JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

MASS SPECTROMETRY-BASED MOLECULAR PROFILING OF SNAKE VENOM SYSTEM: PROTEOMICS, LIPIDOMICS AND MASS SPECTROMETRY IMAGING

Cumulative Dissertation

by

Parviz Ghezellou

prepared at the
Institute of Inorganic and Analytical Chemistry

for the Degree of
Doctor rerum naturalium (Dr. rer. nat.)

Justus Liebig University Giessen
Giessen, 2021

Advisor and first referee: Prof. Dr. Bernhard Spengler
Second referee: Prof. Dr.-Ing. Peter Czermak

Day of oral exam: 24.11.2021

TABLE OF CONTENTS

Page

PUBLICATIONS

1) **A Perspective View of Top-Down Proteomics in Snake Venom Research**

   Parviz Ghezellou, Vannuruswamy Garikapati, Seyed Mahdi Kazemi, Kerstin Strupat, Alireza Ghassempour, and Bernhard Spengler

2) **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 Mahdi Kazemi, Alireza Ghassempour, and Bernhard Spengler

3) **Venom gland mass spectrometry imaging of saw-scaled viper, *Echis carinatus sochureki*, at high lateral resolution***

   Parviz Ghezellou, Sven Heiles, Patrik Kadesch, Alireza Ghassempour, and Bernhard Spengler

* Article used for the chapter II and III of this thesis.

LIST of ABBREVIATIONS

AGC	automatic gain control
AP	atmospheric pressure
CE	capillary electrophoresis
CID	collision-induced dissociation
DHB	2,5-dihydroxybenzoic acid
DDA	data-dependent acquisition
DIA	data-independent acquisition
ECD	electron capture dissociation
ESI	electrospray ionization
GC	gas chromatography
HCD	higher-energy collisional dissociation
HPLC	high-performance liquid chromatography
LC-MS/MS	liquid chromatography coupled with tandem mass spectrometry
MALDI	matrix-assisted laser desorption/ionization
MS	mass spectrometry
MSI	mass spectrometry imaging
m/z	mass-to-charge-number ratio
NCE	normalized collision energy
PC	glycerophosphatidylcholine
PE	glycerophosphatidylethanolamine
pEKW	pyroglutamate-lysine-tryptophan
PLA2	phospholipase A2
PRM	parallel-reaction monitoring
PSM	peptide-spectrum match
QIT	quadrupole ion trap
SVMP	snake venom metalloprotease
TOF	time of flight
TIC	total ion current

ABSTRACT

Animal venoms are rich natural sources of molecules with a broad spectrum of biological and pharmacological activities. Their biochemical and toxic properties have fascinated humankind for epochs. While these toxins have primarily evolved to facilitate prey capture and defense against enemies, they have potential value for translation into human therapeutics. Ten FDA-approved toxin-based medications are currently marketed according to the specificity and potency for particular molecular targets. Continuous improvement and development of bioanalytical techniques, particularly mass spectrometry, enable accurate qualitative and quantitative determination and even spatial localization of biomolecules in various biological systems. These technological advances have significantly increased our knowledge of the structure and function of venom components and their biological interactions, recognizing their potential benefits, e.g., novel biomedical and diagnostic tools.

Here, we performed different studies related to the medically important saw-scaled or carpet vipers (genus *Echis*), which are considered to cause higher global snakebite mortality than any other snake. Among all species of *Echis* genus, *Echis carinatus sochureki* (ECS) is a widely distributed snake species. The species is also found across the thirteen provinces of Iran, where it is assumed to be responsible for most snakebite envenomings. We collected the Iranian specimens of ECS from three different geographically distinct populations, investigated food habits, and performed toxicity assessment and venom proteome profiling to understand the viper life better. Our results show that the prey items most commonly found in all populations were arthropods, with scorpions from the family *Buthidae* particularly well represented. LD_{50} (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 similar venom compositions among the different populations. By combining all proteomics data, we identified 22 protein families, 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

ABSTRACT

(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.

Moreover, we employed autofocusing atmospheric-pressure scanning microprobe matrix-assisted laser desorption/ionization (AP-SMALDI) mass spectrometry imaging (MSI) to investigate endogenous biomolecular localizations and distribution patterns in the venom glands of ECS. For this reason, fresh-freezing and formalin-fixating sample preparations were tested for the gland to obtain data from the morphologically "intact state" of the gland sections. Subsequently, MSI was conducted with 12 μm pixel resolution for both types of preparations, and the lateral distributions of the metabolites were identified. Experiments revealed that lipids belonging to the classes of PC (phosphatidylcholines), SM (sphingomyelins), PE (phosphatidylethanolamines), PS (phosphatidylserines), PA (phosphatidic acids), and TG (triglycerides) are present in the venom gland. PC (32:0) and SM (36:1) were found to be specifically located in the areas where cells are present. The snake venom metalloprotease inhibitor pEKW (m/z 444.2233) was identified in the venom by top-down LC−MS/MS and localized by AP-SMALDI MSI in the gland across secretory epithelial cells. The peptide can inhibit the venom's enzymatic activity during long-term storage within the venom gland and thus protect its tissue. Finally, with a high degree of spectral similarities, we concluded that formalin-fixed tissue, in addition to its high ability to preserve tissue morphology, can be considered as an alternative method to fresh-frozen tissue in the case of lipid and peptide MS imaging in venom gland tissues.

ZUSAMMENFASSUNG

Tiergifte sind reiche natürliche Quellen von Molekülen mit einem breiten Spektrum an biologischen und pharmakologischen Aktivitäten. Ihre biochemischen und toxischen Eigenschaften faszinieren die Menschheit seit Epochen. Obwohl diese Toxine in erster Linie entwickelt wurden, um das Fangen von Beute und die Abwehr von Feinden zu erleichtern, haben sie einen potentiellen Wert für die Übertragung in menschliche Therapeutika. Zehn von der FDA zugelassene Medikamente auf Toxinbasis werden derzeit entsprechend der Spezifität und Wirksamkeit für bestimmte molekulare Ziele vermarktet. Die kontinuierliche Verbesserung und Weiterentwicklung bioanalytischer Methoden, insbesondere der Massenspektrometrie, ermöglichen eine genaue qualitative und quantitative Bestimmung und sogar die räumliche Lokalisierung von Biomolekülen in verschiedenen biologischen Systemen. Diese technologischen Fortschritte haben unser Wissen über die Struktur und Funktion von Giftkomponenten und ihre biologischen Wechselwirkungen erheblich erweitert und ihre potenziellen Vorteile identifiziert, z. B. als neuartige biomedizinische und diagnostische Werkzeuge.

In dieser Arbeit haben wir verschiedene Studien zu den medizinisch wichtigen Säge- oder Teppichvipern (Gattung *Echis*) durchgeführt, von denen angenommen wird, dass sie eine höhere weltweite Sterblichkeit durch Schlangenbisse verursachen als jede andere Schlangenartgattung. Innerhalb der Gattung *Echis* ist *Echis carinatus sochureki* (ECS) eine weit verbreitete Schlangenart. Die Art kommt auch in den dreizehn Provinzen des Iran vor, wo angenommen wird, dass sie für die meisten Schlangenbissvergiftungen verantwortlich ist. Wir sammelten die iranischen ECS-Exemplare aus drei verschiedenen geografisch unterschiedlichen Populationen, untersuchten die Ernährungsgewohnheiten und führten eine Toxizitätsbewertung und ein Gift-Proteom-Profiling durch, um das Leben der Viper besser zu verstehen. Unsere Ergebnisse zeigen, dass die in allen Populationen am häufigsten gefundenen Beutetiere Arthropoden waren, wobei Skorpione aus der Familie Buthidae besonders stark vertreten waren. Die LD50-Werte (Median Lethal Dose) des Rohgifts zeigen sehr ähnliche Gifttoxizitäten bei Säugetieren. In Übereinstimmung mit diesem Ergebnis zeigte die Giftcharakterisierung mittels Top-Down- und Bottom-Up-Proteomik, die sowohl auf Rohgifte als auch auf chromatographische Fraktionen mit Größenausschluss-Chromatographie angewendet wurde, sehr ähnliche

ZUSAMMENFASSUNG

Giftzusammensetzungen zwischen den verschiedenen Populationen. Durch die Kombination aller Proteomikdaten identifizierten wir 22 Proteinfamilien, darunter die am häufigsten vorkommenden Schlangengift-Metalloproteinasen (SVMPs, 29-34%); Phospholipase A2 (PLA2s, 26-31%); Schlangengift-Serinproteinasen (SVSPs, 11–12%); l-Aminosäureoxidasen (LAOs, 8–11%), c-Typ-Lectine/Lectin-like (CTLs, 7–9%) Proteinfamilien und viele neu entdeckte, z. B. Renin-ähnliche Asparagin-Proteasen (RLAPs), Fibroblasten-Wachstumsfaktoren (FGFs), Peptidyl-Prolyl-cis-trans-Isomerasen (PPIs) und Gift-Vasodilator-Peptide (VVPs). Darüber hinaus identifizierten und charakterisierten wir methylierte, acetylierte und oxidierte Proteoformen, die sich auf die PLA2- und Disintegrin-Toxin-Familien beziehen, sowie die Position ihrer Modifikationen. Es scheint daher, dass posttranslationale Modifikationen (PTMs) von Toxinen, insbesondere von Ziellysinresten, eine wesentliche Rolle bei den strukturellen und funktionellen Eigenschaften von Giftproteinen spielen und die therapeutische Reaktion von Gegengiften beeinflussen könnten, die in zukünftigen Studien untersucht werden sollen.

Darüber hinaus verwendeten wir bildgebende Atmosphärendruck-rasternde-Mikrosonden-Matrix-unterstütze Laserdesorptions/-ionisations-Massenspektrometrie (AP-SMALDI MSI), um endogene biomolekulare Lokalisierungen und Verteilungsmuster in den Giftdrüsen von ECS zu untersuchen. Aus diesem Grund wurden Frischgefrier- und Formalinfixierungs-Probenpräparationen für die Drüse getestet, um Daten über den morphologisch "intakten Zustand" der Drüsenschnitte zu erhalten. Anschließend wurde für beide Präparatetypen AP-SMALDI MSI mit 12 µm Pixelauflösung durchgeführt und die laterale Verteilung der Metaboliten identifiziert. Experimente zeigten, dass in der Giftdrüse Lipide der Klassen PC, SM, PE, PS, PA und TG vorhanden sind. Es wurde festgestellt, dass PC (32:0) und SM (36:1) spezifisch in den Bereichen lokalisiert sind, in denen Gewebezellen vorhanden sind. Der Schlangengift-Metalloprotease-Inhibitor pEKW (m/z 444.2233) wurde im Gift durch top-down LC-MS/MS identifiziert und durch AP-SMALDI-MSI in der Drüse in sekretorischen Epithelzellen nachgewiesen. Das Peptid kann die enzymatische Aktivität des Giftes während der Langzeitlagerung in der Giftdrüse hemmen und so dessen Gewebe zu schützen. Schließlich kamen wir aufgrund des hohen Grades an spektralen Ähnlichkeiten zwischen formalinfixiertem Gewebe zu frisch

gefrorenen Gewebe zu dem Schluss, dass Formalinfixierung von Gewebe, zusätzlich zu ihrer positiven Eigenschaft, die Gewebemorphologie zu erhalten, als wertvolle alternative Methode für die bildgebende Lipid- und Peptid-Massenspektrometrie von Giftdrüsengewebe in Betracht gezogen werden kann.

1. INTRODUCTION

1.1. A Brief Prologue about Mass Spectrometry

Mass spectrometry (MS) serves as a unique analytical tool to measure the mass-to-charge-number ratios (*m/z*) of ions for determining their molecular weight (MW). It dates from the early 1900s following the studies of physicists Wilhelm Wien and Joseph John Thomson on cathode rays to measure the masses of charged atoms in the gas phase.[1,2] Afterwards, other physicists such as Alfred O. C. Nier upgraded the MS instrumentation to improve the accuracy of measuring the mass and relative abundance of the elements and their isotopes, which spread the application of MS in many different scientific fields beyond the realm of physics. The methodological experiments of three chemists, Fred McLafferty, Klaus Biemann, and Carl Djerassi, regarding fragmentation processes of various classes of organic molecules, enabled chemists to determine the structure of unknown molecules by mass spectrometry, paving the way for modern biological MS.[3] In the 1980s, small organic molecules were routinely analyzed by MS. However, the analysis of macromolecules, e.g., proteins, nucleic acids and complex carbohydrates, was still challenging at that time because ionization relied on collisions in the gas phase between the analyte and the charged particle, and scientists had not yet figured out how to transfer large molecules into the gas phase without severe fragmentation and decomposition.[3] Later, in 1988, two powerful techniques (Figure 1A), electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), were introduced almost at the same time by John Fenn[4,5] and Franz Hillenkamp/Michael Karas,[6,7] respectively, to produce intact macromolecular ions in the gas phase. Briefly, in ESI, by employing a robust electric field under atmospheric pressure to an analyte-containing liquid in an open capillary tube, charge accumulation at the liquid surface results in formation of highly charged droplets. The process involves solvent evaporation, which happens when the released droplets pass through a curtain of heated inert gas or when the capillary is heated for efficient desolvation.[8] In MALDI instead, the analyte molecules are embedded in an energy-absorbing crystalline matrix and ionized by being protonated or deprotonated via pulsed laser irradiation under proper pulse energy and wavelength.[8]

Mass spectrometry has improved dramatically since the pioneering work of scientists. Numerous Nobel Prizes were awarded for design, methodology, and application of mass

spectrometry as mentioned in the following (www.nobelprize.org), enabling MS to become a remarkable method employed in many scientific disciplines.

- 1906   J.J. Thomson is awarded the Nobel Prize in Physics; "in recognition of the great merits of his theoretical and experimental investigations on the conduction of electricity by gases."

- 1922   Francis W. Aston is awarded the Nobel Prize in Chemistry; "for his discovery, by means of his mass spectrograph, of isotopes, in a large number of non-radioactive elements, and for his enunciation of the whole-number rule."

- 1939   Ernest O. Lawrence is awarded the Nobel Prize in Physics; "for the invention and development of the cyclotron and for results obtained with it, especially with regard to artificial radioactive elements."

- 1989   Wolfgang Paul is awarded the Nobel Prize in Physics; "for the development of the ion trap technique."

- 2002   John B. Fenn and Koichi Tanaka are awarded the Nobel Prize in chemistry; "for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules"

In general, the process of measuring *m/z* of ions by MS is a simple basic plan that includes three steps; 1) conversion of molecules into gas-phase ions by an ionization source; 2) ion separation by their *m/z* values via magnetic or electric fields within the mass analyzer; and 3) detection of the separated ions as electric signals equal in intensities to their abundance at their respective *m/z* values.[8] Currently, the two soft ionization techniques, ESI and MALDI, are able to ionize the most labile biological compounds, and are still the preferred methods for analysis of biomolecules in the field of biology and medicine. The easier online coupling of ESI to separation techniques such as High-Performance Liquid Chromatography (HPLC), has made the method a central part of biological MS applications. However, a remarkable and recent MALDI methodology visualizes the spatial distribution of molecules in mass spectrometry imaging (MSI) studies[9], while another MALDI method is employed for routine rapid detection of microorganisms in clinical or industrial microbiology laboratories.[10] While the ionization process determines the detectable classes of substances, the combination of MS analyzers and detectors determines the analytical quality and reliability conclusively.

