EuPA Open Proteomics 8 (2015) 144–156

Contents lists available at ScienceDirect

EuPA Open Proteomics

journal homepage: www.elsevier.com/locate/euprot

Seeing beyond the tip of the iceberg: A deep analysis of the venome of the Brazilian Rattlesnake, *Crotalus durissus terrificus*

Rafael D. Melani[a], Gabriel D.T. Araujo[a], Paulo C. Carvalho[b], Livia Goto[a], Fábio C.S. Nogueira[a], Magno Junqueira[a], Gilberto B. Domont[a,*]

[a] *Proteomics Unit, Rio de Janeiro Proteomics Network, Departamento de Bioquímica. Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*
[b] *Laboratory for Proteomics and Protein Engineering, Carlos Chagas Institute, Fiocruz, Paraná, Brazil*

ARTICLE INFO

Article history:
Received 28 February 2015
Received in revised form 18 May 2015
Accepted 19 May 2015
Available online 6 June 2015

Keywords:
Crotalus durissus terrificus venom
In-solution digestion
IEF
CPLL
Bottom-up approaches
Shotgun

ABSTRACT

The complete characterization of the snake venom protein components is a requirement for a systems-wide understanding of their biological context. In this work, we provide a deep proteomic characterization of *Crotalus durissus terrificus* venom using different bottom-up approaches. We identified more than five times more protein families than the sum of all identifications previously reported. For the first time in this sub-species, we report the identification of three new toxin families: CRISP, phospholipase-B, and SVVEGF. This work also describes proteins involved in regulation of toxin synthesis and processing that are present in venom.

Introduction

Venomous animals rely on an extensive array of toxins for prey capture and defense. Venom toxins represent a huge and under explored reservoir of bioactive components that could be used as drug library, [1,2] because toxins in venom are directed against a wide variety of pharmacological targets [3–5]. Studies of venoms allow: (I) a better understanding of the mechanism of venoms and toxins actions; (II) development of specific research tools; (III) new improvements in therapies to treat snake bite envenomation; and (IV) development of new drug candidates [5].

A great number of snake toxins have been used in vivo for pharmacological tests and several are currently being developed as novel experimental therapeutics [6]. An example of a new drug candidate is the cobra venom factor used for treatment of diseases of the complement system [7]. A good example of a toxin used as a drug is the bradykinin-potentiating peptide; it also aided in the understanding cardiovascular physiology and development of Captopril®, the first active-site directed inhibitor of angiotensin-converting enzyme, used worldwide to treat human hypertension. This discovery was made possible by studies of *Bothrops jararaca* venom [8–10].

Comprehensive analysis of venom proteomes became possible at the end of the twentieth century together with the emerging "omics" technologies based on 2D electrophoresis combined with Edman degradation and/or mass spectrometry (MS) analysis [11,12]. Soon after, venome analysis relying completely in LC-MS mass fingerprint, RP-HPLC and Edman sequencing was introduced [13]. In 2004 Calvete's group developed a pipeline to explore snake venom proteomes and coined it "snake venomics". This pipeline is one of the most widely-used venomics approaches and is based on RP-HPLC venom fractionation in C18 columns followed by fraction characterization by N-terminal Edman sequencing, SDS–PAGE in reduced and non-reduced conditions, cysteine (—SH and S—S) content verification, and molecular mass determination by MS. Afterwards, electrophoretic bands are excised, trypsinized and submitted to MS analysis; all data are used to identify and quantify protein relative abundance. The results are displayed as a pie chart of protein families found in the venom [14,15].

The following years showed a fast and steady increase in snake venom proteome publications with a variety of technical combinations (see reviews [4,16,17]). The use of shotgun proteomics with hybrid high resolution mass spectrometers (LTQ-FT) was disclosed by Fox and coworkers [18]. The first venom study using a bottom-up proteomic approach combined

* Corresponding author.
E-mail address: gilberto@iq.ufrj.br (G.B. Domont).

http://dx.doi.org/10.1016/j.euprot.2015.05.006

with peptide spectral matching, spectral alignment, and de novo sequencing analyses, aiming to improve protein coverage and identification in western diamondback rattlesnake venom, was published in 2007 [19].

One of the shortcomings in most widely adopted venomics protocols is that it is difficult to identify low abundance proteins in very complex mixtures with wide dynamic range. One way to access this "hidden proteome" is to use a combinatorial peptide ligand library (CPLL) [20]. When this equalizer technology was applied to explore the venome of *Crotalus atrox* and *Bitis arietans,* several low abundance proteins were identified. These low abundance proteins are not likely to be directly involved in generic killing and/or digestive functions, instead they may be component of cellular debris, proteins related to toxin sorting and processing, and/or proteins under neutral selection to became new toxins [21,22]. Therefore, decomplexing venoms prior to MS analysis has shown to be a valuable tool for exploring venom complexity.

Moreover, proteomic approaches are mostly database-dependent and the lack of sequenced snake genomes and transcriptomes are crucial problems for high throughput and efficient identification of venom proteins. However, *de novo* sequencing of high quality MS/MS data followed by sequence similarity search to homologous proteins has proven to be an excellent alternative for non-genome sequenced organisms [23,24]. Some of the tools available to perform automated *de novo* sequencing with high accuracy in large LC–MS/MS data are: PEAKS [25], PepNovo [26], pNovo+ [27]. Protein identification by similarity sequence search can be provided by MS-BLAST [28] or PepExplorer, a new similarity search tool that uses artificial neural networks to disclose protein identification [29]. It must be stressed that the use of different approaches is complementary and the choice will depend on the proposed problem.

Here we propose a pipeline to create a reliable, comprehensive resource of *Crotalus durissus terrificus* (Cdt) venom that is also able to explore low abundance proteins. Cdt venon is one of the most studied snake venoms in the world with more than 390 papers listed in PubMed (15th January 2015). Early studies in the late nineteenth and early twentieth centuries reported distinct pharmacological effects of *Crotalus* and *Bothrops* venom as well as the specificity of respective antivenom sera [30]. The envenomation induced by Cdt bite predominantly precipitates neurotoxic and myotoxic [31] effects because of the presence of high amounts of crotoxin, the first isolated animal venom toxin [32]. In 2010 two distinct papers exploring the Cdt venome were published. Calvete et al. (2010) using venomics pipeline showed venom variations of Central American rattlesnake *Crotalus simus* and South American *Crotalus durissus* complexes, indicating the presence of seven toxin classes in Cdt venom [33]. The second study was based on 2D electrophoresis and seven protein families were identified [31], but only three were also found in Calvete's study. In our pipeline, we combined three bottom-up approaches: venom in-solution digestion (ISD), peptide pre-fractionation by isoelectric focusing (IEF), and combinatorial peptide ligand library (CPLL) techniques. MS data were searched and analyzed by peptide spectrum match (PSM) and *de novo* sequencing coupled with similarity sequence search (SSS) approaches to identify proteins according to their similarity.

We were able to show that, even taking one of the most studied snake venom, our combined approach to Cdt venom identified 77% more protein families than reported in all previous studies, with the identification of new toxin families present in this venom as well as non-toxin protein families. Overall, we demonstrate that a combination of pre-fractionation approaches in combination with a LC–MS/MS platform can provide deep characterization of protein components in venom.

Material and methods

Venom sample

A pool of lyophilized *Crotalus durissus terrificus* (Cdt) venom from adult specimens was obtained from the Butantan Institute, São Paulo – Brazil, as described by Giannotti et al. [34], and generously donated by Prof Ana Maria Moura da Silva from the Laboratory of Immunopathology – Butantan Institute. Venom protein concentration was determined fluorometrically using Qubit 2.0 protein assay kit following manufacturer's instructions (Invitrogen).

Combinatorial peptide ligand library enrichment

We employed a combinatorial hexapeptide ligand library (CPLL) approach using three different pHs [21,22]. Three aliquots of crude venom (20 mg each) were dissolved at room temperature in 5 mL of three distinct pH buffers (buffer pH 3.0, 25 mM sodium phosphate monobasic and 50 mM potassium chloride; buffer pH 7.0, 25 mM sodium phosphate dibasic and 50 mM potassium chloride; buffer pH 11.0, 25 mM CAPS and 50 mM potassium chloride) containing complete ultra-tablets protease inhibitor cocktail (Roche). Each venom sample was centrifuged at 10,000×*g* for 10 min and the supernatant was incubated for 3 h under gentle shaking at room temperature with 30 μL of CPLL beads (generously donated by Dr. Pier Giorgio Righetti, Politecnico di Milano, Italy). CPLL beads were previously hydrated and equilibrate for 1 h with the respective pH buffer. The suspensions were centrifuged at 2000×*g* for 5 min, the supernatants containing the non-captured proteins were reserved and the beads slurry were washed three times with 1 mL of the respective pH buffer to eliminated unbound or non-specific bonds. Captured proteins yet bond to the beads [35] and flow through proteins were submitted to trypsin digestion, separately.

Protein digestion

Four aliquots of one hundred micrograms of venom protein, three different flow through of CPLL containing 100 μg of protein each, and three distinct pH CPLL beads with captured proteins were suspended in urea 7 M and thiourea 2 M. DTT was added to a final concentration of 10 mM and reacted for 1 h at room temperature. Cysteine residues were carboxamidomethylated with addition of 10 mM iodoacetamide for 30 min in the dark. Samples were diluted to 1 M urea with *Tris*-HCL 100 mM, pH 8.5, and MS grade trypsin (Promega) was added (1:50 protease/substrate (w/w)) for overnight digestion at 37 °C. Proteolyses was stopped by adding formic acid to a 1% final concentration. Peptide solutions were centrifuged 10,000×*g* for 5 min to pellet non-soluble materials and/or CPLL beads. Supernatants were desalted in C18 Micro SpinColumns (Harvard Apparatus) following manufacturer's instructions.

Isoelectric focusing (IEF)

In duplicate, 100 μg of digested peptides dissolved in 200 μL of 8 M urea/2 M thiourea were loaded into strip holders to which a 3–10 pH IPG strip (18 cm, GE Healthcare Life Sciences) was laid over and covered with mineral oil. Isoelectric focusing was performed in the IPGphor system (GE Healthcare Life Sciences) for 24 h employing rehydration for 12 h at 20 °C and 50 μA/strip, focusing 0–200 V for 1 h, 200–500 V for 1 h, 500–1000 V for 1 h, 1000 V for 1 h, 1000–8000 V for 30 min, 8000 V for 5 h, and 500 V for 3 h totaling 46000 Vh. After focusing, strips were washed in petroleum ether, divided in 5 pieces and covered with peptide extraction

solution (95% ACN 5% formic acid 0.1) for 24 h, at room temperature. Peptides were dried and submitted to clean up in C18 Micro SpinColumns (Harvard Apparatus) following manufacturer's instructions.

