>	cDNA, MS	[15, 20]
14	VN <u>CC</u> GID <u>ES</u> CCS	Vc5.3	<i>C. victoriae</i>	cDNA	[20]
15	<u>CC</u> QTFYW <u>CC</u> CGO ^a	vc5b/Vc5.4	<i>C. victoriae</i>	cDNA	[15, 20]
16	F <u>CC</u> RTQEV <u>CC</u> EA ^a IKN ^a	GlaMrIII	<i>C. marmoreus</i>	Edman, MS	[21]
17	<u>CC</u> ITFES <u>CC</u> EF ^a DL	GlaMrIV	<i>C. marmoreus</i>	Edman, MS	[21]
18	<u>CC</u> PGWEL <u>CC</u> EW ^a DEW	Mr5.1a	<i>C. marmoreus</i>	cDNA	[19]
19	<u>CC</u> PGWEL <u>CC</u> EWDDGW	Mr5.1b	<i>C. marmoreus</i>	cDNA	[19]
20	F <u>CC</u> RTQEV <u>CC</u> EA ^a IKNG	Mr5.2	<i>C. marmoreus</i>	cDNA	[19]
21	<u>CC</u> ITFES <u>CC</u> EF ^a DLK	Mr5.3	<i>C. marmoreus</i>	cDNA	[19]
22	<u>CC</u> QVMPQ <u>CC</u> EW ^a N	Mr5.4a	<i>C. marmoreus</i>	cDNA	[19]
23	<u>CC</u> QIVPQ <u>CC</u> EW ^a N	Mr5.4b	<i>C. marmoreus</i>	cDNA	[19]
24	IINW <u>CC</u> LIFY <u>Q</u> CC	sr5a	<i>C. spurius</i>	Edman, MS ^b	[14]
25	<u>SC</u> CPQEFL <u>CC</u> LYLVK	Lp5.1	<i>C. leopardus</i>	cDNA	[18]
26	GSV <u>CC</u> KVDT <u>SC</u> CSN	Lp5.2	<i>C. leopardus</i>	cDNA	[18]
27	Z <u>CC</u> L <u>TL</u> PE <u>CC</u> R <u>L</u> ^a	ViVA	<i>C. virgo</i>	MS	[2]
28	Z <u>CC</u> ITI <u>PE</u> <u>CC</u> R <u>I</u> ^a	Vi1359	<i>C. virgo</i>	MS ^b	This study
29	Z <u>CC</u> PTM <u>PE</u> <u>CC</u> R <u>I</u> ^a	Vi1361	<i>C. virgo</i>	MS ^b	This study
30	Z <u>CC</u> PTL <u>PE</u> <u>CC</u> R <u>V</u> ^a	ViVB	<i>C. virgo</i>	MS	[2]
31	I <u>CC</u> YPN <u>W</u> <u>CC</u> D	TeA31	<i>C. textile</i>	cDNA	[17]
32	VN <u>CC</u> GID <u>ES</u> CCS	TeAr154	<i>C. textile</i>	cDNA	[17]
33	V <u>CC</u> RPVQD <u>CC</u> SGK	TeAr151	<i>C. textile</i>	cDNA	[17]

E: gamma-carboxyglutamic acid; X: 6-bromotryptophan; T: glycosylated Thr; O: 4-hydroxyproline.

Z: pyroglutamic acid; L: leucine or isoleucine. Edman: Edman method of peptide sequencing; MS: Mass Spectrometry; NMR: nuclear magnetic resonance; cs: Chemical synthesis; cDNA: cDNA sequencing.

^aAmidated C-terminus.

^bDisulfide connectivity: Cys1-Cys3, Cys2-Cys4.

remaining 27 sequences have the CC-x (5)-CC motif. Recently, Quinton et al. have reported two peptide sequences ViVA ($M_r = 1358.5$) and ViVB ($M_r = 1328.5$) from *Conus virgo*. Of these, ViVA is identical to Vi1359. However, the report of Quinton et al., which used FTICR did not describe the disulfide assignment and the differentiation of Leu/Ile residues. While the study of Quinton et al. exploited the high mass accuracy of FTICR-MS to establish the presence of an N-terminus pyroglutamyl residue, the present report describes a simple chemical modification protocol that can be used in conjunction with normal MALDI mass spectrometers. It is also evident from Table 1 that the majority of sequences determined, thus far, are derived from cDNA sequences, which do not provide an unambiguous primary structure for the mature toxins that arise as a consequence of post-translational processes. Edman procedures need to be necessarily complemented by mass spectrometry in the case of post-translationally mod-

ified peptides. The potential of de novo mass spectrometric sequencing is established in this study. Of the 33 sequences of T-1 conotoxins listed, the disulfide connectivity has been established only in the case of seven peptides, inclusive of the two peptides investigated in this study [14, 16, 22–24]. The peptides from *Conus virgo* determined in this study and by Quinton et al. are the first cases reporting the occurrence of pyroglutamic acid in the T-superfamily of conotoxins. In the context of conotoxins, thus far, pyroglutamyl residue has only been established in a seven residue conopeptide from *Conus imperialis* [28]. In that report, Craig et al. used enzymatic digestion with pyroglutamate aminopeptidase, to confirm the occurrence of this unique post-translationally modified residue. To the best of our knowledge, the present report is the first instance of the identification of pyroglutamic acid by chemical hydrolysis followed by mass spectrometry in conotoxins.

Leu/Ile differentiation has been achieved in the present study by high-energy CID spectra, which per-

mit the observation of w_n (w_7 and w_9) and d_n (d_{12}) type ions arising from side-chain fragmentation, in addition to backbone fragmentation. Recently, the task of Leu/Ile differentiation has also been accomplished unambiguously by (hot) electron capture dissociation (ECD/high-energy CID), as this dissociation mode gives rise to abundant w_n type ions [29–31].

Conclusions

MALDI MS/MS methods coupled with chemical modification reactions permit complete sequence determination and establishment of disulfide bridge arrangements in *Conus* peptides possessing post-translationally modified pyroglutamyl residues.

Acknowledgments

The authors gratefully acknowledge the help of Y. K. Subrahmanyam Prakash in conducting MS experiments. This research was supported by a grant from the Department of Biotechnology, Government of India. AKM and MRSR are recipients of postdoctoral fellowships from the Department of Biotechnology, Government of India. VS is the recipient of a Senior Research Fellowship of the Council of Scientific and Industrial Research, Government of India. The mass spectrometry facilities at the Indian Institute of Science, Bangalore, are supported by Proteomics Program Grant from the Department of Biotechnology, Government of India.

References

1. Sabareesh, V.; Hanumae-Gowd, K.; Ramasamy, P.; Sudarslal, S.; Krishnan, K. S.; Sikdar, S. K.; Balaram, P. Characterization of Contryphans from *Conus loroisi* and *Conus amadis* that Target Calcium Channels. *Peptides* **2006**, *27*, 2647–2654.
2. Quinton, L.; Le Caer, J.-P.; Vinh, J.; Gilles, N.; Chamot-Rooke, J. Fourier Transform Mass Spectrometry: A Powerful Tool for Toxin Analysis. *Toxicon* **2006**, *47*, 715–726.
3. Gowd, K. H.; Sabareesh, V.; Sudarslal, S.; Iengar, P.; Franklin, B.; Fernando, A.; Dewan, K.; Ramaswami, M.; Sarma, S. P.; Sikdar, S.; Balaram, P.; Krishnan, K. S. Novel Peptides of Therapeutic Promise from Indian Conidae. *Ann. N.Y. Acad. Sci.* **2005**, *1056*, 462–473.
4. Sudarslal, S.; Singaravelivelan, G.; Ramasamy, P.; Ananda, K.; Sarma, S. P.; Sikdar, S. K.; Krishnan, K. S.; Balaram, P. A Novel 13 Residue Acyclic Peptide from the Marine Snail, *Conus morni*, Targets Potassium Channels. *Biochem. Biophys. Res. Commun.* **2004**, *317*, 682–688.
5. Jakubowski, J. A.; Sweedler, J. V. Sequencing and Mass Profiling Highly Modified Conotoxins Using Global Reduction/Akylation Followed by Mass Spectrometry. *Anal. Chem.* **2004**, *76*, 6541–6547.
6. Peng, J.; Gygi, S. P. Proteomics: The Move to Mixtures. *J. Mass Spectrom.* **2001**, *36*, 1083–1091.
7. Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. *Nature* **2003**, *422*, 198–207.
8. Sarma, S. P.; Senthil Kumar, G.; Sudarslal, S.; Iengar, P.; Ramasamy, P.; Sikdar, S. K.; Krishnan, K. S.; Balaram, P. Solution Structure of δ -Am2766: A Highly Hydrophobic δ -Conotoxin from *Conus amadis* that Inhibits Inactivation of Neuronal Voltage Gated Sodium Channels. *Chem. Biodiv.* **2005**, *2*, 535–556.
9. Sudarslal, S.; Majumdar, S.; Ramasamy, P.; Dhawan, R.; Pal, P. P.; Ramaswami, M.; Lala, A. K.; Sikdar, S. K.; Sarma, S. P.; Krishnan, K. S.; Balaram, P. Sodium Channel Modulating Activity in a δ -Conotoxin from an Indian marine snail. *FEBS Lett.* **2003**, *553*, 209–212.
10. Kim, T. Y.; Brun, Y. V.; Reilly, J. P. Effects of Tryptic Peptide Esterification in MALDI Mass Spectrometry. *Anal. Chem.* **2005**, *77*, 4185–4193.
11. Craig, A. G.; Badyopadhyay, P.; Olivera, B. M. Post-Translationally Modified Neuropeptides from *Conus* venoms. *Eur. J. Biochem.* **1999**, *264*, 271–275.
12. Kinter, M.; Sherman, N. E. *Protein Sequencing and Identification Using Tandem Mass Spectrometry*; Wiley-Interscience on Mass Spectrometry: New York, 2000; p 149.
13. Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass Spectrometer: Differentiation of Leucine and Isoleucine. *Anal. Chem.* **1987**, *59*, 2621–2625.
14. Aguilar, M. B.; Lezama-Monfil, L.; Maillo, M.; Pedraza-Lara, H.; Lopez-E, Heimer-de la Cotera, E. P. A Biologically Active Hydrophobic T-1-Conotoxin from the Venom *Conus spurius*. *Peptides* **2006**, *27*, 500–505.
15. Jakubowski, J. A.; Kelly, W. P.; Sweedler, J. V. Screening for Post-Translational Modifications in Conotoxins Using Liquid Chromatography/Mass Spectrometry: An important component of conotoxin discovery. *Toxicon* **2006**, *47*, 688–699.
16. Nair, S. S.; Nilsson, C. L.; Emmett, M. R.; Schaub, T. M.; Gowd, K. H.; Thakur, S. S.; Krishnan, K. S.; Balaram, P.; Marshall, A. G. De Novo Sequencing and Disulfide Mapping of a Bromotryptophan-Containing Conotoxins by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2006**, *78*, 8082–8088.
17. Luo, S.; Zhangsun, D.; Zhang, B.; Chen, X.; Feng, J. Direct cDNA Cloning of Novel Conotoxins of the T-Superfamily from *Conus textile*. *Peptides* **2006**, *27*, 2640–2646.