Nowadays, mass spectrometry is recognized as a powerful analytical tool and technique, allowing for a wide range of applications in, e.g., environmental, food, anti-doping, biopharmaceutical, forensic, intact-protein, and even virus analysis. While MS instrumentation has benefitted from various innovations in ionization, separation, and data processing, it has undergone several fundamental changes regarding mass analyzer novelties and instrument architecture over time. Regarding the physics of mass analysis, several generic analyzers (beam-type and trapping-type mass analyzers) were introduced and widely used for analytical mass spectrometry, including time-of-flight (TOF), linear quadrupole (Q), ion trap (IT), Fourier transform ion cyclotron resonance (FT-ICR) and Orbitrap instruments.[21] The awareness of the complexity of biological samples resulting from MS studies in proteomics and metabolomics, has also revealed limitations of analytical tools. Reliable identification of metabolite and protein contents in complex biological samples require high mass resolution, mass accuracy, sensitivity and dynamic range of state-of-the-art mass spectrometers.[11] Besides, tandem mass spectrometry (MS/MS) is also crucial for determining and elucidating unknown and known molecular structures along with protein and peptide sequencing and post-translational modification (PTM) analysis. Analysis of both, analytes (MS^1) and their fragments (MS^n) with accurate details therefore suggested the combination of several MS analyzers. These bioanalytical demands are addressed by well-designed hybrid MS analyzers, including the most popular ones such as triple quadrupole (QqQ), quadrupole/time-of-flight (QTOF), linear quadrupole ion trap/Fourier transform ion cyclotron resonance (LIT/FT-ICR) and quadrupole-Orbitrap (Q-Orbitrap) instruments. However, each mass analyzer has its unique characteristics and applications and its benefits and limitations.[12] Therefore, the choice of mass analyzers required for the application is affected by instrumental costs and the desired performance. Despite the initial use of QTOF and QqQ for rapid sequencing of peptides, they were used later mostly in other research areas, such as analyzing metabolites, nucleic acids, and glycoproteins. While QTOF has the high compound fragmentation capability of quadrupole filters and nearly the high mass resolution capacity of TOF instruments, its sensitivity is lower than that of the other hybrid analyzers due to the non-continuous delivery of ions into the detector.[13] Many of the limitations were eliminated with the LIT/FT-ICR MS hybrid introduction. It is the most advanced mass analyzer based

on the higher mass resolution, sensitivity, and wide dynamic range with sub parts per million (ppm) mass accuracy.[14] However, its large and heavy superconducting magnet, its complexity, and its high price prevent it from becoming a popular laboratory instrument. The technological gaps have been mostly filled by the introduction of the Q-Orbitrap hybrid mass spectrometers, which are more convenient in size, practicability and price. With its high mass resolution (up to 480,000 at m/z 200), excellent mass accuracy (<1 ppm), small size (compact benchtop unit), and relatively lower price compared to FT-ICR instruments, this type of hybrid MS has proven to be an essential analytical tool with a wide range of applications.[15] Besides, high-speed acquisition offered by Orbitrap mass systems provides the conditions needed for liquid chromatography (LC)-compatible MS, ensuring sufficient data records (full scan and data-dependent/independent MS/MS scans) on an LC timescale across narrow chromatographic peaks.[16] The combination of high-resolution accurate-mass (HRAM) analysis with higher acquisition rates also boosted the number of identifications and more in-depth insight into complex mixed spectra.[15]

## 1.2. Mass Spectrometry-based Omics Technologies

Living organisms, as dynamic and complex systems, are composed of more than a trillion cells.[17] Each cell has a large variety of organelles and structures which regulate with an enormous number of molecules. It is assumed that there are more than a hundred thousand various proteins, over a billion nitrogenous base pairs, and a highly complex network of metabolites.[11] The development of omics-based technologies (genomics, transcriptomics, proteomics, and metabolomics) aims to identify, quantify and characterize the biological molecules engaged in the structure, function, and dynamics of a cell, tissue, or organism at a specific point in time. Recent prominent MS technology developments meet the criteria to face a series of challenging tasks such as high sensitivity, selectivity, resolution, and throughput to analyze complex biological samples, making MS-omics one of the highest priorities in the application of Systems Biology approaches. It may provide detection and quantification of thousands of proteins and biologically active metabolites from down to ultra-trace amounts of samples by implementing ''global'' (hypothesis-free) or ''targeted'' (hypothesis driven) modes.[18] It is assumed that the high achievement of MS technology, linked with automatic data handling, soon be successful in the demand for a better

understanding of human diseases, resulting in new molecular biomarkers, hence affecting drug target development and therapies.

Given the complexity of the molecular contents of biological specimens, different separation methods such as capillary electrophoresis (CE), gas chromatography (GC), and HPLC are usually used to de-complex samples before MS analysis.[11,18] CE as an electrokinetic method efficiently separates analytes during the migration of ions through the submillimeter diameter capillaries under the influence of an electric field.[19,20] In GC, volatile compounds are separated by injecting a gaseous or liquid sample into the mobile phase and passing the carrier gas (e.g., hydrogen, nitrogen, or helium) through a stationary phase (microscopic layer of viscous liquid on the column's inner surface), controlled by temperature.[18] HPLC is a technique used to separate analytes, depending on pumps to pass a pressurized liquid solvent carrying the sample through a column filled with a solid adsorbent material. Based on the chemical structure of the analyte, the molecules interact differently with the adsorbent material while passing the stationary phase. Currently, liquid chromatography coupled to electrospray ionization high-resolution mass spectrometry (LC-ESI-HRMS) is considered the method of choice for untargeted and targeted proteomics and metabolomics workflows.[11,18,19]

Despite the goal of almost all MS-based omics studies to identify and quantify whole molecules within biological samples, three main modes were developed and used to achieve this demand in mass spectrometers: data-dependent acquisition (DDA), targeted selected reaction monitoring (SRM), and data-independent acquisition (DIA).[22] These strategies differ in how *m/z* information of precursor and fragment ions is acquired and how the recorded data is analyzed. In DDA mode (Figure 1B), during the first step, a mass analyzer generates a full-scan mass spectrum (MS1) to determine the *m/z* values of the molecular species present in the sample, followed by fragmentation of the N most intense ions and detection of their fragment ions (MS/MS). In this setting, the MS-recorded data can be used simultaneously for downstream data analysis, allowing protein/peptide or metabolite identification and quantification in the same sample analysis. Since in this mode the criterion for precursor ion selection is signal intensity, low-abundant molecules may never be selected, and their information will not be available for identification as a consequence. In addition, since MS devotes most of its acquisition time to generate MS/MS

spectra, the signal intensity for observing MS1 may decrease, making it difficult to detect low-abundant molecules in the analyte. Common instrument configurations for DDA workflows include several Orbitrap hybrids and quadrupole time-of-flight (TOF) devices.[22] SRM is another technique known as a targeted acquisition method in which predefined precursor ions are repeatedly selected for fragmentation over time. Therefore, pre-knowledge about the targeted precursor ions and their fragmentation behavior is needed. The technique is applied to increase the quantification accuracy and reproducibility of the presence of targeted molecular species, rather than to explore new molecules within the samples. Both qualitative and quantitative knowledge is directly based on the resulting fragment ion chromatographic signals of individual precursor ion species. Different instrument types have been used for SRM studies, often carried out on triple quadrupole instruments. To improve mass resolution on the MS2 level, instruments in which the third quadrupole is replaced by a HRAM analyzer have also been used with the method, then named parallel reaction monitoring (PRM), or MS/MS[ALL], recording high-resolution MS/MS signals of predefined sets of precursor ions.[22] The DIA mode has recently been introduced, representing a combination of DDA and SRM. The mode theoretically aims to produce MS2 for all detected precursors ions, thus being able to identify more molecules at low concentrations and overcome the limitation of above techniques.[22] In DIA mode, the precursor ions within a narrow m/z window are selected and fragmented for each cycle. This mass window is then moved over the entire m/z range, orderly collecting MS/MS data from every mass and all detected precursors. Although the technique provides an attractive strategy for proteomics and metabolomics studies, figuring out the connection between the precursor ions and their fragments is a major challenge of MS data handling and interpretation.[11,18,22]
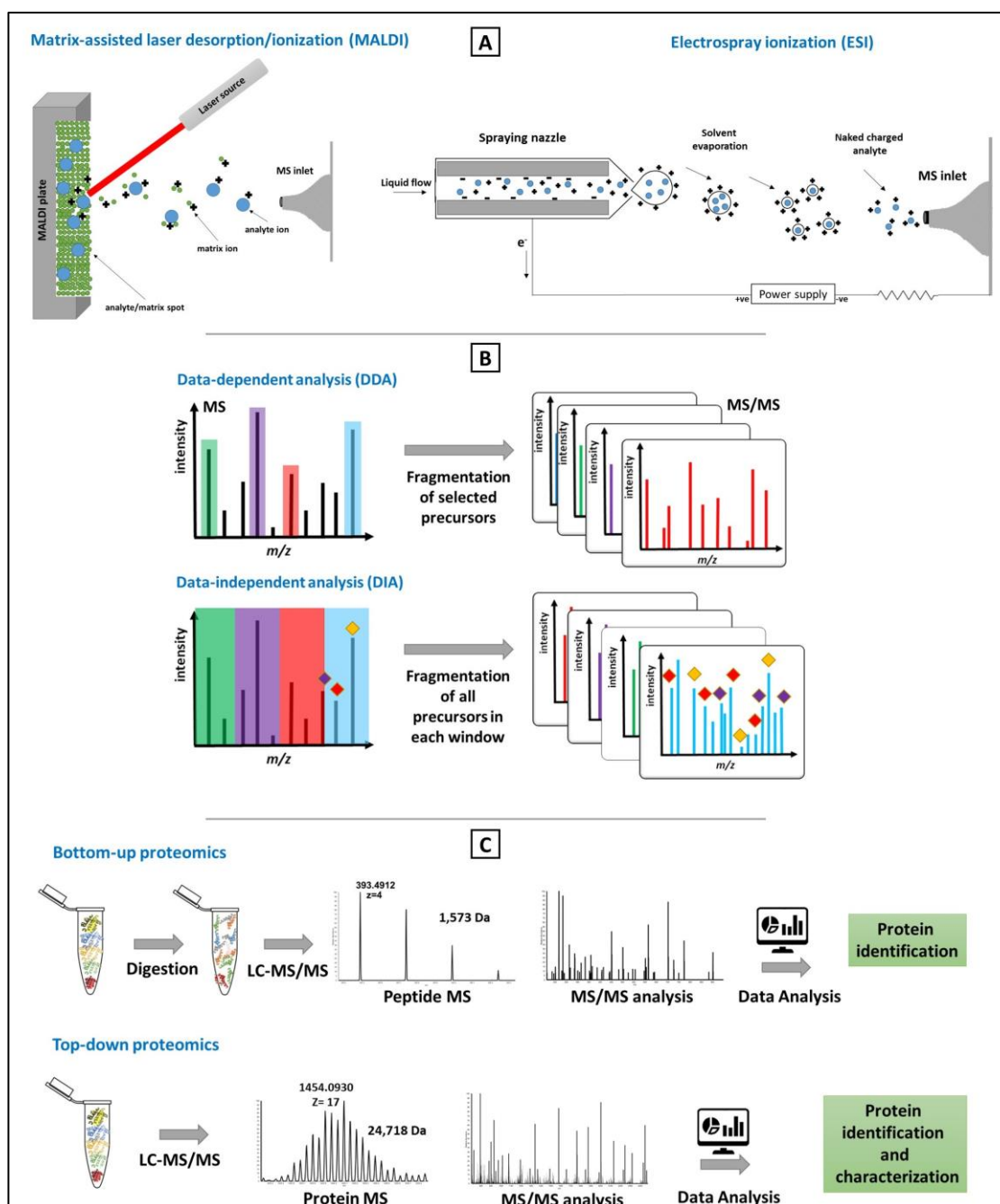
Figure 1. Schematic representation of A) ESI and MALDI ionization methods, B) data-dependent and data-independent acquisition modes, and C) bottom-up vs. top-down proteomics strategies.

1.2.1. MS-based Proteomics

The induction of ESI and MALDI in association with DNA/RNA sequence information, revolutionized the analysis of organisms' proteinous content through MS-based proteomics methods. Proteins play a vital role in organisms' lives. They are composed of peptides, a chain of amino acids derived from mRNA translation. Although DNA typically encodes the common 20 amino acids to determine proteins' sequence, translated proteins may contain more than 140 different residues due to post-translational modifications (PTMs).[23] PTMs are chemical modifications of the polypeptide chain, happening during or after translating RNA into proteins. Their presence can influence structural folding, conformational stability, and finally the biological activity of proteins. Thus, proteomes are significantly more complex than expected from the encoding gene analysis and are not easily predictable. In general, there are also fundamental changes in protein amount or the level of a special PTM upon cells' response to external or internal stimuli.[24] This reality has aided in understanding how biological systems react to change, which has driven the development of markers of change that have improved insights in medicine and biology.

The two main proteomics strategies are broadly employed for protein identification and quantification, known as bottom-up (BU) and top-down (TD) as shown in Figure 1C.[25] In the typical BU workflow, proteins are enzymatically digested into peptides which are then separated either by liquid chromatography or directly subjected for ionization via ESI or MALDI before MS measurement. The BU method is described by peptide mass fingerprinting (PMF) and tandem mass inspection. PMF is a method adopted for protein identification, relying on mass spectra acquisition, primarily by MALDI-TOF MS, to measure tryptic peptide masses and match them to the theoretical peptide masses generated from a protein or even genome sequence in, e.g., Uniport or Swissport databases. The best overlaps between the experimental and theoretical, *in silico*, peptide masses are statistically analyzed to find and report the best match. However, liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based proteomics, termed shotgun proteomics,[26] is now the leading BU proteomics approach.[27] In this technique, the protein sample is digested into peptides using proteases like trypsin which cleave proteins specifically at the C-termini of arginine and lysine residues. The obtained peptides are then separated by LC and eluted from the column right in front of the mass spectrometer in an ESI source. Intact peptide

ions are then detected based on their *m/z* values in a mass spectrometer (MS1). The peptide ions cannot be identified unequivocally by MS1 alone. However, it is possible to fragment peptides, mainly at the weakest bonds (peptide bonds), by colliding them with inert gases in the collision cell of the mass spectrometer. The generated fragment ions, which in case of peptide-bond fragmentation are conventionally called b-ions when containing the peptide's N-terminus and y-ions when containing the C-terminus, are detected in an MS/MS spectrum and used to identify the amino acid sequence by calculating the differences in *m/z* values, a process called *de novo* sequencing.[27] However, manual *de novo* peptide sequencing is typically time-consuming and not an ideal method due to often incomplete fragmentation ladders. The method is thus not practical for most experiments. Investigated samples usually belong to organisms whose probable protein sequences are available in databases. So instead of having to *de novo* sequence all the resulting MS2 spectra, one only needs to figure out the ones that match with known amino acid sequences. To achieve the goal, theoretical MS2 spectra can be created for these peptides, based on all b and y ions. By comparing the observed masses of MS1 and MS2 spectra with the theoretical spectra, one can obtain the best agreement to assign the obtained spectra. Fortunately, various algorithms for automatic data analysis have been developed based on several machine learning approaches which can reliably assign spectra to peptides/ proteins as spectrum-centric database searching.[27]

BU proteomics can help to determine the relative change of protein contents between two or more different samples and the absolute amount of each of the proteins in a mixture. There are two main approaches of performing MS quantitatively, the label-based and the label-free methods.[28] The label-based methods rely on labeling or tagging peptides with stable heavy isotopes (e.g., ^{15}N, ^{13}C, ^{2}H) either by chemical derivatization (e.g., Isobaric Tags for Relative and Absolute Quantitation [iTRAQ] and Tandem Mass Tag [TMT]) or by metabolic labeling (e.g., Stable Isotope Labeling by Amino acids in Cell culture [SILAC]).[28] Then the ratio of peak intensities between the unlabeled "light" forms of the sample and the labeled "heavy" ones represents the ratio of abundances. In contrast, label-free quantitation is based on comparing precursor ion peak intensities between multiple MS runs. It is performed using extracted ion chromatograms (XIC) and features alignment (retention time, *m/z* and charge state). It has to be noted that the quantification of proteins

with label-free methods has the capability of comparing multiplexed MS runs but is less accurate than the label-based one.[18]

BU proteomics takes advantage of digestion of proteins into small peptides, including an increased separation efficiency, appropriate ionization due to a limited number of charges, increased sample homogeneity, easy fragmentation processes, and predictable fragmentation patterns. As mentioned earlier, proteins mainly undergo various processes during their expressions (genetic variation, alternative mRNA splicing, single-nucleotide polymorphism [SNP], and post-translational modifications [PTMs]) that may alter their composition compared to their original forms.[24] Therefore, the major challenge for the BU proteomics is the limited sequence coverage and the "peptide-to-protein inference" problem which results in a lack of information and low protein sequence coverage for the identification of protein isoforms or proteoforms as well as loss of the location of PTMs.[25] Alternatively, TD proteomics methods investigate proteins in their intact state rather than measuring enzymatically digested peptides. Direct analysis of intact proteins offers the richest data, covering high protein sequence, and the ability to achieve proteoform-resolved molecular detail.

