

Nanoscale characterization of spider venom peptides by high-resolution LC-MS/MS analysis.

Claire Dauly (1) Pierre Escoubas (2), Graham M. Nicholson (3), Glenn F. King (4), Martin Hornshaw (5)

(1) Thermo Fisher Scientific, 16 Avenue du Quebec, 91963 Courtaboeuf cedex, France, (2) University of Nice Sophia Antipolis, Institute of Molecular and Cellular Pharmacology, Valbonne, France, (3) Department of Medical & Molecular Biosciences University of Technology Sydney, Sydney, Australia, (4) Institute for Molecular Biosciences, University of Queensland, Brisbane, Australia, (5) Thermo Fisher Scientific, Boundary Way, Hemel Hempstead, UK



Overview

This study used high-resolution LC-MS/MS with a combination of CID, HCD and ETD fragmentation to characterize endogenous peptides in spider venom.



Introduction

Animal venoms are natural libraries of biologically active peptides. They encompass a wide variety of structures and pharmacological activities and represent an enormous resource of novel molecules to be used as insecticide, therapeutic and drug models. However, the obstacle of sample size is daunting as many venomous species are of a size too small for classical bioassay-guided fractionation and biochemical characterization. High-resolution mass spectrometry can be used as an alternative technology for peptide sequence determination and sequence tag generation to permit the use of cDNA libraries. Venom profiling can be used for species identification and to indicate the presence of potentially unknown toxins. Here we demonstrate that *de novo* sequencing at the nanoscale level is applicable to venomomics research.

Methods

Crude venoms from the hexathelid Australian funnel-web spider *Hadroryche infensa*: Orchid beach (Hi:OB) and a small jumping spider (Salticidae, Sp004, body size 4 mm) were first analyzed by high-resolution nano-LC/MS for intact mass determination. Analysis of reduced and alkylated venom improved fragmentation in a systematic MS/MS analysis of fractions. Full-scan spectra were deconvoluted to obtain the exact masses of native and alkylated venom peptides. Detected peptides were selected for higher-energy collisional dissociation (HCD) and electron transfer dissociation (ETD) fragmentation. *De novo* sequencing was performed using MS/MS information and sequence tags were used to search for homologies in specific venom peptide databases.

LC/MS

HPLC System: Thermo Scientific Accela MS pump with a flow splitter
Autosampler: Accela™ autosampler
Columns: C18 (Reprosil) 100um id x 15cm packed tip column (Nikkyo Technos Co. Ltd) with C18 trapping cartridge (Captrap, Michrom Bioresources)
Mobile Phase: A: Water / 0.1% formic acid
 B: Acetonitrile / 0.1% formic acid
Flow Rate: 300 nL/min, post split



Mass Spectrometer: Thermo Scientific LTQ Orbitrap XL with ETD
Spray Voltage: 2.2 kV
Capillary Temp: 200 °C
Dynamic Exclusion™: Repeat count 1
 Exclusion list size 500
 Exclusion duration 120 s

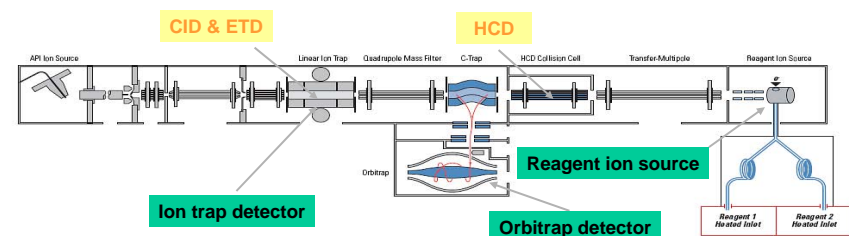
MS/MS Scans:

Detector: Orbitrap™
Resolution: 100000
Reagent AGC Target: 1e6
AGC Target: 3e5

Collision Energy: 35% for HCD scans
Charge state dependant activation time for ETD scans
Data-dependent selection of Top 3 peaks for MSMS

Data Processing

Thermo Scientific Proteome Discoverer software using SEQUEST® database search algorithm, ArachnoServer spider toxin database, and PEAKS Studio 4.5 (Bioinformatics Solutions Inc.) for *de novo* analysis.