LC–MS/MS data acquisition

Peptide samples were resuspended in 0.5% formic acid (FA) and fractionated using a nanoHPLC system Easy-nLC II (Proxeon) on an in-house packed 2 cm × 150 μm i.d. pre-column (Reprosil-Pur C18-AQ, 5 μm, 120 Å, Dr. Maisch), and 20 cm × 75 μm i.d. column (Reprosil-Pur C18-AQ, 3 μm, 120 Å, Dr. Maisch) coupled to a LTQ Velos Orbitrap mass spectrometer (Thermo Scientific). Chromatography was performed at 300 nL/min flow rate with 95% water, 5% ACN and 0.1% FA as mobile phase A and 95%, 5% water and 0.1% FA as phase B. The column was equilibrated with buffer A for 10 min and the runs were performed in optimized gradient of 180 min (5–45% B over 160 min, followed by 45–90% B over 15 min, and 90% B for 5 min).

The Orbitrap mass spectrometer was controlled by Tune 2.6.0 and Xcalibur 2.1 software and was set to operate in data dependent acquisition (DDA) mode to automatically switch between full scan MS (60,000 resolution, 500 ms accumulation time, AGC 1×10^6 ions, range 300–2000 m/z) and MS/MS acquisition. The top 10 most intense ions were selected for fragmentation by low resolution CID–MS/MS (35 normalized collision energy; 10 ms activation time; 100 ms accumulation time; AGC 10^4 ions).

Data analysis

Database

Protein sequences from suborder Serpente were downloaded from Uniprot on March 03, 2013 and EST sequences from suborder Serpente were obtained in NCBI on March 03, 2013. Nucleotides sequences were translated to amino acid sequences using EMBOSS Transeq – EMBL-EBI (http://www.ebi.ac.uk/Tools/st/emboss_transeq/) [36] and all translated frames were incorporated to Uniprot database as well as 127 protein sequences of common MS contaminants. For all inputs, reversed decoy sequences were generated and included in database, totalizing 147,956 sequences.

Peptide spectrum matching (PSM)

Data were analyzed comparing experimental MS2 against theoretical spectra generated from an in-house database using ProLuCID v1.3 search engine [37]. Searches were performed with the following parameters: carboxamidomethylation of cysteines as fixed modification, oxidation of methionine and glutamate at N-terminal position to pyroglutamate as variables modifications, full

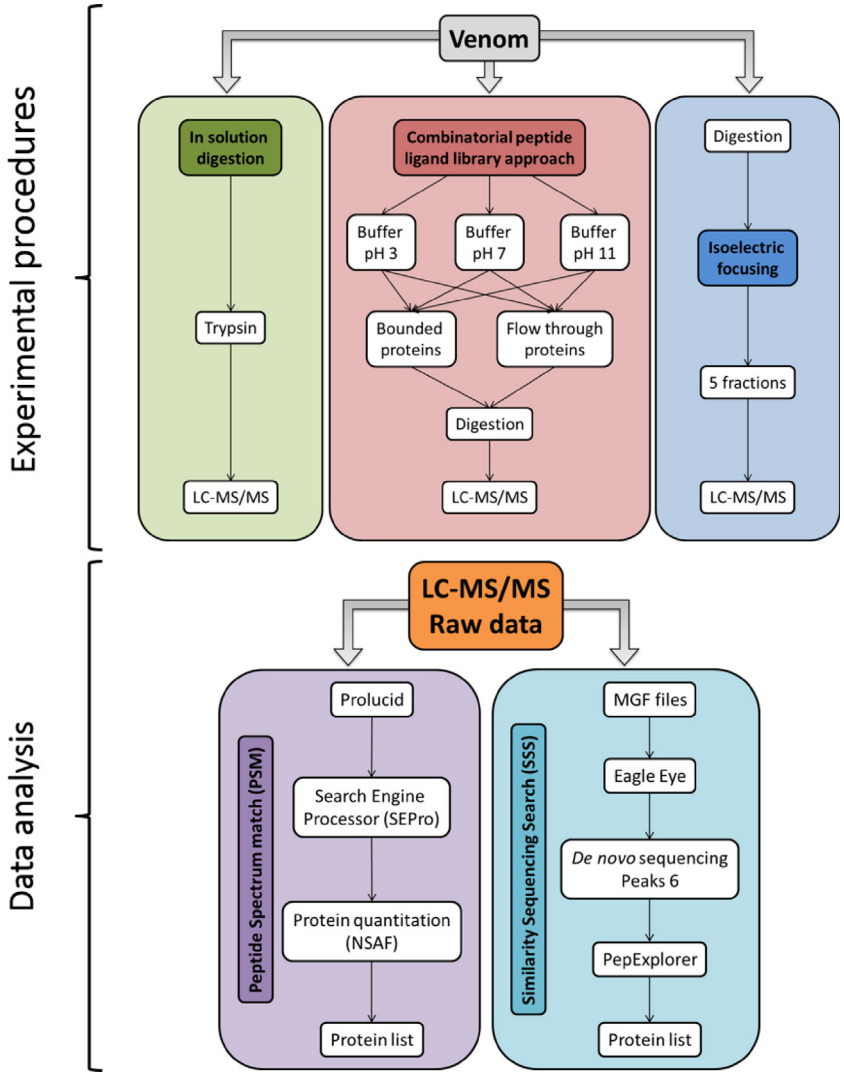

Fig. 1. Schematic representation of proposed experimental and data analysis pipeline for the venom of *C. durissus terrificus*.

Table 1
Protein families identified by peptide spectral matching (PSM) and similarity sequencing searching (SSS) reported in maximum parsimony mode (PTN) and protein group (PG) for each experimental condition: venom in solution digestion (ISD), proteins bounded to combinatorial peptide ligand library (CPLL) in all pHs (Bound), flow through proteins to CPLL in all pHs (FT), and venom digested peptides isoelectrical focused (IEF). Protein names of identified protein family subgroups are italicized.

Protein family	ISD				CPLL								IEF			
					Bound				FT							
	PSM		SSS		PSM		SSS		PSM		SSS		PSM		SSS	
	PTN	PG	PTN	PG	PTN	PG	PTN	PG	PTN	PG	PTN	PG	PTN	PG	PTN	PG
Toxin																
Bradykinin-potentiating and C-type natriuretic peptides	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
Cysteine-rich secretory protein – CRISP	1	1	1	1	0	0	0	0	1	1	2	2	2	1	1	1
Crotamine	1	1	1	1	0	0	0	0	0	0	1	1	1	1	6	1
Ecto-5'-nucleotidase	1	1	2	1	3	1	3	1	1	1	2	1	1	1	2	1
Hyaluronidase	3	1	2	1	4	1	3	1	4	1	2	1	2	1	3	1
L-amino-acid oxidase	6	1	3	1	11	2	11	1	8	1	4	1	6	1	10	1
Phosphodiesterase	3	2	2	2	3	3	3	3	5	2	4	1	4	3	4	3
Phospholipase A2 – Crotoxin acid and basic subunit	20	1	30	1	29	2	29	1	36	2	39	1	25	1	66	1
Phospholipase B	1	1	1	1	2	1	2	1	1	1	1	1	1	1	2	1
Snake venom C-type lectin	4	4	3	3	4	4	6	4	4	4	3	3	4	4	4	2
Convulxin subunit alpha/beta	*2*	*2*	*2*	*2*	*2*	*2*	*4*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*4*	*2*
Crotocetin	*2*	*2*	*1*	*1*	*2*	*2*	*2*	*2*	*2*	*2*	*1*	*1*	*2*	*2*	*0*	*0*
Snake venom metalloproteinase	5	1	14	1	32	1	22	1	21	1	24	1	8	1	14	1
Snake venom nerve growth factor	2	1	2	2	0	0	1	1	0	0	4	1	7	1	7	1
Snake venom serine proteinase	27	1	17	1	15	2	15	2	22	2	19	2	31	1	51	1
Snake venom vascular endothelial growth factor	4	2	4	2	1	1	0	0	5	2	2	1	4	2	2	1
Non-toxin																
Acid alpha glucosidase	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Acid ceramidase	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1
Actin	0	0	0	0	5	1	1	1	0	0	0	0	0	0	0	0
Aminopeptidase N	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Angiotensin-converting enzyme	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Annexin	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Calreticulin	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Carboxypeptidase E	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cathepsin	1	1	0	0	2	2	2	2	0	0	2	2	1	1	2	2
Cellular repressor of E1A-stimulated genes 1 – CREG1	1	1	0	0	1	1	1	1	0	0	0	0	1	1	1	1
Cystatin	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Deoxyribonuclease-2-alpha	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
Dipeptidase 2	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Dipeptidylpeptidase 4a	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Endoplasmic reticulum aminopeptidase 1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
Endoplasmic reticulum resident protein	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Exostosin	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Ganglioside GM2 activator	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Glia-derived nexin	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Glutaminyl-peptide cyclotransferase	2	1	2	1	2	1	2	1	4	1	1	1	1	1	3	1
Glutathione peroxidase	0	0	0	0	4	1	1	1	1	1	0	0	1	1	1	1
Glutathione S-transferase	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Growth/differentiation factor 5-like protein	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1
Heat shock protein	1	1	0	0	1	1	1	1	77	1	1	1	1	1	0	0
Immunoglobulin M heavy chain constant region	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0
Lipase A	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Lysosomal alpha-mannosidase-like protein	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Myosin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
N-acylsphingosine amidohydrolase	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Neural proliferation differentiation/control protein 1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nucleobindin-2	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Protein disulfide isomerase	1	1	1	1	2	2	3	2	1	1	0	0	1	1	0	0
Peptidyl-prolyl cis-trans isomerase	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0
Peroxiredoxin-4-like	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Phospholipase A2 inhibitor subunit B	1	1	1	1	3	1	0	0	0	0	0	0	1	1	2	1
Plasma alpha-L-fucosidase	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Proactivator polypeptide	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
Prolyl 4-hydroxylase subunit alpha-1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Protein FAM3C	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Protein-L-isoaspartate(D-aspartate) *O*-methyltransferase protein	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Pyruvate dehydrogenase complex – component X	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Renin-like aspartic protease	0	0	0	0	1	1	1	1	1	1	0	0	1	1	1	1
Serine carboxypeptidase CPVL	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1
Serum albumin	1	1	0	0	0	0	1	1	0	0	0	0	1	1	1	1
Snake venom C3 complement	0	0	0	0	7	1	4	1	0	0	0	0	0	0	0	0
Sodium-and chloride-dependent amino acid transporter B	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Sphingomyelin phosphodiesterase	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
Transferrin	2	2	1	1	0	0	1	1	0	0	1	1	2	1	1	1
Triosephosphate isomerase	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Villin-1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Total	96	35	94	29	161	59	142	58	198	29	116	26	117	39	193	34

tryptic hydrolysis, maximum of two missed cleavages, 70 ppm precursor (MS) mass tolerance, 550 ppm fragments (MS/MS) mass tolerance, XCorr as primary score, and ZScore as secondary score.

