

Rafael Donadelli Melani, Livia Goto-Silva, Fábio César Sousa Nogueira, Magno Junqueira, and Gilberto Barbosa Domont

Contents

Introduction	368
Shotgun in Venomics	369
Frontier Methodologies in Shotgun Venomics	370
Conclusion and Future Directions	377
Cross-References	377
References	377

Abstract

Shotgun proteomics relies on the identification, quantification, and characterization of proteins in complex samples. Recent advances in instrumentation allow for sensitive and comprehensive shotgun protein analysis in a high-throughput manner. Combination of shotgun techniques to novel analytical strategies opens interesting possibilities for the implementation of new approaches and methodologies in the frontiers of venom biology. Examples are (i) identification of proteins in low abundance, using combinatorial ligand peptide libraries; (ii) relative and absolute protein quantitation; and (iii) identification of posttranslational modifications. The full potential of shotgun analysis in venomics is yet to be explored. Some of the pioneer works in the field will be reviewed.

R.D. Melani • L. Goto-Silva • F.C.S. Nogueira • M. Junqueira • G.B. Domont (✉)
Proteomics Unit, Institute of Chemistry, Federal University of Rio de Janeiro,
Rio de Janeiro, RJ, Brazil
e-mail: melanirafael@yahoo.com.br; liviagoto@gmail.com; fabiocsn@gmail.com;
fabiocsn@ufrj.br; magnojunqueira@gmail.com; gilberto@iq.ufrj.br;
gilbertodomont@gmail.com

Introduction

Venoms are complex mixtures of proteins, peptides, and other molecules that display diverse biological activities in different organisms. First overviews of venom complexity, mostly studying snakes, were performed at the end of the twentieth century with emerging “omics” technologies based in two-dimensional (2D) electrophoresis conjugated to mass spectrometry (MS) and Edman degradation (Li et al. 2004; Nawarak et al. 2003). Widespread utilization of 2D technology for mapping venom proteins was due to its reproducibility, accessibility, and relatively high resolving power, enabling the detection of dozens to hundreds of components in a single gel.

Later on, a pipeline termed “snake venomics” introduced by Calvete’s group (Calvete et al. 2007) consisted in fractionation of crude venom by reverse-phase LC and subsequent characterization combining N-terminal sequencing, SDS-PAGE, and mass spectrometric determination of the molecular masses and cysteine/cystine (SH and S-S) content. Using this approach venoms from more than 80 snakes have been characterized (Calvete 2014). More recently, a complementary approach based on solid-phase combinatorial peptide ligand library (CPLL) has been applied to this pipeline, which was useful to reveal low-abundance proteins within a mixture (Calvete et al. 2009).

Development of faster and more sensitive mass spectrometers allied with improvements in nanoflow high-performance liquid chromatographic (nLC-MS/MS) and computational power allowed characterization of complex mixtures of peptides coming from different proteins in the same analysis. On that basis, shotgun proteomics was set to directly analyze complex protein mixtures, rapidly generating a global profile of proteins within a mixture. It has been intensively applied to proteome profiling, protein quantitation, posttranslational modifications, and protein – protein interaction [for review see (Gelpi 2008, 2009; Yates 2013)].

Shotgun proteomics relies on the MS analysis of peptides obtained from complex samples after tryptic digestion, i.e., bottom-up approach. Complex mixtures of peptides are separated by liquid chromatography (LC) prior to MS analysis in direct interface to a mass spectrometer (LC-MS/MS). Originally, multidimensional protein identification technology (MudPIT) was established as one of the most important and powerful methods for peptide separation. MudPIT combines strong cation-exchange chromatography (SCX) followed by reverse-phase chromatography (RP), on line, with a mass spectrometer analyzer. Peptides are loaded and separated in the SCX column by a series of increasing salt steps. Each eluted step is individually transferred to the RP column, usually C18, and the eluted peptides are directly analyzed by MS (Link et al. 1999).

Reverse-phase chromatography and nano-liquid chromatography (nano-LC) are alternatives to MudPIT for peptide separation. RP is widely used for peptide separation in proteomics. Peptides in the mobile phase interacting with the stationary phase are gradually eluted with increasing concentrations of organic solvents. Nano-LC comes as a sophistication of this technique; it is performed in longer columns, using smaller particle size, and under controlled temperature allowing for greater

chromatographic resolution and larger dynamic range of peptide identification (Hebert et al. 2014; Thakur et al. 2011).

In both cases, peptides are ionized using electrospray (ESI), which is advantageous over matrix-assisted desorption/ionization (MALDI), as it produces multiply charged ions. This reduces the m/z of larger molecules and extends the mass range of the analyzer. Several setups of mass analyzers can be used for shotgun proteomics including scanning or ion beam type, ToF (time of flight) and quadrupole, or trapping, IT (ion trap), orbitrap, and FT-ICR (Fourier transform ion cyclotron resonance). Inside these instruments peptide fragmentation is performed to generate complementary ion pieces to cover protein sequence and enable identification [reviewed in (Nogueira and Domont 2014)]. Despite the excellent results obtained with use of shotgun proteomics in different biological areas, application of shotgun to venom analysis is still in its infancy. Developments in the field, technical challenges and perspectives, will be introduced.