Results

Venom peptide masses ranged from 1000 to 15000 Da with about 75% of the peptides in the 4000 to 8000 Da range. Each crude spider venom contains a group of peptide in the 3000-6000 mass range and second group ranging from approximately 7000-8000 for Hi:OB and from approximately 8000-9000 for Sp004 (Figure 2). 479 peptide masses were detected in Hi:OB venom and 377 peptides in small jumping spider venom. In the Hi:OB sample, numerous toxins could be identified based on an exact match of the mass with masses of known toxins.

FIGURE 1. Intact mass analyses with high resolution LC-MS. Crude venom of Hi:OB was analyzed at 100000 resolution for intact mass determination. Full scan spectra were deconvoluted with Xtract algorithm.

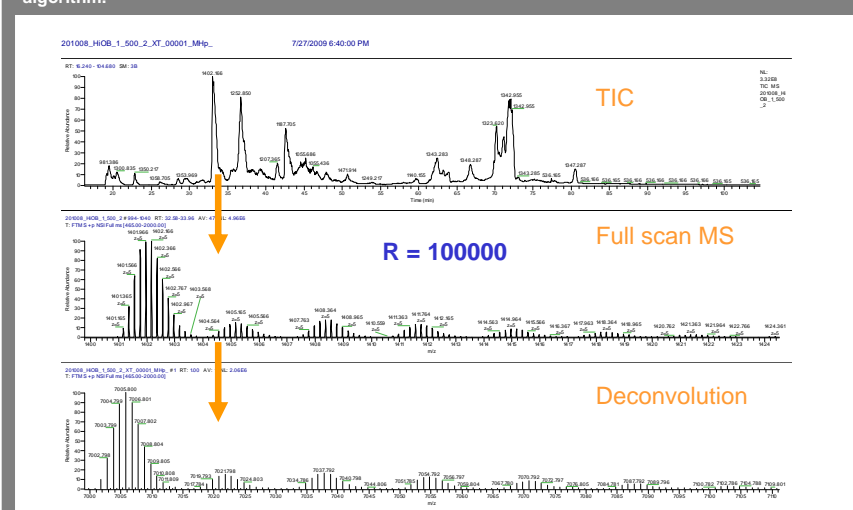
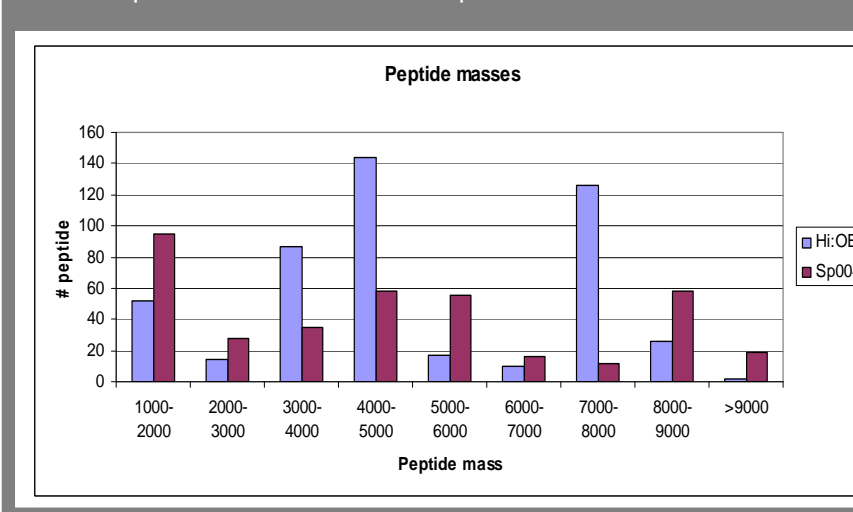
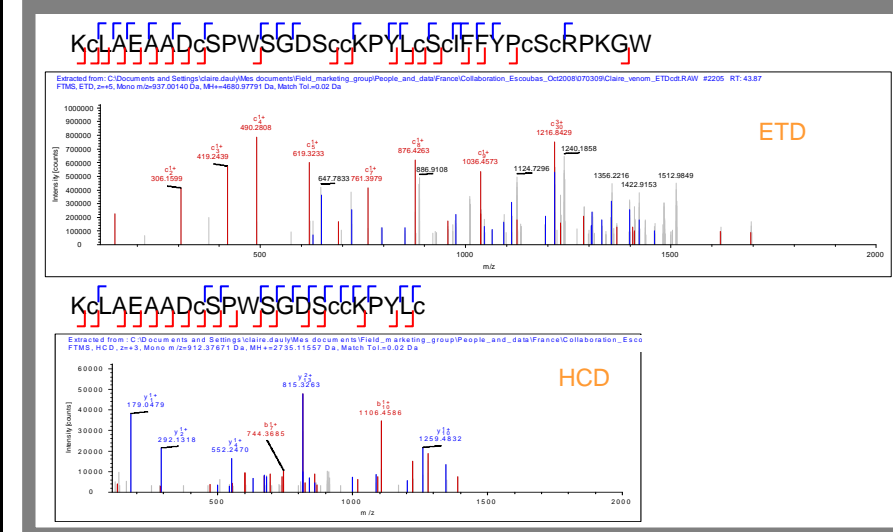


