Integrating Top-Down and Bottom-Up Mass Spectrometric Strategies for Proteomic Profiling of Iranian Saw-Scaled Viper, *Echis carinatus sochureki*, Venom

*Parviz Ghezellou,† Wendell Albuquerque,‡ Vannuruswamy Garikapati,† Nicholas R. Casewell,§ Seyed M. Kazemi,¶ Alireza Ghassempour,¶ Bernhard Spengler,\*,†*

† Institute of Inorganic and Analytical Chemistry, Justus Liebig University Giessen, Germany

‡ Institute of Food Chemistry and Food Biotechnology, Justus Liebig University Giessen, Germany

§ Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, United Kingdom

¶ Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, Tehran, Iran

**Corresponding Author**

Prof. Dr. Bernhard Spengler

E-mail: [Bernhard.Spengler@anorg.chemie.uni-giessen.de](file:///C:\Users\SpenglerUni\Library\Containers\com.apple.mail\Data\Library\Mail%20Downloads\957B760D-DF8F-49BD-9310-16B67A51FCC2\Bernhard.Spengler@anorg.chemie.uni-giessen.de)

**ABSTRACT**

Saw-scaled or carpet vipers (genus *Echis*) are considered to cause a higher global snakebite mortality than any other snake. *Echis carinatus sochureki* (ECS) is a widely distributed snake species, also found across the thirteen provinces of Iran, where it is assumed to be responsible for the most snakebite envenomings. Here, we collected the Iranian specimens of ECS from three different geographically distinct populations, investigated food habits, performed toxicity assessment and venom proteome profiling to better understand saw-scaled viper life. Our results show that the prey items most commonly found in all populations were arthropods, with scorpions from the family Buthidae particularly well represented. LD50 (median lethal dose) values of the crude venom demonstrate highly comparable venom toxicities in mammals. Consistent with this finding, venom characterization via top-down and bottom-up proteomics, applied to both crude venoms and size-exclusion chromatographic fractions, revealed highly comparable venom compositions among the different populations. By combining all proteomics data, we identified 22 protein families from 102 LC-MS/MS raw files, including the most abundant snake venom metalloproteinases (SVMPs, 29-34%); phospholipase A2 (PLA2s, 26-31%); snake venom serine proteinases (SVSPs, 11-12%); L-amino acid oxidases (LAOs, 8-11%), C-type lectins/lectin-like (CTLs, 7-9%) protein families, and many newly detected ones, e.g. renin-like aspartic proteases (RLAPs), fibroblast growth factors (FGFs), peptidyl-prolyl cis-trans isomerases (PPIs) and venom vasodilator peptides (VVPs). Furthermore, we identified and characterized methylated, acetylated, and oxidized proteoforms relating to the PLA2 and disintegrin toxin families and the site of their modifications. It thus seems that post-translational modifications (PTMs) of toxins, particularly target lysine residues, may play an essential role in the structural and functional properties of venom proteins and might be able to influence the therapeutic response of antivenoms, to be investigated in future studies.

**Keywords:** Serpentes:Viperidae:*Echis carinatus sochureki*, saw-scaled viper, middle east, venom, top-down venomics, bottom-up venomics

**INTRODUCTION**

Snake venoms comprise a highly complex mixture of proteins and peptides (90-95%), along with minor amounts of other compounds such as salts, carbohydrates, nucleosides, amines, lipids and free amino acids.1 The proteinaceous components of venoms have been shown to vary among species and to exhibit a wide range of biological functions.2 While these toxins have primarily evolved to facilitate prey capture, they are also utilized defensively in cases of human snakebite.2 Snakebite envenoming is a global public health crisis that results in more than 100,000 deaths per year from at least 1.8-2.7 million annual envenomings.3 While few reliable statistics exist relating to non-fatal outcomes of snakebite, many snake venoms are known to cause permanent physical (e.g. via local tissue necrosis or kidney injury) and psychological disabilities.4 Snakebites also disproportionately affect the rural impoverished populations of the tropics, and thus perpetuates poverty, resulting in the World Health Organization (WHO) recently recognizing snakebite envenoming as a priority “neglected tropical disease”.5 Snakebites are therefore a significant and serious medical problem, but it is worth noting that the same pathological toxins found in venom represent a rich natural biological resource with potential value for translation into human therapeutics.6

The most medically important snake species belong to the families Viperidae (e.g. vipers, pit vipers) and Elapidae (e.g. cobras, mambas, sea snakes, kraits, etc.).7 Generally, the clinical systemic symptoms induced by elapid snakebites are neurotoxic, neuromuscular paralysis and respiratory arrest resulting in mortality. Contrastingly, the clinical features observed after envenoming by viperids usually relate to haemotoxicity and/or cytotoxicity, these include local tissue damage such as oedema, blistering, haemorrhage, dermo- and myonecrosis, and also systemic alterations such as haemorrhage, coagulopathy, cardiovascular disturbances and renal damage.7,8 Among all venomous snakes, the saw-scaled or carpet vipers (family Viperidae; genus *Echis*) are thought to be responsible for a higher global snakebite mortality than any other snake genus.9 In part, this is due to their broad geographical distribution – from the regions of west, northern and east Africa, to west, south and east Arabia, through parts of Iran and Afghanistan north to Uzbekistan, and to Pakistan, India and Sri Lanka.11 The genus *Echis* is categorized into four main clades, containing *E. carinatus*, *E. coloratus*, *E. ocellatus* and *E. pyramidum*.10 The *E. carinatus* group has been further divided in to two subspecies; *E. carinatus carinatus* and *E. carinatus sochureki* (ECS), the latter of which is distributed across the thirteen provinces of Iran.12

Iran is located between three zoogeographical realms (Palearctic, Oriental and Afrotropical), which have helped shape its high habitat diversity, and in turn support diverse herpetofauna.12 There are approximately 37 species of venomous snakes (26 front-fang and 11 rear-fang species) and they are distributed across all ecoregions of the country.12,13 Despite a paucity of reliable snakebite burden data, it is estimated that ~4500-6500 individuals are envenomed in Iran annually, resulting in at least 3-9 deaths.14 It is also assumed that ECS is responsible for the majority of serious envenomings, and the WHO lists this species as Category 1 of “highest medical importance”.15 The clinical profile mainly observed, following bites by this species, include local manifestations such as pain, swelling, blistering, necrosis, and systemic manifestations like spontaneous bleeding, intravascular hemolysis, venom-induced consumption coagulopathy (VICC; historically referred to as disseminated intravascular coagulation (DIC)), which sometimes leads to acute renal failure (ARF) and acute pancreatitis.16–18 The administration of RAZITM polyvalent F(ab’)2 antivenom, which is produced from hyper-immune plasma, collected from horses which were immunized with a mixture of various snake venoms (from *Echis carinatus*, *Macrovipera lebetina*, *Pseudocerastes persicus*, *Gloydious halys*, *Montivipera albicornata* and *Naja oxiana*), continues to be the only treatment for systemic snakebite envenoming in Iran.19 Although there is some controversy regarding the efficacy of this antivenom,20 recently Monzavi et al.16 reported its clinical effectiveness in treating bites by ECS envenoming in northeast Iran.

While antivenom therapies save thousands of lives each year, their efficacies are typically restricted to the species of snake whose venom was used in its manufacture,21 as a result of inter- and intraspecific variation in venom composition.22 Such venom variation can be the result of evolutionary history, climatic factors, ontogeny, or adaptation towards different prey.22-25 Therefore, knowledge of snake venom composition, and associated variation, among conspecific populations can provide important information for predicting the likely efficacy of an existing antivenom, along with influencing the design of more effective immunizing mixtures for future antivenom production.26 Consequently, over the past decade, significant research efforts have been made towards profiling venom proteomes using mass spectrometry (MS)-based proteomics and venom gland transcriptomic approaches.27

The current most widely used proteomics method for identifying venom-expressed proteins is bottom-up (BU) venomics, in which venom proteins are digested into peptide fragments prior to MS interrogation and de novo peptide sequencing. As an example, this approach was previously used to characterize the venom of ECS from United Arab Emirates (UAE), in combination with venom gland transcriptomics.22,28 These studies resulted in the identification of snake venom metalloprotease (SVMP), snake venom serine protease (SVSP), phospholipase A2 (PLA2), L-amino acid oxidase (LAAO), cysteine rich secretory protein (CRiSP), C-type lectin (CTL) and disintegrin (DIS) protein families, among others, in the venom of ECS. However, while BU analysis has a high throughput, is sensitive and robust,29 it is predominately only capable of identifying the representative protein for each expressed gene,30 and does not provide information on the proteoforms, genetic variation and PTMs, associated with the sample.31,32 Thus, a top-down (TD) strategy, based on measurement of an intact protein, is a valuable approach for analyzing venoms at the proteoform-level, as genes often encode several isoforms and proteins with different modifications.33 Steady advances in mass spectrometry technologies has facilitated improvements in TD proteomics, enabling quick and accurate investigation of intact toxin families and their proteoforms (see 33, 34). However, this technique still has limitations in providing full sequence coverage of large (>30 kDa), as well as low-abundant, intact proteins. In this case, the application of a denaturing TD approach, in particular for viper venoms that mostly contain larger protein families (e.g. SVMPs, LAAOs, hyaluronidase, etc.), limits detection to part of a sequence, and requires the use of native MS which is experimentally and bioinformatically challenging.35 Nevertheless, the development of various MS-based proteomics strategies has verified the technique as an essential technology for achieving sequence information for protein identification and the interpretation of post-translational modifications.