Peptides from search results were processed and evaluated using the Search Engine Processor (SEPro) v3.2.0.2 [38]. Based on the scores generated by PSM search, a cutoff score was set to accept 1% of false-discovery rate (FDR), at protein level, based on the total number of identified labeled decoys. In addition, only peptides with more than six amino acid residues were accepted. In a post process step only peptides with less than 10 ppm deviation from theoretical peptide precursor were considered and identified proteins shown in maximum parsimony mode [39].

Sequence similarity search (SSS)

Raw data were extracted to .mgf format using Proteome Discoverer 1.4 (Thermo Scientific) and submitted to a cleanup of common contaminant spectra against a background library of usual LC–MS contaminants using Eagle Eye software [40]. Filtered spectra present in each sample were loaded in Peaks Studio 6 software (Bioinformatics Solutions Inc.) to perform de novo sequencing. The following parameters were applied: carboxamidomethylation of cysteines as fixed modification, oxidation of methionine as variable modification, and no enzyme digestion. We then used PepExplorer [29] to process all peptides (Peaks ALC score higher than 50 and sequences with 6 or more amino acid residues) and perform a sequence similarity search against a database consisting of Uniprot entries from suborder Serpente plus 127 most common MS contaminants. For all gene products a reverse decoy sequence was created yielding a final database containing 33,004 sequences. The similarity search used the PAM30MS substitution matrix to score the alignments and 13 and 5 were set for open gap and extended gap penalties, respectively. Finally, PepExplorer grouped the identifications according to the maximum parsimony approach [39]. Only peptides having alignments with at least 65% of sequence identity against a sequence from the database were accepted and proteins were filtered at 1% of FDR.

Protein quantification

Label free protein quantitative analysis was performed according to the normalized spectral abundance factor (NSAF) [41] provided by Search Engine Processor (SEPro) [38].

Results and discussion

Proteomic approaches

The aim of this work was to perform a deep exploration of the venom of the Brazilian rattle snake; for this we introduce an extended venomics pipeline that combines different bottom-up approaches (in-solution digestion, peptide pre-fractionation by isoelectric focusing, and combinatorial peptide ligand library technique) and two orthogonal types of protein identification (peptide spectrum match and de novo sequencing coupled with similarity sequence search) in a unique pipeline (Fig. 1).

In solution digestion and quantification

Triplicate analyses of digested Cdt venom performed by nano LC–MS/MS in DDA mode using an LTQ Velos Orbitrap led to the identification of a total of 96 proteins, by PSM, in maximum parsimony mode. Proteins could be clustered in 35 groups and 28 protein families, 14 of them classified as snake toxins. The SSS identified 94 proteins, which were clustered into 29 protein groups, and 24 protein families (Table 1). All 14 toxin families were identified by both SSS and PSM techniques. Four protein families were detected only by PSM (cathepsin, CREG1, HSP and neural proliferation differentiation and control protein 1) and one only by SSS (acid ceramidase) summing up 29 protein families (Table 1). PSM is the proteomics gold standard for protein identification; it is much more sensitive than de novo sequencing, therefore, obtaining more identifications by PSM was already expected.

With a single in-solution venom digestion and LC–MS/MS analysis, we were able to pinpoint 29 protein families, which represents 163.6% more identifications than the sum of protein families described in previously proteomics studies performed by Calvete et al. [33] and Georgieva et al. [31] with Cdt venom, summing up 11. We attribute this to the methodological differences between the three studies and possibly from venom profile variations. Venom intraspecific variation was previously reported in Cdt [42], and it is vastly recorded in snake toxinology literature for other species. Different venom composition was observed among individuals of the same species and genders [43], with different feeding diets [44,45], between juveniles and adults [33,46] with diverse geographical location [47,48], and amid venom glands of a single individual [49].

The NSAF was used to obtain rough protein quantity estimates for each toxin family in Cdt venom. NSAF has been shown to be effective in accounting for variation from run to run [50] and for reporting rough absolute quantitation estimates. This is because larger proteins tend to contribute more peptides and therefore more spectra; NSAF's formula accounts for this by considering the number of spectral counts (Sc) from a given protein divided by the number of amino acids (L), then dividing this value by the sum of Sc/L for all identified proteins. As such, this normalization strategy significantly improves analysis of proteins from multiple independent assays compared to spectral counting alone [51]. According to this method, 98.5% of the venom proteome is composed of toxins while only 1.5% are considered non-toxins. The major protein family detected was phospholipase A2 (PLA2) including crotoxin, basic and acid subunits, responsible for 69.3% of total venom proteins (Fig. 2). This result is aligned with Calvete et al.'s [33] and Georgieva et al.'s [31] studies that also reported crotoxin as the major venom component (Table 2). The elevated amounts of crotoxin in Cdt venom is one of the reasons for its high toxicity (LD_{50} 0.06–0.09 mg of venom/g of mice) and for causing neurological disturbances when injected in mice [52] or in human envenomation [31].

The second most abundant protein family detected was snake venom serine proteinase (SVSP) with 13.4% of total venom's NSAF, the same value recorded by Georgieva et al. [31]. On the other hand, Calvete et al.'s study [33] found crotamine to be the second most abundant venom component and SVSP as the third most intense; this discrepancy can occur by the presence of different concentrations of crotamine in the venom, which can be crotamine positive or negative [42] or even to the different strategies for inferring absolute quantitation. Again, NSAF is a rough estimate for absolute quantitation. Another reason could be that Calvete et al.'s group studied a crotamine-enriched venom whereas Georgieva et al. and this study used crotamine-less or negative venoms.

In addition of the two most intense components, 6 toxin families with a middle range concentration between 1 and 3% of total venom proteins were detected: L-amino-acid oxidase (LAAO) 3.0%, snake venom metalloproteinase (SVMP) 2.9%, snake venom vascular endothelial growth factor (SVVEGF) 2.8%, snake venom C-type lectin (SnaCLec) 2.7% composed of two distinct proteins, convulxin, alpha and beta chain (2.1%), and crotocetin, (0.6%), phosphodiesterase (PDE), 1.2%, and crotamine (CRO) 1.0%. The recently described hyaluronidase (HYA) family for *Crotalus* genus was identified in Cdt venom and represents 0.8% of total venom proteins. This protein class is closely related to the increase in crotoxin toxicity [53] and was not identified in previous proteomics studies.

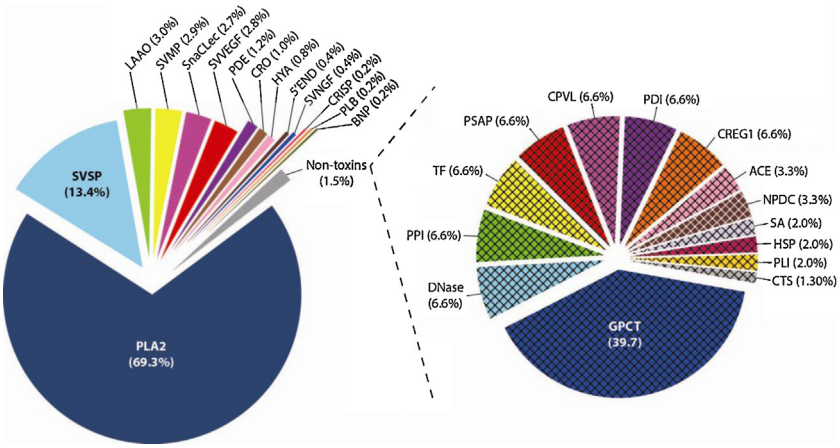

Fig. 2. Overall composition of *C. durissus terrificus* venom according to protein families (in percentages). Phospholipase A2 (PLA2 including crotoxin), snake venom serino proteinase (SVSP), L-amino-acid oxidase (LAAO), snake venom metalloproteinase (SVMP), snake venom C-type lectin (SnaCLec), snake venom vascular endothelial growth factor (SVVEGF), phosphodiesterase (PDE), crotamine (CRO), hyaluronidase (HYA), ecto-5′-nucleotidase (5′END), snake venom nerve growth factor (SVNGF), cysteine-rich secretory protein (CRISP), phospholipase B (PLB), bradykinin-potentiating and C-type natriuretic peptides (BNP), glutaminyl-peptide cyclotransferase (GPCT), deoxyribonuclease-2-alpha (DNase), peptidyl-prolylcis-trans isomerase (PPI), transferrin (TF), proactivator polypeptide (PSAP), serine carboxypeptidase CPVL (CVLP), protein disulfide isomerase (PDI), cellular repressor of E1A-stimulated genes 1 (CREG1), angiotensin-converting enzyme (ACE), neural proliferation differentiation/control protein 1 (NPDC), serum albumin (SA), heat shock protein (HSP), phospholipase A2 Inhibitor (PLI), and cathepsin (CTS).

Table 2
Overview of the relative occurrence and comparison with previously studies of protein families identified by in solution digestion of Cdt venom and quantified by normalized spectral abundance factor – NSAF (in percentage of the total proteins). Protein names of identified protein family subgroups are italicized.

Protein family	Percentage of total venom proteins		
	This study	Calvete et al. [33]	Georgieva et al. [31]
Toxin			
Phospholipase A2 – Crotoxin acid and basic subunit – PLA2	69.3	59.5	48.5
Snake venom serine proteinase – SVSP	13.4	8.2	25.3
L-amino-acid oxidase – LAAO	3.0	4.5	–
Snake venom metalloproteinase – SVMP	2.9	4.8	3.9
Snake venom vascular endothelial growth factor – SVVEGF	2.8	–	–
Snake venom C-type lectin – SnaCLec	2.7	1.7	–
Convulxin alpha/beta subunits	*2.1*	–	–
Crotocetin	*0.6*	–	–
Phosphodiesterase – PDE	1.2	–	1.9
Crotamine – CRO	1.0	19.0	–
Hyaluronidase – HYA	0.8	–	–
Ecto-5′-nucleotidase – 5′END	0.4	–	7.8
Snake venom nerve growth factor – SVNGF	0.4	–	1.9
Cysteine-rich secretory protein – CRISP	0.2	–	–
Phospholipase B – PLB	0.2	–	–
Bradykinin-potentiating and C-type natriuretic peptides – BNP	0.2	2.3	–
Non-toxin			
Glutaminyl-peptide cyclotransferase – GPCT	0.6	–	1.0
Deoxyribonuclease-2-alpha – DNase	0.1	–	–
Peptidyl-prolylcis-trans isomerase – PPI	0.1	–	–
Transferrin – TF	0.1	–	–
Proactivator polypeptide – PSAP	0.1	–	–
Serine carboxypeptidase CPVL – CPVL	0.1	–	–
Protein disulfide isomerase – PDI	0.1	–	–
Cellular repressor of E1A-stimulated genes 1 – CREG1	0.1	–	–
Angiotensin-converting enzyme – ACE	0.05	–	–
Neural proliferation differentiation/control protein 1 – NPDC	0.05	–	–
Serum albumin – SA	0.03	–	–
Heat shock protein – HSP	0.03	–	–
Phospholipase A2 Inhibitor – PLI	0.03	–	–
Cathepsin – CTS	0.02	–	–
Non identified sequences	–	–	9.7

Several low-abundant proteins were identified and quantified such as 5 toxin families with less than 0.5% of the total venom proteins. They are: ecto-5′-nucleotidase (5′END), 0.4%, snake venom nerve growth factor (SVNGF), 0.4%, cysteine-rich secretory protein (CRISP), 0.2%, phospholipase B (PLB), 0.2%, and bradykinin-potentiating and C-type natriuretic peptides (BNP), 0.2%. While the majority of toxin components act directly on killing/paralyzing the prey or in defense, low-abundance proteins may be remnant toxins that are related to evolutionary or ecological timescale plasticity where mutations can drive toxins functional innovations [22]. One example is the relation of hyaluronidase and crotoxin, where the first, present in low-amount, acts as potentiating agent in synergism with the crotoxin [53].