FIGURE 2. Peptide mass distribution for Hi:OB and Sp004.



A combination of HCD and ETD was used to further characterize unknown venom peptides. HCD spectra were of particular use since they contain immonium ions that are useful for *de novo* sequencing. ETD was expected to be very efficient for the fragmentation of large peptides and small proteins. For this study, it allowed the generation of numerous sequence tags, that could be matched to peptide and nucleotide sequences from venoms (Figure 3).

FIGURE 3. w-Hexatoxin-Hi1a identification by ETD and HCD. The intact peptide (m = 4680.98 Da) could be identified from ETD fragmentation whereas a truncated form (m = 2735.12 Da) was observed by HCD.



Sequence homologies were also obtained for both venom samples demonstrating the efficiency of the technique for the characterization of long peptides from very small amounts of biological material, and the feasibility of "nano-venomics" (Figure 4, 5 and 6).

FIGURE 4. w-Hexatoxin-Hi1b identification by sequence homology search after *de novo* sequencing of ETD MSMS spectra.

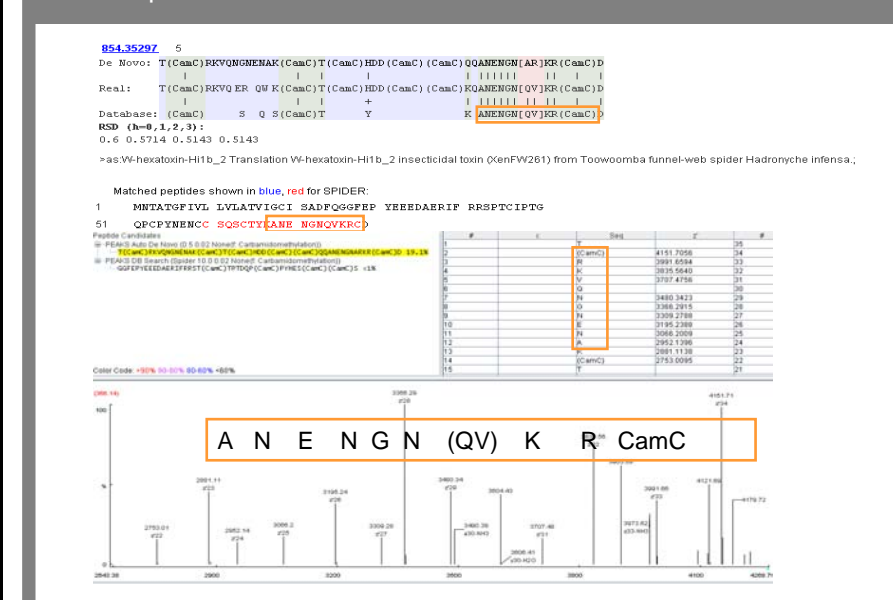


FIGURE 5. List of peptides in Hi:OB venom matched by sequence homology searches. 10 MS/MS spectra could be matched to w-hexatoxin-Hi2b after *de novo* sequencing.