Here, we performed different studies related to one of the medically most important snakes, ECS, which is illustrated schematically as an overview in Figure 1 and explained in the following sub-sections. Briefly, at the beginning of the study, individual snakes were systematically identified and collected for compression from three populations located in the South (Hormozgan province), Southeast (Sistan and Baluchestan province) and East (South Khorasan province) of Iran. It seems likely that the sampled areas are geographically exposed to the high-risk places for snakebite incidence and envenomation, particularly in ECS bite cases. We continued our study with milking the venom of individuals and pooled venom within the populations, followed by an investigation on the feeding habits of individuals within each community. Next, the median lethal toxicity (LD50) of each population-representative venom was evaluated by administration of various dosages of crude venoms to laboratory mice. Subsequently, we used a combination of BU and TD proteomics approaches to identify and characterize in detail venom protein compositions and overview of conspecific venom variation of Iranian ECS.

|  |
| --- |
| **Figure 1**. Overview of the proposed experimental workflow used for interrogation of food habits, toxicity assessment and venom proteome profiling of three Iranian populations of saw-scaled viper, Echis carinatus sochureki. HO: Hormozgan; SK: South Khorasan; SB: Sistan and Bluchestan; LD50: median lethal dose. |

**EXPERIMENTAL SECTION**

**Sample Collection and Venom Milking**

The saw-scaled viper, *Echis carinatus sochureki*, specimens were collected in 2015-2017 between March and October in three locations (South Khorasan [SK], Sistan and Bluchestan [SB], Hormozgan [HO] provinces) of Iran (Figure. 1). Venom extractions (milking) were performed by encouraging the snakes to bite down on parafilm-covered hygiene beakers without exerting pressure on the venom glands, followed by immediate flash-freezing of the samples with liquid nitrogen. In total, venom was milked from 59 (SK), 86 (SB) and 67 (HO) snake specimens from each location. The samples were pooled by region, lyophilized and stored at -80 °C for future research.

**Snake Diet and Venom Lethality**

After milking, the specimens were investigated for dietary habits by checking the stomach contents as well as faecal pellets of the living captured specimens. The belly of the snakes from each population was gently palpated to detect prey items (particularly large prey items). The specimens with gut contents were forced to regurgitate the ingested bolus according to the ethically sound method described by Kjaergaard.36 In addition, the snakes were kept in a cage with water ad libitum until collection of faecal pellets. The prey items were mostly identified to wide taxonomic classification such as phylum, family and genus by using sample evidences e.g. few scales and/or clumps of feathers (birds), hair (mammals), telson and stinger (scorpion) as well as field observations. The snakes were released to their original capture sites after the collection of dietary samples.

The LD50 (median lethal dose) of the three regional ECS venoms were determined by intraperitoneal (IP) administration of varying doses of the three crude venoms (in 0.9% NaCl; total volume 200 µl) to albino mice (20±2 g). Mortality was recorded 24 h after the injections and the LD50 calculated according to the Reed and Muench method.37 All animals used in this study were maintained under standard conditions, and all experiments were performed according to the international guiding principles involving animals for scientific research38 and under approvals granted by the Ethics Committee of Shahid Beheshti University.

**Protein Estimation**

All crude venoms and isolated fractions (see later) were stored at –80 °C and the protein concentrations were determined before each proteomics analysis using a standard Bradford assay (Biorad, Hercules, CA, USA), with bovine serum albumin (BSA) used as reference. Absorbance was measured spectrophotometrically at 595 nm on a BioTek Synergy 2 plate reader (BioTek, Winooski, VT, USA) with Gen5 software (version 2.01).

**Chemicals**

All chemicals were purchased from Sigma-Aldrich (Sigma-Aldrich, MO, USA), and MS-grade solvents (acetonitrile and water) were purchased from VWR chemicals (VWR international, Darmstadt, Germany). RapiGest SF surfactant was purchased from Waters (Waters Corporation, MA, USA). The mass spec grade Trypsin/Lys-C mix was purchased from Promega (Promega, Mannheim, Germany). Zip Tip C18 was purchased from Millipore (Bedford, MA, USA).

**SDS-PAGE and Gel Filtration Separation**

Each of the regional ECS venoms were individually applied to SDS-PAGE, as previously described,39 as well as a HiLoad 16/60 Superdex 200 prep grade gel-filtration chromatography column (GE Healthcare Bio-Sciences, Freiburg, Germany) in a Fast Protein Liquid Chromatography (FPLC) system (Bio-rad NGCTM Quest Plus) running Bio-Rad ChromLab software (version 3.1). For FPLC, a constant flow rate of 0.7 ml/min was used with a buffer consisting of 40 mM Tris-HCL and 150 mM NaCl (pH 7.4). The resulting six fractions were determined by a UV detector (280 nm), and collected automatically by a fraction collector (BioFracTM, Bio-Rad). After size exclusion separation, the collected fractions were dialyzed against distilled water using 2,000 MWCO Slide-A-Lyzer Dialysis Cassettes (Thermo Fisher Scientific) at 4 °C to remove the salts, and then freeze-dried using an Alpha 1-2 LD plus freeze dryer (Martin Christ, Osterode, Germany).

**Bottom-up Proteomics**

Equal amounts of the whole venom and collected fractions (20 µg) were diluted with 50 mM ammonium bicarbonate buffer, containing 0.1% RapiGest, and incubated for 15 min in a thermomixer at 80 °C (Eppendorf Thermomixer C, Hamburg, Germany) to complete proteome solubilisation. The denatured samples were reduced and alkylated with 100 mM dithiothreitol (DTT) at 56 °C for 15 min, and 200 mM iodoacetamide (IAA) at room temperature (dark place) for 30 min, respectively. The digestion was performed with mass spec grade Trypsin/Lys-C mix (1:25 enzyme to proteins ratio) at 37 °C. The reaction was stopped after 16 h by adding concentrated formic acid and incubating at 37 °C for 10 min prior to centrifugation. The peptide samples were desalted before MS measurements by using ZipTip C18 and then concentrated by using Eppendorf Concentrator Plus (Eppendorf, Hamburg, Germany) and finally stored at –80 °C for future use. The digested crude venoms and fractions were separated using an UltiMate 3000 RSLC UHPLC system (Ultra High Performance Liquid Chromatography, Thermo Fisher Scientific) on a Kinetex C18 (2.1 × 100 mm, 2.6 µm 100 A° particle size) column (Phenomenex, CA, USA) coupled to a Q Exactive HF-X and Q Exactive (QE) Orbitrap (Thermo Scientific, Bremen, Germany). Chromatographic analysis was performed at 250 µl/min flow rate with water/0.1% formic acid (mobile phase A) and acetonitrile/0.1% formic acid (mobile phase B). The optimized gradient elution of 90 and 120 min were applied for fractions and whole venoms respectively as follows: isocratically (2% B) for 5 min, followed by 2-40% B over 70 min (30 min decreased for fractions), 40-50% B over 5 min, 50-98% B over 2 min, and re-equilibration in 2% B. The mass spectrometers were operated in data-dependent acquisition (top-10 DDA) with the following parameters in full MS scans: mass range of *m/z* 350 to 1800 in QE HF-X (mass resolution of 120,000, AGC target of 3e6, IT of 50 ms) and QE (mass resolution of 70,000, AGC target of 1e6, IT of 120 ms), and MS/MS scans: mass range of *m/z* 200 to 2000 in QE HF-X (mass resolution of 30,000, AGC target of 1e5, IT of 120 ms, isolation window *m/z* 1.3 and dynamic exclusion of 60 s) and QE MS (mass resolution of 17,500, AGC target of 1e6, IT of 120 ms, isolation window D(*m/z)* 3 and dynamic exclusion of 15 s).