The most abundant non-toxin component noted was glutaminyl-peptide ciclotransferase (GPCT) that, according to NSAF, represents 0.6% of total venom protein composition, and corresponds to 39.7% of total amount of non-toxin components recorded (Fig. 1 and Table 2). The 60.3% remaining non-toxin components, representing 0.9% of total venom proteins, are composed of 13 protein families. The concentrations of proteins in these 13 families range from 0.1% to 0.02%, showing that this method is more sensitive than Calvete et al.'s study [33] that found 0.5% of bradykinin potentiating peptides and Georgieva et al.'s work [31] with 1.0% of GPCT. Calvete group's snake venomics approach is based on RP-HPLC fractionation followed by SDS with protein bands identified by mass spectrometry. Calvete's method is less effective at protein identification than in solution digestion as shown by Margres et al. [54]. This result may be explained by loss of low-abundant proteins in the diverse fractionation and manipulation steps that require large amounts of proteins, as well as the use of mass spectrometers with different accuracies and resolutions. Therefore, in-solution digestion coupled to LC–MS/MS in FT-Orbitrap instruments requires less protein than conventional snake venomics methodology and is a more sensitive approach.

While considering that bottom-up strategies increased protein family identification in the Cdt venome, we point out that this approach is not capable of efficiently identifying isoforms and proteoforms [18]. A single amino acid change or a post-translational modification in a peptide may result in a new toxin with different biological action. In order to identify isoforms/proteoforms, it is possible to use top-down proteomics or for bottom-up approaches, alternative pre-fractionation and LC–MS/MS methods are necessary.

Isoelectric focusing (IEF) pre-fractionation.

IEF pre-fractionation of digested venom peptides allowed the identification of 117 proteins by PSM, which were grouped in 39 protein groups and clustered into 34 protein families. A total of 193 proteins were identified by SSS, which were then clustered into 34 protein groups and 30 protein families. IEF allowed for the identification of 14 toxin families observed in the in-solution digestion analysis and 21 non-toxin families, a total of 35 identifications, comprising 6 more different protein families not related to toxins: myosin (only detected by SSS), acid ceramidase-like, endoplasmic reticulum aminopeptidase 1 protein, glutathione peroxidase, renin aspartic protease, and immunoglobulin M heavy chain constant region (Table 1). The use of IEF before LC–MS/MS analysis increased the number of protein families identified by more than 20%. This highlights the advantage of pre-fractionation prior to LC–MS/MS venom analysis, and corroborates previous findings in which 1D PAGE protein pre-fractionation increased the number of protein identifications from *C. atrox* venom [18].

Combinatorial peptide ligand library approach (CPLL)

Exploration of proteomes with a large dynamic range, like the venom of Cdt – which has an overwhelming amount of crotoxin – is challenging, as abundant proteins overshadow the identification of those in lower abundance [21,55]. One way to access this "hidden proteome" is the use of a CPLL approach combining distinct pH buffers [22], in-bead digestion [35], and the analysis of bound proteins and flow through proteins [22]. By combining these three strategies, we identified 62 protein families, 28 more protein families than in solution digestion and IEF experiments. SSS approach provided 7 and PSM 6 new protein family identifications being the remainder 15 protein families identified by both search approaches (Table 1). Protein families identified only by the CPLL approach were classified as non-toxins and probably are present in less than 0.02% of total venom proteins, representing a dynamic concentration range equal or greater than 4 orders of magnitude. These protein families could only be detected thanks to the depletion of the major components, mostly PLA2, in the sample as well as the enrichment in concentration of these low abundance molecules [21]. Our bottom-up proteomics CPLL approach increased the total number of protein families by more than 77%, owing to in-solution digestion and IEF. The gel-based CPLL approach utilized by Calvete and collaborators to explore the proteome of *C. atrox* identified only two new low abundance protein classes (peroxiredoxin and glutaminyl cyclase), which corresponded to 22% of increase compared to the previous *C. atrox* venom prospection [21]. A similar increase, about 20% in protein families identified, was obtained by us with a simple pre-fractionation of venom peptides in pH 3-10 IEF strips.

Even after using three different pH buffers aiming to maximize the number of proteins, we were still unable to identify crotamine and CRISP toxins as bound to CPLL by PSM or SSS. These toxins are present in the venom at concentrations greater than 0.2% of total venom protein and were detected by PSM and SSS at in-solution digestion and IEF, and by SSS in CPLL flow through. Fasoli et al. used a similar CPLL approach with three pH buffers to analyze the venom of *Bitis arietans* and were unable to identify the 5′END, and cystatin toxin families, an isoform of kunitz-type inhibitor, and the disintegrin bitistatin in the bound proteins [22]. Moreover, Calvete and colleagues were unable to capture CRISP and LAAO toxin

Table 3
Comparative of each different bottom-up experimental and search approaches utilized. Combinatorial peptide ligand library (CPLL), peptide spectral matching (PSM), and similarity sequencing searching (SSS).

	In solution digestion	CPLL		Isoelectric focusing
		Bound proteins	Flow through	
Total protein amount	100 μg	20 mg per pH	100 μg	
Number of fractions	1	3 different pHs	5	
Total runs	4	9	9	20
Total time	12 h	27 h	27 h	60 h
Total MS/MS PSM	45,617	117,360	121,031	195,952
Total PSM	3870	7733	5024	14,486
Total peptides	878	1,312	758	983
Total proteins	96	161	198	117
Total protein families SSS	35	59	29	39
Total matching alignments	5375	6003	7878	50,732
Total peptides	1179	1795	1397	3405
Total proteins	94	142	116	193
Total protein families	29	58	26	34

families when using two different CPLL libraries to explore the venome of *C. atrox* [21]. There are three possible explanations for this: (I) presence of post-translational modifications that interfere target proteins binding to the peptide library, (II) there is no peptide ligand for these proteins in the library, and/or (III) the ligand site present in CPLL suffered rapid saturation by other more abundant competing protein. Neither neural proliferation differentiation/control protein 1 nor myosin were detected in bounded CPLL or flow through but were identified by PSM at in-solution digestion and by SSS at IEF experiments, respectively.

Experimental approaches comparison

CPLL required the greatest amount of protein (i.e., 20 mg per pH of capture buffer) (Table 3); 200 fold more than in solution digestion or IEF experiments. Due to the large quantities of input protein, CPLL enables enrichment and detection of lower-abundance proteins, therefore resulting in the highest number of protein and protein family identifications of any method we tested. IEF was the most time consuming experiment due to the high number of fractions for analysis by mass spectrometry, although it provided the highest number of MS/MS spectra and PSMs (Table 3). On the other hand, the fastest and less resource consuming approach was in-solution digestion, which generated 878 peptides identifications by PSM, only 105 less peptides than the IEF approach, and identified the same number of toxin families as the other two experimental approaches (Table 3).

We noted a similar number of unique peptides in CPLL flow through and in-solution digestion of the whole venom (Fig. 3A). However, CPLL flow through did not have more low-abundance proteins than those bound to CCPL. By analyzing a second Venn diagram constructed with protein families identified from each experimental approach, we verified that the unique peptides from CPLL flow through did aid in identifying new protein families; they originated from protein classes already identified by other methods (Fig. 3B). Yet, these peptides helped in improving protein coverage/confidence.

More than 100 unique peptides were found in the IEF experiment compared to the venom in-solution digestion showing that better peptide fractionation before LC–MS/MS and more runs increase the number of denoted peptides and protein coverage. IEF proved to be an efficient fractionation step by increasing protein identification compared to methods that do not involve pre-fractionation. Actual enrichment of peptides and new protein families could be observed in CPLL bound proteins: 598 unique peptides representing 24 new unique protein families. Adding protein families and peptides identified in all experiments gives a total of 2097 unique peptides representing 64 unique protein families, indicating complementarity of the experimental approaches applied.

In summary, our results indicate that for single venom proteome analysis the use of in-solution digestion or peptides IEF pre-fractionation prior to LC–MS/MS is recommended. For deep venom analysis the use of CPLL coupled to in-solution digestion analysis is better alternative providing the identification of a large number of protein families in a non-redundant and time saving manner.

Peptide spectrum match (PSM) × Similarity sequencing search (SSS)

The use of PSM and sequence similarity search combined allowed an increase of protein identifications. The first identifies peptides in protein databases while the second permits the identification of new proteins that have some sequence similarity with other proteins in the database. It is an important search tool for high variance sequences present in proteins with high number of isoforms, as snake toxins or unsequenced organisms. On the other hand, so far it has remained the most error-prone approach [56]. So, these two approaches are complementary and orthogonal and can be used to validate protein identifications with borderline statistical confidence hits [57].

In this work, all 14 toxin families were identified in each distinct experimental group by both PSM and SSS with high scores, high quality de novo interpretations and alignments. Similar results were obtained for the 33 major non-toxin families representing 73.4% of total identifications. Eight protein families were identified only by PSM (12.5%) and 9 (14.1%) only by SSS.

Protein identification

Out of the 64 protein families identified, 14 (21.9%) of the identifications reports toxins and 50 (78.1%) non-toxin families. Although cDNA libraries have been analyzed from the venom glands of many different snakes, we have used proteomics to identify novel toxin and non-toxin protein components of Cdt venom [58,59].