18. Chen, W.-H.; Han, Y.-H.; Wang, Q.; Miao, X.-W.; Ou, L.; Shao, X.-X. cDNA Cloning of Two Novel T-Superfamily Conotoxins from *Conus leopardus*. *Acta Biochim. Biophys. Sin.* **2006**, *38*, 287–291.
19. Han, Y. H.; Wang, Q.; Jiang, H.; Miao, X. W.; Chen, J. S.; Chi, C. W. Sequence Diversity of T-Superfamily Conotoxins from *Conus marmoreus*. *Toxicon* **2005**, *45*, 481–487.
20. Jakubowski, J. A.; Keays, D. A.; Kelly, W. P.; Sandall, D. W.; Bingam, J.-P.; Livett, B. G.; Gayler, K. R.; Sweedler, J. V. Determining Sequences and Post-Translational Modifications of Novel Conotoxins in *Conus victoriae* Using cDNA Sequencing and Mass Spectrometry. *J. Mass Spectrom.* **2004**, *39*, 548–557.
21. Hansson, K.; Furie, B.; Furie, B. C.; Stenflo, J. Isolation and Characterization of Three Novel Gla-Containing *Conus marmoreus* Venom Peptides, One with a Novel Cysteine Pattern. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 1081–1087.
22. Kalume, D. E.; Stenflo, J.; Czerwic, E.; Hambe, B.; Furie, B. C.; Furie, B.; Roepstorff, P. Structure Determination of Two Conotoxins from *Conus Textile* by a Combination of Matrix Assisted Laser Desorption/Ionization Time of Flight and Electrospray Ionization Mass Spectrometry and Biochemical Methods. *J. Mass Spectrom.* **2000**, *35*, 145–156.
23. Rigby, A. C.; Lucas-Meunier, E.; Kalume, D. E.; Czerwic, E.; Hambe, B.; Dahlqvist, I.; Fossier, P.; Baux, G.; Roepstorff, P.; Baleja, J. D.; Furie, B. C.; Furie, B.; Stenflo, J. A. Conotoxin from *Conus Textile* with Unusual Post-Translation Modifications Reduces Presynaptic Ca^{2+} Influx. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5758–5763.
24. Walker, C. S.; Steel, D.; Jacobsen, R. D.; Lirazan, M. B.; Cruz, L. J.; Hooper, D.; Shetty, R.; DelaCruz, R. C.; Nielsen, J. S.; Zhou, L. M.; Bandopadhyay, P.; Craig, A. G.; Olivera, B. M. The T-Super Family of Conotoxins. *J. Biol. Chem.* **1999**, *274*, 30664–30671.
25. McIntosh, J. M.; Corpuz, G. O.; Layer, R. T.; Garrett, J. E.; Wagstaff, J. D.; Bulaj, G.; Vyzovkina, Yoshikami, D.; Cruz, L. J.; Olivera, B. M. Isolation and Characterization of a novel *Conus* Peptide with Apparent Antinociceptive activity. *J. Biol. Chem.* **2000**, *275*, 32391–32397.
26. Balaji, R. A.; Ohtake, A.; Sato, K.; Gopalakrishnakone, P.; Kini, R. M.; Seow, K. T.; Bay, B. H. λ -Conotoxins, a New Family with Unique Disulphide Pattern and Protein Folding: Isolation and Characterization from the Venom of *Conus marmoreus*. *J. Biol. Chem.* **2000**, *275*, 39516–39522.
27. Sharpe, I. A.; Gehrmann, J.; Loughnan, M. L.; Thomas, L.; Adams, D. A.; Atkins, A.; Palant, E.; Craik, D. J.; Alewood, P. F.; Lewis, R. J. Two New Classes of Conopeptides Inhibit the α -1 Adrenoceptor and Noradrenaline Transporter. *Nat. Neurosci.* **2001**, *4*, 902–907.
28. Craig, A. G.; Jimenez, E. C.; Dykert, J.; Nielsen, D. B.; Gulyas, J.; Abogadie, F. C.; Porter, J.; Rivier, J. E.; Cruz, L. J.; Olivera, B. M.; McIntosh, J. M. A Novel Post-Translational Modification Involving Bromination of Tryptophan. Identification of the Residue, L-6-Bromo-tryptophan, in Peptides from *Conus imperialis* and *Conus radiatus* venom. *J. Biol. Chem.* **1997**, *272*, 4689–4698.
29. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. Dissociative Capture of Hot (3–13 eV) Electrons by Polypeptide Polycations: An Efficient Process Accompanied by Secondary Fragmentation. *Chem. Phys. Lett.* **2002**, *356*, 201–206.
30. Kjeldsen, F.; Haselmann, K. F.; Sorensen, E. S.; Zubarev, R. A. Distinguishing of Ile/Leu Amino Acid Residues in the PP3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2003**, *75*, 1267–1274.
31. Medzihradzsky, K. F. Peptide Sequence Analysis. *Methods Enzymol.* **2005**, *402*, 209–244.