**Bottom-up Data Analysis**

The raw files were processed using Proteome Discoverer version 2.2 (Thermo Scientific) with SEQUEST and MS Amanda (a peptide identification algorithm for high-accuracy and high-resolution mass spectrometry)40 algorithms against the UniProtKB flat file databases (downloaded on July 01, 2018), taxonomically set to the Serpentes (taxon ID # 8570) and *Echis carinatus* (Taxon ID # 40353). The parameters were set to two missed cleavage sites of trypsin digestion, minimum peptide length of 6, MS1 and MS2 tolerances of 10 ppm and 0.5 Da respectively. Dynamic modification was set to oxidation (+15.995 Da [M]) and static modification to carbamidomethyl (+57.021 Da [C]). In addition, acetylation (+42.011 Da), mono-methylation (+14.016 Da), tri-methylation (+42.047 Da) and phosphorylation (+79.966 Da) were set as dynamic modifications for the *Echis carinatus* database search. Percolator41 node was used to validate identified PSMs and filter the data with parameters of a strict Target FDR (false discovery rate) of 0.01 and a relaxed Target FDR of 0.05. The MaxQuant contaminant database was used to mark contaminants in the results file. Peptides and proteins were filtered with only high confidence and master proteins in the final results. Additionally, relative quantitation of protein abundance was achieved based on normalized spectral abundance factor (NSAF) manually.42,43

**Top-down Proteomics**

The Q Exactive HF-X Orbitrap (Thermo Scientific, Bremen, Germany) was equipped with a heated electrospray (HESI) probe installed in the Ion Max source, coupled with an UltiMate 3000 RSLC UHPLC system (Thermo Scientific). Chromatographic analysis was performed using a Jupiter C18 (4.6 × 250 mm, 3 µm 300 A° particle size) column (Phenomenex, CA, USA). The venom samples were dissolved in 1% formic acid and reduced with TCEP (tris (2-carboxyethylphosphine)) according to the protocol described by Liu et al.44 The samples were eluted with water/0.1% formic acid (mobile phase A) and acetonitrile/0.1% formic acid (mobile phase B). The optimized gradient elution of 290 min at a flow rate of 0.5 ml/min was applied for the whole venom separation as follows: isocratically (5% B) for 10 min, followed by 5-15% B over 45min, 15-40% B over 170 min, 40-70% B over 25 min, 70% B for 10 min and re-equilibration in 5% B. For the fractions, a 90 min gradient was used as follows: isocratically (5% B) for 5 min, followed by 5-55% B over 60 min, 55-95% B over 10 min, 95% for 5 min and re-equilibration in 5%. The mass spectrometer operated in data dependent acquisition (DDA) mode with external calibration and positive-ionization mode. The capillary temperature was set to 380 °C, and the spray voltage to 3.5 kV. The auxiliary gas and sheath gas flow rate were set to 20 and 60 respectively. Full MS scans were collected from *m/z* 300 to 3000 with a mass resolution of 120,000 (@ *m/z* 200) and an AGC target of 1e6. MS/MS scans were collected in the Orbitrap with the TopN 2 method (DDA mode) at a mass resolution of 60,000 @ *m/z* 200, an AGC target of 1e6, an (N)CE of 30, an isolation window of *m/z* 4 and dynamic exclusion of 60s.

**Top-down Data Analysis**

TopPIC Suite (<http://proteomics.informatics.iupui.edu/software/toppic>) software package 45 (version 1.1.1; released November 2017) was used to convert the .raw files to MzXML files, to deconvolute the mass spectra and to identify PrSM (Proteoform Spectrum Matches) via MSConvertGUI, TopFD and TopPIC\_GUI, respectively. In the deconvolution process, the maximum charge was set to 60, signal to noise ratio was set to 2, precursor window size set to *m/z* = 4, and *m/z* error to 0.02 u. Intact protein masses and MS/MS spectra were searched against the taxonomy settings of Serpentes (Taxon ID # 8570) and *Echis carinatus* (Taxon ID # 40353) in the UniProtKB database (downloaded on July 01, 2018), with the following parameter settings: mass error tolerance of 10 ppm and decoy database set to filter spectrum and proteoform level with a false discovery rate (FDR) cut-off of 0.01.

**Data Availability**

The mass spectrometry-based TD and BU proteomics data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository46 with the dataset identifier PXD021183 and project name “Saw-scaled viper, *Echis carinatus sochureki*, venomics”.

**RESULTS AND DISCUSSION**

**Dietary Composition and Venom Lethality**

In total, more than two hundred live snake specimens were examined to investigate the dietary composition of the three Iranian saw-scaled viper (*Echis carinatus sochureki*; ECS) communities. We detected prey items in the stomach and faecal contents of 96 live snakes, as detailed in Supplementary Information Table S1. The prey (in total 125 items) most commonly found in all three populations were arthropods (n=77, 61.6%) which most frequently assigned to members of the scorpion family Buthidae (n=73, 58.4%). The other frequently eaten species were from family Bufonidae (n=11, 8.8%), Lacertidae [n=13, 10.4%] and Gekkonidae [n=15, 12%]. We also recorded prey items from the following groups: insecta (n=4, 3.2%), arthropoda (araneae [n=2, 1.6%] and diplopoda [n=2, 1.6%]), aves (only one item, 0.8%) and mammalia (n=4, 3.2%). Regarding the scorpion fauna in the targeted area, two families, Buthidae and Scorpionidae, are distributed over the sampling regions,47,48 but family Buthidae is the predominant prey of ECS. Thus, it could be supposed that the food preference is strongly correlated with the availability and abundance of the Buthidae prey species. Although vertebrate diets are the ancestral characteristic of vipers, and most extant species remain vertebrate predators,49 our study shows that arthropods (particularly scorpions) make up over three-quarters of the diet of Iranian ECS. These findings are consistent with a previous study by Barlow et al.49 who reported that *E. carinatus* feeds on arthropods to a greater extent than vertebrates. In addition, due to an assumed connection between prey composition and venom toxicity, the venom of *E. carinatus* was shown to be more toxic to scorpions than the venom of infrequently arthropod-feeding saw-scaled vipers such as *E. ocellatus*.50 Here, we also assessed the lethal venom potencies of all three regional ECS venoms to determine the approximate LD50 dose to laboratory mice. The results showed highly comparable venom LD50 doses, corresponding to 32.28 µg/mouse for the SK population and 33.50 µg/mouse for SB and HO populations. Thus, in this case, highly comparable dietary composition appears to correlate well with highly comparable venom toxicity. However, further investigation into the venom toxicity on natural prey items is required to gain a deeper understanding of the relationship between diet and venom function at the intra-specific level.

**SDS-PAGE and Size Exclusion Chromatography Separation of The Crude Venoms**

To investigate venom composition, first the overall protein composition of crude venoms from three populations were screened by utilizing 1D SDS-PAGE and intact RP-UHPLC separations (Supplementary Information Figure S1). Evidence from multiple gel band patterns and the UHPLC chromatograms strongly suggested the complex proteome profile of ECS venom. The venom's complexity was expected by referring to the recently analyzed venoms of Pakistani and Indian *E. carinatus*.51-53 In addition, Coomassie blue-staining of the resulting gels revealed a high degree of similarity in the electrophoretic profiles of the three venoms, with all exhibiting polypeptide molecular weight ranges from <10 kDa to around 120 kDa in size. SEC was applied to de-complex the ECS venom proteome prior to MS measurements, and six fractions were detected and collected from each population of ECS and visualized by SDS-PAGE (Supplementary Information Figure S2). Subsequently, the resulting fractions were desalted and subjected to denaturing TD and in-solution digestion BU proteomics for protein identification.

**Bottom-up Proteomics of ECS Venom Proteomes**

To profile the venom proteomes of the three Iranian ECS populations by BU, the whole crude venoms, alongside with size-exclusion-separated fractions, were subjected to in-solution digestion before reverse-phase LC-ESI-HR-MS/MS measurements. All recorded spectra were searched against protein sequences of the Serpentes database entries in UniprotKB (contains all annotated and reviewed protein sequences of *Echis* species) by Sequest/Amanda/Percolator algorithms under the Proteome Discoverer software platform. The false discovery rate (FDR) was restricted to 1% at both protein and peptide levels using the target-decoy strategies and set to only high-confidence sequence matches. The data acquired by BU from the crude venoms (three populations in triplicate) were merged to allow identification of peptides corresponding to 54 (HO), 49 (SK) and 53 (SB) unique protein entries, which belong to 12, 11 and 13 snake protein families respectively (Table 1 and Supporting Information Table S2-4). However, BU measurements of SEC venom fractions (six fractions per population in triplicate) via the same LC-MS/MS method (in DDA mode) led to the identification of 101 (HO), 99 (SK) and 91 (SB) unique protein entries that were clustered in 16, 18 and 14 protein families respectively (Table 1 and Supporting Information Table S5-7). Thus, it is important to note that pre-digestion protein fractionation by SEC substantially increased the BU MS/MS protein quantification by nearly two-fold. The results indicate that the three populations shared 14 protein families, belonging to the snake venom metalloproteinase (SVMP), group-II phospholipase A2 (PLA2), snake venom serine proteinase (SVSP), L-amino acid oxidase (LAO), C-type lectin/lectin-like (CTL), cysteine-rich secretory protein (CRISP), snake venom nerve growth factor (NGF), phospholipase B (PLB), disintegrin (DIS), 5’-nucleotidase (5’NTD), glutaminyl-peptide cyclotransferase (GC), renin-like aspartic protease (RLAP), aminopeptidase (AP), and phosphodiesterase (PDE) families. Furthermore, the venoms from HO and SK populations also exhibited the presences of hyaluronidase (HYAL) and vascular endothelial growth factor (VEGF) protein families. Two additional protein families were detected only in the SK population, namely acetylcholinesterase (AChE) and dipeptidyl peptidase (DPEP).