Cysteine rich secretory protein (CRISP)

CRISP family was firstly described as a snake venom component in 1997 in the venom of *Trimeresurus mucrosquamatus* [60]. CRISP toxins were widely distributed and were purified from venoms of Elapidae, Colubridae and Viperidae snakes families [61–63]. The

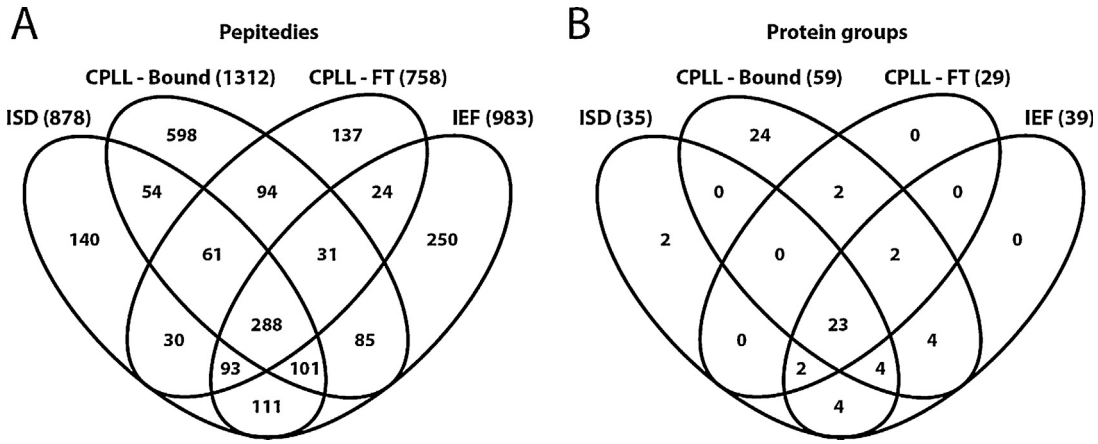

Fig. 3. Comparison of the peptides (A) and protein groups (B) identified by peptide spectral matching between the different bottom-up approaches. Venom in solution digestion (ISD), proteins bounded to combinatorial peptide ligand library (CPLL) in all pHs (Bound), flow through proteins to CPLL in all pHs (FT), and venom digested peptides isoelectrical focused (IEF).

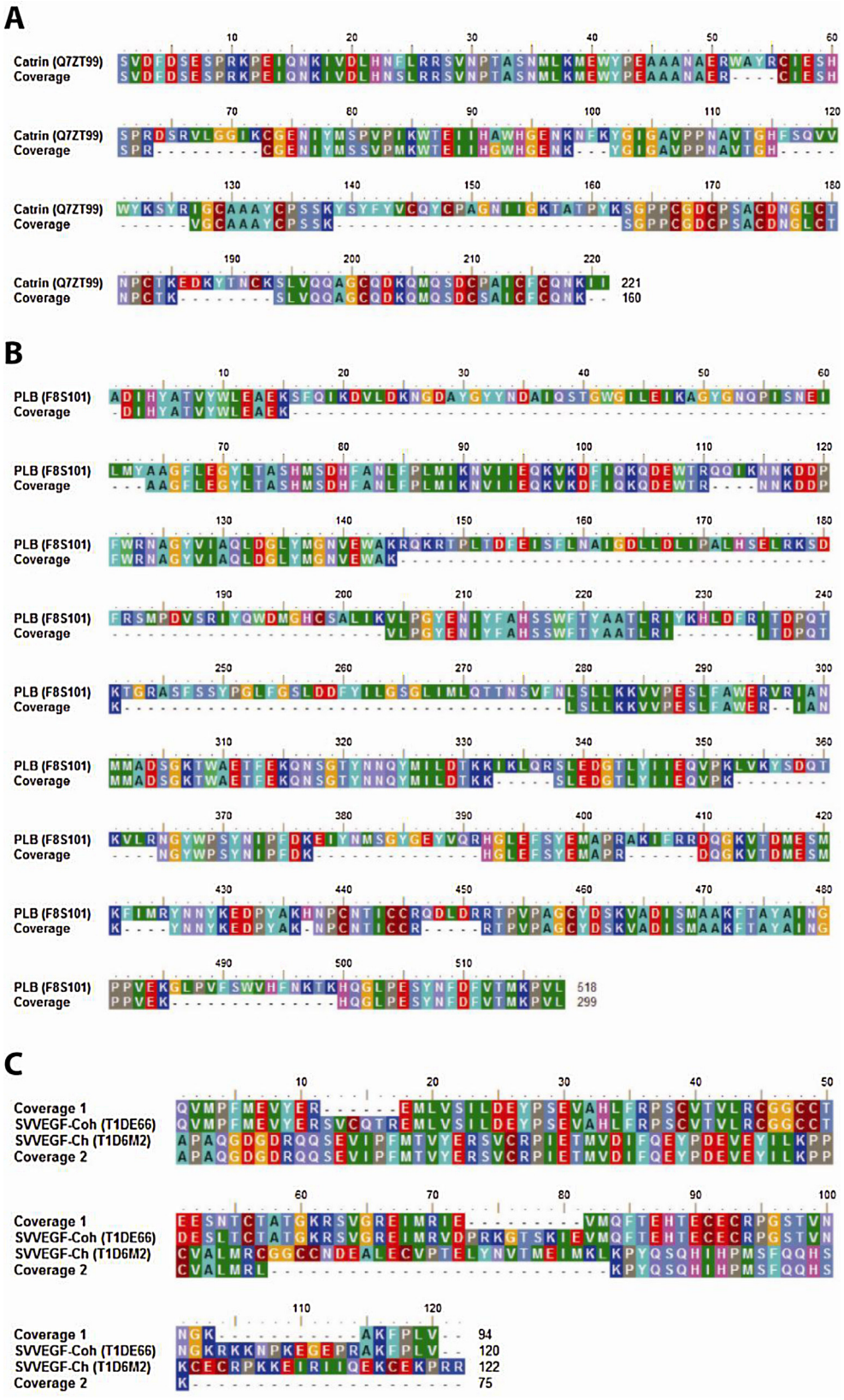

Fig. 4. Sequence coverage of new toxin identifications for Cdt venom. (A) CRISP peptides aligned to Catrin 1–2 sequence from *C. atrox* (Q7ZT99). (B) Phospholipase B (PLB) peptides aligned to PLB sequence from *C. adamanteus* (F8S101). (C) Snake venom vascular endothelial growth factor (SVVEG) peptides aligned to SVVEGF sequence from *C. oreganus helleri* (T1DE66) and from *C. horridus* (T1D6M2).

presence of CRISP in the venom of snakes of *Crotalus* genus was well documented by many studies using different techniques. By proteomics approaches this family was identified in the venom of *C. simus simus*,*C. simus culminatus* [64], *C. tigris* [65], *C. durissus collilineatus*, and *C. durissus cascavella* [47], by transcriptomics approaches in the venom of *C. viridis* (Uniprot, unpublished), and *C. horridus* [58], by proteomic/transcriptomic approaches in *C. adamanteus* [54,59] *C. oreganus helleri* [66] and by biochemistry approaches catrin 1 and 2 were isolated from the venom of *C. atrox* and chemically/pharmacologically characterized by Yamazaki et al. [62].

Even with a large distribution among *Crotalus* genus, CRISP family has never been identified in Cdt. In this work, we describe for the first time, by proteomics approaches, the presence of CRISP in this subspecies. We report 15 non-redundant peptides from CRISP family by PSM and/or SSS in in-solution digestion, flow through from CPLL and IEF (Table S1) experiments that correspond to 72.4% coverage of mature catrin 1–2 toxin (Q7ZT99) (Fig. 4A). Members of CRISP family represent a single polypeptide chain with a molecular weight range of 23–26 kDa. They have 16 cysteine residues strictly conserved, which form 8 disulfide bridges and 10 of these cysteines residues are clustered in the C-terminal of the protein [62]. CRISP toxins have been described to actuate against different types of ion channels as cyclic nucleotide-gated channels [61], voltage-gated Ca^{2+} channels [62,63], high-conductance calcium-activated potassium channel [67] and voltage-gated potassium channel Kv1.3 [68]. Taken together, these results indicate that these proteins act as neurotoxins.

Phospholipase B (PLB)

PLB toxin family was identified in all experimental procedures performed and a total of 42 non-redundant peptides were annotated by PSM and/or SSS (Table S2). They accounted for 57.7% of PLB sequence coverage from *C. adamanteus* (F8S101) (Fig. 4B). This toxin family was never identified in Cdt venom before it was first described in the species and in South American rattlesnakes. As venom component, PLB was firstly described in 1964 in *Naja naja*, *Pseudechis porphyriacus*, and *Agkistrodon piscivorus* venoms [69]. Years later, PLB was isolated from the venom of *Pseudechis colletti* and partially pharmacologically and enzymatically characterized [70]. Recently, PLB was detected by transcriptomic and proteomic techniques in the venom of *Drysdalia coronoides* [71], *C. adamanteus* [54,59], *Ovophis okinavensis* and *Protobothrops flavoviridis* [72] and only by transcriptome in *C. horridus* [58] and *Micrurus fulvius* [73] venoms. Unfortunately, little is known about this toxin family and it mature protein sequence, and its three-dimensional structure remain unknown.

The PLB isolated from *P. colletti* exhibit strong hemolytic activity *in vitro* for rabbit and human erythrocytes. In vivo, tests revealed that 120 μg of PLB was lethal for mice and death was preceded by myoglobinuria [70].The main role of PLB in envenomation is unclear, however it hydrolysis of cell membrane phospholipids liberate intracellular contents and these intracellular components can indirectly produce toxic effect.

Snake venom vascular endothelial growth factor (SVVEGF)

SVVEGF was previously identified in the venom of *C. durissus cascavella* and *C. durissus collilineatus* [47], but never in the venom of the subspecies Cdt. Toxin family was detected in all our experimental conditions in a total of 58 non-redundant peptides denoted by PSM and/or SSS (Table S3). Identified SVVEGF proteins belong to two distinct groups distinguished at the C terminal region of the proteins. Identified peptides covered 78.3% of SVVEGF sequence from *C. oreganus helleri* (T1DE66) and 61.5% from *C. horridus* (T1D6M2) (Fig. 4C).

SVVEGFs are one of the most structurally diverse snake toxin families, especially in the C-terminal, supposedly a co-receptor-binding region as well as near the receptor-binding loops 1 and 3, whereas the sequences of mammalian tissue VEGFs is highly conserved [74]. SVVEGF can bind with distinct affinities to VEGF receptors (VEGFRs): VEGFR1 and/or VEGFR2, and heparin. Presumably, SVVEGF facilitates the access of other venom components to their target by promoting capillary permeability [74]. However, the complete understanding of the role it plays in the envenomation remains unclear.