Shotgun in Venomics

Advantages in the use of shotgun approaches for venom studies are obvious. First, parallelization and shortage of analysis time allows investigation of multiple samples. In this way, one can study, with high sample-to-sample reproducibility, several venom samples in which animal sex, age, and environmental differences could be compared on equal basis. Second, high sensitivity and dynamic range of today's equipment setups present the capacity to identify over 4,000 proteins from a complex mixture in a single run and proteome profile an organism in one hour analysis (Hebert et al. 2014; Thakur et al. 2011). Third, background coming from human keratins and trypsin is reduced compared to in-gel protocols because of lesser sample manipulation. And fourth, shotgun is compatible with peptide fractionation/ enrichment and quantification methods.

Early contributions from shotgun approaches to venomics relied on the combination of liquid chromatography coupled to mass spectrometry (LC-MS/MS). The application of shotgun to *Acanthophis* (death adder) venoms revealed a novel panel of venom molecules not contemplated in previous studies (Fry et al. 2002). Following this trend, Li and colleagues (Li et al. 2004) utilized multiple approaches in an attempt to fully characterize the proteome of *Naja naja atra* and *Agkistrodon halys* venoms, including in-solution digestion followed by HPLC coupled with an ion-trap MS. The first snake venom work completely based on shotgun proteomics and using high-resolution FT-ICR MS was published in 2006 by Fox and colleagues. They unraveled the proteome of *Bothrops jararaca* and *Crotalus atrox* identifying hundreds of proteins belonging to 12 and 9 protein families, respectively (Fox et al. 2006).

Shotgun proteomics has been successfully employed in recent publications. Rokyta and co-workers used information from venom gland transcriptome of *Crotalus adamanteus* as a database to catalog its venom system (Margres et al. 2014). Using nanospray LC/MSE they were able to identify 52 of the 78 unique putative toxin transcript clusters, including 44 of the 50 most highly

expressed transcripts. In another study high-throughput profiling of snake venom gland transcriptomes and proteomes of *Ovophis okinavensis* and *Protobothrops flavoviridis* were achieved using shotgun proteomics (Aird et al. 2013). In this study 100 % of transcripts that occurred at higher than contaminant levels had their corresponding peptides identified. Analytical strategy included digestion with three enzymes: trypsin, chymotrypsin, and Glu-C. Remarkably, shotgun proteomics was used in annotation of king cobra (*Ophiophagus hannah*) genome along with transcriptome data (Vonk et al. 2013). Vonk and co-workers were able to annotate open reading frames of 12 venom toxin gene families after genome assembly. Transcriptome analysis revealed 20 toxin families from which 14 protein families were identified using shotgun proteomics.

Application of MudPIT approach interfaced with MS for venome analysis was reported for *Naja kaouthia* (Kulkeaw et al. 2007). This approach allowed identification of 61 proteins classified in 12 groups. Authors compare shotgun analysis with identification of 24 visible spots in two-dimensional electrophoresis (2-DE) and highlight the limitations of 2-DE-based MS related to sample complexity, lower recovery of membrane proteins, and problems to separate highly acidic or basic proteins (Wolters et al. 2001). MudPIT approach in combination with multi-tissue transcriptomic analysis was utilized together for the first time by Haney and colleagues to explore the venome of Western black widow spider *Latrodectus hesperus*. Combining these two powerful techniques, the authors were able to identify 61 proteins including latrotoxins, inhibitor cystine knot (IKC) toxins, cysteine-rich secretory proteins (CRISPs), hyaluronidases, chitinases, serine proteases, metalloproteinases, leucine-rich repeat proteins, and latroductins (Haney et al. 2014).

Compatibility of shotgun proteomics to multiple fractionation methods opens interesting analytical possibilities. The use of combinatorial peptide ligand library (CPLL), for instance, allows identification of low-abundance proteins in complex samples (Righetti et al. 2011). Van Vaerenbergh and co-workers report the use of combinatorial peptide ligand library to disclose honeybee venom proteome using sample pretreatment to enrich for minor components followed by LC-MS/MS analysis. This strategy revealed an unexpectedly rich venom composition: 102 proteins and peptides were found; 83 were newly described in bee venom samples (Van Vaerenbergh et al. 2014).

Frontier Methodologies in Shotgun Venomics

Protein Identification

Peptide spectrum match (PSM) is the gold standard for protein identification by MS. This method relies on comparing experimental MS/MS mass spectra obtained by peptide fragmentation using collision-induced dissociation (CID), higher-energy collisional dissociation (HCD), electron-transfer dissociation (ETD), or others methods, to theoretical spectra in silico generated from a database containing

protein sequences. In general, the identification process follows from the sequence whose theoretical spectrum yields the highest matching score according to some probabilistic (e.g., false-discovery rate – FDR) or empirical (e.g., Washburn criterion) function. Computational tools, which apply PSM, are SEQUEST, ProLuCID, Mascot, MaxQuant, and others. Given database dependency, PSM is very efficient and sensitive for proteins coming from organisms with a large number of protein sequences deposited in databases. A drawback of this method is that discrepancies between the experimental data and predicted peptide sequences can avoid protein identification. Mismatches between theoretical peptide fragmentation and experimental mass spectrum can be due to unexpected posttranslational modification, amino acid substitution, or unusual fragmentation or merely because the sequence of the peptide is not in the database.

In general, the field of venom proteins faces the problem of working with un-sequenced organisms and, consequently, with fewer protein sequences in databases and limited representation of proteoforms. Thus, identification of proteins coming from venom samples must surpass incompleteness of database information. This scenario has been progressively changing over the last years with several projects involving the sequence of snake genomes (Castoe et al. 2013; Vonk et al. 2013), the use of transcriptomes (Valente et al. 2009), and more recently, the next-generation sequence (Calvete 2014). These works permitted more complete characterization of the genes expressed in active venom glands, improving the possibility of protein identification by MS.

Considering that there are numerous homologous proteins in each venom that share significant extents of identical sequence information, different strategies based in de novo sequencing combined to sequence similarity search have been developed for a more efficient identification of proteins coming from un-sequenced organisms by MS. De novo sequencing consists in the reconstruction of the original peptide sequence, manually or automatically, not taking in account protein sequence databases. Automatic interpretation of MS/MS spectra using algorithms, such as Lutefisk, Sherenga, Peaks, and PepNovo, results in a list of de novo candidate sequences. Those sequences are searched against multiple protein databases for cross-species identification using sequence alignment algorithms that perform sequence similarity search, such as MS-BLAST, FASTS, and ProBLAST. Sequence similarity search is error tolerant; this way, multiple mismatches between compared sequences are allowed, which increases cross-species identification, compared to conventional strategies. Computational simulations suggested that MS-BLAST allows an efficient cross-species identification of peptides down to 50 % of the sequence identity. Since sequence similarity search employs peptide sequence candidates rather than raw MS/MS as in the conventional database search, it can be considered an orthogonal strategy for the identification of proteins. It can be used for error-tolerant search on top of conventional searches or validate statistically borderline hits obtained by conventional database searches.

Duan and co-workers combined venom gland transcriptome and 2-DE-shotgun approach to study the venom of the Chinese orb-weaving spider *Araneus ventricosus*. Using PSM and manual de novo sequencing followed by sequence

similarity search to identify proteins and toxins, they were able to find more than 86 % of toxins on the EST database. A total of 130 of 150 nonredundant toxin sequences, including twelve sequences with a posttranslational modification correspondent to methyl esterification of glutamic acid, were unambiguously identified only by manual de novo sequencing (Duan et al. 2013). Guercio et al. have applied an automated pipeline combining MASCOT conventional search with de novo sequencing using BioAnalyst QS software and MS-BLAST-based sequence similarity search to the characterization of *Bothrops atrox* venom (Guercio et al. 2006). 2D spots containing proteins from three different stages of maturation: juveniles, subadults, and adults were submitted to this pipeline. The identification of these proteins substantiated the conclusion that snake venom is subjected to ontogenetic variations. Recently, Tashima et al. utilized an approach for efficient de novo sequencing of peptides previously uncharacterized belonging to bradykinin-potentiating peptides, poly-histidine-poly-glycine peptides, and L-amino acid oxidase fragments in the venom of two rare snake species, *Bothrops cotiara* and *B. fonsecai* and the Brazilian pit viper *B. Jararaca* (Tashima et al. 2012).

Bandeira et al. developed a sophistication of this method using the overlap of peptides coming from digestion with different proteases, which was named shotgun protein sequencing. Those overlapping fragments were used to generate accurate de novo reconstructions of various proteins in western diamondback rattlesnake venom (Bandeira et al. 2007). Recent approaches combine multiple protease digestion with triplet fragmentation method, CID, HCD, and ETD, to extract protein sequence directly from MS2 spectra leading to up to 99 % sequence accuracy (Guthals et al. 2013).

To date, the aforementioned methods for protein identification by MS were applied in different venom studies contributing to identification of countless number of new toxins. Nevertheless, a large diversity of new organisms and new toxin mixtures are still to be explored to achieve a more complete database of toxin sequences coming from various “noncanonical” organisms. The widespread implementation of next-generation sequencing platforms and facilities can propel the database repertoire for un-sequenced organisms. This will increase identification using conventional search methods. In this scenario, de novo sequencing and sequence similarity search consolidate as robust orthogonal identification methods, which will be relevant to new posttranslational modifications identification, amino acid switches, and unexpected proteolysis. On top of that, they emerge as database-independent tool for protein sequencing.

Posttranslational Modification and Venomics

Posttranslational modifications (PTMs) are covalent processing events in proteins mainly by attachment of chemical moieties and also by proteolytic processing (Portes-Junior et al. 2014). These modifications are abundant during all cellular events and may determine protein conformation, function, localization, turnover, and interaction. Over the last two decades, MS proved to be a suitable tool for

PTM identification. Advantages of MS rely upon its very high sensitivity, ability to identify sites and discover novel PTMs, capability to identify modified peptides from complex protein mixtures, and ability to quantify PTM site occupancy. Most studied PTMs are phosphorylation, glycosylation, acetylation, and ubiquitination (Jensen 2006). Additionally other modifications are found in proteins from venoms like proteolysis, hydroxylation, and carboxylation (Buczek et al. 2005).

PTM analysis is challenging mainly because they are located at specific amino acid residues in proteins, usually present in substoichiometric levels and, some of them, dynamically regulated, such as, phosphorylation and acetylation. In addition, some PTMs are labile during MS and MS/MS. Furthermore, some of the modifications increase hydrophobicity, which complicates sample processing, may affect the cleavage efficiency of proteases, such as trypsin, and reduce ionization and detection efficiency in MS. For all these reasons, it is necessary to employ complementary approaches to have success in PTM analysis, comprising the use of high-resolution mass spectrometers, combination of multiple protease treatments and shotgun proteomic approach, enrichment of modified proteins and/or peptides prior to MS, application of specific MS methods, and the use of particular bioinformatics tools.

Strategies employed to enrich modified peptides are specific to each kind of PTM. For phosphorylated peptides, the main strategies are immunoprecipitation with anti-phosphoserine (pS), anti-phosphothreonine (pT), or anti-phosphotyrosine (pY); affinity chromatography using immobilized metal affinity chromatography (IMAC), titanium dioxide (TiO₂), or SIMAC, which is a sequential elution of IMAC combining TiO₂; chemical derivatization where phosphate groups are β-eliminated and subsequently undergo Michael addition; and affinity purification of modified peptides. In the case of glycopeptide enrichment, the key strategies are lectin affinity chromatography, hydrophilic interaction liquid chromatography (HILIC), hydrazine chemistry, and TiO₂ for sialylated peptides. For other modifications, the enrichment methods are based mainly in immunoaffinity experiments applying specific antibodies. These strategies to enrich modified peptides can be used in tandem, reaching better specificity (for a review see (Olsen and Mann 2013).

PTM characterization is a current topic in venom analysis. In the last decades, PTM strategies have been focused mainly in proteins isolated by reverse-phase liquid chromatography from different venoms. Another common strategy is critically dependent on two-dimensional electrophoresis (2-DE) for venom analysis and staining with specific PTM dyes, for example, ProQ Diamond for phosphoproteins and ProQ Emerald for glycoproteins (Birrell et al. 2006, 2007). Shotgun proteomic and PTM characterization strategies employed to venomics are still in their infancy, and only few cases are found in the literature, such as Verano-Braga and co-workers' (Verano-Braga et al. 2013) in which a large-scale venom analysis was performed in combination with phospho- and glycopeptide enrichments for PTM identification and localization of scorpion toxins, from *Tityus serrulatus*. In a pipeline combining transcriptome and shotgun approach to study multidomain toxins in centipedes, Undheim and colleagues identified posttranslational processing localization of the mature toxins at centipedes' venom gland (Undheim et al. 2014).

Venom proteins and peptides undergo a plethora of PTM variations that are important for their action, and, for this reason, each modification is fine-tune adjusted, as reported by Resende and co-workers (Resende et al. 2013). The authors identified phosphorylation sites in two major toxins, icarapin and melittin, from Africanized and European honeybee venoms. Pharmacological tests demonstrated that melittin phosphorylated at 18Ser was less toxic compared to the native peptide revealing toxicity alteration by PTM (Resende et al. 2013). The use of shotgun proteomics allied to strategies for PTM enrichment, such as TiO₂ and IMAC for phosphorylation as well as lectin affinity and hydrophilic chromatography for glycosylation, can project venomics to a new level, allowing identification of new PTMs, site occupancy, and modification stoichiometry. Besides, new advances in proteome analysis, e.g., middle- and top-down strategies, are able to identify different modification in venom toxins, making possible characterization of cross talk among PTMs and their roles in toxin action and function.

Quantitative Proteomics

Protein quantitation in shotgun proteomics may provide precise information about complex sample composition. In venomics this information is vital to compare venoms as well as to link protein composition with pharmacological or pathological effects, ultimately setting the ground for new treatments and development of antivenoms [for details, see the review (Nogueira and Domont 2014)].

Two different types of protein quantitation can be performed in shotgun proteomics experiments: relative protein quantitation, most commonly applied, based on protein ratio between two or more samples, or absolute protein quantitation that determines protein concentrations or copy numbers present in each sample. Relative quantitation is basically performed using label-based or label-free methods. The first is based on labeling proteins with stable isotopes *in vivo* (metabolically) or *in vitro* (chemically or enzymatic) with molecules composed of light and heavy isotopes of ²H, ¹³C, ¹⁵N, and ¹⁸O or isobaric tags. On the other hand, label-free approaches use normalized ion intensity of identical peptides, extracted ion chromatogram, or normalized spectral counts of each protein to compare protein abundance in distinct samples.

Absolute quantitation is based in selected reaction monitoring (SRM) or multiple reaction monitoring (MRM) methods. It requires the use of standard, isotopically labeled synthetic proteotypic peptides, to target selected peptides combined with internal standards.

ICAT (isotope-coded affinity tag) was introduced in 1999 by Steven Gygi and colleagues. It is based on the labeling of sulfhydryl lateral chain of cysteine residues using light and heavy versions of the reagent (Gygi et al. 1999). Many snake venom proteins are extremely cysteine rich, making them excellent targets for this approach. Advantage of the ICAT approach is that enrichment of peptides containing cysteine reduces complexity of the sample. On the other hand, it also reduces the number of identified peptides per protein and cannot detect proteins/peptides that do not contain cysteine residues. Other disadvantages are shift on

retention time of light- and heavy-labeled peptides and increase in difficulty to analyze MS2 spectra because many observed fragments come from the marker instead of the polypeptide chain.

A variant of ICAT original method termed SoPIL (polymer-based isotope labeling) uses soluble polymers instead of solid phase to avoid heterogeneous reaction conditions and nonlinear kinetics. This approach was utilized to compare type A and B venom of *Crotalus scutulatus scutulatus* and the venom of two regionally distinctive snakes *Crotalus oreganus helleri* and *Bothrops colombiensis*. Both comparisons were able to quantify and identify ca. 100 unique peptides representing over 30 venom proteins. Quantitative data were in accord with pharmacological/biological tests and supported the theory that there are intraspecific variation in venoms of some species and interspecies from different geographical locations. This was the first quantitative proteomics analysis of snake venom based on stable isotope labeling (Galan et al. 2008).

Isobaric mass tags are widely utilized nowadays in shotgun proteomics. Isobaric mass tags available are iTRAQ (isobaric tags for relative and absolute quantification) and TMT (tandem mass tags). Both tags rely on protein chemistry targeting primary amines, N-terminal and ϵ -amino and groups, to label peptides. Each plex reagent has the same mass achieved by a combination of ^{13}C , ^{15}N , and ^{18}O in the reporter and balance groups. Reporter subunit is responsible for peptide quantitation; when fragmented from the iTRAQ reagent, distinct reporter ion masses are recorded in the m/z spectrum lower than 150 Da. A clear-cut difference to other approaches is the possibility of multiplexing with iTRAQ, 4 or 8-plex, and TMT, 2, 6, or 10-plex. Derived peptides have similar retention times in LC and appear as a single isobaric peak in MS without increase sample complexity at this level.

iTRAQ was used for the first time in a quantitative analysis of venom proteins by Zelanis and co-workers. Venom of newborns and adults of *Bothrops jararaca* were analyzed to ontogenetic variation. Isobaric label approach was able to identify and quantify 29 venom proteins at different live stages. Thus, snake venom metalloproteinases (SVMP) were detected as more abundant toxin family in the newborn venom; snake venom serine proteinases (SVSP) were found as major toxin family in the adult venom; cysteine-rich secretory proteins (CRISP) and snake venom vascular endothelial growth factor (svVEGF) are more abundant in newborns and L-amino acid oxidases (LAAO) in adults. No clear differential expression was detected in phospholipases A2 (PLA2) toxin family. These differences suggest a shift of the main components upon transition from newborn to maturity (Zelanis et al. 2011).

Label-free quantification workflows have become extremely popular in shotgun experiments, and as the title suggests, this method quantifies proteins/peptides without any additional labeling step. Quantification is already present in collected LC-MS/MS data, an attractive simple feature for its simplicity, cost, and reproducibility, deleting steps, reducing sample complexity, and allowing a larger number of samples for comparison than multiplex experiments. However, it is not as accurate as isotope or isobaric label, and quantification quality is strongly dependent on reproducibility of the LC-MS/MS data and on bioinformatics tools for processing. There are two main concepts applied for sample comparison using label-free: the first uses the comparison

of peak area (extracted ion chromatogram) generated by mass spectrometric signal intensity for any given peptide, and the second is based on spectra counts, the number of acquired spectra matching to a peptide/protein.