|  |
| --- |
| **Figure 2. Overview composition and relative abundance of protein families from venom proteomes of three Echis carinatus sochureki populations.** Pie charts representing the relative abundance of protein families based on the normalized spectral abundance factor (NSAF), identified in six size exclusion fractions of each venom community by bottom-up proteomics approach. SVMP: snake venom metalloproteinases; PLA2: group II phospholipase A2; CTL: C-type lectin; DIS: disintegrin; SP: serine protease; LAO: L-amino oxidase; CRISP: cysteine-rich secretory proteins; VEGF: vascular endothelial growth factors; NGF: nerve growth factor; AP: aminopeptidase; PDE: phosphodiesterase; HYAL: hyaluronidase; RLAP: renin-like aspartic protease; PLB: phospholipase B; GC: glutaminyl-peptide cyclotransferase; 5’NTD: 5’-nucleotidase; DPEP: dipeptidyl peptidase; AChE: acetylcholinesterase. |

The major protein classes identified in all ECS venoms are SVMP, PLA2, svSP, LAO and CTL, which is consistent with previous analyses of UAE ECS venom composition.22,28 Interestingly, the relative concentration of the major protein families detected in the three venom populations are highly similar as shown in Figure 2. Among them, SVMP and PLA2 are the most abundant toxin families, accounting for ~29-34% and ~26-31% of the venom components, respectively. These findings are also consistent with the main consequences of the local (edema, swelling, haemorrhage and pain) and systemic (blood coagulation) manifestations of snakebite by ECS in Iran.16 Previous venomics studies of E. carinatus from India revealed that SVMP, CTL, and PLA2 are the three dominant toxin families,52,53 corroborating our results. However, in Iranian ECS venom, the relative abundance of SVSP was higher than that of the C-type lectin, so that it has taken the third dominant place.

In all three venoms, both P-II and P-III sub-classes of SVMPs were detected, although the latter (more haemorrhagic than the P-II form) were more abundant, with a three- to six-fold increase over the P-IIs. In the case of the PLA2 protein family, our results showed the presence of two major sub-subgroups: Asp(49)- and Ser(49)-PLA2s were detected in the venom of all three Iranian ECS populations (Supplemental Information Table S2-11). Although the Ca++-dependent D(49)-PLA2 isoform is the plesiotypic form of PLA2 molecules in viperids, showing potent esterolytic activity,54 our data demonstrated that the Ca++-independent S(49)-PLA2 isoform, which exhibits low enzymatic activity,54 is the dominant PLA2 toxin. The next most abundant toxin family secreted in all targeted ECS venoms is SVSP, which exhibits highly consistent abundances across the three populations (11.5-12.5%). This multigene enzyme protein family has been shown hemostatic system disturbances and affects hemostasis, blood coagulation, platelets and fibrinolytic system in snakebite victims. Noteworthy, the SVSP toxin family was previously reported to be a minor toxin family in the venom gland and venom of ECS from UAE and India.22,52 As shown in Figure 2, other major toxin families, LAOs and CTLs, were present in lower but highly similar abundance across the three populations. It is assumed that they exert distinct multifunctional actions through snakebite envenoming such as edema, anticoagulation, platelet aggregation inhibiting, cytotoxicity, and disruption of hemostasis mechanisms through focusing on core elements involved in the blood-coagulation cascade.51,55,56 Further interesting toxin family with an even lower abundance (~5%) is CRiSP, which was recently isolated from *Bothrops jararaca* and reported to dictate the induction of pro-inflammatory responses that provoke the production of interleukin (IL)-6, also targeting the complement system.57,58 These records are also in good agreement with the clinical observations of Iranian ECS envenomation in human patients.16

In addition to those toxin families described above, our multidimensional (MD)-LC/MS-based BU proteomics approach obtained data on peptides that confirmed the existence of a wide variety of low-abundance protein families (e.g. <1%), such as 5’NTD, GC, RLAP, AP, PLB, PDE, HYAL, NGF, VEGF, AChE, and DPEP, in Iranian ECS venoms (Figure 2, Table 1 and Supporting Information Table S2-7). Notably, few of them have been reported in EC venom proteomes so far. This trend suffers mainly from the distribution of high-abundance proteins in complex samples (such as snake venom) which can prevent the detection of low-abundant ones, and also searching the MS data against incomplete sequence information in an existing database. In this case, using physicochemical protein separation before RP-LC/HR-MS/MS is capable of better resolving the complexity of venom proteome and helps to detect higher numbers of polypeptides.59 In addition to the proteomics approach, a previous transcriptomic study of the ECS venom gland detailed the expression of these low abundant components, with the exception of PLB and GC.28 PLB is a catalytic enzyme which cleaves phospholipids from sn-1 and sn-2 position,60 and exhibited a potent haemolytic activity on both rabbit and human erythrocytes.61 GC is also known to play a role in catalyzing N-terminal pyroglutamate (pGlu) formation of amino-terminal glutamate residues of some snake venom polypeptides.62 The modification function is still poorly understood but seems to be essential for the maturation and conformation of some toxins and maybe for protection from degradation by peptidases such as aminopeptidase.63 Additionally, we found very trace quantities of AChE enzyme in venom of the SK population that functionally hydrolyzes esters of choline. This enzyme has been reported mostly in the venom of Elapidae family with significant amounts but not in viperidae exclusively.64 However, weak AChE activity was detected for the venom of Pakistani *E. carinatus* recently.51 It should be noted that, although some studies identified and reported low abundance components of snake venom, their biological and pathophysiological functions are not well understood, and further studies are necessary to confirm whether many of these components actually contribute to pathology.

As already mentioned, despite the large geographical distances between each ECS population (~400 km), venom components of populations seem highly similar. However, each population shows some distinct properties, in terms of both quantitative and qualitative venom composition (Figure 2). Intraspecific venom variation is a known phenomenon65 and has previously been correlated with natural selection pressure for regional feeding habits65,66 among several other factors.23-26 Thus, the remarkable similarity in venom composition correlates with our dietary survey, and suggests that consistent foraging preferences (e.g. scorpions as the dominant prey item) may underpin the remarkable similarity of venom compositions of three ECS populations.

|  |
| --- |
| **Table 1.** Identified protein families in the whole and SEC fractions of three Echis carinatus sochureki venom populations by bottom-up and top-down proteomics. |

**Top-down Proteomics of ECS Venom Proteomes**

In addition to BU mass spectrometry of digested venom proteomes, we applied TD proteomics to provide in-depth profiling of the three ECS venoms and to facilitate the characterization of proteoforms with post-translational modifications (PTMs). All TD experimental spectra were initially searched against a decoy UniprotKB database containing Serpentes proteins, using a 1% spectrum-and proteoform-level FDR for stringent filtering identification with TopPIC. Data extracted from the triplicate TD LC-MS/MS analysis of the mixed crude venom (non-reduced and reduced form), revealed a total of 6910 identified PrSMs (Supplemental Information Table S12), which mapped to 170 proteins with unique accession numbers (Supplemental Information Table S8). In addition, duplicate LC-MS/MS analyses of each SEC fraction in non-reduced and reduced forms generated 5489 identified PrSMs (Supplemental Information Table S13-15), and resulted in 102 (HO), 114 (SB) and 133 (SK) proteins sorted by unique accession number (Supplemental Information Table S9-11). An overview of the TD results obtained from these various measurements revealed that all the detected proteins were grouped into 20 protein families, namely: 5'NTD, PLA2, CRISP, SVMP, DIS, CTL, GC, LAO, SVSP, RLAP, AP, PLB, PDE, HYAL, VEGF, BPPs/pHpG/C-NP (Bradykinin-potentiating/poly-His-poly-Gly/C-type natriuretic peptides), DPEP (Dipeptidyl peptidase), FGF (Fibroblast growth factor), PPI (Peptidyl-prolyl cis-trans isomerase) and VVP (Venom vasodilator peptide) (see Table 1). Comparisons of the BU and TD data indicated that differences in protein family detection related predominately to low-abundant protein families. For example, we detected FGF, PPI and VVP protein families in the TD datasets (mixed crude venom), while these were absent from BU MS measurements. In addition, these three protein classes were not observed in TD analysis of the venom fractions, and were only detected in the sample containing a mixture of the three venom population pools. Contrastingly, both NGF and AChE were only detected in the BU datasets (Table 1). It seems likely that the lack of detection of these protein families via proteomics measurements is the result of low-abundance loss during sample digestion, coupled with purification steps before MS measurements.67–69 Upon merging the TD and BU data, in total 22 protein families were identified through the series of 102 LC-MS/MS data files. Supplemental Tables S12-15 detail the identified PrSMs obtained from LC-ESI-HR-TD proteomics. These data highlighted that the largest number of identified PrSMs related to the SVMP, PLA2, and LAO toxin families, which is largely consistent with findings from BU. Ultimately, this comparative approach reveals that BU and TD proteomics complement each other and facilitate thorough identification of the diversity of snake venom components.