Non-toxin proteins

Out of the 64 protein families identified, 50 families were classified as non-toxins (78.1%). These are low abundance proteins and they represent less than 1.5% of proteins contained in the venom. Whose presence can be due to different reasons: (I) participation in toxin maturation, (II) regulation and self-protection processes, (III) release of cellular proteins in the gland lumen as a consequence of glandular physiology and venom production, (IV) venom extraction artifacts, and (V) proteins that can become toxins.

The most abundant non-toxin protein identified was GPCT, found in higher concentrations than some toxins; it is related to toxin maturation. GPCT plays a crucial role in the maturation process of many bioactive peptides/proteins present into the venom by N-terminal pyroglutamate (pyrrolidone carboxylic acid, pGlu) formation from glutamyl precursor. Pyroglutamate is a common and important post-translational modification (PTM) found in venom toxins, some examples are BPPs [75,76], metalloproteinase inhibitor peptides [77], some SVMP [78,79] and B and C chains of the acidic subunit of crotoxin [80,81]. The GPCT high concentration found in the venom compared to other non-toxin components indicates that PTMs may be processed in venom vesicles and GPCT is released together with mature toxins in venom gland lumen or it may be directly released into the gland lumen separated from immature toxins. A recent study detected the presence of GPCT interacting with SVMP in venom gland extracts but not in extracted venom supporting the first hypothesis. On the other hand, the same study showed that SVMP are activated in the venom gland lumen, which requires the presence of specific proteins for toxin maturation in the venom [82].

Other classes of proteins identified are directly involved into self-protection and regulation processes. Examples are glutathione peroxidase (GPX), peroxiredoxin-4-like (PRX4), cystatin, and cathepsin. GPX has peroxidase activity and plays an important role in eliminating peroxides generated during metabolism [83]. PRPX4 was identified only as CPLL-bound protein, as well as by Calvete and coworkers [21], and it plays a similar role as GPX. Cystatins are cysteine–protease inhibitors that may have as primary function protect cells from unregulated proteolytic activity. Snake cystatins are known to have specific interactions with lysosomal cysteine proteases cathepsins [84]. Both proteins were identified in this study, supporting the idea that cystatin is not a toxin family but a regulatory protein present in the venom [84].

The exact location and process of protein maturation is still not clear. However, heat shock protein (HSP), protein disulfide isomerase (PDI), and peptidyl–prolyl *cis*-trans isomerase (PPI), which are intracellular proteins located in the endoplasmic reticulum and play important role in protein folding process, were identified in the venom. The release of these proteins may be related to cytosol extravasation during the venom vesicles fusion to cellular membrane and/or they may be present inside the venom vesicles.

No evidence of venom protein turnover in the glandular lumen or ducts has been presented until now and the presence of cellular

debris in venom suggest that cell particles of replaced old glandular cells are not reabsorbed [85]. These observations associated with the glandular physiology and venom production could explain the identification of some proteins related to housekeeping and the processing and sorting of toxins.

Snake handling and manual pressure exerted on the venom glands during venom extraction process depending on its intensity and frequency can cause cellular or tissue damage releasing cellular content or body fluids in the venom gland lumen or other venom apparatus compartments resulting in venom contamination [34]. Some non-toxins identified in Cdt venom, such as PLA2 inhibitor, serum albumin, IgM, transferrin, and snake venom C3 complement were previously mapped by bottom-up experiments in *Bothrops jararaca* plasma analysis [29], a strong indication of plasma contamination or plasma presence inside the venom gland.

Many of the identified non-venom proteins are cellular or tissue proteins expressed in a wide variety of body tissues and may not play a significant role at the envenomation process. However, they are relevant to a better understanding of venom components and molecular pathways. Moreover, it has been recently proposed that some of these proteins present in the venom may have become source of new biological activity molecules by gene duplication or sub functionalization generating new toxins [86].

Conclusions

Here we present an up to date panel of Cdt venom components analyzed by three different bottom-up methodologies. Although Cdt venom is among the best studied snake venoms, our approach allowed for the identification of three classes of toxin never described in this venom as well as 50 non-venom proteins, most of which had never been detected in snake venom samples. These new identifications show the power of bottom-up proteomics approaches in snake venom studies, and uncovered the complexity of venom components.

Classic works in toxinology characterize one or a few of the most abundant venom proteins based on assay-guided techniques which are time and resource consuming, yet indispensable for the identification of completely unknown toxins and their pharmacological characterization. On the other hand, classical biochemistry or protein chemistry methods are rapidly being replaced by proteomic approaches capable to generate massive data volume, without compromising high-mass accuracy, and in short periods of time with lower amounts of venom. Large scale application of the bottom-up approaches described here in the characterization of snake venoms dig deep into complexity of venom, and broaden our access to the therapeutic potential of snake venoms and the identification of new toxins, as well as provide a better understanding of the molecular complexity of different venoms and the mechanisms of venom production and regulation.

Acknowledgements

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Grant No. 063/2010 – Edital Toxinologia) and the FAPERJ/Rio de Janeiro Proteomics Network. The authors would like to thank for the generous donation of venom samples and the CPLL beads Dr. Ana Maria Moura da Silva, Laboratory of Immunopathology, Butantan Institute, Brazil, and Dr. Pier Giorgio Righetti, Politecnico di Milano, Italy, respectively. We are also grateful to A.W. Goering for his help with English editing of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.euprot.2015.05.006.

References

[1] J.J. Calvete, L. Sanz, Y. Angulo, B. Lomonte, J.M. Gutiérrez, Venoms, venomics, antivenomics, FEBS Lett. 583 (2009) 1736–1743, doi:http://dx.doi.org/10.1016/j.febslet.2009.03.029.
[2] P. Escoubas, G.F. King, Venomics as a drug discovery platform, Expert Rev. Proteomics 6 (2009) 221–224, doi:http://dx.doi.org/10.1586/epr.09.45.
[3] J.W. Fox, S.M.T. Serrano, Snake toxins and hemostasis, Toxicon 45 (2005) 949, doi:http://dx.doi.org/10.1016/j.toxicon.2005.04.007.
[4] D. Georgieva, R.K. Arni, C. Betzel, Proteome analysis of snake venom toxins: pharmacological insights, Expert Rev. Proteomics 5 (2008) 787–797, doi:http://dx.doi.org/10.1586/14789450.5.6.787.
[5] R.J.R. McCleary, R.M. Kini, Non-enzymatic proteins from snake venoms: a gold mine of pharmacological tools and drug leads, Toxicon 62 (2013) 56–74, doi:http://dx.doi.org/10.1016/j.toxicon.2012.09.008.
[6] I. Vetter, J.L. Davis, L.D. Rash, R. Anangi, M. Mobli, P.F. Alewood, et al., Venomics: a new paradigm for natural products-based drug discovery, Amino Acids 40 (2011) 15–28, doi:http://dx.doi.org/10.1007/s00726-010-0516-4.
[7] C.-W. Vogel, D.C. Fritzinger, Cobra venom factor: structure, function, and humanization for therapeutic complement depletion, Toxicon 56 (2010) 1198–1222, doi:http://dx.doi.org/10.1016/j.toxicon.2010.04.007.
[8] Silva M.R e, W.T. Beraldo, G. Rosenfeld, Bradykinin, a Hypotensive and smooth muscle stimulating factor released from plasma globulin by snake venoms and by trypsin, Am. J. Physiol. – Leg Content 156 (1949) 261–273.
[9] S.H. Ferreira, L.J. Greene, V.A. Alabaster, Y.S. Bakhle, J.R. Vane, Activity of various fractions of Bradykinin potentiating factor against angiotensin I converting enzyme, Nature 225 (1970) 379–380, doi:http://dx.doi.org/10.1038/225379a0.
[10] D.W. Cushman, H.S. Cheung, E.F. Sabo, M.A. Ondetti, Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids, Biochemistry (Moscow) 16 (1977) 5484–5491, doi:http://dx.doi.org/10.1021/bi00644a014.
[11] V. Rioux, M.-C. Gerbod, F. Bouet, A. Ménez, A. Galat, Divergent and common groups of proteins in glands of venomous snakes, Electrophoresis 19 (1998) 788–796, doi:http://dx.doi.org/10.1002/elps.1150190531.
[12] A.S. Kamiguti, R.D.G. Theakston, N. Sherman, J.W. Fox, Mass spectrophotometric evidence for P-III/P-IV metalloproteinases in the venom of the Boomslang (*Dispholidus typus*), Toxicon 38 (2000) 1613–1620, doi:http://dx.doi.org/10.1016/S0041-0101(00)00089-1.
[13] B.G. Fry, J.C. Wickramaratna, W.C. Hodgson, P.F. Alewood, R.M. Kini, H. Ho, et al., Electrospray liquid chromatography/mass spectrometry fingerprinting of Acanthophis (death adder) venoms: taxonomic and toxinological implications, Rapid Commun. Mass Spectrom. 16 (2002) 600–608, doi:http://dx.doi.org/10.1002/rcm.613.
[14] P. Juárez, L. Sanz, J.J. Calvete, Snake venomics: characterization of protein families in *Sistrurus barbouri* venom by cysteine mapping, N-terminal sequencing, and tandem mass spectrometry analysis, Proteomics 4 (2004) 327–338, doi:http://dx.doi.org/10.1002/pmic.200300628.
[15] J.J. Calvete, P. Juárez, L. Sanz, Snake venomics. Strategy and applications, J. Mass Spectrom. 42 (2007) 1405–1414, doi:http://dx.doi.org/10.1002/jms.1242.
[16] J.W. Fox, S.M.T. Serrano, Exploring snake venom proteomes: multifaceted analyses for complex toxin mixtures, Proteomics 8 (2008) 909–920, doi:http://dx.doi.org/10.1002/pmic.200700777.
[17] R.D. Melani, L. Goto-Silva, F.C.S.N. Nogueira, M. Junqueira, G.B. Domont, Shotgun approaches for venom analysis, in: P. Gopalakrishnakone, J.J. Calvete (Eds.), Venom Genomics Proteomics, Springer, Netherlands, 2014, pp. 1–12.
[18] J.W. Fox, L. Ma, K. Nelson, N.E. Sherman, S.M.T. Serrano, Comparison of indirect and direct approaches using ion-trap and Fourier transform ion cyclotron resonance mass spectrometry for exploring viperid venom proteomes, Toxicon 47 (2006) 700–714, doi:http://dx.doi.org/10.1016/j.toxicon.2006.01.022.
[19] N. Bandeira, K.R. Clauser, P.A. Pevzner, Shotgun protein sequencing assembly of peptide tandem mass spectra from mixtures of modified proteins, Mol. Cell Proteomics 6 (2007) 1123–1134, doi:http://dx.doi.org/10.1074/mcp.M700001-MCP200.
[20] P.G. Righetti, E. Boschetti, L. Lomas, A. Citterio, Protein Equalizer™ Technology: the quest for a democratic proteome, Proteomics 6 (2006) 3980–3992, doi:http://dx.doi.org/10.1002/pmic.200500904.
[21] J.J. Calvete, E. Fasoli, L. Sanz, E. Boschetti, P.G. Righetti, Exploring the venom proteome of the Western Diamondback Rattlesnake, *Crotalus atrox*, via Snake Venomics and Combinatorial Peptide Ligand Library Approaches, J. Proteome Res. 8 (2009) 3055–3067, doi:http://dx.doi.org/10.1021/pr900249q.
[22] E. Fasoli, L. Sanz, S. Wagstaff, R.A. Harrison, P.G. Righetti, J.J. Calvete, Exploring the venom proteome of the African puff adder, *Bitis arietans*, using a combinatorial peptide ligand library approach at different pHs, J. Proteomics 73 (2010) 932–942, doi:http://dx.doi.org/10.1016/j.jprot.2009.12.006.
[23] A. Shevchenko, C.-M. Valcu, M. Junqueira, Tools for exploring the proteomosphere, J. Proteomics 72 (2009) 137–144, doi:http://dx.doi.org/10.1016/j.jprot.2009.01.012.

[24] Ma B., Johnson R.. De Novo Sequencing and Homology Searching Mol Cell Proteomics 2012; 11: O111.014902. doi:10.1074/mcp.O111.014902

[25] B. Ma, K. Zhang, C. Hendrie, C. Liang, M. Li, A. Doherty-Kirby, et al., PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry, Rapid Commun. Mass Spectrom. 17 (2003) 2337–2342, doi:http://dx.doi.org/10.1002/rcm.1196.

[26] A. Frank, P. Pevzner PepNovo, De novo peptide sequencing via probabilistic network modeling, Anal. Chem. 77 (2005) 964–973, doi:http://dx.doi.org/10.1021/ac048788h.

[27] H. Chi, R.-X. Sun, B. Yang, C.-Q. Song, L.-H. Wang, C. Liu, et al., pNovo: de novo peptide sequencing and identification using HCD, Spectra, J. Proteome Res. 9 (2010) 2713–2724, doi:http://dx.doi.org/10.1021/pr100182k.

[28] A. Shevchenko, S. Sunyaev, A. Loboda, A. Shevchenko, P. Bork, W. Ens, et al., Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching, Anal. Chem. 73 (2001) 1917–1926, doi:http://dx.doi.org/10.1021/ac0013709.

[29] F.V. Leprevost, R.H. Valente, D.B. Lima, J. Perales, R. Melani, J.R. Yates, et al., PepExplorer: a similarity-driven tool for analyzing de novo sequencing results, Mol. Cell Proteomics 13 (2014) 2480–2489, doi:http://dx.doi.org/10.1074/mcp.M113.037002.

[30] Instituto Vital Brazil, Casa de Vital Brazil, Fundação Butantan, editors. A defesa contra o ophidismo: 100 anos depois: comentarios. 1st ed. Niterói: Instituto Vital Brazil, 2011.

[31] D. Georgieva, M. Ohler, J. Seifert, M. Bergen von, R.K. Arni, N. Genov, et al., Snake venomic of *Crotalus durissus terrificus*—Correlation with pharmacological activities, J. Proteome Res. 9 (2010) 2302–2316, doi:http://dx.doi.org/10.1021/pr901042p.

[32] K.H. Slotta, H. Fraenkel-Conrat, Estudos químicos sobre os venenos ofídicos. 4. Purificação e cristalização do veneno da cobra cascavel, Mem. Inst. Butantan (1938) 505–512.

[33] J.J. Calvete, L. Sanz, P. Cid, P. de la Torre, M. Flores-Di'az, M.C. Dos Santos, et al., Snake venomics of the Central American Rattlesnake *Crotalus simus* and the South American *Crotalus durissus* complex points to neurotoxicity as an adaptive paedomorphic trend along *Crotalus* dispersal in South America, J. Proteome Res. 9 (2010) 528–544, doi:http://dx.doi.org/10.1021/pr9008749.

[34] K.C. Giannotti, A. Sesso, K.F. Grego, W. Fernandes, R.P. Cardoso Jr., G.G. Camargo, et al., Viperid venom glands with defective venom production. Morphological study, Toxicon 70 (2013) 32–43, doi:http://dx.doi.org/10.1016/j.toxicon.2013.03.019.

[35] B.R. Fonslow, P.C. Carvalho, K. Academia, S. Freeby, T. Xu, A. Nakorchevsky, et al., Improvements in proteomic metrics of low abundance proteins through proteome equalization using ProteoMiner prior to MudPIT, J. Proteome Res. 10 (2011) 3690–3700, doi:http://dx.doi.org/10.1021/pr200304u.

[36] M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, et al., A new bioinformatics analysis tools framework at EMBL-EBI, Nucleic Acids Res. 38 (2010) W695–W699, doi:http://dx.doi.org/10.1093/nar/gkq313.

[37] T. Xu, J. Venable, S.K. Park, D. Cociorva, B. Lu, L. Liao, et al., ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program, Mol. Cell Proteomics 5 (2006) S174.

[38] P.C. Carvalho, J.S.G. Fischer, T. Xu, D. Cociorva, T.S. Balbuena, R.H. Valente, et al., Search engine processor: filtering and organizing peptide spectrum matches, Proteomics 12 (2012) 944–949, doi:http://dx.doi.org/10.1002/pmic.201100529.

[39] B. Zhang, M.C. Chambers, D.L. Tabb, Proteomic parsimony through bipartite graph analysis improves accuracy and transparency, J. Proteome Res. 6 (2007) 3549–3557, doi:http://dx.doi.org/10.1021/pr070230d.

[40] M. Junqueira, V. Spirin, T. Santana Balbuena, P. Waridel, V. Surendranath, G. Kryukov, et al., Separating the wheat from the Chaff: unbiased filtering of background tandem mass spectra improves protein identification, J Proteome Res 7 (2008) 3382–3395, doi:http://dx.doi.org/10.1021/pr800140v.

[41] B. Zybailov, A.L. Mosley, M.E. Sardiu, M.K. Coleman, L. Florens, M.P. Washburn, Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*, J. Proteome Res. 5 (2006) 2339–2347, doi:http://dx.doi.org/10.1021/pr060161n.

[42] I.M.B. Francischetti, M.E.C. Gombarovits, J.G. Valenzuela, C.R. Carlini, J.A. Guimarães, Intraspecific variation in the venoms of the South American rattlesnake (*Crotalus durissus terrificus*), Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. 127 (2000) 23–36, doi:http://dx.doi.org/10.1016/S0742-8413(00)00129-8.

[43] M.C. Menezes, M.F. Furtado, S.R. Travaglia-Cardoso, A.C.M. Camargo, S.M.T. Serrano, Sex-based individual variation of snake venom proteome among eighteen *Bothrops jararaca* siblings, Toxicon 47 (2006) 304–312, doi:http://dx.doi.org/10.1016/j.toxicon.2005.11.007.

[44] M. Li, B.G. Fry, R.M. Kini Eggs-Only Diet, Its Implications for the toxin profile changes and ecology of the Marbled Sea Snake (*Aipysurus eydouxii*), J. Mol. Evol. 60 (2005) 81–89, doi:http://dx.doi.org/10.1007/s00239-004-0138-0.

[45] H.L. Gibbs, L. Sanz, J.E. Chiucchi, T.M. Farrell, J.J. Calvete, Proteomic analysis of ontogenetic and diet-related changes in venom composition of juvenile and adult Dusky Pigmy rattlesnakes (*Sistrurus miliarius barbouri*), J. Proteomics 74 (2011) 2169–2179, doi:http://dx.doi.org/10.1016/j.jprot.2011.06.013.

[46] J.L. López-Lozano, M.V. de Sousa, C.A.O. Ricart, C. Chávez-Olortegui, E. Flores Sanchez, E.G. Muniz, et al., Ontogenetic variation of metalloproteinases and plasma coagulant activity in venoms of wild *Bothrops atrox* specimens from Amazonian rain forest, Toxicon 40 (2002) 997–1006, doi:http://dx.doi.org/10.1016/S0041-0101(02)00096-X.

[47] J. Boldrini-França, C. Corrêa-Netto, M.M.S. Silva, R.S. Rodrigues, P. De La Torre, A. Pérez, et al., Snake venomics and antivenomics of Crotalus durissus subspecies from Brazil: assessment of geographic variation and its implication on snakebite management, J. Proteomics 73 (2010) 1758–1776, doi:http://dx.doi.org/10.1016/j.jprot.2010.06.001.

[48] J.J. Calvete, L. Sanz, A. Pérez, A. Borges, A.M. Vargas, B. Lomonte, et al., Snake population venomics and antivenomics of *Bothrops atrox*: paedomorphism along its transamazonian dispersal and implications of geographic venom variability on snakebite management, J. Proteomics 74 (2011) 510–527, doi:http://dx.doi.org/10.1016/j.jprot.2011.01.003.

[49] E.K. Johnson, K.V. Kardong, C.L. Ownby, Observations on white and yellow venoms from an individual southern Pacific rattlesnake (*Crotalus viridis helleri*), Toxicon 25 (1987) 1169–1180, doi:http://dx.doi.org/10.1016/0041-0101(87)90135-8.

[50] M.-Q. Dong, J.D. Venable, N. Au, T. Xu, S.K. Park, D. Cociorva, et al., Quantitative mass spectrometry identifies insulin signaling targets in *C. elegans*, Science 317 (2007) 660–663.

[51] A.C. Paoletti, T.J. Parmely, C. Tomomori-Sato, S. Sato, D. Zhu, R.C. Conaway, et al., Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors, Proc. Natl. Acad. Sci. 103 (2006) 18928–18933, doi:http://dx.doi.org/10.1073/pnas.0606379103.

[52] M.L. Santoro, M.C. Sousa-e-Silva, L.R. Gonçalves, S.M. Almeida-Santos, D.F. Cardoso, I.L. Laporta-Ferreira, et al., Comparison of the biological activities in venoms from three subspecies of the South American rattlesnake (*Crotalus durissus terrificus*, *C. durissus cascavella* and *C. durissus collilineatus*), Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. 122 (1999) 61–73.

[53] K.C.F. Bordon, M.G. Perino, J.R. Giglio, E.C. Arantes, Isolation, enzymatic characterization and antiedematogenic activity of the first reported rattlesnake hyaluronidase from *Crotalus durissus terrificus* venom, Biochimie 94 (2012) 2740–2748, doi:http://dx.doi.org/10.1016/j.biochi.2012.08.014.

[54] M.J. Margres, J.J. McGivern, K.P. Wray, M. Seavy, K. Calvin, D.R. Rokyta, Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*), J. Proteomics 96 (2014) 145–158, doi:http://dx.doi.org/10.1016/j.jprot.2013.11.001.