Using label-free quantification based on data-dependent acquisition (DDA) and extracted ion chromatogram, Resende and colleagues compared the venom proteome of Africanized honeybees (AHB) with that of two European subspecies, *Apis mellifera ligustica* and *A. m. carnica*. From 51 identified proteins, only 11 significantly changed among all samples, and nine of these are enriched in AHB venom. Melittin was the most abundant toxin in all venoms; second major toxin phospholipase A2 and mast cell degranulating peptide are more abundant in AHB species. Melittin polymorphism in each venom sample was also quantified, and an alanine residue in the variable position (15Ala) was the most abundant for all bee species. Considering these results in terms of neutralization of toxic effects, antivenoms specific for AHB should be effective in the treatment of toxic reactions from other European honeybee subspecies (Resende et al. 2013).

Peptide or spectral count approach relies on the empirical observation that most abundant proteins in a sample present more tandem MS spectra and peptide spectrum matches (PSM) than less abundant proteins. Quantitation is performed correlating the number of PSM and sequence coverage to the amount of a protein in a sample, a proportion that exists in data-dependent acquisition experiments.

Using only the number of spectra-matched peptides to infer the relative protein abundance Tayo and colleagues performed a qualitative and quantitative work in the differences of conotoxin components presented in the proximal, central, and distal sections of the *Conus textile* venom duct. King-Kong 2 toxin was found as the most abundant peptide in the entire venom duct followed by conotoxins TxO4 and TxO6. Most of the identified *C. textile* toxins were differentially expressed in the venom duct, like TxMKLT-0223 toxin only identified at the central region or King-Kong 2 present in all parts but more abundant in proximal region. Based in these results, authors suggested that specialization for conotoxin biosynthesis occurs in the different parts of the venom duct (Tayo et al. 2010).

In another paper, venoms from six different species of Bothrops complex were analyzed by shotgun proteomics in a comparative and phylogenetic study. The normalized mean of each protein family spectral counts was utilized to measure the amount of toxin families in venoms. Quantitative and qualitative differences were observed in venom composition of the Bothrops complex, mostly for *B. jararacussu* venom. However there was no apparent significant relationship between phylogeny of the snakes and venom composition (Sousa et al. 2013).

Employing emPAI index (exponentially modified protein abundance index), Li and co-workers compared quantitatively and qualitatively manually extracted glandular venom and venom extracted through the use of electrical stimulation from Italian honeybees (*Apis mellifera ligustica*). Twenty proteins were identified in the samples, and 9 venom toxins showed higher abundance in electrically extracted venom than in gland venom. Further 5 toxins exhibit no significant difference between samples, and other 6 were only identified in one sample. In this fashion, this study demonstrated that venom extracted manually is different

from venom extracted using electrical stimulation, and these differences may be important in their use as pharmacological agents (Li et al. 2013).

Overall, distinct shotgun approaches for quantification were utilized to evaluate toxin/protein content in animal venoms. However, multiplex quantification and cross-species quantification remain a challenge in the field. Furthermore, application of absolute quantitation is still to be explored.

Conclusion and Future Directions

Literature search demonstrates that the study of venoms is still mostly performed using classical proteomics approaches rather than up-to-date rationales and instrumentation. To quickly advance in the field, studies using shotgun proteomics, an integrative action that employs modern mass spectrometers, rationales, approaches and techniques, must be implemented.

In spite of still being a difficult task to accomplish, venomes do not have the same high protein complexity of cell extracts, and the challenge to qualify and quantify their proteomes is a less difficult enterprise. The compelling necessity to understand and compare pathophysiological effects of venoms can be achieved using the modern arsenal of instruments and proteomics techniques. Individual and biologically related venomes can be characterized including posttranslational modifications; their proteins/ toxins identified and quantified. It is imperative, however, that proteomic scientists and cell biologists join efforts to disclose molecular/cellular targets of toxins, specificities of interaction and action, physicochemical and conformational parameters of these interactions, as well as the regulatory pathways involved. This knowledge will help to postulate new strategies to improve efficiency of immunotherapies for treating, principally, snake envenomation and local tissue degradation/regeneration, a neglected disease according to the World Health Organization.

Cross-References

- ▶ [Snake Venom Peptidomics](#)
- ▶ [Snake Venom Proteopeptidomics: What Lies Behind the Curtain](#)

References

- Aird SD, Watanabe Y, Villar-Briones A, Roy MC, Terada K, Mikheyev AS. Quantitative high-throughput profiling of snake venom gland transcriptomes and proteomes (*Ovophis okinavensis* and *Protobothrops flavoviridis*). *BMC Genomics*. 2013;14:790.
- Bandeira N, Clauser KR, Pevzner PA. Shotgun protein sequencing: assembly of peptide tandem mass spectra from mixtures of modified proteins. *Mol Cell Proteomics*. 2007;6:1123–34.

- Birrell GW, Earl S, Masci PP, Jersey JD, Wallis TP, Gorman JJ, Lavin MF. Molecular diversity in venom from the Australian brown snake, *Pseudonaja textilis*. *Mol Cell Proteomics*. 2006;5(2):379–89.
- Birrell GW, Earl STH, Wallis TP, Masci PP, Jersey JD, Gorman JJ, Lavin MF. The diversity of bioactive proteins in Australian snake venoms. *Mol Cell Proteomics*. 2007;6(6):973–86.
- Buczek O, Bulaj G, Olivera BM. Conotoxins and the posttranslational modification of secreted gene products. *Cell Mol Life Sci*. 2005;62:3067–79.
- Calvete JJ. Next-generation snake venomomics: protein-locus resolution through venom proteome decomplexation. *Expert Rev Proteomics*. 2014;11(3):315–29.
- Calvete JJ, Juárez P, Sanz L. Snake venomomics. *Strategy Appl J Mass Spectrom*. 2007;42:1405–14.
- Calvete JJ, Fasoli E, Sanz L, Boschetti E, Righetti PG. Exploring the venom proteome of the western diamondback rattlesnake, *Crotalus atrox*, via snake venomomics and combinatorial peptide ligand library approaches. *J Proteome Res*. 2009;8:3055–67.
- Castoe TA, de Koning AP, Hall KT, Card DC, Schield DR, Fujita MK, Ruggiero RP, Degner JF, Daza JM, Gu W, et al. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proc Natl Acad Sci U S A*. 2013;110:20645–50.
- Duan Z, Cao R, Jiang L, Liang S. A combined de novo protein sequencing and cDNA library approach to the venom analysis of Chinese spider *Araneus ventricosus*. *J Proteomics*. 2013;78:416–27.
- Fox JW, Ma L, Nelson K., Sherman NE, Serrano SMT. Comparison of indirect and direct approaches using ion-trap and Fourier transform ion cyclotron resonance mass spectrometry for exploring viperid venom proteomes. *Toxicon*. 2006; 47: 700–14.
- Fry BG, Wickramaratna JC, Hodgson WC, Alewood PF, Kini RM, Ho H, Wüster W. Electrospray liquid chromatography/mass spectrometry fingerprinting of *Acanthophis* (death adder) venoms: taxonomic and toxicological implications. *Rapid Commun Mass Spectrom*. 2002;16:600–8.
- Galan JA, Guo M, Sanchez EE, Cantu E, Rodriguez-Acosta A, Perez JC, Tao WA. Quantitative analysis of snake venoms using soluble polymer-based isotope labeling. *Mol Cell Proteomics*. 2008;7:785–99.
- Gelpi E. From large analogical instruments to small digital black boxes: 40 years of progress in mass spectrometry and its role in proteomics. Part I 1965–1984. *J Mass Spectrom*. 2008;43:419–35.
- Gelpi E. From large analogical instruments to small digital black boxes: 40 years of progress in mass spectrometry and its role in proteomics. Part II 1985–2000. *J Mass Spectrom*. 2009;44:1137–61.
- Guercio RA, Shevchenko A, Lopez-Lozano JL, Paba J, Sousa MV, Ricart CA. Ontogenetic variations in the venom proteome of the Amazonian snake *Bothrops atrox*. *Proc Natl Acad Sci U S A*. 2006;4:11.
- Guthals A, Clauser KR, Frank AM, Bandeira N. Sequencing-grade De novo analysis of MS/MS triplets (CID/HCD/ETD) from overlapping peptides. *J Proteome Res*. 2013;12:2846–57.
- Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol*. 1999;17:994–9.
- Haney RA, Ayoub NA, Clarke TH, Hayashi CY, Garb JE. Dramatic expansion of the black widow toxin arsenal uncovered by multi-tissue transcriptomics and venom proteomics. *BMC Genomics*. 2014;15:366.
- Hebert AS, Richards AL, Bailey DJ, Ulbrich A, Coughlin EE, Westphall MS, Coon JJ. The one hour yeast proteome. *Mol Cell Proteomics*. 2014;13:339–47.
- Jensen ON. Interpreting the protein language using proteomics. *Nat Rev Mol Cell Biol*. 2006;7:391–403.
- Kulkeaw K, Chaicumpa W, Sakolvaree Y, Tongtawe P, Tapchaisri P. Proteome and immunome of the venom of the Thai cobra, *Naja kaouthia*. *Toxicon*. 2007;49:1026–41.
- Li S, Wang J, Zhang X, Ren Y, Wang N, Zhao K, Chen X, Zhao C, Li X, Shao J, et al. Proteomic characterization of two snake venoms: *Naja naja atra* and *Agkistrodon halys*. *Biochem J*. 2004;384:119–27.

- Li R, Zhang L, Fang Y, Han B, Lu X, Zhou T, Feng M, Li J. Proteome and phosphoproteome analysis of honeybee (*Apis mellifera*) venom collected from electrical stimulation and manual extraction of the venom gland. *BMC Genomics*. 2013;14:766.
- Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates 3rd JR. Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol*. 1999;17:676–82.
- Margres MJ, McGivern JJ, Wray KP, Seavy M, Calvin K, Rokyta DR. Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *J Proteomics*. 2014;96:145–58.
- Nawarak J, Sinchaikul S, Wu C-Y, Liao M-Y, Phutrakul S, Chen S-T. Proteomics of snake venoms from elapidae and viperidae families by multidimensional chromatographic methods. *Electrophoresis*. 2003;24(16):2838–54.
- Nogueira FC, Domont GB. Survey of shotgun proteomics. *Methods Mol Biol*. 2014;1156:3–23.
- Olsen JV, Mann M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol Cell Proteomics*. 2013;12:3444–52.
- Portes-Junior JA, Yamanouye N, Carneiro SM, Knittel PS, Sant'Anna SS, Nogueira FC, Junqueira M, Magalhães GS, Domont GB, Moura-da-Silva AM. Unraveling the processing and activation of snake venom metalloproteinases. *J Proteome Res*. 2014;13:3338–48.
- Resende VMF, Vasilij A, Santos KS, Palma MS, Shevchenko A. Proteome and phosphoproteome of Africanized and European honeybee venoms. *Proteomics*. 2013;13:2638–48.
- Righetti PG, Fasoli E, Boschetti E. Combinatorial peptide ligand libraries: the conquest of the 'hidden proteome' advances at great strides. *Electrophoresis*. 2011;32:960–6.
- Sousa LF, Nicolau CA, Peixoto PS, Bernardoni JL, Oliveira SS, Portes-Junior JA, Mourão RHV, Lima-dos-Santos I, Sano-Martins IS, Chalkidis HM, et al. Comparison of phylogeny, venom composition and neutralization by antivenom in diverse species of *Bothrops* complex. *PLoS Negl Trop Dis*. 2013;7:e2442.
- Tashima AK, Zelanis A, Kitano ES, Ianzer D, Melo RL, Rioli V, Sant'anna SS, Schenberg AC, Camargo AC, Serrano SM. Peptidomics of three *Bothrops* snake venoms: insights into the molecular diversification of proteomes and peptidomes. *Mol Cell Proteomics*. 2012;11:1245–62.
- Tayo LL, Lu B, Cruz LJ, Yates JR. Proteomic analysis provides insights on venom processing in *Conus textile*. *J Proteome Res*. 2010;9:2292–301.
- Thakur SS, Geiger T, Chatterjee B, Bandilla P, Fröhlich F, Cox J, Mann M. Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol Cell Proteomics*. 2011;10:M110.003699.
- Undheim EA, Sunagar K, Hamilton BR, Jones A, Venter DJ, Fry BG, King GF. Multifunctional warheads: diversification of the toxin arsenal of centipedes via novel multidomain transcripts. *J Proteomics*. 2014;102:1–10.
- Valente RH, Guimarães PR, Junqueira M, Neves-Ferreira AG, Soares MR, Chapeaurouge A, Trugilho MR, León IR, Rocha SL, Oliveira-Carvalho AL, et al. *Bothrops insularis* venom: a proteomic analysis supported by transcriptomic-generated sequence data. *J Proteomics*. 2009;72:241–55.
- Van Vaerenbergh M, Debyser G, Devreese B, de Graaf DC. Exploring the hidden honeybee (*Apis mellifera*) venom proteome by integrating a combinatorial peptide ligand library approach with FTMS. *J Proteomics*. 2014;99:169–78.
- Verano-Braga T, Dutra AA, Leon IR, Melo-Braga MN, Roepstorff P, Pimenta AM, Kjeldsen F. Moving pieces in a venom puzzle: unveiling post-translationally modified toxins from *Tityus serrulatus*. *J Proteome Res*. 2013;12:3460–70.
- Vonk FJ, Casewell NR, Henkel CV, Heimberg AM, Jansen HJ, McCleary RJ, Kerckamp HM, Vos RA, Guerreiro I, Calvete JJ, et al. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc Natl Acad Sci U S A*. 2013;110:20651–6.
- Wolters DA, Washburn MP, Yates 3rd JR. An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem*. 2001;73:5683–90.

- Yates JR. The revolution and evolution of shotgun proteomics for large-scale proteome analysis. *J Am Chem Soc.* 2013;135:1629–40.
- Zelanis A, Tashima AK, Pinto AFM, Paes Leme AF, Stuginski DR, Furtado MF, Sherman NE, Ho PL, Fox JW, Serrano SMT. *Bothrops jararaca* venom proteome rearrangement upon neonate to adult transition. *Proteomics.* 2011;11:4218–28.