It is interesting to bring to notice that most abundant snake toxin families (such as PLA2, SVMP and svSP) are encoded by multi-locus gene families.70,71 These gene families generate a range of alternative gene products which are unequal to the fundamental polypeptides. Here, PrSMs derived from TD MS of ECS venoms, searched against the Serpentes database, give rise to many proteoforms with unknown mass discrepancies (Supplemental Table S12-15). While gene-based diversity is much lower than the proteoform variety, these unexplained mass shifts are often the result of primary structure alterations (PSAs), e.g. substitution, insertion or deletion of amino acids (gene product alterations), post-translational modification (PTMs) and terminal truncations.69 However, without a species-specific molecular database, such as that derived from genomic or transcriptomic experiments, it is ambiguous to discern the observed mass discrepancies to the aforementioned alterations/modifications accurately. To guide our interpretations here, we searched TD spectra against an ECS-specific database deposited in UniProtKB (which includes the prior venom gland transcriptome sequences of ECS from UAE) using the MIScore method,70 and choosing four common PTMs (acetylation, methylation, oxidation, and phosphorylation) in TopPIC software. Our resulting data revealed 166 modified full-length isoforms, with enough fragment ions to cover exact masses, across the three population-level venoms (HO=57, SK=43, SB=66), belonging to the PLA2 (157 proteoforms) and disintegrin (9 proteoforms) toxin families (Supplemental Information Table S16). These reported modifications correspond to mass shifts associated with three types of common PTMs; specifically via combinations of 11 methylated, 5 oxidized and 3 acetylated proteoforms in the venom fractions of all three populations (Table 2). Among all of the identified proteoforms, 54 PLA2 proteoforms showed a combination of two modified sites, of which 46 proteoforms can be elucidated by two methylation sites and 8 proteoforms by one acetylation and one oxidation sites. For example, comparison of the TD MS/MS of unmodified PLA2 proteoform (*m/z* 1063.2693 with charge state 13+ and proteoform mass of 13809.41 Da) and its modified proteoform (*m/z* 923.50 with charge state 15+ and proteoform mass of 13837.41 Da) resulted in a mass increase of 28 Da (Figure 3). The tandem mass spectra of the modified proteoform contained fragment ions (b and y ions) with additional mass of +28 and +14 Da compared to the unmodified proteoform, which can be explained by two methylations. Figure 4 illustrates another example relating to a methylated proteoform of the disintegrin protein family. It should be noted that ±1 Da errors are mostly observed in the deconvolution of precursor masses with more than 5 kDa,73 and in this case the deconvoluted MS2 spectra with mass shift of +13 Da instead of +14 Da were mapped and reported as a methylated proteoform (Figure 4).

The TD data also suggested that the methylated sites correlated with N-terminal (S1), Glu (E12), Gln (Q11) and Lys (K7, K15, and K67) amino acid residues (Table 2 and Supplemental Information Figure S3-12). In addition, indirect measurements of proteins at the peptide level by bottom-up MS provide some insight into the types of potential PTMs and their localization, in accordance with TD results (Table 2, Supplemental Information Table S17 and Figure S13).

|  |
| --- |
| **Figure 3. Top-down spectra of HCD fragmentations, and sequence coverages of Phospholipase A2 proteoforms in Echis carinatus sochureki venom from Iran**. A) Xtract-deconvoluted MS/MS spectrum from unmodified protein sequence which was identified as PLA2 protein family, and localization of b and y ions in sequence, B) Xtract-deconvoluted MS2 spectrum generated from Ser-1 and Glu-4 modified proteoform of PLA2 protein family, and its sequence coverage. The masses of fragment peaks (red arrows) are shifted by 14 and 28 Da, which can be explained by a methylation in the two localized sites (red words) with their MIScores of 99.90%. |
| **Figure 4.** **Top-down MS/MS spectra of HCD fragmentations and sequence coverages of unmodified and modified disintegrin proteoforms in Echis carinatus sochureki venom from Iran**. A) Xtract-deconvoluted MS/MS spectrum generated from unmodified protein sequence, identified as disintegrin protein family, and localization of b and y ions in sequence, B) Xtract-deconvoluted MS2 spectrum generated from a Glu-3 modified proteoform of disintegrin protein family, and its sequence coverage. The mass shifts (+14 Da) between product ions (red arrows) can be explained by a methylation in the localized sites (red words) with the MIScores of 92%. | |

|  |
| --- |
| **Table 2.** Proteoform characterization of PLA2 toxin family by top-down mass spectrometry. The TD MS/MS spectra are searched against the UniProtKB Echis carinatus database (#40353) by using TopPIC suite software with employing the four common PTMs (methylation, acetylation, oxidation and phosphorylation) and modification identification score (MIScore). The symbol marked with star indicates modified peptide which was identified by BU approach. Graphic sequence representations were generated by Caititu software,74 and the colors indicate the structure section of helical regions (blue), hydrogen-bonded turns (yellow) and beta strands (green) within the PLA2 protein sequence (UniProtKB entry identifier #P48650). |

Furthermore, data indicate the presence of mass shifts of +42 Da upon some lysine residues of PLA2 proteoforms (Table 2, Supplemental Information Figure S8-10). The mass shift suggests a residue modification by mono-acetylation (∆M = 42.0106 Da) or by tri-methylation (∆M = 42.0470 Da). Distinguishing isobaric tri-methyl and acetyl modifications with 0.0364 Da difference requires a mass-resolving power of nearly 41,000 and a mass accuracy of 25 ppm for a 1500 Da peptide. Our BU results identified peptides with acetylated and methylated modifications of residues in PLA2 isoforms (Table 2 and Supplemental Information Table S17) with a high mass resolution (more than 100 k) and mass accuracy (less than 2 ppm). In this case, the obtained results unambiguously distinguish between acetylated or methylated peptides.

Protein methylation and acetylation are one of the most abundant functional forms of PTMs which can significantly change the structural properties of proteins and in principle can influence any cellular process.71,72 Compared with all modified amino acid sites, lysine (particularly K7 and K15) was predominately targeted for methylation and acetylation in the PLA2 proteoforms (Table 2). The occurrence of acetylated lysine in snake venom proteome has also been reported previously in the minor venom proteins of East African green (*Dendroaspis angusticeps*) and black (*D. polylepis*) mamba by Petras and colleagues.73 Beside above-mentioned PTMs, TD results revealed other types of modifications such as pyroglutamic acid formation at the N-terminus of the disintegrin proteoform (Supplemental Information Figure S12) and oxidation of methionine (as a chemical modification) in PLA2 proteoform.33 It thus seems that PTMs of toxins (particularly lysine modification) may play an important role in structural and functional properties of venom proteins. Characterizing them may therefore prove to be important for better understanding toxin activity and evolution of venom components. Furthermore, toxin neutralization through antibody binding may be influenced by PTMs on epitope’s sites and/or structure of toxin, resulting in weak therapeutic response of antivenoms.

**CONCLUDING REMARKS**

Despite much recent research efforts focusing on the characterization of snake venoms, advances in technological and methodological approaches now enable such investigations with unparalleled resolution. In this study, population TD and BU proteomics of the medically important Iranian saw-scaled viper (*Echis carinatus sochureki*) venoms were performed by applying multidimensional chromatography coupled to high-resolution mass spectrometry. To promote better comparison between the snake venoms, food habits and venom toxicity of the populations were also studied. The proteomics approaches revealed great complexity in the protein composition of all ECS venoms which was not reported before. The data also illustrates the remarkable similarity of venom compositions and toxicity, which suggest close correlation with consistent dietary preferences between the populations. In addition, the application of TD proteomics enabled the identification and characterization of PTMs present in the venom proteome, thereby unravelling a dynamic additional layer of toxin complexity. Thus, this study demonstrates the potential of a combined separation and proteomics approach as an analytical platform for the comprehensive analysis of complex proteome samples such as animal venoms.

**ASSOCIATED CONTENT**

**Supporting Information**

The following supplemental information is available free of charge at <http://pubs.acs.org>.

* Table S1: List of food habits of ECS communities in Iran. Table S2, S3, S4; Identified proteins by bottom-up MS/MS analysis of ECS crude venom from HO, SK and SB populations. Table S5, S6, S7: Identified proteins by bottom-up MS/MS analysis of SEC fractions of ECS venoms from HO, SK and SB populations. Table S8: List of confirmed proteins using top-down MS/MS analysis of ECS crude venom (mixed population). Table S9, S10, S11: List of confirmed proteins using top-down MS/MS analysis of SEC fractions of ECS venoms from HO, SK and SB populations. Table S12: List of proteoforms with four common PTMs in isolated fractions of three ECS population venoms. Table S13, S14, S15: List of all identified PrSM in isolated fractions of HO, SB, and SK population venoms. Table S16: List of all modified proteoform from all SEC fractions. Table S17: List of modified peptides of all SEC fractions using bottom-up MS/MS analysis. (.xlsx)
* Figure S1: Total ion current (TIC) and SDS-PAGE patterns of ECS venoms. Figure S2: SEC elution profiles of all ECS population venoms. Figure S3: An identified PrSM of PLA2 toxin family with one methylation site at K15. Figure S4: methylation site at C26. Figure S5: two methylation sites (Q11; E12). Figure S6: methylation site at S1. Figure S7: methylation site at K67. Figure S8: acetylation site at K7. Figure S9: two modification sites (K7; P17). Figure S10: acetylation site at K15. Figure S11: one oxidation site at K7. Figure S12: An identified PrSM of disintegrin protein family. Figure S13. PTMs identified by bottom-up proteomic approaches in PLA2 toxin family. (.pdf)

**AUTHOR INFORMATION**

**Corresponding Author**

Prof. Dr. Bernhard Spengler

E-mail: [Bernhard.Spengler@anorg.chemie.uni-giessen.de](file:///C:\Users\SpenglerUni\Library\Containers\com.apple.mail\Data\Library\Mail%20Downloads\957B760D-DF8F-49BD-9310-16B67A51FCC2\Bernhard.Spengler@anorg.chemie.uni-giessen.de)

Homepage: <http://www.uni-giessen.de/fbz/fb08/Inst/iaac/spengler>

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

Financial support by the Deutsche Forschungsgemeinschaft (DFG) (INST 162/500-1 FUGG) and by the State of Hesse through LOEWE Center DRUID (Novel Drug Targets against Poverty-Related and Neglected Tropical Infectious Diseases) is gratefully acknowledged. A.G. is thankful for the financial support of National Institute for Medical Research Development (NIMAD, Grant No. 942485) of Iran. V.G. thanks the German Academic Exchange Service (DAAD) for a doctoral fellowship (DAAD-GSSP-2015, ID91566181). N.R.C. acknowledges support from the Wellcome Trust and Royal Society via a Sir Henry Dale Fellowship (200517/Z/16/Z).

**ABBREVIATIONS**

BU: Bottom-up; AGC: automatic gain control; DDA: data dependent acquisition; DTT: dithiothreitol; IAA: Iodoacetamide; FDR: false discovery rate; HCD: higher-energy collisional dissociation; LC-MS/MS: liquid chromatography tandem mass spectrometry; PrSMs: Proteoform spectrum-matches; NSAF: normalized spectral abundance factor; PTM: post-translational modification; TD: top-down; 5'NTD: 5'-nucleotidase; PLA2: Phospholipase A2; CRISP: Cysteine-rich secretory protein; SVMP: Snake venom metalloproteinase; DIS: Disintegrin; CTL: C-type lectine; GC: Glutaminyl-peptide cyclotransferase; LAO: L-amino-acid oxidase; SVSP: Snake venom serine proteinase; RLAP: Renin-like aspartic protease; AP: Aminopeptidase; PLB: Phospholipase B; NGF: Snake venom nerve growth factor; PDE: Phosphodiesterase; HYAL: Hyaluronidase; VEGF: Vascular endothelial growth factor; BPPs/pHpG/C-NP: Bradykinin-potentiating/poly-His-poly-Gly/C-type natriuretic peptides; DPEP: Dipeptidyl peptidase; FGF: Fibroblast growth factor; AChE: Acetylcholinesterase; PPI: Peptidyl-prolyl-cis-trans isomerase; VVP: Venom vasodilator peptide.

**REFERENCES**

(1) Markland, F. S. Snake Venoms and the Hemostatic System. *Toxicon* **1998**, *36* (12), 1749–1800.

(2) Casewell, N. R.; Wüster, W.; Vonk, F. J.; Harrison, R. A.; Fry, B. G. Complex Cocktails: The Evolutionary Novelty of Venoms. *Trends Ecol. Evol.* **2013**, *28* (4), 219–229.

(3) Gutiérrez, J. M.; Calvete, J. J.; Habib, A. G.; Harrison, R. A.; Williams, D. J.; Warrell, D. A. Snakebite Envenoming. *Nat. Rev. Dis. Prim.* **2017**, *3*, 17063.

(4) Williams, S. S.; Wijesinghe, C. A.; Jayamanne, S. F.; Buckley, N. A.; Dawson, A. H.; Lalloo, D. G.; de Silva, H. J. Delayed Psychological Morbidity Associated with Snakebite Envenoming. *PLoS Negl. Trop. Dis.* **2011**, *5* (8), e1255.

(5) Chippaux, J. P. Snakebite Envenomation Turns Again into a Neglected Tropical Disease! J. Venom. Anim. *Toxins Incl. Trop. Dis.* **2017**, *23* (1), 1–2.

(6) King, G. F. Venoms as a Platform for Human Drugs: Translating Toxins into Therapeutics. *Expert Opin. Biol. Ther.* **2011**, *11* (11), 1469–1484.

(7) Gutiérrez, J. M.; Theakston, R. D. G.; Warrell, D. A. Confronting the Neglected Problem of Snake Bite Envenoming: The Need for a Global Partnership. *PLoS Med.* **2006**, *3* (6), 0727–0731.

(8) Warrell, D. A. Snake Bite. *Lancet* **2010**, *375* (9708), 77–88.

(9) Warrell, D. A.; Davidson, N. McD.; Omerod, L. D.; Pope, H. M.; Watkins, B. J.; Greenwood, B. M.; Ried, H. A. Bites by the Saw-Scaled or Carpet Viper (*Echis carinatus*): Trial of Two Specific Antivenoms. *The British Medical Journal.* **1974**, *4* (5942), 427-440.

(10) Pook, C. E.; Joger, U.; Stümpel, N.; Wüster, W. When Continents Collide: Phylogeny, Historical Biogeography and Systematics of the Medically Important Viper Genus *Echis* (Squamata: Serpentes: Viperidae). *Mol. Phylogenet. Evol.* **2009**, *53* (3), 792–807.

(11) Arnold, N.; Robinson, M.; Carranza, S. A Preliminary Analysis of Phylogenetic Relationships and Biogeography of the Dangerously Venomous Carpet Vipers, *Echis* (Squamata, Serpentes, Viperidae) Based on Mitochondrial DNA Sequences. *Amphibia-Reptilia* **2009**, *30* (2), 273–282.

(12) Safaei-mahroo, B.; Ghaffari, H.; Fahimi, H.; Broomand, S.; Yazdanian, M.; Najafi Majd, E.; Yousefkhani, S. S.; Rezazadeh, E.; Hosseinzadeh, M. S.; Nasrabadi, R.; Rajabizadeh, M.; Mashayekhi, M.; Motesharei, A.; Naderi, A.; Kazemi, A. M. The Herpetofauna of Iran: Checklist of Taxonomy, Distribution and Conservation Status. *Asian Herpetol. Res.* **2015**, *6* (4), 257–290.

(13) Rezaie-Atagholipour, M.; Ghezellou, P.; Hesni, M. A.; Dakhteh, S. M. H.; Ahmadian, H.; Vidal, N. Sea Snakes (Elapidae, Hydrophiinae) in Their Westernmost Extent: An Updated and Illustrated Checklist and Key to the Species in the Persian Gulf and Gulf of Oman. *Zookeys* **2016**, *622*, 129–164.

(14) Dadpour, B.; Shafahi, A.; Monzavi, S. M.; Zavar, A.; Afshari, R.; Khoshdel, R. Snakebite Prognostic Factors : Leading Factors of Weak Therapeutic Response Following Snakebite Envenomation. *Asia Pac J Med Toxicol* **2012**, *1* (1), 27–33.

(15) Chippaux, J. P. Guidelines for the Production, Control and Regulation of Snake Antivenom Immunoglobulins. *Biol. Aujourdhui.* **2010**, *204* (1), 87–91.

(16) Monzavi, S. M.; Afshari, R.; Khoshdel, A. R.; Mahmoudi, M.; Salarian, A. A.; Samieimanesh, F.; Shirmast, E.; Mihandoust, A. Analysis of Effectiveness of Iranian Snake Antivenom on Viper Venom Induced Effects Including Analysis of Immunologic Biomarkers in the *Echis carinatus sochureki* Envenomed Victims. *Toxicon* **2019**, *158*, 38–46.

(17) Sagheb, M. M.; Sharifian, M.; Moini, M.; Salehi, O. Clinical Features of Snake Bite in Southern Iran. *Trop. Doct.* **2011**, *41* (4), 236–237.

(18) Ali, G.; Kak, M.; Kumar, M.; Bali, S. K.; Tak, S. I.; Hassan, G.; Wadhwa, M. B. Acute Renal Failure Following *Echis carinatus* ( Saw – scaled Viper ) Envenomation. *Indian J. Nephrol.* **2004**, *14*, 177–181.

(19) Theakston, R. D. G.; Warrell, D. A. Antivenoms: A List of Hyperimmune Sera Currently Available for the Treatment of Envenoming by Bites and Stings. *Toxicon* **1991**, *29* (12), 1419–1470.

(20) Rahmani, A. H.; Jalali, A.; Alemzadeh-Ansari, M. H.; Tafazoli, M.; Rahim, F. Dosage Comparison of Snake Anti-Venom on Coagulopathy. *Iran. J. Pharm. Res.* **2014**, *13* (1), 283-289.

(21) Harrison, R. A.; Hargreaves, A.; Wagstaff, S. C.; Faragher, B.; Lalloo, D. G. Snake Envenoming: A Disease of Poverty. *PLoS Negl. Trop. Dis.* **2009**, *3* (12), e569.

(22) Casewell, N. R.; Wagstaff, S. C.; Wuster, W.; Cook, D. A. N.; Bolton, F. M. S.; King, S. I.; Pla, D.; Sanz, L.; Calvete, J. J.; Harrison, R. A. Medically Important Differences in Snake Venom Composition Are Dictated by Distinct Postgenomic Mechanisms. *Proc. Natl. Acad. Sci.* **2014**, *111* (25), 9205–9210.