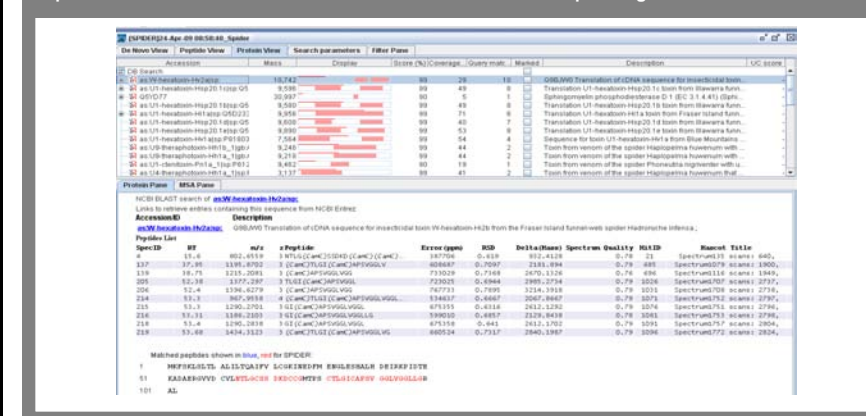
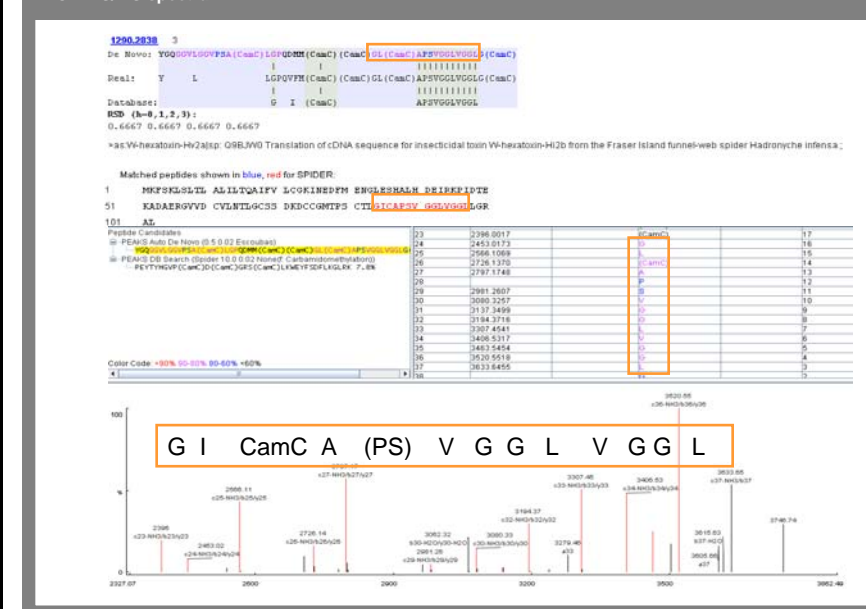


FIGURE 6. w-Hexatoxin-Hi2b identification by sequence homology search after *de novo* sequencing of HCD MS/MS spectra.



Conclusions

High resolution nano-LC-MS/MS can be used to characterize crude spider venom peptides from small amounts of spider venom with high mass accuracy.

- A combination of HCD and ETD fragmentation allowed a large number of peptides to be identified by database searching and/or sequence homology matching after *de novo* sequencing.
- Spider venomics will be further developed to allow deep mining of the enormous biological resource represented by small venomous animals. Peptide libraries based on venoms will be of major importance in the discovery of drugs and insecticides of the future.

SEQUEST is a registered trademark of the University of Washington. PEAKS is a trademark of Bioinformatics Solutions Inc. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries.