Diversity of conotoxin types from *Conus californicus* reflects a diversity of prey types and a novel evolutionary history

C. A. Elliger[1], T. A. Richmond[2], Z.N. Lebaric[1],
N. T. Pierce[1], J. V. Sweedler[2*] and W. F. Gilly[1]

[1]Hopkins Marine Station, Stanford University, Pacific Grove, CA 93950, USA
[2]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Supplementary Material

Contents:

Mass Spectral Analysis

In what follows the details on the mass spectrometric measurements are presented, including the experimental protocols and results.

Experimental.

1. Chemicals and equipment.

Acetonitrile (ACN) used with both liquid chromatography (LC) setups was purchased from Fisher Scientific (Pittsburgh, PA, USA). Dithiothreitol (DTT), trifluoroacetic acid (TFA), iodoacetamide (IAA), MALDI matrix, and all other chemicals were from Sigma (St. Louis, MO, USA). The Micromass capLC (Waters Corp., Milford, MA, USA) was connected to a 300 μm id x 15 cm column with the model 2487 single wavelength detector operating at 220 nm used for absorbance detection. Fractions were collected onto MALDI targets from Bruker Daltonics (Billerica, MA, USA) using a Proteineer fc (Bruker) robot controlled by Excel-based macros. Matrix assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS) data was collected on an Ultraflex II (Bruker) in reflectron mode. A quadrupole (Q)TOF MS instrument was used for electrospray (ESI) experiments. The Waters Q-TOF Ultima was equipped with the same model capLC as listed above flowing into an Atlantis dC18 100 μm x 100 mm HPLC column.

2. *Venom collection*

Conus californicus specimens were collected from Monterey Bay and housed in tanks at Hopkins Marine Station under conditions equivalent to the temperate environment of California's coastal waters. Venom ducts were removed by dissection, and the contents were expelled onto 1 cm square pieces of Parafilm. These Parafilm pieces were then placed in 1.5 ml micro centrifuge tubes and stored at –80°C until used.

3. *LC-MALDI-MS*

3.1. *Analysis of native peptides*

Individual venom samples were prepared for analysis in the following manner. Distilled water (150 μl) was added to the tube containing the venom duct contents on Parafilm. The sample was briefly vortexed, sonicated for 5 min, followed by 20 s of additional vortexing. Insoluble material was removed

by centrifugation at 10,000 rcf for 5 min. Sonication, vortexing, and centrifugation were then repeated on the supernatant prior to chromatography.

Approximately 2.5 µl of native venom extract was injected onto a 300 µm id column using a manual injection valve (VICI Valco Cheminert 04W-0178H). Solvent compositions of A and C were 95: 5: 0.1: 0.01 (water/ACN/acetic acid/TFA) with solvent B being 90: 10: 0.1: 0.08 (ACN/water/acetic acid/TFA). After injection the sample was loaded onto an OptiPak 0.5 µl trap column (Optimize Technologies, Oregon City, OR, USA) at 8 µl /min using solvent C. The elution gradient ran 2.0 µl/min flow of 5% B for the first 5 min (during sample loading) followed by a ramp up to 60% B over the next 45 min. The outlet of the column was connected by a 15 cm length of 50 µm id fused silica to the absorbance detector. The outlet of the detector was connected to a 1 m length of the same capillary installed into the sheath spotting mechanism of the Proteineer fc. Thirty-second spots were collected onto a 384 spot steel MALDI target (Bruker) accompanied by 1 µl of α-cyano-4-hydroxycinnamic acid matrix (10 mg/ml α-cyano in 50/50 ACN/water with 0.1% TFA). Sample spots were collected between 10 and 50 min of each HPLC run (80 spots total). Samples were analyzed within 48 h of collection on the Ultraflex II TOF-TOF MALDI system.

3.2. *Analysis of reduced/alkylated peptides*

Venom duct samples corresponding to each of the above runs were reduced and alkylated for comparison to the native peptides. Ten µl of the originally prepared venom duct extract were dissolved in 100 µl of 50 mM NH_4HCO_3, and 5 µl of 200 mM DTT (in 100 mM NH_4HCO_3) was added followed by incubation at 37°C for 1 hour to reduce the disulfide bonds within the conotoxins. This was followed by alkylation with a 4 µl aliquot of 1.0 M IAA solution prepared fresh for each use by dissolving 37 mg of IAA in 200 µl of 100 mM NH_4HCO_3. Alkylation was allowed to occur in the dark for 1 h at room temperature, after which any remaining IAA was removed by addition of 10 µl of 200 mM DTT, and the mixture was stored at –20° overnight before purification. PepClean C-18 Spin Columns (Pierce Rockford, IL, USA) were used to clean up the reduced/alkylated samples followed by SpeedVac sample volume reduction and then LC-MALDI as previously described.

4. *LC-ESI-MS*

ESI mass spectra were collected in parallel with MALDI analyses to characterize peptides in greater detail. Native and reduced/alkylated samples were prepared as described above. The Q-TOF instrument

was used to collected ESI data on native and reduced/alkylated venom duct material. Gradient elution was used to elute retained analytes from the trap column onto the Atlantis column with solvents that were similar to those used in the LC-MALDI experiment. The data-dependent analysis selected three parent masses for every full scan. The software adjusted the collision energy from 25 to 40 eV depending on the charge state of the selected ion. The detection window in the survey scan was set from 400 to 1950 *m/z* with an acquisition time of 1.0 s. Tandem MS spectra were acquired from 50 to 2000 *m/z* with a time of 1.5 s. Data files were analyzed manually and also turned into PKL files using ProteinLynx Global Server (Waters). PKL files were subjected to database searching and automated *de-novo* analysis with Peaks Studio (Bioinformatics Solutions Inc., Waterloo, Canada).

Results

In total we were able to confirm four conotoxin frameworks in our venom samples by LC-MS (Supplementary Table 2). While many additional peptides were observed, the limited size of the cDNA library prevented identification of further peptides. As an example of our analysis of peptides, the data for the framework VI (cal6) peptides is presented. In total three peptides were sequenced that closely resembled the Cal6.4 sequence. An eluting peptide was identified (*m/z* 714.5^{+4}) that matched the translated cDNA sequence GCWLCLGPNACCRGSVCHDYCPR. By swapping a single amino acid residue (R23 to S23) we obtained an exact match to the intense peaks corresponding to the masses 2880.4 and 2802.6. Supplementary Fig. 1 details the sequences of both peptides, which include a hydroxyproline (O) and in one case a bromotryptophan (W). The peptide peak at 714.54^{+4} that matched the cDNA was approximately 10-fold less intense as compared to the two peaks manually matched to the cDNA with an S at position 23. Additionally, there was no brominated form of the R23 peptide detected. A single peptide yielded a high enough quality fragmentation spectrum post reduction and alkylation that enabled *de novo* sequencing. This peptide eluted around 35 min at *m/z* 836.1^{+4}. It is characterized as a conotoxin containing three disulfide bonds with a native mass of 2993.4 and an alkylated monoisotopic mass of 3341.6^{+1}. Supplementary Fig. 2 displays the sequence analysis of this precursor ion.

Table S1. Primers used in RT-PCR.

Conotoxin	Forward Primer	Reverse primer
Cal1	C86F: ATGCGCTGTCTTCCAGTCTTC	86R: CTGTTGGCCACTTGTTACATCG
Cal5.1	W9F: AGAGGCCACGTTACGACAGG	W9R: CCTTTTCGAAGTCATGATTTCTG
Cal6.1	CGF: GGACATGACAAACATGAGAAGG	CGR: CAGAAGGGAAAGCAAGACGCC
Cal6.2	W59F: GGAACGTCGACTGGCATGAGG	W59R: AGGGCAGGGCAGACACCAGG
Cal6.3	W66F: TGGGGCGATGCCAAGAAACC	W66R: GGCGTGAATTACAGAAGCTCG
Cal9.1	C81F: GCATGATGGCTACTGGAGAAG	C81R: GCACCCTCAGGTCACAAGCG
Cal9.2	W55F: GACGACTGCTGGTCAATTGAAC	W55R: AGCACATGAGCCGTTCGTAGG
Cal12.2	C68F: GGTGCTACTCCATGGCAGAAC	C68R: AGTGCATCGAGCACGGATGG
Cal14.1	W23F: TGTGCGGTCATCATGAAGCTG	W23R: TGTCCCATCCCTCTGTTTGAAG
Cal14.2	W1F: GGAAAACCAACGACCATCAAAATG	W1R: GTTCCTTGCCAAAGTTCCTTGC
Cal16	W33F: CCGCTGACATCATGCGCTGTC	W33R: GAAGTCGTGTCACAGAGTGCG
Cal22	W31F: ATGACCAGAGCACTGCACGG	W31R: ATGAAGTAGGCTGGAGAGTCC

Table S2. Compilation of the completely sequenced peptides using both the *de novo* approach and database search from our cDNA library.

Peptide	Sequence	*m/z*	Ret. time	[M+H]+
cal5a	DAADVKOVARTNEGOGRDOAOCCQHPIETCC	886.40 +4	24	3542.6
cal5b	TNEGOGRDOAOCCQHPIETCC	835.39 +3	21	2504.2
cal5c	TNEGPGRDOAOCCQHPIETCC	830.09 +3	22	2488.2
cal5d	TNEGPGRDPAOCCQHPIETCC	824.70 +3	25	2472.1
cal5e	TNEGPGRDPAPCCQHPIETCC	819.35 +3	27	2456.0
cal6.4a	GCWLCLGPNACCRGSVCHDYCPR	714.54 +4	37	2855.2
cal6.4b	GCWLCLGONACCRGSVCHDYCPS	934.65 +3	38	2802.0
cal6.4c	GCWLCLGONACCRGSVCHDYCPS	960.67 +3	42	2880.2
cal14a	GCPADCPNTCDSSNKCSOGFPamide	1171.93 +2	30	2342.9
cal16a	pQGCVCNANAKFCCGEamide	878.79 +2	32	1756.6

Table S3. GenBank accession numbers for the Cal12 conotoxin sequences.

Conotoxin	GenBank Accession No.
Cal12.1.1a	EF644174
Cal12.1.1b	EF644175
Cal12.1.1c	EF644177
Cal12.1.1d	EF644178
Cal12.1.1e	EF644180
Cal12.1.2a	EF644182
Cal12.1.2b	EF644183
Cal12.1.2c	EF644184
Cal12.1.2d	EF644185
Cal12.1.2e	EF644187
Cal12.1.2f	EF644188
Cal12.1.2g	EF644190
Cal12.1.2h	EF644191
Cal12.1.3a	EF644194
Cal12.1.3b	EF644195
Cal12.2a	EF644196
Cal12.2b	EU022528
Cal12.2c	EU022529
Cal12.2d	EU022530

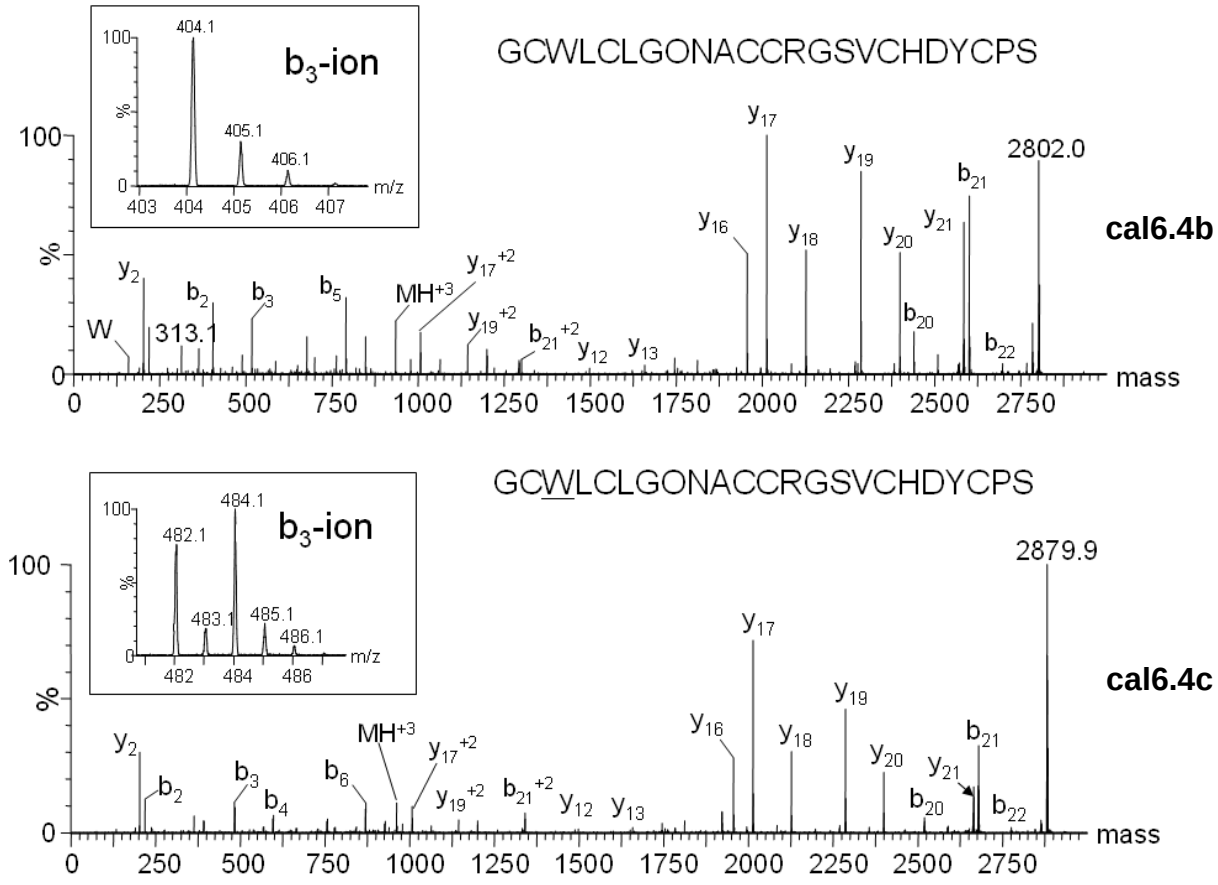

Fig S1. Fragmentation spectra matching the cDNA clone for Cal6.4. These two peptides, cal6.4b/cal6.4c, both contain a hydroxyproline (O8), with the bottom spectrum confirming the presence of the bromotryptophan (W3). Each inset shows the b_3-ion detailing the isotopic distribution.

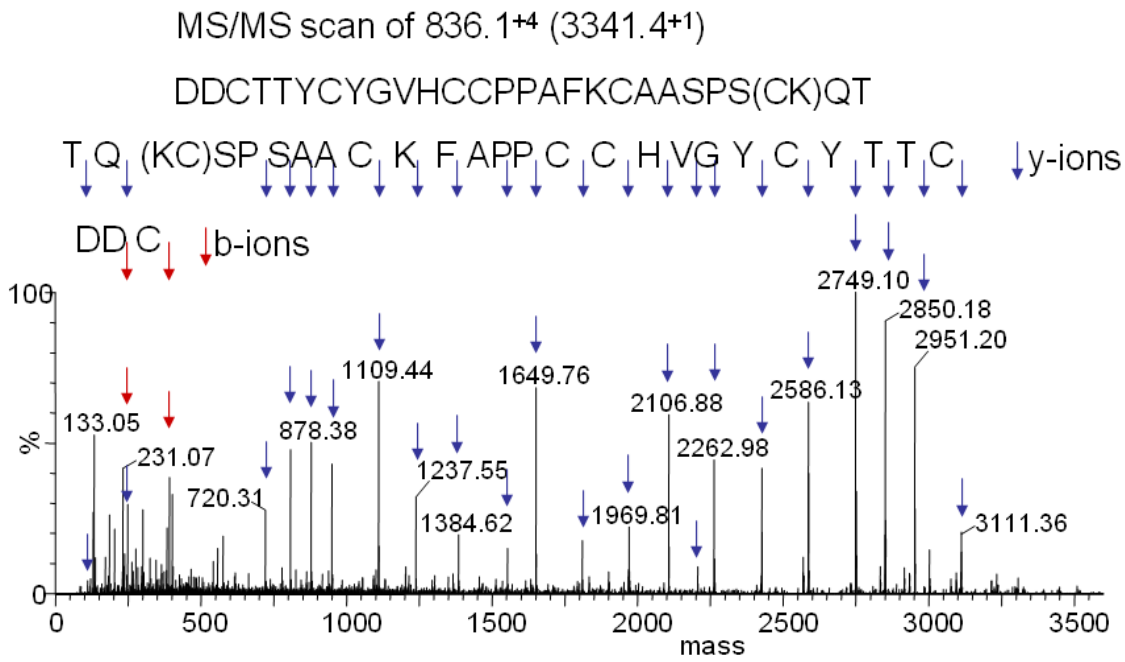

Fig. S2. Characterization of cal6.5a. The MS spectrum shows the analysis of the precursor at 836.1^{+4} (3341.4^{+1}) that eluted at 35 min. The blue arrows indicate y-ions and the red arrows indicate b-ions. The sequence of the peptide is determined through *de novo* analysis and minor ambiguities remain, including the possible transposition of "CK" at position 25-26. The spectrum is deconvoluted to mass rather than *m/z*.