De Novo Sequencing and Disulfide Mapping of a Bromotryptophan-Containing Conotoxin by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

ARTICLE *in* ANALYTICAL CHEMISTRY · JANUARY 2007

Impact Factor: 5.64 · DOI: 10.1021/ac0607764 · Source: PubMed

9 AUTHORS, INCLUDING:

Tanner Schaub
New Mexico State University
56 PUBLICATIONS **1,077** CITATIONS
SEE PROFILE

KS Krishnan
Tata Institute of Fundamental Research
64 PUBLICATIONS **1,514** CITATIONS
SEE PROFILE

Padmanabhan Balaram
Indian Institute of Science
507 PUBLICATIONS **14,222** CITATIONS
SEE PROFILE

Anal. Chem. **2006,** *78,* 8082−8088

De Novo Sequencing and Disulfide Mapping of a Bromotryptophan-Containing Conotoxin by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Sudarslal Sadasivan Nair,[†] Carol L. Nilsson,[†] Mark R. Emmett,[†] Tanner M. Schaub,[†] Konkallu Hanumae Gowd,[‡] Suman S. Thakur,[§] K. S. Krishnan,[‡,||] Padmanabhan Balaram,[§] and Alan G. Marshall*[,†]

National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, Florida 32310-4005, Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India, and Tata Institute of Fundamental Research, Mumbai- 400005, India

T-1-family conotoxins belong to the T-superfamily and are composed of 10−17 amino acids. They share a common cysteine framework and disulfide connectivity and exhibit unusual posttranslational modifications, such as tryptophan bromination, glutamic acid carboxylation, and threonine glycosylation. We have isolated and characterized a novel peptide, Mo1274, containing 11 amino acids, that shows the same cysteine pattern, -CC-CC, and disulfide linkage as those of the T-1-family members. The complete sequence, GNWCCSARVCC, in which W denotes bromotryptophan, was derived from MS-based de novo sequencing. The FT-ICR MS/MS techniques of electron capture dissociation (ECD), infrared multiphoton dissociation, and collision-induced dissociation served to detect and localize the tryptophan bromination. The bromine contributes a distinctive isotopic distribution in all fragments that contain bromotryptophan. ECD fragmentation results in the loss of bromine and return to the normal isotopic distribution. Disulfide connectivity of Mo1274, between cysteine pairs 1−3 and 2−4, was determined by mass spectrometry in combination with chemical derivatization employing tris(2-carboxyethyl)-phosphine, followed by differential alkylation with *N*-ethylmaleimide and iodoacetamide. The ECD spectra of the native and partially modified peptide reveal a loss of bromine in a process that requires the presence of a disulfide bond.

Conotoxins are disulfide-rich peptide toxins isolated from predatory marine snails.[1] The ability of these toxins to selectively block or modify various ion channels and receptors makes them interesting to neuroscientists. Depending on their molecular targets, conotoxins are grouped into seven superfamilies, primarily on the basis of the arrangement of cysteine residues in the mature peptide and homology in the signal sequence region.[2,3] Twenty-five percent of the nearly 200 known conotoxins (from the Swiss-Prot database) contain two disulfide bonds and most of them belong to either A-superfamily (composed of α and ρ family conotoxins) or T-superfamily (contributed by χ, ϵ, λ, and τ family members). Unlike other superfamilies, the T-superfamily members exhibit diverse primary structure, in terms of both the arrangement of cysteine residues and the disulfide connectivity. For example, toxins in the ϵ and τ families (in the current nomenclature, ϵ and τ are collectively called T-1-family[4]) share an identical cysteine framework of CC-CC and a common disulfide connectivity between Cys residues, 1−3 and 2−4,[5] whereas λ and χ toxins (λ and χ collectively belong to T-2-family conotoxins[4]) possess a CC-C-C cysteine pattern and Cys 1−4 and 2−3 disulfide linkages.[6,7] Unlike T-2 conotoxins, the T-1 counterparts show less homology in sequence and exhibit a variety of post-translational modifications, leading to diverse functional attributes.[8,9] As a result, the biological activity exhibited by each conotoxin within a family varies. For example, a 13-residue peptide isolated from the molluscivorous snail, *Conus textile*, specifically acts on presynaptic Ca^{2+} channels;[10] the remaining members of T-1 family conotoxins (earlier named τ-conotoxins) do not share the same specificity.

* To whom correspondence should be addressed. Telephone: 1-850-644-0529. Fax: 1-850-644-1366. E-mail: marshall@magnet.fsu.edu.

[†] Florida State University.

[‡] Tata Institute of Fundamental Research.

[§] Indian Institute of Science.

[||] National Center for Biological Science.

(1) Olivera, B. M.; Rivier, J.; Scott, J. K.; Hillyard, D. R.; Cruz, L. J. *J. Biol. Chem.* **1991**, *266*, 22067−22070.

(2) Terlau, H.; Olivera, B. M. *Physiol. Rev.* **2004**, *84*, 41−68.

(3) Lewis, R. J. *IUBMB Life* **2004**, *56*.

(4) Olivara, B. M. *Annu. Rev. Ecol. Syst.* **2002**, *33*, 25−47.

(5) Walker, C. S.; Steel, D.; Jacobsen, R. B.; Lirazan, M. B.; Cruz, L. J.; Hooper, D.; Shetty, R.; DelaCruz, R. C.; Nielsen, J. S.; Zhou, L. M.; Bandyopadhyay, P.; Craig, A. G.; Olivera, B. M. *J. Biol. Chem.* **1999**, *274*, 30664−30671.

(6) Balaji, R. A.; Ohtake, A.; Sato, K.; Gopalakrishnakone, P.; Kini, R. M.; Seow, K. T.; Bay, B.-H. *J. Biol. Chem.* **2000**, *275*, 39516−39522.

(7) Sharpe, I. A.; Gehrmann, J.; Loughnan, M. L.; Thomas, L.; Adams, D. A.; Atkins, A.; Palant, E.; Craik, D. J.; Adams, D. J.; Alewood, P. F.; Lewis, R. J. *Nat. Neurosci.* **2001**, *4*, 902−907.

(8) Han, Y. H.; Q., W.; Jiang, H.; Miao, X. W.; Chen, J., S.; Chi, C. W. *Toxicon* **2005**, *45*, 481−487.

(9) Aguilar, M. B.; Lezama-Monfil, L.; Maillo, M.; Pedraza-Lara, H.; Lopez-Vera, E.; Heimer de la Cotera, E. P. *Peptides* **2006**, *27*, 500−505.

(10) Rigby, A. C.; Lucas-Meunier, E.; Kalume, D. E.; Czerwiec, E.; Hambe, B.; Dahlqvist, I.; Fossier, P.; Baux, G.; Roepstorff, P.; Baleja, J. D.; Furie, B. C.; Furie, B.; Stenflo, J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 5758−5763.

10.1021/ac0607764 CCC: $33.50
Published on Web 11/03/2006

Another remarkable property of the T-superfamily members is its wide array of post-translational modifications. Besides the disulfide bonds, all T-superfamily conotoxins possess either one or more post-translational modifications, such as C-terminal amidation, glutamic acid carboxylation, proline hydroxylation, threonine glycosylation, or tryptophan bromination.[5,11,12] The first reports on post-translational bromination of tryptophan associated with a natural peptide were published by Olivera and co-workers, who described a special class of *Conus* peptide, called contryphan, harboring an L-6-bromotryptophan residue.[13,14] The exact identity of that modification was established by comparison of HPLC retention time of the native peptide with various synthetic analogues. Alternatively, addition of a synthetic, phenylhydantoin-derivatized bromotryptophan as a standard in Edman degradation revealed the presence of bromotryptophan (in the standard Edman process, 6-bromotryptophan appears as a characteristic late-eluting peak). Steen and Mann[15] used quadrupole time-of-flight mass analysis to identify bromotryptophan associated with biomolecules, through a precursor ion scanning strategy in which the immonium of m/z 237.00, corresponding to ^{79}Br, serves as a specific "marker" for the presence of bromotryptophan.

The presence of post-translational modifications increases the complexity of determining the primary sequence of biomolecules by conventional methods. As mentioned above, Edman degradation coupled with amino acid composition analysis has been used to determine the primary sequence of several conotoxins. However, those techniques require pure material and are not amenable to the identification of all post-translational modifications. DNA analysis is of little help because the gene sequence cannot predict post-translational modifications. Mass spectrometry combined with modern molecular biology techniques can provide the sequence information, including the presence of post-translational modifications.[16] High-resolution Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers enable facile de novo sequencing of peptides from a small amount of sample. The unequaled resolution and mass accuracy conferred by FT-ICR MS,[17] along with its unique fragmentation capabilities, such as infrared multiphoton dissociation (IRMPD) and electron capture dissociation (ECD),[18,19] are especially useful for determination of primary sequence, post-translational modifications, and disulfide connectivity.

Here, we describe mass spectrometry-based de novo sequencing of the conotoxin, Mo1274, isolated from a vermivorous snail, *Conus monile*, found off the coast of India. Electron capture dissociation was used to identify and localize tryptophan bromination. Structural characterization was performed with home-built 7.0-, 9.4-, and 14.5-T electrospray ionization (ESI) FT-ICR mass spectrometers.[20−22] The experimentally determined sequence was validated by chemical derivatization and protease digestion followed by mass spectrometry. We also determined that the two disulfides of Mo1274 connect cysteine pairs 1−3 and 2−4. This systematic approach allows us to classify Mo1274 as a member of the T-1-family of conotoxins.

EXPERIMENTAL METHODS

Materials. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Fluka, and *N*-ethylmaleimide (NEM) and iodoacetamide were from Sigma-Aldrich. HPLC grade water and acetonitrile were obtained from J.T. Baker. All other chemicals used were of analytical grade.

HPLC Purification. The venom ducts dissected from *C. monile* were preserved in ethanol. The extracted venom was concentrated and subjected to HPLC fractionation with an HP 1100 series HPLC system. The venom extract was purified by separation in a semipreparative reversed-phase C_{18} column (Jupiter, Phenomenex, 10 × 250 mm, 4-μm particle size, 90 Å pore size) with solution A (0.1% trifluoroacetic acid (TFA) (v/v) in 10% aqueous acetonitrile) and solution B (0.1% TFA (v/v) in 90% aqueous acetonitrile) as the mobile phase. The flow rate was maintained at 1 mL/min with a linear gradient from 20 to 70% of solution B over 30 min. The absorbance was monitored at 226 nm, and peaks were collected manually. The peptide was further purified over an analytical column (Zorbax, 4.6 × 250 mm, 5-μm particle size, 300-Å pore size). The chemically derivatized peptide was purified with a microbore (1 × 50 mm) reversed-phase C_{18} column coupled to a low-flow HPLC system (Micro-Tech Scientific Inc.).

Reduction and Alkylation. For complete reduction, an aliquot of the HPLC purified peptide (~0.25 nmol) was vacuum-dried (Savant Inc.), resuspended in 5 μL of acetonitrile containing 0.1% formic acid, and mixed with 5 μL of 40 mM TCEP dissolved in 200 mM sodium citrate buffer, pH 3.0, followed by incubation at 37 °C for 1 h. For alkylation, 3 μL of 250 mM NEM suspended in 200 mM sodium citrate buffer, pH 3.0 was added, and the resultant mixture was incubated at room temperature for 1 h followed by HPLC purification. Partial reduction of the peptide was achieved by incubating equal amounts (5 μL) of the stock peptide dissolved in acetonitrile containing 0.1% formic acid and 10 mM TCEP in 200 mM sodium citrate buffer, pH 3.0, at 37 °C for 5 min. The first alkylation step was followed by addition of 3 μL of 100 mM NEM in 200 mM sodium citrate buffer, pH 3.0, and incubation for 1 h at room temperature followed by HPLC fractionation. In the second alkylation step, the partially modified peptide fraction was dried in vacuum, resuspended in 100 mM Tris-HCl buffer, pH 8.0, and incubated with 40 mM TCEP in 100 mM Tris-HCl buffer, pH 8.0, for 1 h at 37 °C. Subsequently, alkylation was

(11) Hansson, K.; Furie, B.; Furie, B. C.; Stenflo, J. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 1081−1087.
(12) Kalume, D. E.; Stenflo, J. P.; Czerwiec, E.; Hambe, B.; Furie, B. C.; Furie, B.; Roepstorff, P. *J. Mass Spectrom.* **2000**, *35*, 145−156.
(13) 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. *J. Biol. Chem.* **1997**, *272*, 4689−4698.
(14) Jimenez, E. C.; Craig, A. G.; Watkins, M.; Hillyard, D. R.; Gray, W. R.; Gulyas, J.; Rivier, J. E.; Cruz, L. J.; Olivera, B. M. *Biochemistry* **1997**, *36*, 989−994.
(15) Steen, H.; Mann, M. *Anal. Chem.* **2002**, *74*, 6230−6236.
(16) Jakubowski, J. A.; Keays, D. A.; Kelley, W. P.; Sandall, D. W.; Bingham, J. P.; Livett, B. G.; Gayler, K. R.; Sweedler, J. V. *J. Mass Spectrom.* **2004**, *39*, 548−557.
(17) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1−35.
(18) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 245−249.
(19) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201−222.
(20) White, F. M.; Marto, J. A.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*.
(21) Senko, M. W.; Hendrickson, C. L.; Pasa-Tolic, L.; Marto, J. A.; White, F. M.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom* **1996**, *10*, 1824−1828.
(22) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; He, Y.; Hendrickson, C. L.; Marshall, A. G.; Griffin, P. R. *Anal. Chem.* **2006**, *78*, 1005−1014.

performed by adding 1 μL of 200 mM iodoacetamide and incubating at room temperature in the dark for 1 h. The modified peptide was desalted by reversed-phase HPLC.

Trypsin Digestion. The reduced and alkylated peptide was mixed with tosyl phenylalanyl chloromethyl ketone-treated trypsin at a ratio of 50:1 (w/w) and incubated at 37 °C for 3 h and analyzed by electrospray ionization (ESI) FT- ICR MS.

Chemical Synthesis. Peptide synthesis was performed by standard solid-phase peptide synthesis protocols with an LKB-Biolynx 4175, semiautomatic peptide synthesizer, employing 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry. The coupling reactions were mediated with pentafluorophenyl (Opfp) esters of the Fmoc-protected amino acids (standard side chain protection) with the exception of the introduction of Fmoc-DL-6-bromotryptophan. Fmoc-DL-6-bromotryptophan was obtained from DL-6-bromotryptophan (Biosynth AG, Staad, Switzerland) following standard procedures. The coupling of Fmoc-DL-6-bromotryptophan was achieved by the in situ activation of its C-terminus with *O*-(1*H*-benzotriazole-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate/diisopropylethylamine/*N*-hydroxybenzotriazole as the activation mixture. The synthesis was performed on a Wang resin (Novabiochem) at a target scale of 0.44 mequiv, corresponding to 400 mg of resin (loading capacity 1.1 mequiv/g), and the extent of coupling was monitored by the Kaiser ninhydrin test. After complete synthesis, the peptide was simultaneously cleaved off the resin and deprotected by TFA/anisole/ethanedithiol (95:4:1). Subsequent to complete deprotection, the resin was filtered, trifluoroacetic acid removed by evaporation in vacuo, and the peptide was precipitated with ether. The precipitate was repeatedly washed with ether and purified over a C_{18} column (9.4 mm × 250 mm, 5−10 μm particle size) by an acetonitrile/H_2O/TFA solvent system. The purified synthetic peptide was characterized by matrix-assisted laser desorption/ionization (MALDI) TOF−TOF mass spectrometry (Ultraflex, Bruker, Bremen, Germany). Monoisotopic mass: synthetic Mo1274, $[M + H]^+_{red.}$ = 1279.9 (calc), 1279.6 (obs)). MALDI MS/MS spectra of both reduced and alkylated natural and synthetic peptides were obtained.

Mass Spectrometry. All high-resolution MS experiments were conducted with a home-built, unshielded 7.0-T FT-ICR mass spectrometer,[23] a passively shielded 9.4-T FT-ICR instrument,[21] or a hybrid linear ion trap-14.5-T FT-ICR MS instrument (LTQ-FT, Thermo Electron Corp., San Jose, CA).[22] All FT-ICR mass spectra were recorded with the 7.0- and 9.4-T instruments, and collision-induced dissociation (CID) was performed in the linear quadrupole ion trap. Samples were infused by microelectrospray[24] or with an automated nanoelectrospray system (Nanomate, Advion Biosciences).[25] Ions were externally accumulated in a storage octopole[26] in the 7.0- and 9.4-T instruments. Ions were selected with a quadrupole mass filter and subsequently accumulated in a second octopole for the 9.4-T system. The ions were then transferred through multiple ion guides and accumulated in an open-ended cylindrical Penning trap by gated trapping.[27]

For IRMPD, quadrupole-isolated precursor ions were photon-irradiated by a 40-W, 10.6-μm CO_2 laser (Synrad, Mukilteo, WA). For ECD, the quadrupole-isolated precursor ions were irradiated with low-energy electrons emitted from an indirectly heated dispenser cathode (Heat Wave, Watsonville, CA). The trap electrodes were kept at 10 V while biasing the cathode at −5 V and pulsing the grid potential to +50 V. Following irradiation, the trap electrode voltage was dropped to +2 V, grid to +5 V, and cathode to +10 V to purge any remaining electrons from the ICR cell. Ions were frequency-sweep excited (72−2884 kHz, at 150 Hz/μs) and detected in direct broadband mode (1 MWord time domain data). Hanning apodization and one zero-fill were followed by Fourier transform and followed by magnitude calculation and frequency-to-mass conversion. The spectrum represents a sum of 25−100 time domain transients. All spectra were externally calibrated with a standard ESI Agilent tuning mix and were analyzed by modular ICR data acquisition system (MIDAS) software.[28] Theoretical masses were calculated from assigned elemental compositions by use of Isopro software, version 3.1.

For CID performed in the linear ion trap of the 14.5-T mass spectrometer, helium was used as the collision gas at a normalized collision energy (Thermo Electron Corp.) setting of 10−15% and a measured pressure of ~650 mTorr. Each spectrum is the average of 20−100 individual scans and was analyzed with Xcalibur software (Thermo).

RESULTS AND DISCUSSION

Peptide Purification and Mass Spectral Characterization. Crude venom, extracted from the ducts of *C. monile* species, was fractionated and purified by reversed-phase chromatography. Each HPLC fraction containing peptides was analyzed by mass spectrometry. A fraction that displayed an isotopic distribution unusual for a peptide (data not shown) was further purified to near-homogeneity on a reversed-phase analytical column. Analysis of the peptide by ESI FT-ICR MS at 9.4 T showed a doubly charged ion distribution (Figure 1, inset), which upon deconvolution yielded a neutral monoisotopic molecular mass of 1274.319 Da and is thus labeled Mo1274. The isotopic distribution for Mo1274 indicated the possible presence of one bromine atom. The near-equal abundance of the peak doublet at m/z 638.159 and 639.138 is characteristic of a molecule containing the equally abundant bromine isotopes, ^{79}Br and ^{81}Br (50.69 and 49.31% relative abundance).[13]

Following reduction with TCEP and alkylation with NEM, the mass of Mo1274 increased to 1778.541 Da, indicating the presence of two disulfide bonds (upon alkylation with NEM, the nominal mass of each modified cysteine is 228 Da, a net increase of 125 Da for each residue). To determine the complete amino acid sequence of Mo1274, we employed a combination of ECD, IRMPD, and CID fragmentation techniques. The complementary information offered by each technique provided the complete sequence information. Both IRMPD and CID (slow heating processes) generated predominantly b- and y-type product ions through a sequence-specific, vibration-induced CO−N backbone cleavage.[29] ECD, on the other hand, generates c- and z-product ions through bond-specific, N−C_α cleavages.[30]

(23) Quenzer, T. L.; Emmett, M. R.; Hendrickson, C. L.; Kelly, P. H.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 1721−1725.
(24) Emmett, M. R.; Caprioli, R. M. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 605−613.
(25) Zhang, S.; Van, Pelt, C. K.; Henion, J. D. *Electrophoresis* **2003**, *24*, 3620−3632.
(26) Senko, M. W.; Hendrickson, C. L.; Emmett, M. R.; Shi, S. D.-H.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 970−976.
(27) Beu, S. C.; Laude, J., D. A. *J. Mass Spectrom.* **1992**, *112*, 215−230.
(28) Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839−1844.

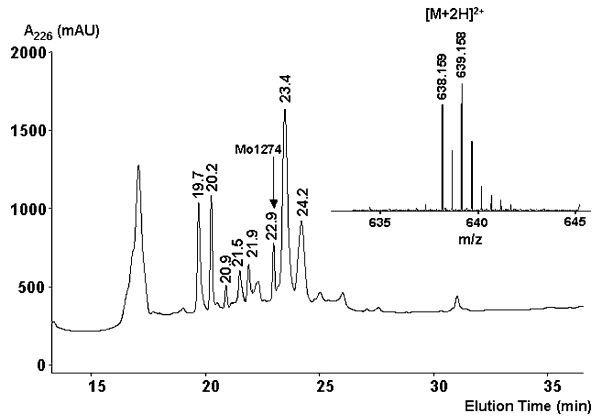

Figure 1. HPLC purification profile of *C. monile* venom extract. Fractionation was carried out with a Phenomenex C_{18} RP column (10 × 250 mm, 4-μm particle size and 90-Å pore size) and eluted at 1 mL min^{-1} with a linear gradient of acetonitrile, containing 0.1% TFA. Inset shows a 9.4-T ESI FT-ICR mass spectral segment for the LC peak at 22.9 min (denoted by an arrow in the chromatogram). The doubly charged peak at *m*/*z* 638.159 after deconvolution yielded a monoisotopic neutral mass of 1274.318 Da, subsequently referred to as Mo1274.

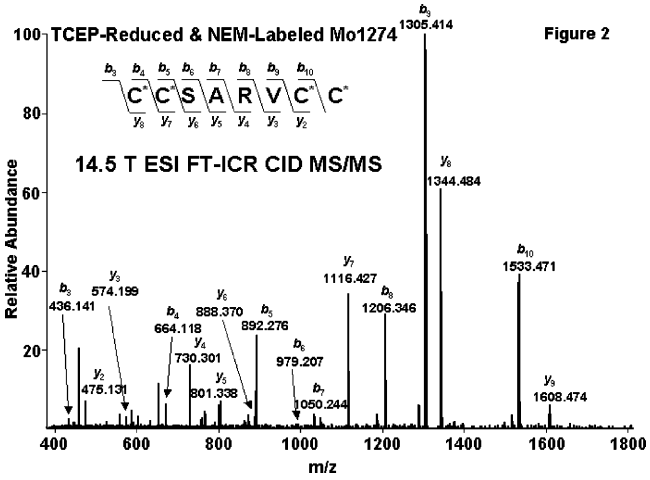

Figure 2. CID mass spectrum of TCEP-reduced and NEM-labeled Mo1274, recorded with a 14.5-T LTQ FT-ICR mass spectrometer. The doubly charged ion at *m*/*z* 890.302 was selected as the precursor ion. Inset shows the sequence derived from this MS/MS spectrum. C* denotes alkylated cysteine.

Figure 2 shows the CID mass spectrum of NEM-labeled Mo1274. The spectrum is complex, and initial sequence assignment was ambiguous due to difficulty in assigning the b_2 ion. Although no b_2 peak is seen in Figure 2, its mass may be inferred from its corresponding y_8 ion. The b_2 ion was later assigned by MALDI-MS/MS (*m*/*z*, 172.24, Figure 4). A closer examination of the isotopic distribution pattern for each product ion helped us to assign a nearly complete sequence of Mo1274. For example, the peaks at *m*/*z* 436.141 (b_3) and 664.118 (b_4) show the unusual isotope pattern, seen also in the parent ion (Figure 1, inset). Those ions were thus tentatively proposed to contain bromotryptophan. Furthermore, the difference in mass of 228 Da between b_3 and b_4 ions identifies an NEM-modified cysteine (the mass of each

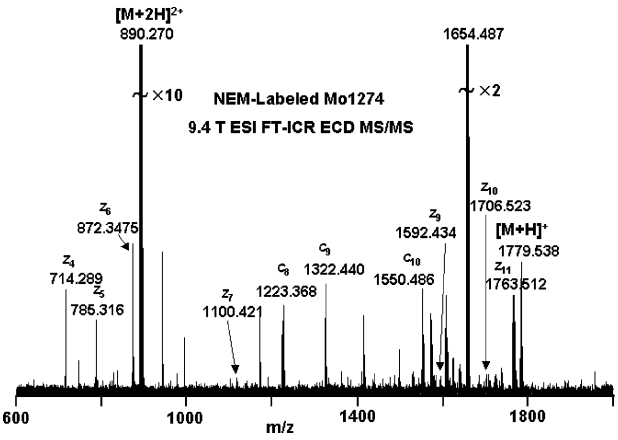

Figure 3. ECD mass spectrum of doubly protonated, fully NEM-labeled Mo1274. The difference in mass of 57 Da between z_{10} and z_{11} ions identifies a glycine residue at position 1 of Mo1274. Similarly, the mass difference of 114 Da obtained from the ion pairs z_9 and z_{10} may be assigned to an asparagine residue at position 2. Note the peak at *m*/*z* 1654.487, corresponding to the previously unreported loss of one NEM moiety from the $[M + 2H]^{+\bullet}$ ion species.

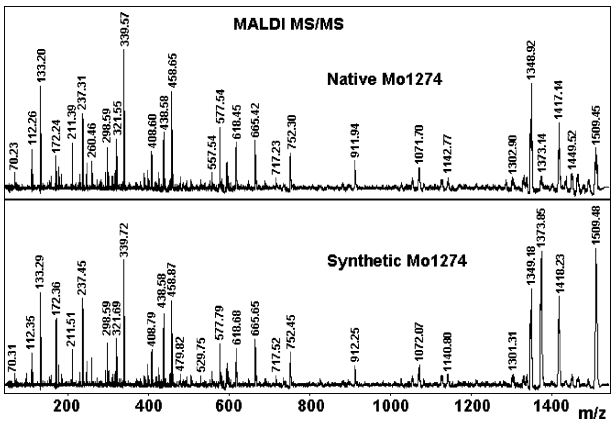

Figure 4. MALDI MS/MS comparison of reduced and alkylated peptides: native Mo 1274 (top) and synthetic Mo 1274 (bottom). The reduction and alkylation were effected by use of dithiothreitol (DTT) and iodoacetamide.

cysteine residues increases by 126 Da upon alkylation by NEM). On the other hand, the product ions that did not contain a bromotryptophan residue showed normal isotope patterns. The normal and anomalous isotope distributions were then analyzed to generate a sequence tag and partial sequence of Mo1274, -CCSARVCC, assigned according to a series of b ions from b_3 to b_{10} and complementary C-terminal y-ions from y_2 to y_9. Successive fragmentation (MS^3) of the ions of *m*/*z* 436.141 produced ions of *m*/*z* 172.000 Da. That difference in mass (264 Da) corresponds to bromotryptophan. Further fragmentation of 172.000 (MS^4) generated only a water loss (18 Da). If a b_2 ion mass of 172 Da is assumed to originate from a combination of two unmodified amino acids, then only two possibilities, GN- or NG-, are feasible.

The ECD spectrum of the linearized peptide (labeled by NEM, Figure 3), however, showed three successive z-ions of *m*/*z* 1763.512 (z_{11}), 1706.523 (z_{10}), and 1592.434 (z_9), the differences between which unambiguously assign glycine as the N-terminal residue followed by an asparagine. The final determined sequence of Mo1274 is GNWCCSARVCC (in which W is bromotryptophan), in good agreement with the FT-ICR detected mass of 1274.319

(29) McLuckey, S. A.; Goeringer, D. E. *J. Mass Spectrom.* **1997**, *32*, 461−474.
(30) Zubarev, R. A. *Curr. Opin. Biotechnol.* **2004**, *15*, 12−16.

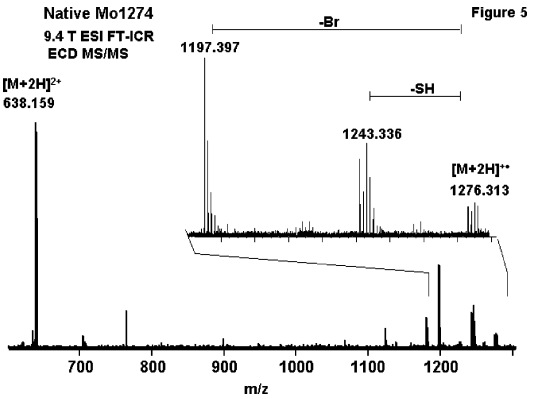

Figure 5. ECD mass spectrum of native Mo1274. The doubly protonated ion at *m*/*z* 638.159 was selected as the precursor ion. The ion of *m*/*z* 1276.313 $[M + 2H]^{+\bullet}$ is formed from the precursor ion $[M + 2H]^{2+}$ by charge neutralization. The prominent ions at *m*/*z* 1243.336 and 1197.397 originated from *m*/*z* 1276.313 through the loss of SH and Br moieties, respectively.

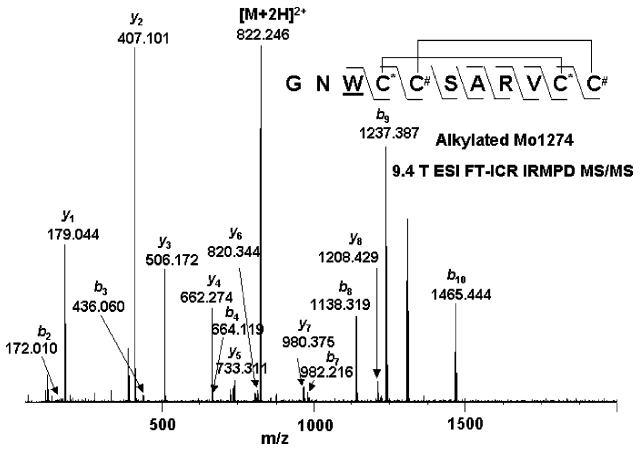

Figure 6. Infrared multiphoton dissociation mass spectrum of differentially alkylated Mo1274. The doubly charged ion at *m*/*z* 822.246 is the precursor ion. Direct sequence analysis shows that the cysteines (labeled as C*) at amino acid positions 4 and 10 are alkylated by NEM and the remaining cysteines (labeled as $C^{\#}$) at positions 5 and 11 are alkylated by iodoacetamide.

Da. The mass analysis reveals that the C-terminus of Mo1274 is not amidated (C-terminal amidation is a common post-translational modification and leads to a 1 Da reduction in mass). The presence of a single carboxylic acid residue in the molecule was confirmed by esterification by use of dry methanol/acetyl chloride resulting in an increase of 14 Da. Prolonged incubation also resulted in hydrolysis of aspargine side chain carboxyamide, followed by subsequent esterification, resulting in further mass increase of 15 Da. These chemical modifications further confirm the presence of one carboxylic and one amide group in the molecule. A second prominent peak at *m*/*z* 1654.487 (from Figure 3) differs from the charge-reduced species of *m*/*z* 1779.538 ($[M + H]^{+}$) by 125.051 Da. A similar loss is observed by ECD after one pair of cysteines of Mo1274 was modified by NEM. These data strongly suggest loss of NEM from the modified cysteine, given that the exact mass of NEM is 125.047 Da. We were not able to observe this loss by IRMPD or CID. Side chain fragmentation of alkylated cysteine residues in ECD mass spectrometry has recently been reported.[31]

To validate the proposed sequence, reduced and NEM-labeled Mo1274 was digested with trypsin. Because the peptide contains one trypsin cleavage site at arginine, two fragments with masses 1223 (N-terminal fragment) and 573 Da (C-terminal fragment) were expected. ESI FT-ICR MS at 14.5 T yielded two peptides: *m*/*z* 612.454, corresponding to the doubly charged N-terminal 1223 Da fragment, and 574.181, the protonated C-terminus (data not shown). Each of those ions was isolated and subjected to CID MS/MS fragmentation, and product ion mass spectra were detected by LTQ FT-ICR MS. The fragment ion data completely agree with the proposed sequence of Mo1274.

To confirm the sequence derived from mass spectrometry, a synthetic peptide containing DL-6 bromotryptophan at amino acid position 3 was prepared. Figure 4 demonstrates the near identity of MALDI MS/MS fragmentation patterns obtained from linearized native and synthetic peptides, in which the free thiol groups have been modified with iodoacetamide. The resulting mass of the derivatized cysteine (C*) residue is 160 Da, an increase of 58 Da over unmodified cysteine. Note that synthetic peptide is a distereomeric mixture containing both D and L residues at position 3. Differences in mass spectra of peptide diastereomers have been noted. The site of bromination on tryptophan has not been directly determined. We infer that the positional isomer occurring in Mo1274 is 6-bromotryptophan following the earlier work of Craig et al., demonstrating 6-bromotryptophan in *Conus imperialis*.[13] Notably, 6-bromotryptophan has also been observed in other marine natural products.[32]

Sequencing of Mo1274 was facilitated by the combined use of CID-, IRMPD-, and ECD-induced fragmentation of the linearized peptide after disulfide bond reduction and subsequent alkylation of free thiol. Interestingly, the ECD spectrum of native Mo1274 shows no cleavage of backbone amide bonds. Figure 5 shows the odd-electron ion species of *m*/*z* 1276.313 ($[M + 2H]^{+\cdot}$) generated from the doubly charged even-electron precursor of *m*/*z* 638.159 ($[M + 2H]^{2+}$). A dominant ion is obtained by the neutral loss (78.904 (experimental) versus 78.918 Da (atomic mass)) of the halogen atom. The ions of *m*/*z* 1243.336 and 1197.397 are formed from $[M + 2H]^{+\cdot}$ through loss of SH and Br. The loss of bromine is demonstrated not only by the mass difference but also in the change in the isotopic distribution. The other three significant peaks in this figure are assigned to neutral losses from the above-mentioned ion species

For peptide disulfides, dissociation of the disulfide bond following single electron capture and subsequent protonation would lead to the formation of a free thiol group and a thiyl radical. The loss of an SH group from such a species can be readily rationalized. The loss of bromine can occur by two possible routes. Direct electron capture by the carbon−bromine bond as for 6-bromotryptophan by population of a delocalized antibonding orbital in the aromatic system is one possibility. Alternatively, electron transfer from the disulfide, which is the initial site of electron capture, to the indole ring is also possible. The absence of bromine loss in the reduced and alkylated peptide argues strongly for the involvement of the disulfide bond in the dissociative process.

(31) Chalkley, R. J.; Brinkworth, C. S.; Burlingame, A. L. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1271−1274.

(32) Jimenez, E. C.; Watkins, M.; Olivera, B. M. *Biochemistry* **2004**, *43*, 12343−12348.

Table 1. Post-Translational Modifications of the T-1 Conotoxins[a]

serial no.	sequence	species	method	ref
1	γ**CC**γDGW **CC**TAAO	*C. textile*	cDNA, MS, NMR	5, 10
2	**CC**PGK**P** **CC**R**I**G*	*C. victoriae*	cDNA, MS	16
3	NA**CC**IVRQ **CC**	*C. marmoreus*	cDNA	8
4	N**CC**RRQI **CC**GRTK*	*C. textile*	cDNA	5, 42
5	I**CC**YPVW **CC**D	*C. textile*	cDNA	5
6	E**CC**EDGW **CC**TAAPLTGR	*C. textile*	cDNA	5
7	GN**W**C**C**SARV **CC***	*C. monile*	MS, CS, #	this work

[a] Abbreviations: γ-carboxyglutamate (γ); 6-Br-tryptophan (W); glycosylation (T*); hydroxyproline (O); amidated C-terminus (*), mass spectrometry (MS), chemical synthesis (CS), disulfide connectivity, Cys1-Cys3, Cys2-Cys4 (#).

Initial work by McLafferty and co-workers demonstrated that electron capture dissociation in multiply charged proteins is favored at disulfide bonds.[33] The mechanism of fragmentation under ECD conditions is a subject of active debate, with the consensus that the initial electron capture events occurred at positively charged sites, with subsequent internal rearrangement, possibly involving the generation of "hot" hydrogen atom species.[34−38] Recent theoretical studies have addressed the electron attachment step in electron capture dissociation and electron-transfer dissociation mass spectrometry. Simons and co-workers have used classical dynamics trajectory simulation and ab initio electronic structure calculations in an attempt to evaluate the energetics of electron capture in a model system containing the cleavable S−S bond and a positively charge NH_3^+ group. Their results suggest that electron capture occurs very much more efficiently at the positive site although they still hold the possibility that Coulomb-stabilized S−S σ^* orbitals may also serve to accept electrons directly.[39] Because ECD was unable to provide informative product ions from which to assign the disulfide connectivity, we followed the traditional approach of chemical labeling, in which the two cysteine pairs were reduced and derivatized with two different alkylating agents, followed by ESI FT-ICR mass analysis.

Disulfide Connectivity. To determine the disulfide connectivity between two cysteine pairs, Mo1274 was partially reduced with TCEP in a 200 mM sodium acetate buffer, pH 3.0, immediately followed by alkylation with NEM in the same buffer. The differentially labeled peptide fractions were separated by reversed-phase HPLC. ESI-MS analysis of the two eluted peaks represented partially modified (mass increase of 252 Da) and fully modified (mass increase of 504 Da) peptides. The experiment was repeated to maximize the yield of the partially modified peptides. The duration of incubation after TCEP addition and the TCEP concentration were varied for maximum yield. The partially labeled peptide, in which one cysteine bridge remains intact, was then completely reduced by TCEP, followed by alkylation with iodoacetamide. That peptide product was purified by HPLC and subjected to MS analysis. Figure 6 shows the IRMPD mass spectrum of the quadrupole-isolated, doubly protonated ion of m/z 822.246 from the differentially labeled Mo1274. Fragmentation by IRMPD predominantly generated b- and y-type ions. Direct sequence analysis indicated that cysteine pairs at positions 1 and 3 were labeled by NEM (mass increase of 126 Da per unmodified cysteine residue), and the remaining pair at positions 2 and 4 were modified by iodoacetamide (mass increase of 58 Da per cysteine). These data confirm that the disulfide pairing in Mo1274 is 1−3 and 2−4.

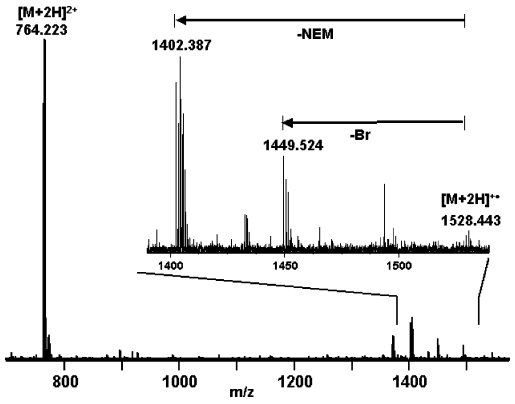

Figure 7. ECD mass spectrum of partially modified Mo1274, in which a pair of cysteines is reduced (DTT) and alkylated with NEM. Ions of m/z 1449.524 and 1402.387 are formed from $[M + 2H]^{+\bullet}$ via Br and NEM losses.

CONCLUSIONS

In this report, we describe the purification and mass spectral characterization of a novel bromotryptophan-containing peptide, Mo1274, isolated from the venom of a vermivorous snail, *C. monile*. The primary structure of the peptide was determined through mass spectrometry-based de novo sequencing. The peptide contains bromotryptophan, the presence of which was initially suggested by the unusual isotopic distribution showing two strong peaks (corresponding to doubly charged ions) with 1-Da separation at m/z 638.159 and 639.158. The presence of this pair of isotopes is a characteristic of a molecule containing ^{79}Br and ^{81}Br atom. A combined approach of ECD and CID confirmed the presence and location of bromotryptophan. Determination of the primary sequence, cysteine framework, and assignment of disulfide linkage map classifies Mo1274 in the T-1-family of

(33) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.
(34) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266.
(35) Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14011−14016.
(36) Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208−224.
(37) Sobczyk, M.; Anusiewicz, I.; Berdys-Kochanska, J.; Sawicka, A.; Skurski, P.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 250−258.
(38) Zubarev, R. A.; Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337−339.
(39) Anusiewicz, I.; Berdys-Kochanska, J.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 5801−5813.

conotoxins. A common feature of this class of conotoxins is the presence of adjacent cysteine residues at the N- and C-termini of the molecule. The CC segments are separated by a short stretch of four or five amino acids. Among the sequences determined thus far, a majority of T1 conotoxins possess five intervening residues. Mo1274 is an addition to the small group (see Table 1) that contains four residues between the CC segments. The present study demonstrates the utility of applying multiple fragmentation modes in de novo sequencing of natural peptides.

ECD fragmentation is particularly useful for precise mapping of post-translational modifications of peptides and proteins because of its ability to retain labile modifications such as phosphorylation and glycosylation.[19,40,41] In this study, reduced and alkylated Mo1274 retains bromine under ECD conditions; but if the disulfide bonds are still intact, ECD results in the loss of bromine. The loss of bromine under different conditions could prove useful for future conotoxin analyses. Note that the presence of a single disulfide bond also enables bromine loss (Figure 7). Our results suggest a role for electron capture by the disulfide followed by electron transfer to the 6-bromotryptophan residue, resulting in dissociative loss of bromine. A future study of model peptides with disulfide bonds and bromotryptophan residues may provide further insight into this novel fragmentation process.

(40) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530−4536.
(41) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355−2361.
(42) Luo, S.; Zhangsun, D.; Zhang, B.; Chen, X.; Feng, J. *Peptides.* In press.

ACKNOWLEDGMENT

The authors thank Christopher L. Hendrickson, John P. Quinn, Kim F. Haselmann and Yury O. Tsybin for useful discussions and suggestions. This work was supported by the NSF National High-Field FT-ICR Mass Spectrometry Facility (DMR 00-84173), the National High Magnetic Field Laboratory, and Florida State University. Research at Bangalore was supported by a grant from the Department of Biotechnology, Government of India.

Received for review April 26, 2006. Accepted September 21, 2006.

AC0607764