(23) Zancolli, G.; Calvete, J. J.; Cardwell, M. D.; Greene, H. W.; Hayes, W. K.; Hegarty, M. J.; Herrmann, H.; Holycross, A. T.; Lannutti, D. I.; Mulley, J. F.; Sanz, L.; Travis, Z. D.; Whorley, J. R.; Wuster, C. E.; Wuster, W. When One Phenotype is not Enough: Divergent Evolutionary Trajectories Govern Venom Variation in a Widespread Rattlesnake Species. *Proc. R. Soc. B* **2019**, *286* (1898), 20182735.

(24) Chippaux, J. P.; Williams, V.; White, J. Snake Venom Variability: Methods of Study, Results and Interpretation. *Toxicon* **1991**, *29* (11), 1279–1303.

(25) Zelanis, A.; Tashima, A. K.; Rocha, M. M. T.; Furtado, M. F.; Camargo, A. C. M.; Ho, P. L.; Serrano, S. M. T. Analysis of the Ontogenetic Variation in the Venom Proteome/Peptidome of *Bothrops jararaca* Reveals Different Strategies to Deal with Prey. *J. Proteome Res.* **2010**, *9*, 2278-2291.

(26) Gutiérrez, J. M.; Sanz, L.; Flores-Díaz, M.; Figueroa, L.; Madrigal, M.; Herrera, M.; Villalta, M.; León, G.; Estrada, R.; Borges, A.; Alape-Giron, A.; Calvete, J. J. Impact of Regional Variation in *Bothrops Asper* Snake Venom on the Design of Antivenoms: Integrating Antivenomics and Neutralization Approaches. *J. Proteome Res.* **2010**, *9* (1), 564–577.

(27) Calvete, J. J. Snake Venomics – from Low-Resolution Toxin-Pattern Recognition to Toxin-Resolved Venom Proteomes with Absolute Quantification. *Expert Rev Proteomics*. **2018**, *15* (7), 555-568.

(28) Casewell, N. R.; Harrison, R. A.; Wüster, W.; Wagstaff, S. C. Comparative Venom Gland Transcriptome Surveys of the Saw-Scaled Vipers (Viperidae: *Echis*) Reveal Substantial Intra-Family Gene Diversity and Novel Venom Transcripts. *BMC Genomics* **2009**, *10*, 564.

(29) Bekker-Jensen, D. B.; Kelstrup, C. D.; Batth, T. S.; Larsen, S. C.; Haldrup, C.; Bramsen, J. B.; Sørensen, K. D.; Høyer, S.; Ørntoft, T. F.; Andersen, C. L.; Nielsen, M. L.; Olsen, V. J. An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell Syst.* **2017**, *4* (6), 587–599.e4.

(30) Meyer, B.; Papasotiriou, D. G.; Karas, M. 100% Protein Sequence Coverage: A Modern Form of Surrealism in Proteomics. *Amino Acids* **2011**, *41* (2), 291–310.

(31) Schlüter, H.; Apweiler, R.; Holzhütter, H.-G.; Jungblut, P. R. Finding One’s Way in Proteomics: A Protein Species Nomenclature*. Chem. Cent. J.* **2009**, *3* (1), 11.

(32) Smith, L. M.; Kelleher, N. L. Proteoform: A Single Term Describing Protein Complexity. *Nat. Methods* **2013**, *10* (3), 186–187.

(33) Ghezellou, P.; Garikapati, V.; Kazemi, S. M.; Strupat, K.; Ghassempour, A.; Spengler, B. A Perspective View of Top-down Proteomics in Snake Venom Research. *Rapid Commun. Mass Spectrom.* **2018**, *32*, 1–8.

(34) Melani, R. D.; Nogueira, F. C. S.; Domont, G. B. It Is Time for Top-down Venomics. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2017**, *23* (1), 1–8.

(35) Melani, R. D.; Skinner, O. S.; Fornelli, L.; Domont, G. B.; Compton, P. D.; Kelleher, N. L. Mapping Proteoforms and Protein Complexes From King Cobra Venom Using Both Denaturing and Native Top-down Proteomics. *Mol. Cell. Proteomics* **2016**, *15* (7), 2423–2434.

(36) Kjaergaard, J. A Method for Examination of Stomach Content in Live Snakes and Some Information on Feeding Habits in Common Viper (*Vipera Uberus*) in Denmark. *Nat. Jutl.* **1981**, *19*, 45–48.

(37) Reed, L. J.; Muench, H. A Simple Method of Estimating Fifty per Cent Endpoints*. Am. J. Hyg.* **1938**, *27* (3), 493–497.

(38) Ostad, N. Toxicity Testing and the Current Situation in IRAN. *Iran. J. Pharm. Res.* **2008**, *7* (1), 1–3.

(39) Nekouei, M.; Ghezellou, P.; Aliahmadi, A.; Arjmand, S.; Kiaei, M. Changes in Biophysical Characteristics of PFN1 Due to Mutation Causing Amyotrophic Lateral Sclerosis. *Metab Brain Dis* **2018**, *33* (6), 1975-1984.

(40) Dorfer, V.; Pichler, P.; Stranzl, T.; Stadlmann, J.; Taus, T.; Winkler, S.; Mechtler, K. MS Amanda, a Universal Identification Algorithm Optimized for High Accuracy Tandem Mass Spectra. *J. Proteome Res.* **2014**, *13* (8), 3679–3684.

(41) Käll, L.; Canterbury, J. D.; Weston, J.; Noble, W. S.; MacCoss, M. J. Semi-Supervised Learning for Peptide Identification from Shotgun Proteomics Datasets. *Nat. Methods* **2007**, *4* (11), 923–925.

(42) Zybailov, B.; Mosley, A. L.; Sardiu, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces Cerevisiae*. *J. Proteome Res.* **2006**, *5* (9), 2339–2347.

(43) Paoletti, A. C.; Parmely, T. J.; Tomomori-Sato, C.; Sato, S.; Zhu, D.; Conaway, R. C.; Conaway, J. W.; Florens, L.; Washburn, M. P. Quantitative Proteomic Analysis of Distinct Mammalian Mediator Complexes Using Normalized Spectral Abundance Factors. *Proc. Natl. Acad. Sci.* **2006**, *103* (50), 18928–18933.

(44) Liu, P.; O’Mara, B. W.; Warrack, B. M.; Wu, W.; Huang, Y.; Zhang, Y.; Zhao, R.; Lin, M.; Ackerman, M. S.; Hocknell, P. K.; Chen, G.; Tao, L.; Rieble, S.; Wang, J.; Wang-Iverson, D. B.; Tymiak, A. A.; Grace, M. J.; Russell, R. J. A Tris (2-Carboxyethyl) Phosphine (TCEP) Related Cleavage on Cysteine-Containing Proteins. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (5), 837–844.

(45) Kou, Q.; Xun, L.; Liu, X. TopPIC: a Software Tool for Top-Down Mass Spectrometry-based Proteoform Identification and Characterization. *Bioinform.* **2016**, *32*, 3495-3497.

(46) Perez-riverol, Y.; Csordas, A.; Bai, J.; Bernal-llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Perez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaino, J. A. The PRIDE Database and Related Tools and Resources in 2019 : Improving Support for Quantification Data. *Nucleic Acids Res.* **2019**, *47*, 442–450.

(47) Nejati, J.; Mozafari, E.; Saghafipour, A.; Kiyani, M. Scorpion fauna and epidemiological aspects of scorpionism in southeastern Iran. *Asian Pac J Trop Biomed.* **2014**, *4*, 217-221.

(48) Mirshamsi, O.; Sari, A.; Hosseinie, S. History of study and checklist of the scorpion fauna (Arachnida: Scorpiones) of Iran. *Prog. Biol. Sci.* **2011**, *1*, 16-28.

(49) Barlow, A.; Pook, C. E.; Harrison, R. A; Wüster, W. Coevolution of Diet and Prey-Specific Venom Activity Supports the Role of Selection in Snake Venom Evolution. *Proc. Biol. Sci.* **2009**, *276* (1666), 2443–2449.

(50) Richards, D. P.; Barlow, A.; Wüster, W. Venom Lethality and Diet: Differential Responses of Natural Prey and Model Organisms to the Venom of the Saw-Scaled Vipers (*Echis*). *Toxicon* **2012**, *59* (1), 110–116.

(51) Hashmi, S. U.; Alvi, A.; Munir, I.; Perveen, M.; Fazal, A.; Jackson, T. N. W.; Ali, S. A. Functional Venomics of the Big-4 Snakes of Pakistan. *Toxicon* **2020**, 179, 60–71.

(52) Patra, A.; Kalita, B.; Chanda, A.; Mukherjee, A. K. Proteomics and Antivenomics of *Echis carinatus carinatus* Venom: Correlation with Pharmacological Properties and Pathophysiology of Envenomation. *Sci. Rep.* **2017**, 7 (1), 1–17.

(53) Bhatia, S.; Vasudevan, K. Comparative Proteomics of Geographically Distinct Saw-Scaled Viper (*Echis carinatus*) Venoms from India. *Toxicon X* **2020**, 7, 100048.

(54) Lomonte, B.; Rangel, J. Snake Venom Lys49 myotoxins: from Phospholipase A2 to non-enzymatic Membrane Disruptors. *Toxicon* **2012**, *60*, 520-530.