[55] P.G. Righetti, E. Fasoli, E. Boschetti, Combinatorial peptide ligand libraries: the conquest of the hidden proteome advances at great strides, Electrophoresis 32 (2011) 960–966, doi:http://dx.doi.org/10.1002/elps.201000589.

[56] S. Pevtsov, I. Fedulova, H. Mirzaei, C. Buck, X. Zhang, Performance evaluation of existing de novo sequencing algorithms, J. Proteome Res. 5 (2006) 3018–3028, doi:http://dx.doi.org/10.1021/pr060222h.

[57] N. Wielsch, H. Thomas, V. Surendranath, P. Waridel, A. Frank, P. Pevzner, et al., Rapid validation of protein identifications with the borderline statistical confidence via de novo sequencing and MS BLAST searches, J. Proteome Res. 5 (2006) 2448–2456, doi:http://dx.doi.org/10.1021/pr060200v.

[58] D.R. Rokyta, K.P. Wray, M.J. Margres, The genesis of an exceptionally lethal venom in the timber rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics, BMC Genom. 14 (2013) 394, doi:http://dx.doi.org/10.1186/1471-2164-14-394.

[59] D.R. Rokyta, A.R. Lemmon, M.J. Margres, K. Aronow, The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*), BMC Genom. 13 (2012) 312, doi:http://dx.doi.org/10.1186/1471-2164-13-312.

[60] T.-Y. Chang, S.-H. Mao, Y.-W. Guo, Cloning and expression of a cysteine-rich venom protein from *Trimeresurus mucrosquamatus* (Taiwan Habu), Toxicon 35 (1997) 879–888, doi:http://dx.doi.org/10.1016/S0041-0101(96)00192-4.

[61] Y. Yamazaki, R.L. Brown, T. Morita, Purification and cloning of toxins from elapid venoms that target cyclic nucleotide-gated ion channels†,‡, Biochemistry (Moscow) 41 (2002) 11331–11337, doi:http://dx.doi.org/10.1021/bi026132h.

[62] Y. Yamazaki, F. Hyodo, T. Morita, Wide distribution of cysteine-rich secretory proteins in snake venoms: isolation and cloning of novel snake venom cysteine-rich secretory proteins, Arch. Biochem. Biophys. (2003) 133–141, doi:http://dx.doi.org/10.1016/S0003-9861(03)00028-6.

[63] Y. Yamazaki, H. Koike, Y. Sugiyama, K. Motoyoshi, T. Wada, S. Hishinuma, et al., Cloning and characterization of novel snake venom proteins that block smooth muscle contraction, Eur. J. Biochem. 269 (2002) 2708–2715, doi:http://dx.doi.org/10.1046/j.1432-1033.2002.02940.x.

[64] E.N. Castro, B. Lomonte, M. del Carmen Gutiérrez, A. Alagón, J.M. Gutiérrez, Intraspecies variation in the venom of the rattlesnake *Crotalus simus* from Mexico: different expression of crotoxin results in highly variable toxicity in the venoms of three subspecies, J. Proteomics 87 (2013) 103–121, doi:http://dx.doi.org/10.1016/j.jprot.2013.05.024.

[65] J.J. Calvete, A. Pérez, B. Lomonte, E.E. Sánchez, L. Sanz, Snake Venomics of *Crotalus tigris*: the minimalist toxin arsenal of the deadliest neartic rattlesnake venom. Evolutionary clues for generating a pan-specific antivenom against crotalid type II venoms, J. Proteome Res. 11 (2012) 1382–1390, doi:http://dx.doi.org/10.1021/pr201021d.

[66] K. Sunagar, E.A.B. Undheim, H. Scheib, E.C.K. Gren, C. Cochran, C.E. Person, et al., Intraspecific venom variation in the medically significant Southern Pacific Rattlesnake (*Crotalus oreganus helleri*): biodiscovery, clinical and evolutionary implications, J. Proteomics 99 (2014) 68–83, doi:http://dx.doi.org/10.1016/j.jprot.2014.01.013.

[67] J. Wang, B. Shen, M. Guo, X. Lou, Y. Duan, X.P. Cheng, et al., Blocking effect and crystal structure of natrin toxin, a cysteine-rich secretory protein from *Naja atra* venom that targets the BKCa channel†,‡, Biochemistry (Moscow) 44 (2005) 10145–10152, doi:http://dx.doi.org/10.1021/bi050614m.

[68] F. Wang, H. Li, M. Liu, H. Song, H. Han, Q. Wang, et al., Structural and functional analysis of natrin, a venom protein that targets various ion channels, Biochem. Biophys. Res. Commun. 351 (2006) 443–448, doi:http://dx.doi.org/10.1016/j.bbrc.2006.10.067.
[69] H.M. Doery, J.E. Pearson, Phospholipase B in snake venoms and bee venom, Biochem. J. 92 (1964) 599–602.
[70] A.W. Bernheimer, R. Linder, S.A. Weinstein, K.-S. Kim, Isolation and characterization of a phospholipase B from venom of collett's snake, *Pseudechis colletti*, Toxicon 25 (1987) 547–554, doi:http://dx.doi.org/10.1016/0041-0101(87)90290-X.
[71] S.T. Chatrath, A. Chapeaurouge, Q. Lin, T.K. Lim, N. Dunstan, P. Mirtschin, et al., Identification of novel proteins from the venom of a cryptic snake *Drysdalia coronoides* by a combined transcriptomics and proteomics approach, J. Proteome Res. 10 (2011) 739–750, doi:http://dx.doi.org/10.1021/pr1008916.
[72] S.D. Aird, Y. Watanabe, A. Villar-Briones, M.C. Roy, K. Terada, A.S. Mikheyev, Quantitative high-throughput profiling of snake venom gland transcriptomes and proteomes (*Ovophis okinavensis* and *Protobothrops flavoviridis*), BMC Genom. 14 (2013) 790, doi:http://dx.doi.org/10.1186/1471-2164-14-790.
[73] M.J. Margres, K. Aronow, J. Loyacano, D.R. Rokyta, The venom-gland transcriptome of the eastern coral snake (*Micrurus fulvius*) reveals high venom complexity in the intragenomic evolution of venoms, BMC Genom. 14 (2013) 531, doi:http://dx.doi.org/10.1186/1471-2164-14-531.
[74] Y. Yamazaki, Y. Matsunaga, Y. Tokunaga, S. Obayashi, M. Saito, T. Morita, Snake venom vascular endothelial growth factors (VEGF-Fs) exclusively vary their structures and functions among species, J. Biol. Chem. 284 (2009) 9885–9891, doi:http://dx.doi.org/10.1074/jbc.M809071200.
[75] R. Joseph, S. Pahari, W.C. Hodgson, R.M. Kini, Hypotensive agents from snake venoms, Curr. Drug Targets – Cardiovasc. Haematol. Disord. 4 (2004) 437–459, doi:http://dx.doi.org/10.2174/1568006043335808.
[76] L.S. Wermelinger, D.L.S. Dutra, A.L. Oliveira-Carvalho, M.R. Soares, C. Bloch, R.B. Zingali, Fast analysis of low molecular mass compounds present in snake venom: identification of ten new pyroglutamate-containing peptides, Rapid Commun. Mass Spectrom. 19 (2005) 1703–1708, doi:http://dx.doi.org/10.1002/rcm.1973.
[77] S.C. Wagstaff, P. Favreau, O. Cheneval, G.D. Laing, M.C. Wilkinson, R.L. Miller, et al., Molecular characterisation of endogenous snake venom metalloproteinase inhibitors, Biochem. Biophys. Res. Commun. 365 (2008) 650–656, doi:http://dx.doi.org/10.1016/j.bbrc.2007.11.027.
[78] F.X. Gomis-Rüth, L.F. Kress, J. Kellermann, I. Mayr, X. Lee, R. Huber, et al., Refined 2.0 Å X-ray crystal structure of the snake venom zinc-endopeptidase adamalysin II: primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin, J. Mol. Biol. 239 (1994) 513–544, doi:http://dx.doi.org/10.1006/jmbi.1994.1392.
[79] L. Watanabe, J.D. Shannon, R.H. Valente, A. Rucavado, A. Alape-Girón, A.S. Kamiguti, et al., Amino acid sequence and crystal structure of BaP1, a metalloproteinase from Bothrops asper snake venom that exerts multiple tissue-damaging activities, Protein Sci. 12 (2003) 2273–2281, doi:http://dx.doi.org/10.1110/ps.03102403.
[80] S.D. Aird, I.I. Kaiser, R.V. Lewis, W.G. Kruggel, Rattlesnake presynaptic neurotoxins: primary structure and evolutionary origin of the acidic subunit, Biochemistry (Moscow) 24 (1985) 7054–7058, doi:http://dx.doi.org/10.1021/bi00346a005.
[81] S.D. Aird, J.R. Yates III, P.A. Martino, J. Shabanowitz, D.F. Hunt, I.I. Kaiser, The amino acid sequence of the acidic subunit B-chain of crotoxin, Biochim. Biophys. Acta BBA – Protein Struct. Mol. Enzymol. 1040 (1990) 217–224, doi:http://dx.doi.org/10.1016/0167-4838(90)90079-U.
[82] J.A. Portes-Junior, N. Yamanouye, S.M. Carneiro, P.S. Knittel, S.S. Sant'Anna, F.C. S. Nogueira, et al., Unraveling the processing and activation of snake venom metalloproteinases, J. Proteome Res. 13 (2014) 3338–3348, doi:http://dx.doi.org/10.1021/pr500185a.
[83] R. Brigelius-Flohé, Tissue-specific functions of individual glutathione peroxidases, Free Radic. Biol. Med. 27 (1999) 951–965, doi:http://dx.doi.org/10.1016/S0891-5849(99)00173-2.
[84] R. Richards, L. St Pierre, M. Trabi, L.A. Johnson, J. de Jersey, P.P. Masci, et al., Cloning and characterisation of novel cystatins from elapid snake venom glands, Biochimie 93 (2011) 659–668, doi:http://dx.doi.org/10.1016/j.biochi.2010.12.008.
[85] S.P. Mackessy, L.M. Baxter, Bioweapons synthesis and storage: The venom gland of front-fanged snakes, Zool. Anz. – J. Comp. Zool. 245 (2006) 147–159, doi:http://dx.doi.org/10.1016/j.jcz.2006.01.003.
[86] A.D. Hargreaves, M.T. Swain, M.J. Hegarty, D.W. Logan, J.F. Mulley, Restriction and recruitment—gene duplication and the origin and evolution of snake venom toxins, Genome Biol. Evol. 6 (2014) 2088–2095, doi:http://dx.doi.org/10.1093/gbe/evu166.