(55) Tan, N. H.; Fry, B. G.; Sungar, K.; Jackson, T. N. W.; Reeks, T.; Fung, S. Y. L-amino Acid Oxidase Enzymes. In Venomous Reptiles and Their Toxins : Evolution, Pathophysiology, and Biodiscovery; Fry, B. G., Ed.; *Oxford University Press*: New York, **2015**; pp 291-298.

(56) Du, X. Y.; Clemetson, K. J. Reptile C-type Lectins. In Handbook of Venoms and Toxins of Reptiles; Mackessy, S. P., Ed.; *CRC Press*: Boca Raton, **2009**; pp 359-375.

(57) Lecht, S.; Chiaverelli, R. A.; Gerstenhaber, J.; Calvete, J. J.; Lazarovici, P.; Casewell, N. R.; Harrison, R.; Lelkes, P. I.; Marcinkiewicz, C. Anti-Angiogenic Activities of Snake Venom CRISP Isolated from Echis carinatus sochureki. Biochim. Biophys. *Acta - Gen. Subj.* **2015**, *1850* (6), 1169–1179.

(58) Lodovicho, M. E.; Costa, T. R.; Bernardes, C. P.; Menaldo, D. L.; Zoccal, K. F.; Carone, S. E.; Rosa, J. C.; Pucca, M. B.; Cerni, F. A.; Arantes, E. C.; Tytgat, J.; Faccioli, L. H.; Pereira-Crott, L. S.; Sampaio, S. V. Investigating Possible Biological Targets of Bj-CRP, the First Cysteine-Rich Secretory Protein (CRISP) Isolated from *Bothrops Jararaca* Snake Venom. *Toxicol. Lett.* **2017**, *265*, 156–169.

(59) Kunalan, S.; Othman, I.; Syed Hassan, S.; Hodgson, W. C. Proteomic Characterization of Two Medically Important Malaysian Snake Venoms, *Calloselasma Rhodostoma* (Malayan Pit Viper) and *Ophiophagus Hannah* (King Cobra). *Toxins* **2018**, *10* (11), 434.

(60) Morgan, C. P.; Insall, R.; Haynes, L.; Cockcroft, S. Identification of Phospholipase B from *Dictyostelium Discoideum* Reveals a New Lipase Family Present in Mammals, Flies and Nematodes, but Not Yeast. *Biochem. J.* **2004**, *382*, 441–449.

(61) Bernheimer, A. W.; Linder, R.; Weinstein, S. A.; Kim, K. S. Isolation and Characterization of a Phospholipase B from Venom of Collett’s Snake, *Pseudechis Colletti*. *Toxicon* **1987**, *25* (5), 547–554.

(62) Pawlak, J.; Manjunatha Kini, R. Snake Venom Glutaminyl Cyclase. *Toxicon* **2006**, *48* (3), 278–286.

(63) Wang, Y. M.; Huang, K. F.; Tsai, I. H. Snake Venom Glutaminyl Cyclases: Purification, Cloning, Kinetic Study, Recombinant Expression, and Comparison with the Human Enzyme. *Toxicon* **2014**, *86*, 40–50.

(64) Fry, B. G. Lesser-Known or Putative Reptile Toxins. In Venomous reptiles and their toxins: evolution, pathophysiology, and biodiscovery; *Oxford University Press*: New York, **2015**; pp 364–407.

(65) Creer, S.; Malhotra, A.; Thorpe, R. S.; Stöcklin, R. S.; Favreau, P. S.; Hao Chou, W. S. Genetic and Ecological Correlates of Intraspecific Variation in Pitviper Venom Composition Detected Using Matrix-Assisted Laser Desorption Time-of-Flight Mass Spectrometry (MALDI-TOF-MS) and Isoelectric Focusing. *J. Mol. Evol.* **2003**, *56* (3), 317–329.

(66) Daltry, J. C.; Wuster, W.; Thorpe, R. S. Diet and Snake Venom Evolution. *Nature* **1996**, *379* (6565), 537–540.

(67) Andersen, J. S.; Wilkinson, C. J.; Mayor, T.; Mortensen, P.; Nigg, E. A.; Mann, M. Proteomic Characterization of the Human Centrosome by Protein Correlation Profiling. *Nature* **2003**, *426*, 570.

(68) Guerrier, L.; Thulasiraman, V.; Castagna, A.; Fortis, F.; Lin, S.; Lomas, L.; Righetti, P. G.; Boschetti, E. Reducing Protein Concentration Range of Biological Samples Using Solid-Phase Ligand Libraries. *J. Chromatogr. B* **2006**, *833* (1), 33-40.

(69) Lu, X.; Zhu, H. Tube-Gel Digestion: a Novel Proteomic Approach for High Throughput Analysis of Membrane Proteins. *Mol. Cell. Proteomics* **2005**, *4* (12), 1948-1958.

(70) Shibata, H.; Chijiwa, T.; Oda-Ueda, N.; Nakamura, H.; Yamaguchi, K.; Hattori, S.; Matsubara, K.; Matsuda, Y.; Yamashita, A.; Isomoto, A.; Mori, K.; Tashiro, K.; Kuhara, S.; Yamasaki, S.; Fujie, M.; Goto, H.; Koyanagi, R.; Takeuchi, T.; Fukumaki, Y.; Ohno, M.; Shoguchi, E.; Hisata, K.; Satoh, N.; Ogawa, T. The Habu Genome Reveals Accelerated Evolution of Venom Protein Genes. *Sci. Rep.* **2018**, *8* (1), 11300.

(71) Vonk, F. J.; Casewell, N. R.; Henkel, C. V; Heimberg, A. M.; Jansen, H. J.; McCleary, R. J. R.; Kerkkamp, H. M. E.; Vos, R. A.; Guerreiro, I.; Calvete, J. J.; Wüster, W.; Woods, A. E.; Logan, J. M.; Harrison, R. A.; Castoe, T. A.; Jason de Koning, A. P.; Pollock, D. D.; Yandell, M.; Calderon, D.; Renjifo, C.; Currier, R. B.; Salgado, D.; Pla, D.; Sanz, L.; Hyder, A. S.; Ribeiro, J. M. C.; Arntzen, J. W.; van den Thillart, G. E.; Boetzer, M.; Pirovano, W.; Dirks, R. P.; Spaink, H. P.; Duboule, D.; McGlinn, E.; Kini, R. M.; Richardson, M. K.. The King Cobra Genome Reveals Dynamic Gene Evolution and Adaptation in the Snake Venom System. *Proc. Natl. Acad. Sci.* **2013**, *110* (51), 20651–20656.

(72) Kou, Q.; Wu, S.; Liu, X. Systematic Evaluation of Protein Sequence Filtering Algorithms for Proteoform Identification Using Top-Down Mass Spectrometry. *Proteomics* **2018**, *18* (3–4), 1–12.

(73) Kou, Q.; Zhu, B.; Wu, S.; Ansong, C.; Tolić, N.; Paša-Tolić, L.; Liu, X. Characterization of Proteoforms with Unknown Post-Translational Modifications Using the MIScore. *J. Proteome Res.* **2016**, *15* (8), 2422–2432.

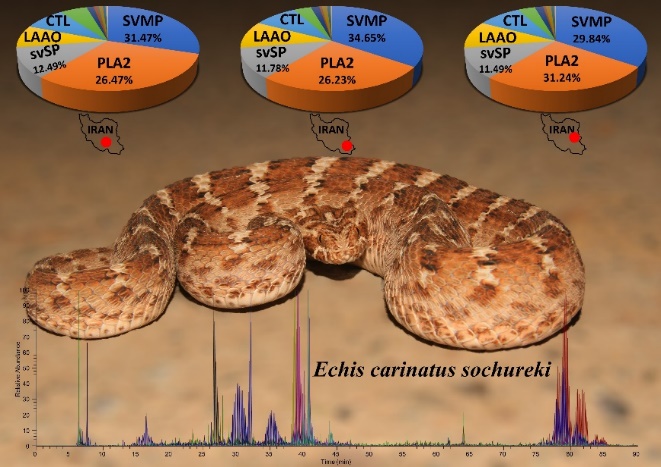
(74) Verdin, E.; Ott, M. 50 Years of Protein Acetylation: From Gene Regulation to Epigenetics, Metabolism and Beyond. *Nat. Rev. Mol. Cell Biol.* **2014**, *16*, 258.

(75) Murn, J.; Shi, Y. The Winding Path of Protein Methylation Research: Milestones and New Frontiers. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 517.

(76) Petras, D.; Heiss, P.; Harrison, R. A.; Süssmuth, R. D.; Calvete, J. J. Top-down Venomics of the East African Green Mamba, *Dendroaspis Angusticeps*, and the Black Mamba, *Dendroaspis Polylepis*, Highlight the Complexity of Their Toxin Arsenals. *J. Proteomics* **2016**, *146*, 148–164.

(77) Carvalho, P. C.; Junqueira, M.; Valente, R. H.; Domont, G. B. Caititu: a Tool to Graphically Represent Peptide Sequence Coverage and Domain Distributioin. *J. Proteomics* **2008**, *71* (4), 486-489.

**For TOC only**

****