Denaturing TD proteomics experiments follow fundamentally the same process as BU proteomics, from sample preparation (e.g., protein extraction and solubility using detergent-containing buffers) to labeling and label-free methods in the case of quantification, separation of the proteome, followed by mass spectrometric analysis and data interpretation. However, in each of these steps, several issues such as protein solubility, dynamic range, proteome complexity, and data analysis still interrupt the routine use of TD proteomics.[29] Sample preparation is a crucial concern, depending on temperature, pH, salt concentration, surfactants, and elution solvents, which have to take care of before starting the experiment. For example, the conventional detergents (e.g., SDS; sodium dodecyl sulfate) are not MS-compatible, resulting in signal suppression and lack of protein detection. In addition, the complexity of proteomes, especially the number of different proteoforms, requires powerful protein separation techniques regarding the similar physicochemical properties of the isoforms. Therefore, the ultimate intention is to perform a single proteoform elution, which is hard to overtake at this moment, even by employing multidimensional separation plans. The detection and identification of intact

proteins rely also on high-performance mass spectrometers. High mass resolution (>50 K), mass accuracy (<3 ppm), and high transmission speed (>5 Hz) are critical for separating and accurately assigning spectral peaks derived from complex precursor spectra of multiple intact proteoforms or MS/MS spectra comprising hundreds of fragment ions.[29,30] Therefore, time-of-flight (TOF), Fourier transform ion cyclotron resonance (FT-ICR) and Fourier transform orbital trapping (FT-Orbitrap) mass analyzers, particularly in hybrid form, are typically used. A fragmentation (MS/MS) technique is another key requirement to cleave proteins and to localize modification sites.[30] In general, beam-type collision-induced dissociation (CID) or high-energy collision dissociation (HCD) producing b- and y-type product ions are the most common methods in proteomics experiments. Electron capture dissociation (ECD) and electron transfer dissociation (ETD) are different methods that can maintain labile PTMs by generating c- and z-type ions. Furthermore, ultraviolet photodissociation (UVPD) has been shown to generate a broader array of fragment ions (b−/y-,c−/z-,a−/x-,d-,v- and w-) which improved the sequence coverage of intact proteins and boosted the confidence of protein identification.[29] Along with qualifying intact proteins and considering the alterations of proteoforms, information at the protein expression level and modification are essential. Although label-free and labeling quantitative approaches in BU proteomics showed significant advances, quantitation of proteoforms, especially in large-scale TD measurements, requires large amounts of research and time.[30]

In recent years, with the simultaneous development of high throughput data analysis platforms, it has become possible to identify and quantify thousands of proteoforms by the denaturing TD method. However, due to technical limitations related to mass spectrometry and separation, we can now comprehensively analyze proteins with a molecular mass only below 50 kDa. Using the future generation MS (e.g., 21T FT-ICR)[31] or innovative data acquisition procedures can help address the challenges of high mass protein detection by increasing the S/N of large proteoforms in short time transients during denaturing TD experiments. An alternative solution is using native TD mass spectrometry, in which even protein assemblies as big as megaDalton such as viruses can be transferred into the gas phase and determined by MS.[29,30]

1.2.2. MS-based Metabolomics

While the identification of proteomes by proteomics techniques reveals one of the functional aspects of the cell, metabolites profiling can provide valuable information about cell physiology and biology. Metabolites represent biochemical compounds of low molecular weight (~30-1500 Da) that belong to various compound classes, such as amino acids, fatty acids, nucleotides, etc.[32] They are an intermediate or end product of cell metabolisms present in a high dynamic range of concentrations. Metabolomics is the method of analyzing the metabolome within cells, biofluids, tissues, or organisms. The workflows (targeted and untargeted metabolomics) are commonly applied to two analytical platforms to identify and quantify metabolites: nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry.[33] Since NMR as a general technique for investigating metabolites still suffers from low sensitivity and complexity of spectra, MS is more sensitive, can therefore identify more compounds at low concentrations. However, the inability to distinguish metabolite isomers is a weakness of MS technology.[33]

Selecting and preparing biological samples for metabolomics analysis significantly affects the recorded data and its quality. Therefore, full attention should be considered to the experimental design of sample collection, extraction, and metabolite storage prior to analysis. It is worth mentioning that lipidomics, as a subdivision of metabolomics, focuses on identifying alterations of concentration levels of lipid species in biological samples. It can be considered as a targeted metabolomics strategy since it involves studying a subset of specific metabolites (lipids). However, lipidomics is classified as targeted (study of specific lipids) or untargeted (global exploratory analyses) because of the complexity of lipids. Sample preparation is the only difference between a metabolomic and a lipidomic experiment.[34] In lipidomics, it is necessary to include a lipid extraction step, usually by liquid-liquid extraction or solid-phase extraction, before NMR or MS analysis.

In recent years, with the advancement of MS technology, we have witnessed the increasing growth of metabolomics technique in the area of system biology. MS-based metabolomics is currently being applied to investigate the metabolite composition and screen the molecular changes during metabolic processes in numerous disciplines, including food, environmental, plant, and toxicological sciences as well as drug development and medicine. Because of the complexity of biological samples, comprising several hundred to

thousands of metabolites, it is often required to isolate metabolites before MS acquisition. Thus, hyphenated separation techniques combined with MS have become a very effective tool for analyzing small molecules. The main chromatographic methods that are usually paired with MS and widely used are HPLC and GC.[35] Over the past decade, each of these approaches has seen tremendous growth. Such advances, along with the development of new software packages and databases, now make it possible to quantitatively analyze several hundred metabolites in automation mode. GC-MS obtains better metabolite separation than LC and generally avoids ion suppression due to performing ionization in the gas phase in the mass spectrometer. However, unlike LC, GC typically requires chemical derivation of metabolic species before GC-MS analysis.[35] Therefore, LC-MS has seen a major uptake in this field because it detects a more extensive reservoir of metabolites without the need for a chemical derivatization step. Reversed phase chromatography is commonly used to separate non-polar to medium polar molecules. In contrast, the HILIC mode (hydrophilic interaction liquid chromatography) has become the technique of choice for very strong to slightly polar metabolites.[36]

Ionization is one of the most critical steps in measuring metabolites. The degree of ionization determines the ability to detect and quantify a metabolite. ESI is the ionization technique of choice for LC-MS for several reasons. It ionizes enough molecules in the liquid phase and can be used globally for small molecules (>1000 Da) as well as for large molecules such as peptides and proteins.[11] In addition, ESI is a soft ionization technique, so it does not induce significant segmentation of molecular ions. The disadvantage of using ESI is that its ionization efficiency is adversely affected by the presence of salts, so chromatographic methods are limited to using only volatile buffers such as ammonium acetate or ammonium formate. Electron ionization (EI) is the ionization method of choice for GC-MS analysis where the gas-phase molecules are bombarded by accelerated electrons in a high-vacuum ion source.[11] The ionization condition of EI is highly energetic that may generate considerable fragmentation of the molecular ion.[37] Nevertheless, based on the extent of fragmentation, little to no molecular ion may be detected, likely making a reliable identification more difficult.[37] It is also possible to achieve a metabolomics analysis by MALDI-MS and MALDI mass spectrometry imaging (MALDI-MSI), notably for screening tissues, cells, and compartments.[9]

Sampling of cells and tissues to extract metabolites or proteins regularly involves tissue homogenization and cell lysis, eliminating quantitative and qualitative information about the distribution of biomolecules at the cell or tissue surface. However, the location of biomolecules by imaging them can provide valuable and essential information about the objects in clinical and biological studies. Recently, MS-imaging as powerful analytical techniques (e.g., MALDI-, DESI-MS, and SIMS) have provided information on the spatial distribution of metabolites and proteins/peptides by rasterizing across tissue slices and collecting a mass spectrum at each pixel of the sample.[9] The mass spectral image is a composite of all the pixels, and through computational analysis, the distribution of different metabolites can be visualized.

Metabolomics technologies generate plenty of complex data, and handling them is a significant step for identifying and quantifying metabolites. Typically, data handling in metabolomics can be grouped into two steps: raw data processing and data analysis.[35,38] The raw data processing step starts with extracting spectral information (m/z, retention time and intensity) from the actual vendor file format (e.g., RAW, WIFF and BAF) and storing it in open file format (e.g., mzML, mgf and mzData), which can be used in different operating systems and third-party software.[18] In addition to this primary step, the data processing needs more tasks, including filtering (removing effects of noise and baseline), feature detection (detect representations of measured ions from the raw signal), alignment (cluster measurements across different samples), and normalization (removes unwanted systematic variation between samples).[38] Regarding data processing, several open source software packages (e.g., MZmine and XCMS) are available, facilitating progress in metabolomic data processing.[18] The next step to data processing is peak identification by comparing the data with known metabolite libraries. The process can be made manually for smaller numbers of metabolites. However, with larger sets of metabolites, data processing software needs to include a library (e.g., NIST) or database (e.g., METLIN, HMDB, and Lipid maps) search function and assign a similarity score to each peak, finally annotating them with names of metabolites or compound classes.[38] For data analysis, the processed data can be directly subjected to univariate (e.g., t-test and ANOVA) or multivariate (e.g., PCA and PLS-DR) statistical analysis types. Furthermore, different network-based approaches such as KEGG (http://www.genome.jp/kegg/) and MetaCyc

(https://metacyc.org/) may be used to infer molecular pathways and components via integrative analysis of metabolite features.[38]

1.3. Venomous Snakes and Their Venoms

Venomous animals are widely distributed in the animal kingdom, with more than 100,000 different species.[29] The production of toxic compounds by these creatures has fascinated humankind for epochs. The venoms of these animals comprise complex mixtures of peptides and proteins with incredible biological specificities. These venom components play a crucial role in various pathophysiological processes such as apoptosis, neurotransmission, hemostasis and signal transduction.[29] Given the diversity of structures and functions of venom toxins, it is not surprising that these substances are used as pharmacological tools and as prototypes in drug development.[39,40] In the quest for treatment of diseases which do not respond to presently available therapies, venoms represent an essential, yet unexplored reservoir of bioactive compounds. There is an increasing number of examples for the identification and development of venom proteins/peptides into pharmaceuticals. According to the specificity and potency for particular molecular targets and structural architectures of venom components, there are currently six U.S. Food and Drug Administration (FDA)-approved drugs derived from venom peptides or proteins, and more than ten are in clinical trials.[40,41]

Venomous snakes, belonging to the superfamily Colubroidae, include the largest group of venomous animals. Among all snakes, vipers (family Viperidae) and elapids (family Elapidae) are the medically most important groups, collectively causing the majority of snakebite envenoming and fatalities in humans and their domestic animals.[42] There are at least 1.8–2.7 million snakebites worldwide annually, resulting in more than 100,000 deaths.[43] Therefore, and because it disproportionately affects poor people and perpetuates poverty, the World Health Organization (WHO) has recognized snakebite envenoming as a priority "neglected tropical disease".[44]

Venoms produced by these snakes contain medically significant toxins that play a key role in the pathophysiology of human victims of envenoming. The clinically most relevant effect of elapid snake (e.g., cobra, sea snake, mamba, krait and Australian elapid) envenoming is neuromuscular paralysis leading to respiratory arrest and death.[45] The

clinical features observed after envenoming by viperids are usually more complex and include local tissue damage such as oedema and blistering, dermonecrosis and myonecrosis, and also systemic alterations like haemorrhage, coagulopathy, cardio-vascular disturbances and renal damage.[43,45] It is noteworthy that there are important quantitative and qualitative differences in venom composition between and within species of venomous snakes.[46] In addition to significant geographic and ontogenetic variations in snake venom composition, individual variations have also been reported.[47] However, more closely related species of venomous snakes tend to have venoms with more similar composition than more distantly related species. The parenteral administration of antivenom (purified immunoglobulin or antibody fragments), which aims to neutralize venom-induced toxic effects and reverse pathological symptoms of envenoming, continues to be the only treatment of systemic snakebite envenoming. Better knowledge of the snake specimens used for venom collection and the composition of their venom samples could greatly improve the rational design, therapeutic effectiveness and production of antivenoms. Knowledge-driven development of new and improved antivenoms could also help to reduce the incidence and severity of adverse reactions associated with antivenom administration.

Over the past decade, significant research efforts have been made toward profiling venom proteomes using mass spectrometry (MS)-based proteomics and venom gland transcriptomic approaches.[48] 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 and MS/MS interrogation for peptide sequencing. However, while BU analysis has a high throughput, is sensitive and robust, it is predominately only capable of identifying the representative protein for each expressed gene, and does not provide information on the proteoforms, genetic variation, and PTMs, associated with the sample. 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.[29,30] Steady advances in mass spectrometry technologies have facilitated improvements in TD proteomics, enabling quick and accurate investigation of intact toxin families and their proteoforms.[29] However, this technique still has limitations in providing

full sequence coverage of large (>30 kDa) and 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, requiring the use of native MS, which is experimentally and bioinformatically challenging.[29] 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.

In addition to the study of venom components, it is necessary to note that snake venoms are produced in a pair of venom glands located directly below the snake's eyes. The morphology of the glands has been revealed by classical histology and microscopic studies. However, knowledge about the gland's cellular secretory and functional processes is still incomplete and has so far been neglected by the omics disciplines. Therefore, molecular distribution topography could help to explain the mechanisms behind tissue activation and toxin production on the cellular level. Trying to meet the goals have recently led to using MSI as an ideal imaging technique to interrogate the spatial distribution of venom components inside a few venomous animal glands, e.g., centipedes (*Thereuopoda longicornis*, *Scolopendra morsitans*, and *Ethmostigmus rubripes*),[49] honeybee (*Apis mellifera*),[50] sea anemones (*Oulactis muscosa* and *Actinia tenebrosa*),[51,52] the brown forest cobra (*Naja subfulva*),[53] and the fire ant (*Solenopsis invicta*).[54]

1.4. Study Plan

Here, we performed different studies regarding mass spectrometry-based molecular profiling of venoms and venom glands of the medically most important snakes, *Echis carinatus sochureki* or saw-scaled viper (ECS). Briefly, individual snakes were systematically identified and collected from three different populations located in the South (Hormozgan province; HO), Southeast (Sistan and Baluchestan province; SB), and East (South Khorasan province; SK) 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. Therefore, we continued our study by milking the venom of individuals and pooled their venoms within the populations, followed by an investigation

of individuals' feeding habits within each community. Subsequently, the median lethal toxicity (LD50) of each population-representative venom was evaluated by administering various dosages of crude venoms to laboratory mice. In the subsequent step, the combination of bottom-up and top-down proteomics approaches was used to identify and characterize in detail venom protein compositions and to provide an overview of the conspecific venom variation of Iranian saw-scaled viper.

Furthermore, we developed a workflow to investigate the spatial arrangement of metabolites within the venom gland tissue using mass spectrometry-based imaging. For this purpose, two sample preparation strategies, formalin-fixation and fresh-freezing, were implemented for the venom glands of ECS. The resulting tissue sections were imaged with lateral resolutions down to 12 μm, allowing localizing and identifying a bioactive metalloprotease inhibitor peptide and small metabolites from venom gland tissue. Besides, by comparing the MSI data obtained from both types of sample preparations as well as the supporting data achieved by untargeted lipidomics of the fixed tissue lipidome, we found that the formalin-fixation method, with its high ability to preserve tissue morphology, provides data regarding lipids and peptide imaging comparable to those obtained from the fresh-frozen tissues of the venom gland.

2. RESULTS and DISCUSSION

2.1. Top-Down and Bottom-Up Proteomics Guided Venom Profiling of the Iranian Population of Saw-Scaled Viper, *Echis carinatus sochureki*

The whole crude and size-exclusion-separated fractions of all three populated venoms were subjected either as peptide forms (after in-solution tryptic digestion) for BU or intact forms for TD proteomics analysis. In total, upon merging the TD and BU data, 22 protein families were identified in the venoms of ECS, including 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), phosphodiesterase (PDE), hyaluronidase (HYAL), vascular endothelial

growth factor (VEGF), acetylcholinesterase (AChE), dipeptidyl peptidase (DPEP), bradykinin-potentiating/poly-His-poly-Gly/c-type natriuretic peptides (BPPs/pHpG/C-NP), dipeptidyl peptidase (DPEP), fibroblast growth factor (FGF), peptidyl-prolyl cis-trans isomerase (PPI) and venom vasodilator peptide (VVP) protein families.
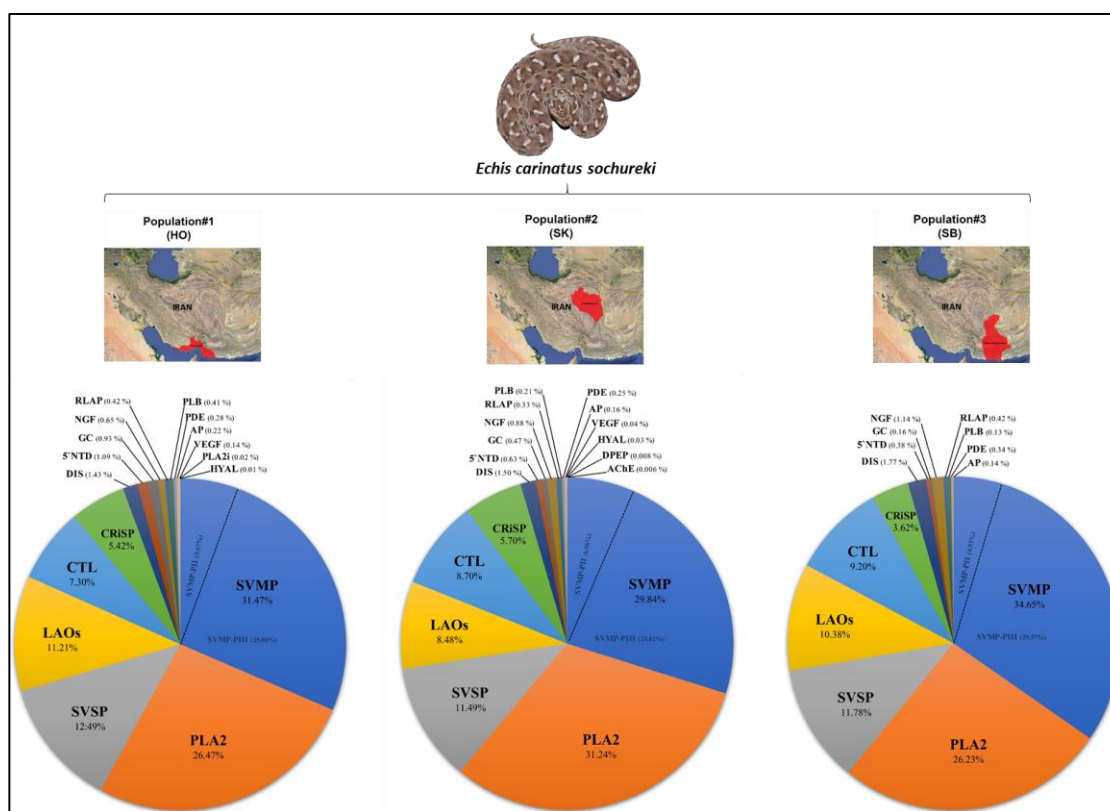
Figure 2. Overview composition and relative abundance of protein families from venom proteomes of three *E. carinatus sochureki* populations. SVMP: snake venom metalloproteinases; PLA2: group-II phospholipase A2; CTL: C-type lectin; DIS: disintegrin; SP: serine protease; LAO: L-amino oxidase; CRISP: 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; and AChE: acetylcholinesterase.

Despite the large geographical distances between each ECS population (~400 km), the relative concentration of the major protein families detected in all venom populations were highly similar as shown in Figure 2. However, each population shows some distinct properties, in terms of both quantitative and qualitative venom composition. 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. The major protein classes identified in all ECS venoms are SVMP, PLA2, SVSP, LAO and CTL toxin families. It is consistent with previous transcriptomics and proteomics analyses of ECS venom and venom gland from United Arab Emirates[55,56], and recent venomics studies of *E. carinatus* from India.[57,58] 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 snakebites by ECS in Iran,[59] whose envenoming syndrome is mainly attributed to SVMPs and PLA2. Another interesting toxin family with an even lower abundance (~5%) is CRiSP, which was recently isolated from *Bothrops jararaca* and was reported to dictate the induction of pro-inflammatory responses that provoke the production of interleukin (IL)-6, also targeting the complement system.[60,61] These records are also in good agreement with the clinical observations of Iranian ECS envenomation in human patients.[59]

In addition to those toxin families described above, our multidimensional (MD)-LC-MS/MS 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). Notably, only few of them have been reported in EC venom proteomes before and are thus functionally poorly studied so far.

It is worth noting that the most abundant snake toxin families (such as PLA2, SVMP and SVSP) are encoded by multi-locus gene families.[62,63] These gene families generate a range of alternative gene products which are unequal to the fundamental polypeptides. Here, proteoform-spectrum matches (PrSMs) derived from TD MS, give rise to many proteoforms with unknown mass discrepancies. 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.[64] Our TD data revealed 166 modified full-length isoforms, with enough fragment ions to cover exact masses, across the three population-level venoms belonging to the PLA2 (157

proteoforms) and disintegrin (9 proteoforms) toxin families. These reported modifications correspond to mass shifts associated with three types of common PTMs. 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 site. As an 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. 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.

Table 1. Proteoform characterization 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).

Protein families	Theor. Mass	Obser. mass	Modified protein sequence	PTM type	PTM Location site	MIScore[a]	BU sequence motif
PLA2	13,809.41 Da	13,823.41 Da	.SVVELGKMIIQETGKSPFPSYTSYG(C)[Methyl]FCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	C26	99.90%	-
PLA2	13,809.41 Da	13,823.41 Da	.SVVELGKMIIQETG(K)[Methyl]SPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	K15	99.90%	IIQETG**k**SPFPSY
PLA2	13,809.41 Da	13,823.43 Da	.SVVELGKMIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKY]RENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	K67	58.50%	-
PLA2	13,809.41 Da	13,823.41 Da	.SVVELGKMII(Q)[Methyl]ETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	Q11	92.00%	LGKMII**q**ETGKSP
PLA2	13,809.41 Da	13,837.41 Da	.SVVELGKMII(Q)[Methyl](E)[Methyl]TGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl;Methyl	Q11;E12	74.30%;71.40%	GKMIIQ**e**TGKSPF
PLA2	13,809.41 Da	13,823.41 Da	.(S)[Methyl]VVELGKMIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	S1	99.80%	**s**VVELGK
PLA2	13,809.41 Da	13,837.41 Da	.(S)[Methyl]VV(E)[Methyl]LGKMIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl;Methyl	S1;E4	99.90%;99.90%	-
PLA2	13,809.41 Da	13,822.42 Da	.SVV(E)[Methyl]LGKMIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	E4	92.00%	SVV**e**LGKMII
PLA2	13,809.41 Da	13,823.41 Da	.SVVELG(K)[Methyl]MIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl	K7	99.70%	SVVELG**k**MIIQET
PLA2	13,809.41 Da	13,837.40 Da	.SVV(E)[Methyl]LG(K)[Methyl]MIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Methyl;Methyl	E4;K7	47.90%;49.90%	-
PLA2	13,809.41 Da	13,851.37 Da	.SVVELGKMIIQETG(K)[Acetyl]SPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Acetyl	K15	99.90%	IIQETG**k**SPFPSY
PLA2	13,809.41 Da	13,851.42 Da	.SVVELG(K)[Acetyl]MIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Acetyl	K7	99.90%	SVVELG**k**MIIQET
PLA2	13,809.41 Da	13,867.42 Da	.SVVELGKMIIQETG(K)[Acetyl]S(P)[Oxidation]FPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Acetyl;Oxidation	K15;P17	99.60%;99.60%	-
PLA2	13,809.41 Da	13,825.41 Da	.SVVELG(K)[Oxidation]MIIQETGKSPFPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Oxidation	K7	99.90%	-
PLA2	13,809.41 Da	13,825.41 Da	.SVVELGKMIIQETGKSPFPSYTSYG(C)[Oxidation]FCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Oxidation	C26	91.10%	-
PLA2	13,809.41 Da	13,825.41 Da	.SVVELGKMIIQETGKS(P)[Oxidation]FPSYTSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Oxidation	P17	95.60%	-
PLA2	13,809.41 Da	13,825.41 Da	.SVVELGKMIIQETGKSPFPS(Y)[Oxidation]TSYGCFCGGGERGPPLDATDRCCLAHSCCYDTLPDCSPKTDRYKYKRENGEIICENSTSCKKRICECDKAVAVCLRKNLNTYNKKYTYYPNFWCKGDIEKC.	Oxidation	Y21	91.80%	-
Short DIS	5,420.34 Da	5,433.32 Da	.QC(E)[Methyl]SGPCCRNCKFLKEGTICKRARGDDMDDYCNGKTCDCPRNPHKGP.A	Methyl	E3	92.00%	-

[a] Probability-based PTM localization score with reference to Bayesian models.

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. 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 1).

Furthermore, data indicate the presence of mass shifts of +42 Da upon some lysine residues of PLA2 proteoforms (Table 1). The mass shift suggests a residue modification by mono-acetylation ($\Delta M = 42.0106$ Da) or by tri-methylation ($\Delta M = 42.0470$ 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 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.[65] Compared with all modified amino acid sites, lysine (particularly K7 and K15) was predominately targeted for methylation and acetylation in the PLA2 proteoforms (Table 1). 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.[66] Beside above-mentioned PTMs, TD results revealed other types of modifications such as pyroglutamic acid formation at the N-terminus of the disintegrin proteoform and oxidation of methionine (as a chemical modification) in PLA2 proteoform.[29] 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.

2.2. Mass Spectrometry Imaging of the Snake Venom Gland

As shown in figure 3, the MSI workflow starts by removing the venom gland and cutting the tissue into 20 µm-thick longitudinal sections, followed by image recording via a light microscope. The venom gland is soft and fragile, making the task of sample preparation very challenging for MSI and also for histological studies. Therefore, we applied two sample preparation strategies, fresh-frozen and formalin-fixation, to collect valid data from the morphologically "original state" of the venom gland tissue. In the next step, the DHB-matrix was deposited onto the tissue by pneumatic spraying, and subsequently, the sample surfaces were scanned with a 12 µm step size in positive-ion mode.

For data analysis, we initially employed unsupervised signal annotation using METASPACE online data searching against HMDB and LipidMaps databases. Mainly, lipids belonging to the classes of phosphatidylcholine (PC), sphingomyelin (SM), phosphatidylethanolamine (PE), phosphatidic acid (PA), and triglyceride (TG) were detected. Besides assigning known metabolites automatically by METASPACE, MS spectra were also interpreted manually to evaluate recorded data using the MIRION software package. Representative ion images are shown in Figure 4 for both, fresh-frozen (up) and formalin-fixed (down) tissue sections. The MS ion at m/z 734.5695 (blue), was assigned to the protonated molecular species of PC (32:0), a phosphatidylcholine with a total number of 32 carbon atoms in the fatty acyl chains and no double bond. PCs are common structural lipids of eukaryotic cell membranes. The associated signal has a distinct distribution, visualizing all cell walls of secretory epithelium within the venom gland. The signal at m/z 753.5881 (red) was assigned to a sphingomyelin lipid species SM (36:1) as sodiated molecular ion. Sphingomyelins are also known as a major constituent of the cellular membranes of animal tissues and play a critical role in signal transduction, cell differentiation and metabolism.

As observed by optical microscopy, snap-freezing damaged the gland, leading to cell-wall tearing in the tissue section while cutting (Figure 4). Consequently, large amounts of cell debris were widely observed in most fresh-frozen tissue areas. This is consistent with MS images of structural lipids that appear to be located in empty cavities for venom storage. However, in the formalin-fixed tissue images, they were found only in the tubules' areas where the secretory epithelial cells are located. In an RGB overlay, the co-localization of the aforementioned ions becomes more apparent.
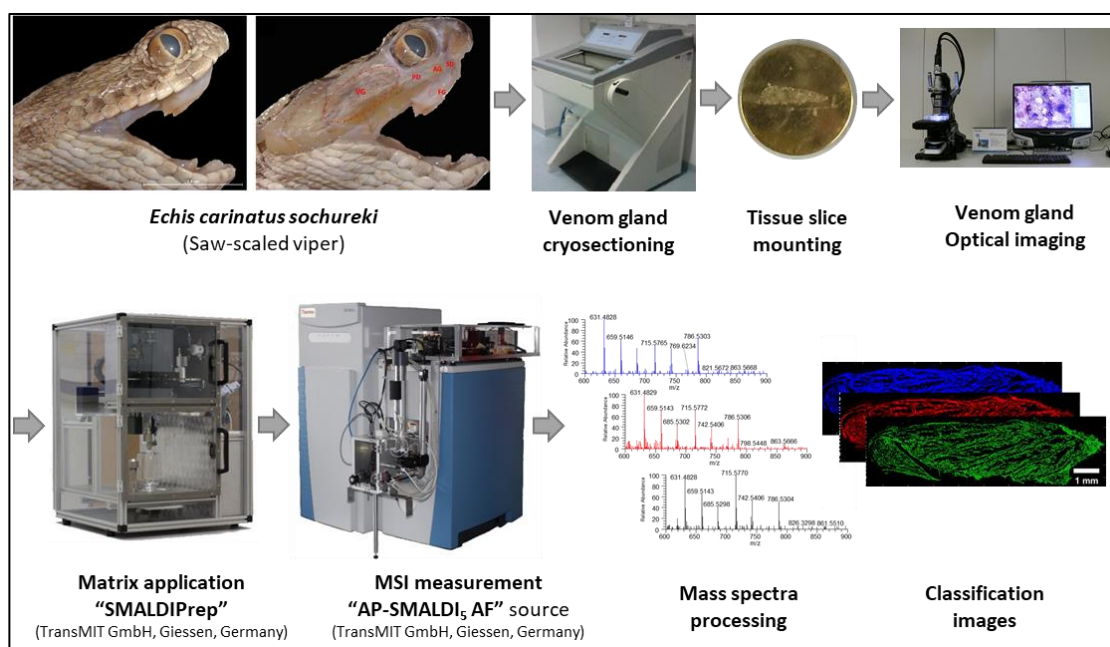
Figure 3. Workflow for snake venom gland mass spectrometry imaging (sVG-MSI).
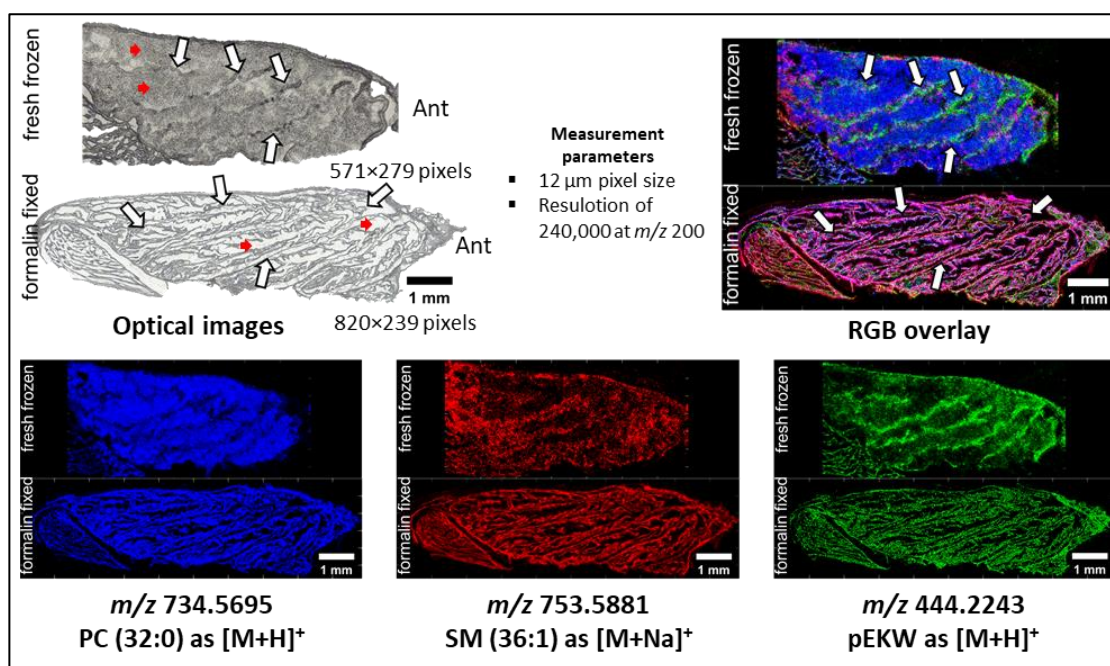
Figure 4. Snake venom gland mass spectrometry imaging (sVG-MSI). Positive-ion AP-SMALDI of phosphatidylcholine (PC), sphingomyelin (SM), and peptide from venom gland tissue sections of a saw-scaled viper (*Echis carinatus sochureki*) species. White arrows indicate the tubules' areas where the secretory epithelial cells are located and red arrows show empty cavities for venom storage.

Apart from lipid screening, we studied the snake venom proteome and the venom gland MSI for peptide profiling and the possible influence of the two tissue preparation methods

on peptide imaging. In this respect, the whole and the fractionated crude venom of *E. carinatus sochureki*, were investigated by LC-MS/MS, verifying the existence of an endogenous tri-peptide pyroglutamate-lysine-tryptophan (pEKW, m/z 444.2233 [M+H]+). This short peptide was also detected previously from closely related species, African saw-scaled viper (*Echis ocellatus*), venom and has been demonstrated to be a metalloprotease inhibitor.[67] Generally, venomous snakes are adapted to endure long periods without food, in which case their venom may even be stored in the gland for a long time without use.[68] Given that snake venom contains a variety of degradative enzymes such as phospholipase A2, metalloprotease, and serine proteinase, their enzymatic activities seem likely to be inhibited during storage in the venom gland, to prevent self-intoxication. For this reason, various physiological features like high concentration of citrate, ionic strength, pH and the production of endogenous peptides are thought to work together to inhibit venom enzymes, and therefore protect against auto-toxicity.[69] Recent studies have shown that the venom of vipers comprises significant amounts of short peptides known as pyroglutamic tripeptide inhibitors, pEKW (m/z 444.22), pEQW (m/z 444.18), pERW (m/z 472.22) and pENW (m/z 430.17).[70] These endogenous peptides serve as snake venom metalloprotease inhibitors (SVMPIs). Snake venom metalloproteinases (SVMPs) are among the most abundant Viperidae snake venom components, ergo severe haemorrhage and local tissue damage appear in the victims envenomated by vipers. Here, we also found that SVMPs are the major toxin classes in the venom of Iranian *E. carinatus sochureki*, jointly with the phospholipase A2 (PLA2) and snake venom serine protease (SVSP).

In present study, positive-ion AP-SMALDI MSI revealed the presence of pEWK, m/z 444.2243, in both types of venom gland sections (Figure 4, in green). To prove the identity of the tri-peptide directly from the venom gland, on-tissue AP-SMALDI MS^2 measurements were conducted, fragmenting the precursor ion (m/z 444.22 ± 0.25 Da) across both, fresh-frozen and formalin-fixed tissues. The tandem mass spectra obtained by on-tissue MS/MS indicated product ions comparable to those observed by LC-MS/MS of the crude venom, and thus provide confirmation of the tri-peptide, pEKW within the determined topography. Here MALDI imaging demonstrates that the tri-peptide, pEKW, is spatially distributed across the tissue in close proximity to secretory cells (Figure 4), thus supporting the idea that the peptide protects secretory tissue from enzymatic proteolysis by

SVMPs. This view can be expanded to give an idea of using the tri-peptide protease inhibitor as a neutralizing agent to prevent SVMP-induced haemorrhage in viper-bite victims.

To evaluate the general quality of the mass spectra, peak intensities and signal-to-noise ratios were compared between the recorded spectra of both tissue preparation methods, fresh-freezing vs. formalin-fixation. We found that the intensities of lipid and peptide peaks were closely similar in the spectra independent of preparation types. However, the corresponding signal-to-noise ratios of lipids and peptide in the formalin-fixed gland section were roughly two-fold higher than in spectra of fresh-frozen tissue. This can be explained by ion suppression in freshly frozen tissue due to more numerous components,[71] and dissolving and wash-out of some metabolites (mostly polar) in formalin solution, leading to dilution of compounds. Additionally, since recent studies reported that no PE and PS were deteced in formalin-fixed brain tissue[73] due to cross-linking with proteins and metabolites,[72] we expected complete loss of MS detection of these species in the formalin-fixed venom gland. However, our results showed that it is possible to detect these lipids in the fixed tissue, although we experienced some reduction in their ion signal intensities but not a complete loss. To support MSI annotations of PE and PS in the formalin-fixed tissue, we also applied untargeted high-resolution LC-MS/MS of the extracted lipidome from the fixed venom gland tissue. Accordingly, with such a high degree of spectral similarities and its high ability to preserve tissue morphology, formalin-fixation can be proposed as an advantageous method for lipid and peptide imaging in venom gland tissues.

3. REFERENCES

1. Thomson, J. J. XL. Cathode Rays. *Lond. Edinb. Dubl. Philos. Mag.* **1897**, 44, 293–316.

2. Wien, W. Untersuchungen über die elektrische Entladung in verdünnten Gasen. *Ann. Phys.* **1901**, 310, 421–435.

3. Griffiths, J. A Brief History of Mass Spectrometry. *Anal. Chem.* **2008**, 80, 5678-5683.

4. Mann, M., Meng, C.K. Fenn, J.B. Interpreting Mass Spectra of Multiply Charged Ions. *Anal. Chem.* **1989**, 61, 1702-1708.

5. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science*, **1989**, 246,64-71.

6. Karas, M., Bachmann, D., Hillenkamp, F. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal. Chem.* **1985**, 57, 2935-2939.

7. Hillenkamp, F., Karas, M. Matrix-Assisted Laser Desorption/Ionisation, an Experience. *Int. J. Mass Spectrom.* **2000**, 71-77.

8. de Hoffmann, E., Stroobant, V. Mass Spectrometry: Principles and Applications. John Wiley & Sons, Chichester. **2007**.

9. Spengler, B. Mass Spectrometry Imaging of Biomolecular Information. *Anal. Chem.* **2015**, 87,64−82.

10. Welker, M.; Van Belkum, A.; Girard, V.; Charrier, J.-P.; Pincus, D. An Update on the Routine Application of MALDI-TOF MS in Clinical Microbiology. *Expert Rev. Proteomics* **2019**, 16 (8), 695–710.

11. Girolamo, F.; Lante, I.; Muraca, M.; Putignani, L. The Role of Mass Spectrometry in the "Omics" Era. *Curr. Org. Chem.* **2013**, 17 (23), 2891–2905.

12. Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annu. Rev. Biomed. Eng.* **2009**, 11 (1), 49–79.

13. Ibáñez, M.; Guerrero, C.; Sancho, J. V; Hernández, F. Screening of Antibiotics in Surface and Wastewater Samples by Ultra-High-Pressure Liquid Chromatography Coupled to Hybrid Quadrupole Time-of-Flight Mass Spectrometry. *J. Chromatogr. A* **2009**, 1216 (12), 2529–2539.

14. Onzo, A.; Pascale, R.; Acquavia, M. A.; Cosma, P.; Gubitosa, J.; Gaeta, C.; Iannece, P.; Tsybin, Y.; Rizzi, V.; Guerrieri, A.; Ciriello, R.; Bianco, G. Untargeted Analysis of Pure Snail Slime and Snail Slime-induced Au Nanoparticles Metabolome with MALDI FT-ICR MS. *J. Mass Spectrom.* **2021**, 56 (5), e4722.

15. Eliuk, S.; Makarov, A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu. Rev. Anal. Chem.* **2015**, 8 (1), 61–80.

16. Makarov, A.; Scigelova, M. Coupling Liquid Chromatography to Orbitrap Mass Spectrometry. *J. Chromatogr. A* **2010**, 1217 (25), 3938–3945.

17. Bianconi, E.; Piovesan, A.; Facchin, F.; Beraudi, A.; Casadei, R.; Frabetti, F.; Vitale, L.; Pelleri, M. C.; Tassani, S.; Piva, F.; Perez-Amodio, S.; Strippoli, P.; Canaider, S. An Estimation of the Number of Cells in the Human Body. *Ann. Hum. Biol.* **2013**, 40 (6), 463–471.

18. Smith, R.; Mathis, A. D.; Ventura, D.; Prince, J. T. Proteomics, Lipidomics, Metabolomics: A Mass Spectrometry Tutorial from a Computer Scientist's Point of View. *BMC Bioinformatics* **2014**, 15 (S7), S9.

19. Dakna, M.; He, Z.; Yu, W. C.; Mischak, H.; Kolch, W. Technical, Bioinformatical and Statistical Aspects of Liquid Chromatography–Mass Spectrometry (LC–MS) and Capillary Electrophoresis-Mass Spectrometry (CE-MS) Based Clinical Proteomics: A Critical Assessment. *J. Chromatogr. B* **2009**, 877 (13), 1250–1258.

20. Mischak, H.; Coon, J. J.; Novak, J.; Weissinger, E. M.; Schanstra, J. P.; Dominiczak, A. F. Capillary Electrophoresis–Mass Spectrometry as a Powerful Tool in Biomarker Discovery and Clinical Diagnosis: An Update of Recent Developments. *Mass Spectrom. Rev.* **2009**, 28 (5), 703–724.

21. Glish, G. L.; Vachet, R. W. The Basics of Mass Spectrometry in the Twenty-First Century. *Nat. Rev. Drug Discov.* **2003**, 2 (2), 140–150.

22. Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem. Rev.* **2013**, 113 (4), 2343–2394.

23. Uy, R.; Wold, F. Posttranslational Covalent Modification of Proteins. *Science* **1977**, 198 (4320), 890–896.

24. Oliveira, A. P.; Sauer, U. The Importance of Post-Translational Modifications in Regulating Saccharomyces Cerevisiae Metabolism. *FEMS Yeast Res.* **2012**, 12 (2), 104–117.

25. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999**, 121 (4), 806–812.

26. Wolters, D. A.; Washburn, M. P.; Yates, J. R. An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics. *Anal. Chem.* **2001**, 73 (23), 5683–5690.

27. Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annu. Rev. Biomed. Eng.* **2009**, 11 (1), 49–79.

28. Schubert, O. T.; Röst, H. L.; Collins, B. C.; Rosenberger, G.; Aebersold, R. Quantitative Proteomics: Challenges and Opportunities in Basic and Applied Research. *Nat. Protoc.* **2017**, 12 (7), 1289–1294.

29. 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.* **2019**, 33 (S1), 20–27.

30. Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. Top Down Proteomics: Facts and Perspectives. *Biochem. Biophys. Res. Commun.* **2014**, 445 (4), 683–693.

31. He, L.; Rockwood, A. L.; Agarwal, A. M.; Anderson, L. C.; Weisbrod, C. R.; Hendrickson, C. L.; Marshall, A. G. Top-down Proteomics—a near-Future Technique for Clinical Diagnosis? *Ann. Transl. Med.* **2020**, 8 (4), 136–136.

32. Clish, C.B. Metabolomics: an Emerging but Powerful Tool for Precision Medicine. *Cold Spring Harb Mol Case Stud.* **2015**;1(1):a000588.

33. Gulati, K.; Sarkar, S.; Poluri, K. M. Metabolomics Analysis of Complex Biological Specimens Using Nuclear Magnetic Resonance Spectroscopy; Wood, P. L., Ed.; Springer US: New York, NY, **2021**; pp 155–171.

34. Züllig, T.; Trötzmüller, M.; Köfeler, H. C. Lipidomics from Sample Preparation to Data Analysis: A Primer. *Anal. Bioanal. Chem.* **2020**, 412 (10), 2191–2209.

35. Yao, L.; Sheflin, A. M.; Broeckling, C. D.; Prenni, J. E. Data Processing for GC-MS- and LC-MS-Based Untargeted Metabolomics; D'Alessandro, A., Ed.; Springer New York: New York, NY, **2019**; pp 287–299.

36. Tang, D.-Q.; Zou, L.; Yin, X.-X.; Ong, C. N. HILIC-MS for Metabolomics: An Attractive and Complementary Approach to RPLC-MS. *Mass Spectrom. Rev.* **2016**, 35 (5), 574–600.

37. Nagana Gowda, G. A.; Djukovic, D. Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. *Methods Mol. Biol.* **2014**, 1198, 3–12.

38. Katajamaa, M., Orešič, M. Data processing for Mass Spectrometry-based Metabolomics. *J. Chrom. A.* **2007**, 1158, 318-328.

39. Lewis RJ, Garcia ML. Therapeutic potential of venom peptides. *Nat Rev Drug Discov.* **2003**, 2(10), 790-802.

40. King GF. Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin Biol Ther.* **2011**, 11(11), 1469-1484.

41. Vetter I, Davis JL, Rash LD, et al. Venomics: A new paradigm for natural products-based drug discovery. *Amino Acids.* **2011**, 40(1), 15-28.

42. Warrell DA, Gutiérrez JM, Calvete JJ, Williams D. New approaches & technologies of venomics to meet the challenge of human envenoming by snakebites in India. *Indian J Med Res.* **2013**, 138, 38-59.

43. Gutiérrez JM, Calvete JJ, Habib AG, Harrison RA, Williams DJ, Warrell DA. Snakebite envenoming. *Nat Rev Dis Prim.* **2017**, 3, 17063-17079.

44. Gutiérrez JM, Lomonte B, León G, et al. Snake venomics and antivenomics: Proteomic tools in the design and control of antivenoms for the treatment of snakebite envenoming. *J Proteomics.* **2009**, 72(2), 165-182.

45. Warrell DA. Snake bite. *Lancet.* **2010**, 375(9708), 77-88.

46. Chippaux JP, Williams V, White J. Snake venom variability: methods of study, results and interpretation. *Toxicon.* **1991**, 29(11), 1279-1303.

47. Alape-Giron A, Sanz L, Flores-diaz M, Madrigal M, Sasa M, Calvete JJ. Snake venomics of the lancehead pitviper *Bothrops asper*: geographic, individual, and ontogenetic variations. *J Proteome Res.* **2008**, 7(8), 3556-3571.

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

49. Undheim, E. A. B.; Hamilton, B. R.; Kurniawan, N. D.; Bowlay, G.; Cribb, B. W.; Merritt, D. J.; Fry, B. G.; King, G. F.; Venter, D. J. Production and Packaging of a Biological Arsenal: Evolution of Centipede Venoms under Morphological Constraint. *Proc. Natl. Acad. Sci.* **2015**, 112, 4026−4031.

50. Seppälä, U.; Francese, S.; Turillazzi, S.; Moneti, G.; Clench, M.; Barber, D. In Situ Imaging of Honeybee (*Apis Mellifera*) Venom Components from Aqueous and Aluminum Hydroxide−Adsorbed Venom Immunotherapy Preparations. *J. Allergy Clin. Immunol.* **2012**, 129, 1314−1320.

51. Mitchell, M. L.; Hamilton, B. R.; Madio, B.; Morales, R. A. V.; Tonkin-Hill, G. Q.; Papenfuss, A. T.; Purcell, A. W.; King, G. F.; Undheim, E. A. B.; Norton, R. S. The Use of Imaging Mass Spectrometry to Study Peptide Toxin Distribution in Australian Sea Anemones. *Aust. J. Chem.* **2017**, 70, 1235−1237.

52. Madio, B.; Peigneur, S.; Chin, Y. K. Y.; Hamilton, B. R.; Henriques, S. T.; Smith, J. J.; Cristofori-Armstrong, B.; Dekan, Z.; Boughton, B. A.; Alewood, P. F.; Tytgat, J.; King, G. F.; Undheim, E. A. B. PHAB Toxins: A Unique Family of Predatory Sea Anemone Toxins Evolving via Intra-Gene Concerted Evolution Defines a New Peptide Fold. *Cell. Mol. Life Sci.* **2018**, 75, 4511−4524.

53. Hamilton, B. R.; Marshall, D. L.; Casewell, N. R.; Harrison, R. A.; Blanksby, S. J.; Undheim, E. A. B. Mapping Enzyme Activity on Tissue by Functional Mass Spectrometry Imaging. *Angew. Chem., Int. Ed.* **2020**, 59, 3855−3858.

54. Das, T.; Alabi, I.; Colley, M.; Yan, F.; Griffith, W.; Bach, S.; Weintraub, S. T.; Renthal, R. Major Venom Proteins of the Fire Ant *Solenopsis Invicta*: Insights into Possible Pheromone-Binding Function from Mass Spectrometric Analysis. *Insect Mol. Biol.* **2018**, 27, 505−511.

55. 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. U.S.A.* **2014**, 111, 9205−9210.

56. 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.

57. 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, No. 17119.

58. Bhatia, S.; Vasudevan, K. Comparative Proteomics of Geographically Distinct Saw-Scaled Viper (*Echis carinatus*) Venoms from India. *Toxicon X* **2020**, 7, No. 100048.

59. 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.

60. 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, 1169−1179.

61. 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.

62. 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, No. 11300.

63. 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. U.S.A.* **2013**, 110, 20651−20656.

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

65. Kou, Q.; Wu, S.; Liu, X. Systematic Evaluation of Protein Sequence Filtering Algorithms for Proteoform Identification Using Top-Down Mass Spectrometry. *Proteomics* **2018**, 18, No. 1700306.

66. 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.

67. Wagstaff, R.; Favreau, P.; Cheneval, O.; Laing, G. D.; Wilkinson, M. C.; Miller, R. L.; Stocklin, R.; Harrison, R. A. Molecular Characterizatioin of Endogenous Snake Venom Metalloproteinase Inhibitors. *Biochem. Biophys. Res. Commun.* **2008**, 365, 650−656.

68. Munekiyo, S. M.; Mackessy, S. P. Presence of Peptide Inhibitors in Rattlesnake Venoms and Their Effects on Endogenous Metalloproteases. *Toxicon* **2005**, 45, 255−263.

69. Francis, B.; Kaiser, I. I. Inhibition of Metalloproteinases in *Bothrops asper* Venom by Endogenous Peptides. *Toxicon* **1993**, 31, 889−899.

70. Aoki-Shioi, N.; Koh, C. Y.; Kini, R. M. Natural Inhibitors of Snake Venom Metalloproteinases. *Aust. J. Chem.* **2020**, 73 (4), 277− 286.

71. Pietrowska, M.; Gawin, M.; Polańska, J.; Widłak, P. Tissue Fixed with Formalin and Processed without Paraffin Embedding Is Suitable for Imaging of Both Peptides and Lipids by MALDI-IMS. *Proteomics* **2016**, 16 (11−12), 1670−1677.

72. Carter, C. L.; Jones, J. W.; Farese, A. M.; MacVittie, T. J.; Kane, M. A. Inflation-Fixation Method for Lipidomic Mapping of Lung Biopsies by Matrix Assisted Laser Desorption/Ionization-Mass Spectrometry Imaging. *Anal. Chem.* **2016**, 88, 4788−4794.

73. Gaudin, M.; Panchal, M.; Ayciriex, S.; Werner, E.; Brunelle, A.; Touboul, D.; Boursier-Neyret, C.; Auzeil, N.; Walther, B.; Duyckaerts, C.; Laprevote, O. Ultra-Performance Liquid Chromatography-Mass Spectrometry Studies of Formalin-induced Alterations of Human Brain Lipidome. *J. Mass Spectrom.* **2014**, 49, 1035−1042.

CHAPTER II

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 Mahdi Kazemi, Alireza Ghassempour, and Bernhard Spengler

Article

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 Mahdi Kazemi, Alireza Ghassempour, and Bernhard Spengler*

Cite This: *J. Proteome Res.* 2021, 20, 895−908 | Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

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, and 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. LD$_{50}$ (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 liquid chromatography and tandem mass spectrometry (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*
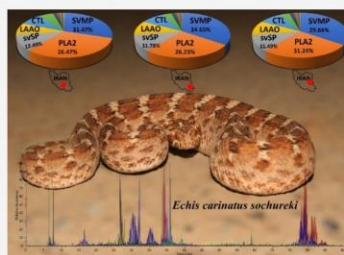
■ INTRODUCTION

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 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 nonfatal 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
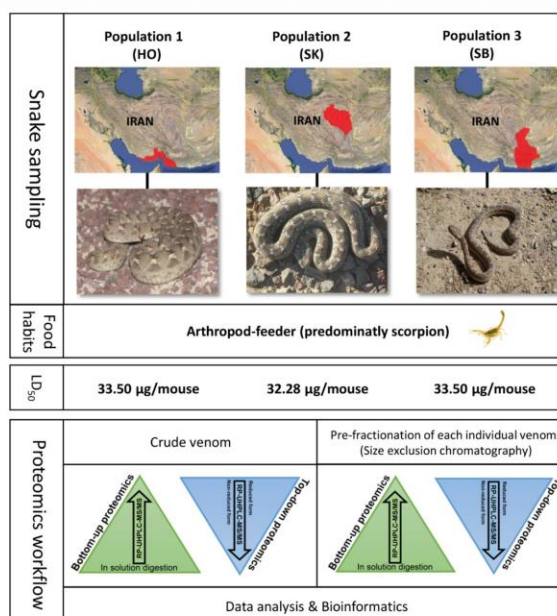
Figure 1. Overview of the proposed experimental workflow used for interrogation of food habits, toxicity assessment, and venom proteome profiling of three Iranian populations of saw-scaled viper, *E. carinatus sochureki*. HO: Hormozgan; SK: South Khorasan; SB: Sistan and Baluchestan; LD_{50}: median lethal dose.

rich natural biological resource with a 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 hemotoxicity and/or cytotoxicity, and these include local tissue damage such as edema, blistering, hemorrhage, dermo-, and myonecrosis, and also systemic alterations such as hemorrhage, 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 *Echis carinatus*, *Echis coloratus*, *Echis ocellatus*, and *Echis pyramidum*.[10] The *E. carinatus* group has been further divided into two subspecies; *Echis carinatus carinatus* and *Echis 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, and 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 RAZI 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 *E. carinatus*, *Macrovipera lebetina*, *Pseudocerastes persicus*, *Gloydius 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 toward different preys.[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 toward 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 the 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 predominantly 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 have facilitated improvements in TD proteomics, enabling quick and accurate investigation of intact toxin families and their proteoforms (see refs 33 and 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 subsections. 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 (LD$_{50}$) of each population-representative venom was evaluated by the 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 an overview of the conspecific venom variation of Iranian ECS.

■ EXPERIMENTAL SECTION

Sample Collection and Venom Milking

The saw-scaled viper, *E. carinatus sochureki*, specimens were collected in 2015−2017 between March and October in three locations (South Khorasan [SK], Sistan and Baluchestan [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 fecal 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 the collection of fecal pellets. The prey items were mostly identified from wide taxonomic classification such as phylum, family, and genus using sample evidence, 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 LD$_{50}$ (median lethal dose) values 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 LD$_{50}$ values were 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 research[38] 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 (Bio-rad, Hercules, CA), with bovine serum albumin (BSA) used as a reference. Absorbance was measured

44

spectrophotometrically at 595 nm on a BioTek Synergy 2 plate reader (BioTek, Winooski, VT) with Gen5 software (version 2.01).

Chemicals

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

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Gel Filtration Separation

Each of the regional ECS venoms was 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 α 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 solubilization. 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 a 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 using ZipTip C18 and then concentrated 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 (Ultrahigh-Performance Liquid Chromatography, Thermo Fisher Scientific) on a Kinetex C18 (2.1 mm × 100 mm, 2.6 μm 100 Å particle size) column (Phenomenex, CA) coupled to a Q Exactive HF-X and Q Exactive (QE) Orbitrap (Thermo Scientific, Bremen, Germany). Chromatographic analysis was performed at a 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 was 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: a mass range of m/z 350−1800 in QE HF-X (mass resolution of 120 000, AGC target of 3×10^6, IT of 50 ms) and QE (mass resolution of 70 000, AGC target of 1×10^6, IT of 120 ms), and MS/MS scans: a mass range of m/z 200−2000 in QE HF-X (mass resolution of 30 000, AGC target of 1×10^5, 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 1×10^6, IT of 120 ms, isolation window $\Delta(m/z)$ 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 E. 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. The 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), monomethylation (+14.016 Da), trimethylation (+42.047 Da), and phosphorylation (+79.966 Da) were set as dynamic modifications for the E. carinatus database search. A percolator[41] node was used to validate identified PSMs and filter the data with the 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 the 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 Å particle size) column (Phenomenex, CA). The venom samples were dissolved in 1% formic acid and reduced with TCEP (tris (2-carboxy-ethylphosphine)) 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 45 min, 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 a data-dependent acquisition (DDA) mode with external calibration and the 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 rates were set to 20 and 60,

Table 1. Identified Protein Families in the Whole and SEC Fractions of Three *E. carinatus sochureki* Venom Populations by Bottom-Up and Top-Down Proteomics[a]

bottom-up						top-down				
whole venom			SEC fractions			whole venom	SEC fractions			
HO	SK	SB	HO	SK	SB	mixed	HO	SK	SB
5′NTD	5′NTD	5′NTD	5′NTD	5′NTD	5′NTD	5′NTD	5′NTD	5′NTD	5′NTD
PLA2	PLA2	PLA2	PLA2	PLA2	PLA2	PLA2	PLA2	PLA2	PLA2
CRISP	CRISP	CRISP	CRISP	CRISP	CRISP	CRISP	CRISP	CRISP	CRISP
SVMP	SVMP	SVMP	SVMP	SVMP	SVMP	SVMP	SVMP	SVMP	SVMP
DIS	DIS	DIS	DIS	DIS	DIS	DIS	DIS	DIS	DIS
CTL	CTL	CTL	CTL	CTL	CTL	CTL	CTL	CTL	CTL
GC	GC	GC	GC	GC	GC	GC	GC	GC	GC
LAO	LAO	LAO	LAO	LAO	LAO	LAO	LAO	LAO	LAO
SVSP	svSP	svSP	svSP	svSP	suSP	svSP	suSP	svSP	SVSP
RLAP			RLAP	RLAP	RLAP	RLAP	RLAP	RLAP	RLAP
	AP	AP	AP	AP	AP	AP	AP	AP	AP
PLB		PLB	PLB	PLB	PLB	PLB	PLB	PLB	PLB
NGF	NGF	NGF	NGF	NGF	NGF				
	PDE	PDE	PDE	PDE	PDE		PDE	PDE	PDE
			HYAL	HYAL		HYAL			
			VEGF	VEGF				VEGF	
				AChE					
				DPEP				DPEP	
						BPPs/pHpG/C-NP	BPPs/pHpG/C-NP	BPPs/pHpG/C-NP	BPPs/pHpG/C-NP
						FGF			
						VVP			

[a]5′-Nucleotidase (5′NTD), phospholipase A2 (PLA2), cysteine-rich secretory protein (CRISP), snake venom metalloproteinase (SVMP), disintegrin (DIS), C-type lectin (CTL), glutaminyI-peptide cyclotransferase (GpCT), L-amino-acid oxidase (LAO), snake venom serine proteinase (SVSP), renin-like aspartic protease (RLAP), aminopetidase (AP), phospholipase B (PLB), snake venom nerve growth factor (NGF), phosphodiesterase (PDE), hyaluronidase (HYAL), vascular endothelial growth factor (VEGF), bradykinin-potentiating and C-type natriuretic peptides (BPPs/C-NP), dipeptidyl peptidase (DPEP), fibroblast growth factor (FGF), snake venom metalloproteinase inhibitor (SVMPi), acetylcholinesterase (AChE), peptidyl-prolyl-cis-trans isomerase (PPI), and venom vasodilator peptide (VVP).

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 1×10^6. 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 60 s.

Top-Down Data Analysis

The 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, the signal-to-noise ratio was set to 2, the precursor window size set to $\Delta m/z = 4$, and the 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 *E. 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 the decoy database set to the 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 repository[46] with the dataset identifier PXD021183 and project name "Saw-scaled viper, *E. 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 (*E. carinatus sochureki*; ECS) communities. We detected prey items in the stomach and fecal contents of 96 live snakes, as detailed in Supporting Information Table S1. The prey (in total 125 items) most commonly found in all three populations were arthropods ($n = 77, 61.6\%$), which are 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
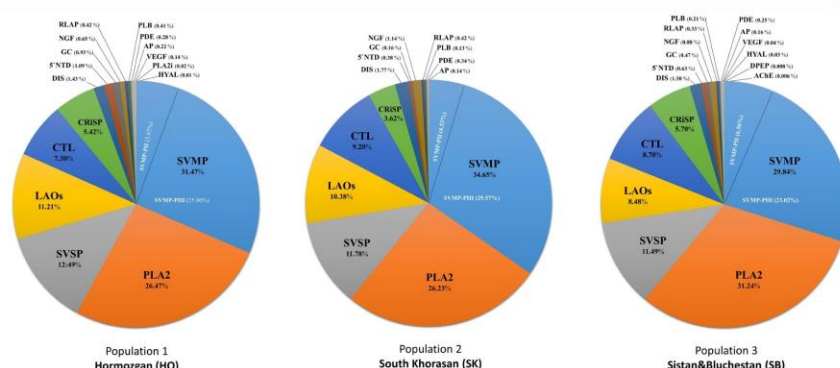
Figure 2. Overview composition and relative abundance of protein families from venom proteomes of three *E. carinatus sochureki* populations. Pie charts representing the relative abundance of protein families based on the normalized spectral abundance factor (NSAF), identified in six size-exclusion fractions of each venom community by a 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; and AChE: acetylcholinesterase.

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 LD_{50} dose to laboratory mice. The results showed highly comparable venom LD_{50} 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, the 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 the diet and venom function at the intraspecific 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 was screened by utilizing 1D SDS-PAGE and intact reversed phase ultra performance liquid chromatography (RP-UHPLC) separations (Supporting 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. Size-exclusion chromatography (SEC) was applied to decomplex 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 (Supporting 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 liquid chromatography electrospray ionization high resolution mass spectrometry (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 the 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 Tables S2−S4). However, BU measurements of SEC venom fractions (six fractions per population in triplicate) via the same liquid chromatography and tandem mass spectrometry (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 Tables S5−S7). Thus, it is important to note that predigestion protein fractionation by SEC substantially increased the BU MS/MS protein quantification by nearly 2-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 presence 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).

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 concentrations 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, hemorrhage, 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; hence, it has taken the third dominant place.

In all three venoms, both P-II and P-III subclasses of SVMPs were detected, although the latter (more hemorrhagic than the P-II form) were more abundant, with a 3- to 6-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 (Supporting Information Tables S2−S11). 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 to be hemostatic system disturbances and affects hemostasis, blood coagulation, platelets, and fibrinolytic system in snakebite victims. Note-worthily, 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 proinflammatory

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 Tables S2−S7). 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 in detecting 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 that cleaves phospholipids from sn-1 and sn-2 positions[60] and exhibits a potent hemolytic 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 the venom of the SK population that functionally hydrolyzes esters of choline. This enzyme has been reported mostly in the venom of the 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 any 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 phenomenon[65] and has previously been correlated with natural selection pressure for regional feeding habits[65,66] 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.

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 mod-

Table 2. Proteoform Characterization of PLA2 Toxin Family by Top-Down Mass Spectrometry

[a] Probability-based PTM localization score with reference to Bayesian models. [b] Populations: Hormozgan (HO), South Khorasan (SK), and Sistan and Baluchestan (SB). [c] The TD MS/MS spectra are searched against the UniProtKB Echis carinatus database (#40353) using TopPIC suite software by employing the four common PTMs (methylation, acetylation, oxidation, and phosphorylation) and modification identification score (MIScore). The symbol marked with a star indicates a modified peptide, which was identified by the BU approach. Graphic sequence representations were generated using Caititu software[77] and the colors indicate the structure section of helical regions (blue), hydrogen-bonded turns (yellow), and β strands (green) within the PLA2 protein sequence (UniProtKB entry identifier #P48650).

ifications (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 (nonreduced and reduced form) revealed a total of 6910 identified PrSMs (Supporting Information Table S12), which mapped to 170 proteins with unique accession numbers (Supporting Information Table S8). In addition, duplicate LC-MS/MS analyses of each SEC fraction in nonreduced and reduced forms generated 5489 identified PrSMs (Supporting Information Tables S13−S15), and resulted in 102 (HO), 114 (SB), and 133 (SK) proteins sorted by unique accession number (Supporting Information Tables S9−S11). An overview of the TD results obtained from these various measurements revealed that all of 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 predominantly 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. Supporting Tables S12−S15 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 the 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 multilocus 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

49
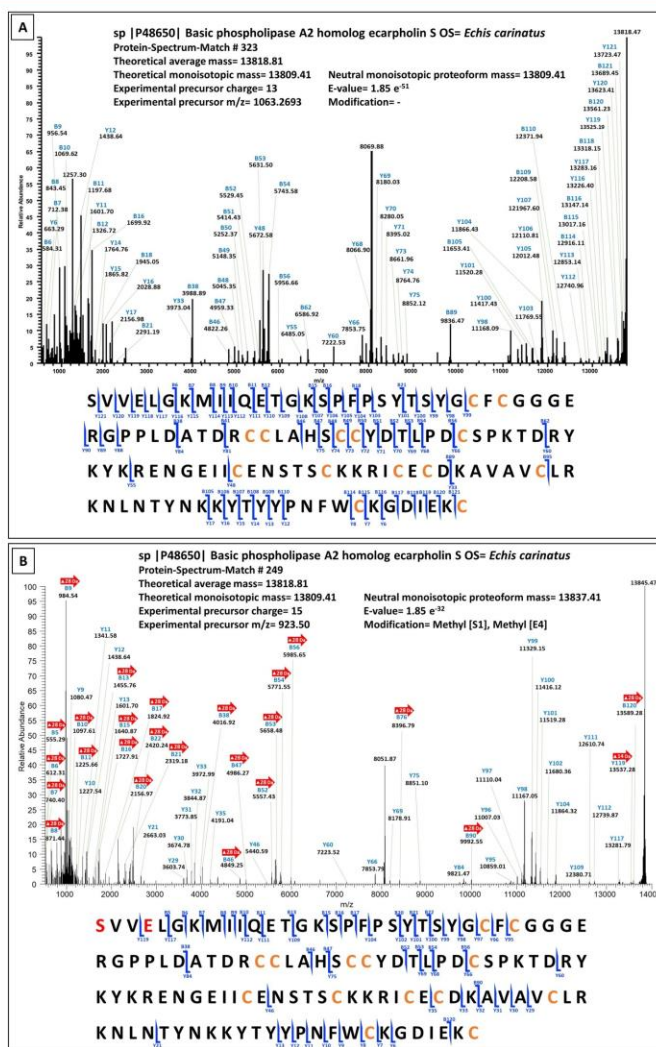
Figure 3. Top-down spectra of HCD fragmentations and sequence coverages of Phospholipase A2 proteoforms in *E. carinatus sochureki* venom from Iran. (A) Xtract-deconvoluted MS/MS spectrum from an unmodified protein sequence, which was identified as the PLA2 protein family, and localization of b and y ions in sequence, (B) Xtract-deconvoluted MS² spectrum generated from the Ser-1 and Glu-4 modified proteoform of the 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 methylation in the two localized sites (red words) with their MIScores of 99.90%.

rise to many proteoforms with unknown mass discrepancies (Supporting Table S12–S15). 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.[72] 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
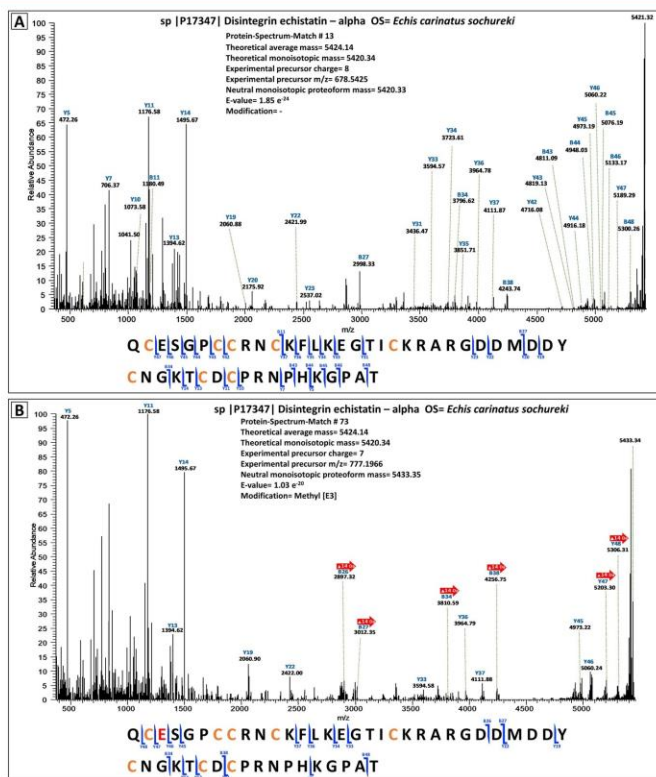
Figure 4. Top-down MS/MS spectra of HCD fragmentations and sequence coverages of unmodified and modified disintegrin proteoforms in *E. carinatus sochureki* venom from Iran. (A) Xtract-deconvoluted MS/MS spectrum generated from unmodified protein sequence, identified as the disintegrin protein family, and localization of b and y ions in the sequence, (B) Xtract-deconvoluted MS² spectrum generated from a Glu-3 modified proteoform of the disintegrin protein family, and its sequence coverage. The mass shifts (+14 Da) between product ions (red arrows) can be explained by methylation in the localized sites (red words) with the MIScores of 92%.

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,[73] 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 (Supporting 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, a comparison of the TD MS/MS of the 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 an 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 a 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 Supporting Information Figures S3–S12). 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, Supporting Information Table S17 and Figure S13).

Furthermore, data indicate the presence of mass shifts of +42 Da upon some lysine residues of PLA2 proteoforms (Table 2, Supporting Information Figures S8–S10). The mass shift suggests a residue modification by monoacetylation (ΔM = 42.0106 Da) or by trimethylation (ΔM = 42.0470 Da). Distinguishing isobaric trimethyl 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 Supporting 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.[74,75] 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 the snake venom proteome has also been reported previously in the minor venom proteins of East African green (*Dendroaspis angusticeps*) and black (*Dendroaspis polylepis*) mamba by Petras and colleagues.[76] Besides the above-mentioned PTMs, TD results revealed other types of modifications, such as pyroglutamic acid formation at the N-terminus of the disintegrin proteoform (Supporting Information Figure S12) and oxidation of methionine (as a chemical modification) in the PLA2 proteoform.[33] It thus seems that PTMs of toxins (particularly lysine modification) may play an important role in the structural and functional properties of venom proteins. Characterizing them may therefore prove to be important for a better understanding of 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 a 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 (*E. carinatus sochureki*) venoms were performed by applying multidimensional chromatography coupled to high-resolution mass spectrometry. To promote a 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 suggests a 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 unraveling 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

Total ion current (TIC) and SDS-PAGE patterns of ECS venoms (Figure S1); SEC elution profiles of all ECS population venoms (Figure S2); identified PrSM of PLA2 toxin family with one methylation site at K15 (Figure S3); methylation site at C26 (Figure S4); two methylation sites (Q11; E12) (Figure S5); methylation site at S1 (Figure S6); methylation site at K67 (Figure S7); acetylation site at K7 (Figure S8); two modification sites (K7; P17) (Figure S9); acetylation site at K15 (Figure S10); one oxidation site at K7 (Figure S11); an identified PrSM of the disintegrin protein family (Figure S12); PTMs identified by bottom-up proteomic approaches in the PLA2 toxin family (Figure S13) (PDF)

List of food habits of ECS communities in Iran (Table S1); identified proteins by bottom-up MS/MS analysis of ECS crude venom from HO, SK, and SB populations (Tables S2, S3, and S4); identified proteins by bottom-up MS/MS analysis of SEC fractions of ECS venoms from HO, SK, and SB populations (Tables S5, S6, and S7); list of confirmed proteins using top-down MS/MS analysis of ECS crude venom (mixed population) (Table S8); list of confirmed proteins using top-down MS/MS analysis of SEC fractions of ECS venoms from HO, SK, and SB populations (Table S9, S10, and S11); list of proteoforms with four common PTMs in isolated fractions of three ECS population venoms (Table S12); list of all identified PrSM in isolated fractions of HO, SB, and SK population venoms (Table S13, S14, and S15); list of all modified proteoforms from all SEC fractions (Table S16); and list of modified peptides of all SEC fractions using bottom-up MS/MS analysis (Table S17) (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Bernhard Spengler − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University Giessen, Giessen 35392, Germany;* ⓞ orcid.org/0000-0003-0179-5653; Email: Bernhard.Spengler@anorg.chemie.uni-giessen.de

Authors

Parviz Ghezellou − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University Giessen, Giessen 35392, Germany*

Wendell Albuquerque − *Institute of Food Chemistry and Food Biotechnology, Justus Liebig University Giessen, Giessen 35392, Germany*

Vannuruswamy Garikapati − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University Giessen, Giessen 35392, Germany*

Nicholas R. Casewell − *Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Liverpool L3 5QA, U.K.*

Seyed Mahdi Kazemi − *Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, Tehran 1983963113, Iran*

Alireza Ghassempour − *Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, Tehran 1983963113, Iran;* orcid.org/0000-0002-6435-9915

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jproteome.0c00687

Notes

The authors declare no competing financial interest.

■ **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, ohospholipase A2; CRISP, Cysteine-rich secretory protein; SVMP, snake venom metalloproteinase; DIS, disintegrin; CTL, C-type lectin; 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, 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, 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. Primers* **2017**, 3, No. 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 Neglected Trop. Dis.* **2011**, 5, No. e1255.

(5) Chippaux, J. P. Snakebite Envenomation Turns Again into a Neglected Tropical Disease! *J. Venomous Anim. Toxins Incl. Trop. Dis.* **2017**, 23, No. 38.

(6) King, G. F. Venoms as a Platform for Human Drugs: Translating Toxins into Therapeutics. *Expert Opin. Biol. Ther.* **2011**, 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, e150.

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

(9) Warrell, D. A.; Davidson, N. M.; 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. *Br. Med. J.* **1974**, 4, 437−440.

(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, 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, 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, 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, 27−33.

(15) Chippaux, J. P. Guidelines for the Production, Control and Regulation of Snake Antivenom Immunoglobulins. *Biol. Aujourd'hui* **2010**, 204, 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, 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, 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, 283−289.

(21) Harrison, R. A.; Hargreaves, A.; Wagstaff, S. C.; Faragher, B.; Lalloo, D. G. Snake Envenoming: A Disease of Poverty. *PLoS Neglected Trop. Dis.* **2009**, 3, No. 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. U.S.A.* **2014**, *111*, 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*, No. 20182735.

(24) Chippaux, J. P.; Williams, V.; White, J. Snake Venom Variability: Methods of Study, Results and Interpretation. *Toxicon* **1991**, *29*, 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*, 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*, 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*, 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*, 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*, No. 11.

(32) Smith, L. M.; Kelleher, N. L. Proteoform: A Single Term Describing Protein Complexity. *Nat. Methods* **2013**, *10*, 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.* **2019**, *33*, 20−27.

(34) Melani, R. D.; Nogueira, F. C. S.; Domont, G. B. It Is Time for Top-down Venomics. *J. Venom. Anim. J. Venomous Anim. Toxins Incl. Trop. Dis.* **2017**, *23*, No. 44.

(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*, 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. Jutlandica.* **1981**, *19*, 45−48.

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

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

(39) Nekouei, M.; Ghezellou, P.; Aliahmadi, A.; Arjmand, S.; Kiaei, M.; Ghassempour, A. Changes in Biophysical Characteristics of PFN1 Due to Mutation Causing Amyotrophic Lateral Sclerosis. *Metab. Brain Dis* **2018**, *33*, 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*, 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*, 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*, 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. U.S.A.* **2006**, *103*, 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*, 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**, No. btw398.

(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*, D442−D450.

(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*, S217−S221.

(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. R. Soc. B* **2009**, *276*, 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*, 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*, No. 17119.

(53) Bhatia, S.; Vasudevan, K. Comparative Proteomics of Geographically Distinct Saw-Scaled Viper (*Echis carinatus*) Venoms from India. *Toxicon X* **2020**, *7*, No. 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*, 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*, 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*, 547−554.

(62) Pawlak, J.; Manjunatha Kini, R. Snake Venom Glutaminyl Cyclase. *Toxicon* **2006**, *48*, 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*, 317−329.

(66) Daltry, J. C.; Wuster, W.; Thorpe, R. S. Diet and Snake Venom Evolution. *Nature* **1996**, *379*, 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*, 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*, 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*, No. 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. U.S.A.* **2013**, *110*, 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*, No. 1700306.

(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*, 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.* **2015**, *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*, 486−489.

Supplementary Information

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 Mahdi Kazemi, Alireza Ghassempour, and Bernhard Spengler

Table of contents

Figure S-5. Total ion current (TIC) of *Echis carinatus sochureki* (Iran) venom from A) non-reduced and B) reduced form of the pooled populations, alongside the one-dimension gel electrophoresis (12% SDS-PAGE) of reduced form of each population venom (HO: Hormozgan; SB: Sistan&Bluchestan; SK: South Khorasan).
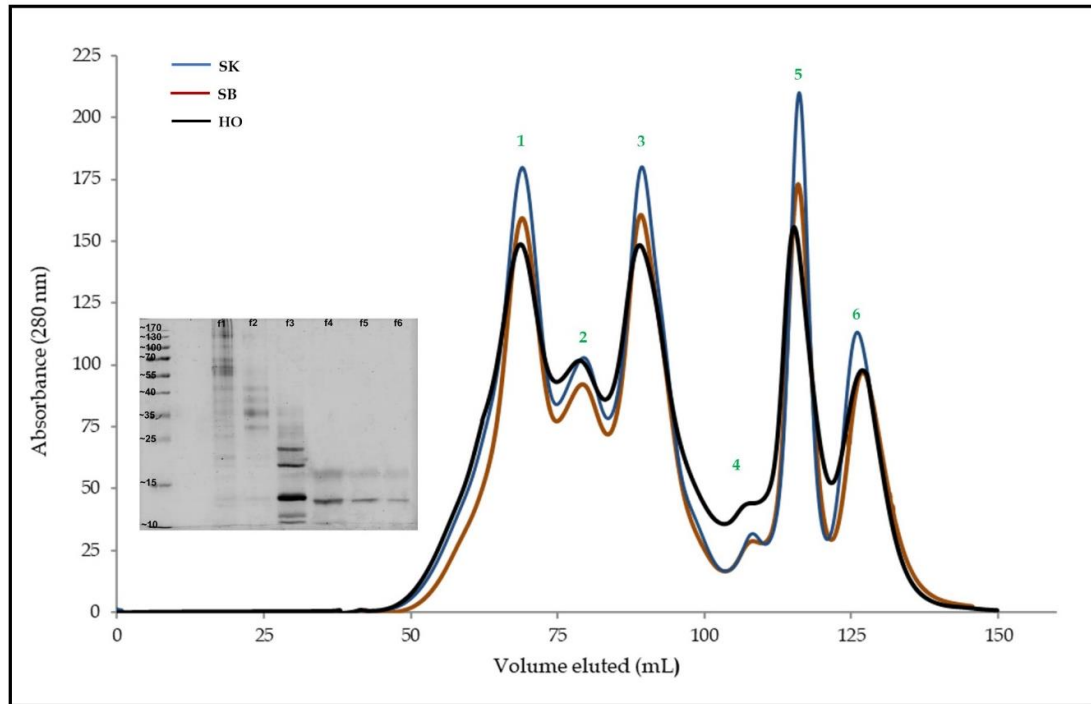
Figure S-6. SEC fractionation of each population (HO: Hormozgan; SB: Sistan&Bluchestan; SK: South Khorasan) of *Echis carinatus sochureki* (Iran) by using HiLoad 16/60 Superdex gel-filtration chromatography column (GE Healthcare Bio-Sciences, Freiburg, Germany) in a Fast Protein Liquid Chromatography (FPLC) system (Bio-rad NGC^TM Quest Plus).
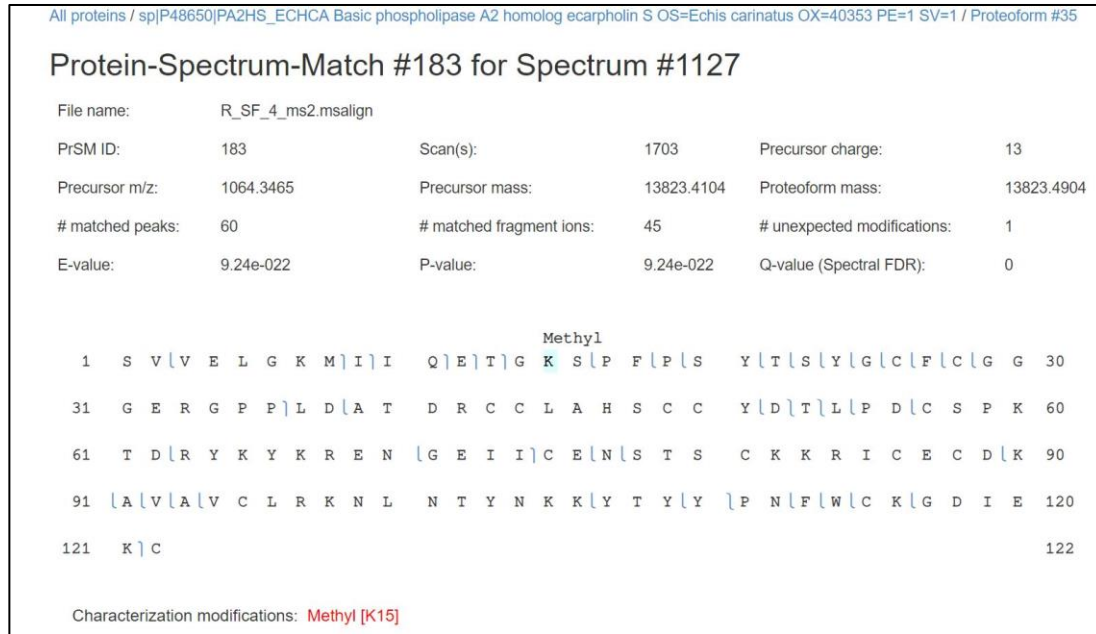
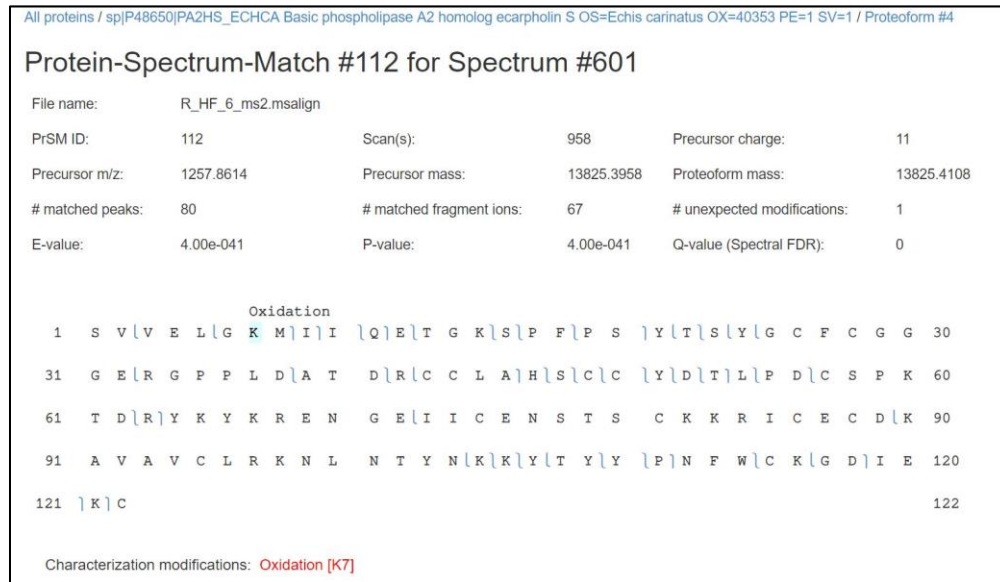**Figure S-7.** A methylated proteoform of PLA2 protein family (UniProt ID # P48650) with one methylation site (K15), identified by TopPIC software.
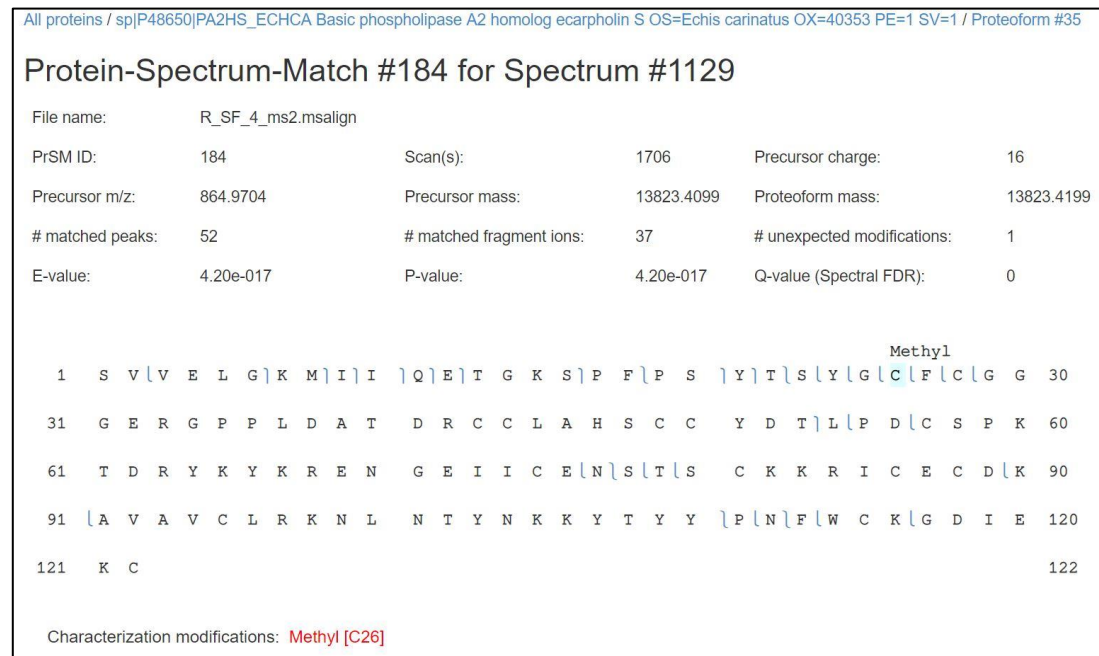

Figure S-8. A methylated proteoform of PLA2 protein family (UniProt ID # P48650) with one methylation site (C26), identified by TopPIC software.
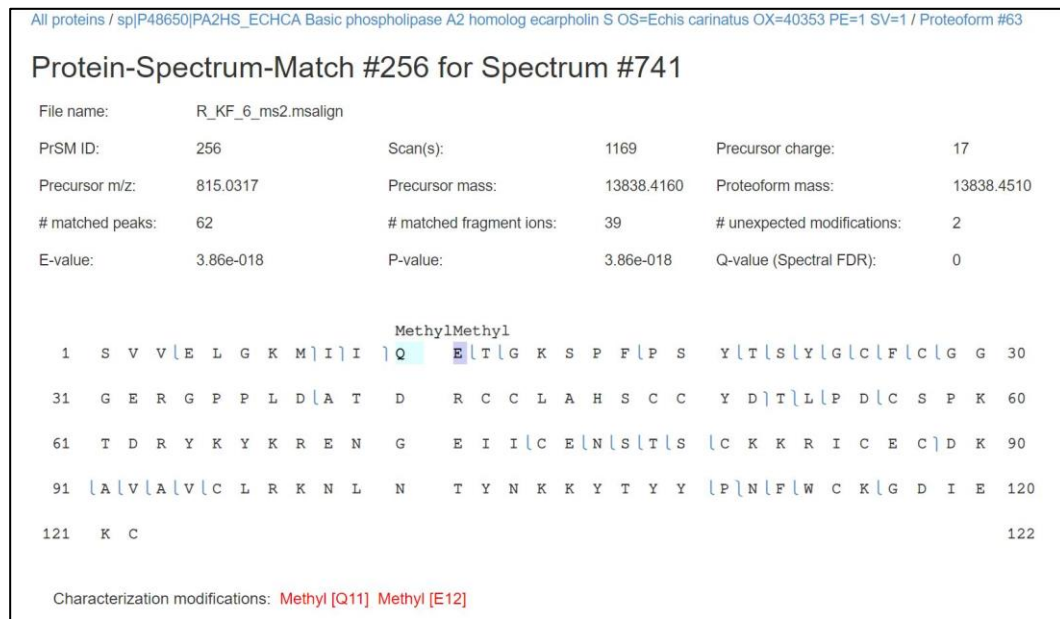
Figure S-9. A methylated proteoform of PLA2 protein family (UniProt ID # P48650) with two methylation site (Q11;E12), identified by TopPIC software.
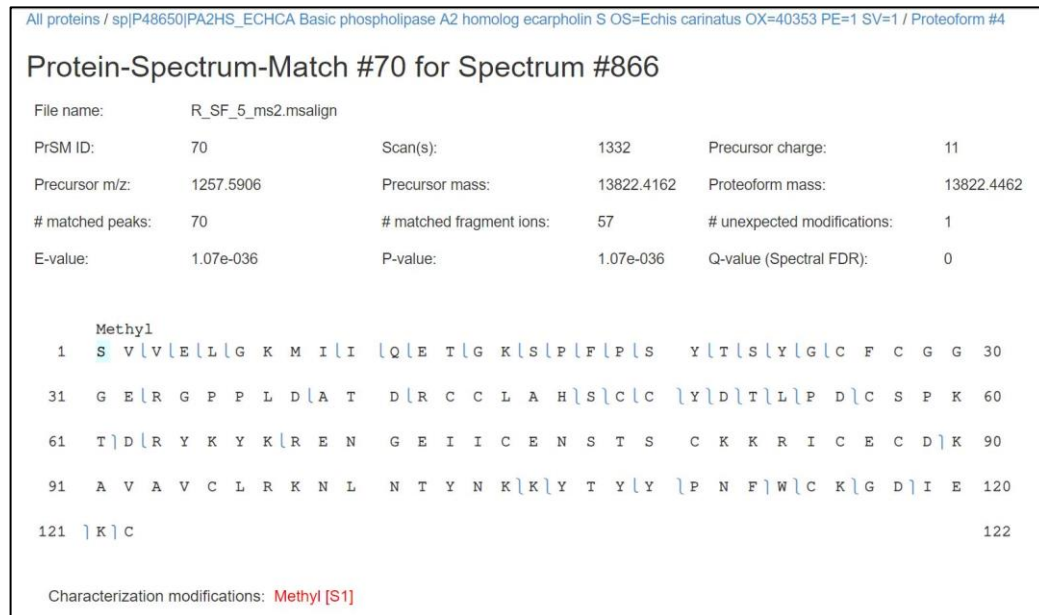
Figure S-10. A methylated proteoform of PLA2 protein family (UniProt ID # P48650) with one methylation site (S1), identified by TopPIC software.

All proteins / sp|P48650|PA2HS_ECHCA Basic phospholipase A2 homolog ecarpholin S OS=Echis carinatus OX=40353 PE=1 SV=1 / Proteoform #35

Protein-Spectrum-Match #185 for Spectrum #1130

File name:	R_SF_4_ms2.msalign				
PrSM ID:	185	Scan(s):	1707	Precursor charge:	18
Precursor m/z:	768.9748	Precursor mass:	13823.4159	Proteoform mass:	13823.4309
# matched peaks:	34	# matched fragment ions:	28	# unexpected modifications:	1
E-value:	6.28e-018	P-value:	6.28e-018	Q-value (Spectral FDR):	0

```
  1   S   V   V   E   L]  G   K   M]  I]  I   ]Q   E   T   G   K   S]  P   F]  P   S   ]Y   T   S   Y   G   C   F   C[  G   G   30

 31   G   E   R   G   P   P   L   D   A   T       D   R   C   C   L   A   H   S   C   C       Y   D   T   L[  P   D   C   S   P   K   60
                             Methyl
 61   T   D   R   Y   K   Y[  K ] R   E   N       G   E   I   I   C[  E[  N[  S[  T[  S       C   K   K   R   I   C   E   C   D   K   90

 91   A   V[  A[  V[  C   L   R   K   N   L       N   T   Y   N   K   K[  Y   T   Y   Y       [P[  N[  F[  W[  C[  K[  G   D[  I   E   120

121   K   C                                                                                                                      122
```

Characterization modifications: Methyl [K67]

Figure S-11. A methylated proteoform of PLA2 protein family (UniProt ID # P48650) with one methylation site (K67), identified by TopPIC software.

All proteins / sp|P48650|PA2HS_ECHCA Basic phospholipase A2 homolog ecarpholin S OS=Echis carinatus OX=40353 PE=1 SV=1 / Proteoform #24

Protein-Spectrum-Match #208 for Spectrum #1073

File name:	R_SF_5_ms2.msalign				
PrSM ID:	208	Scan(s):	1642	Precursor charge:	16
Precursor m/z:	866.7840	Precursor mass:	13852.4275	Proteoform mass:	13852.4325
# matched peaks:	70	# matched fragment ions:	46	# unexpected modifications:	1
E-value:	4.62e-022	P-value:	4.62e-022	Q-value (Spectral FDR):	0

```
                         Acetyl
  1   S   V   V   E   L   G   K   M   I[  I   ]Q[  E ] T] G   K   S   P   F[  P   S   ]Y[  T] S[  Y[  G[  C[  F[  C[  G   G   30

 31   G   E   R   G   P   P   L   D[  A   T       D   R   C   C   L   A   H[  S   C   C       Y   D] T] L[  P[  D[  C   S   P   K   60

 61   T   D   R   Y   K   Y   K   R   E   N       G   E   I   I[  C   E[  N[  S[  T[  S       C   K   K   R   I   C   E   C   D   K   90

 91   [A   V[  A   V[  C   L   R   K   N   L       N   T   Y]  N   K   K[  Y[  T   Y   Y       [P[  N[  F[  W   C   K[  G   D   I   E   120

121   K   C                                                                                                                      122
```

Characterization modifications: Acetyl [K7]

Figure S-12. An acetylated proteoform of PLA2 protein family (UniProt ID # P48650) with one acetylation site (K7), identified by TopPIC software.
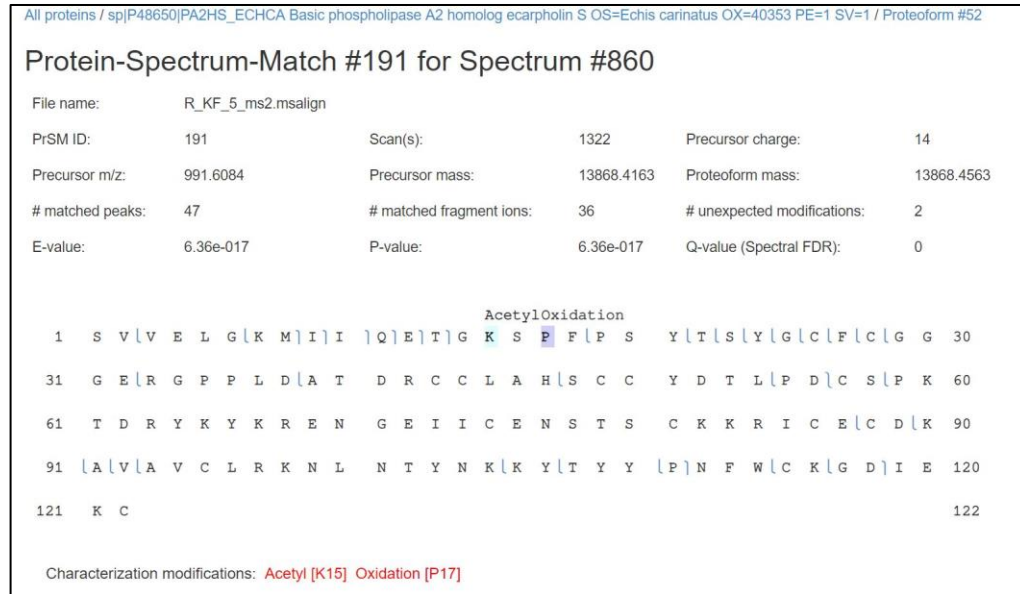
Figure S-13. A proteoform of PLA2 protein family (UniProt ID # P48650) with two modification sites (K7;P17), identified by TopPIC software.
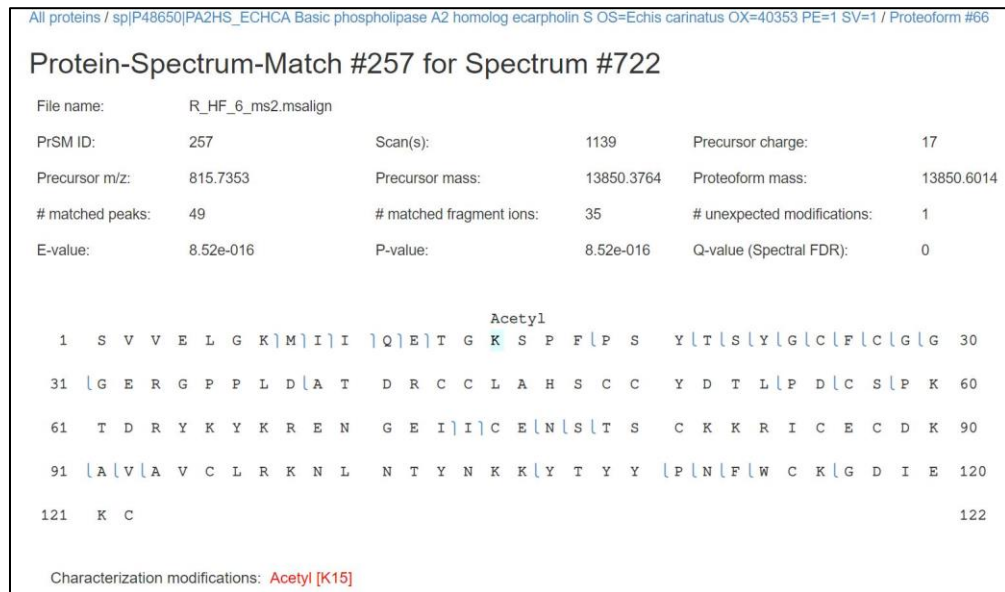
Figure S-14. An acetylated proteoform of PLA2 protein family (UniProt ID # P48650) with one acetylation site (K15), identified by TopPIC software.

Figure S-15. An oxidized proteoform of PLA2 protein family (UniProt ID # P48650) with one oxidation site (K7), identified by TopPIC software.
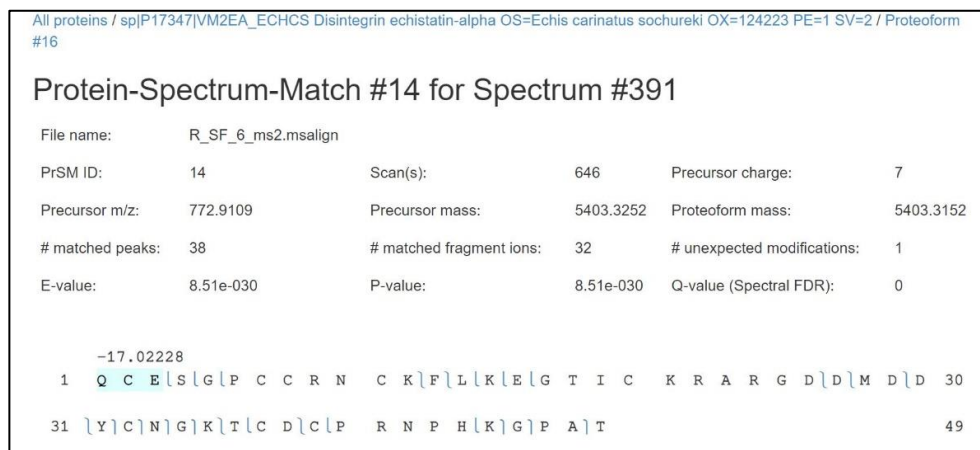
Figure S-16. A modified proteoform of disintegrin protein family (UniProt ID # P17347) with molecular mass loss of -17.02, identified by TopPIC software. The mass shift can be explained by the formation of pyroglutamic acid at the N-terminal (Q residue) of disintegrin proteoform.
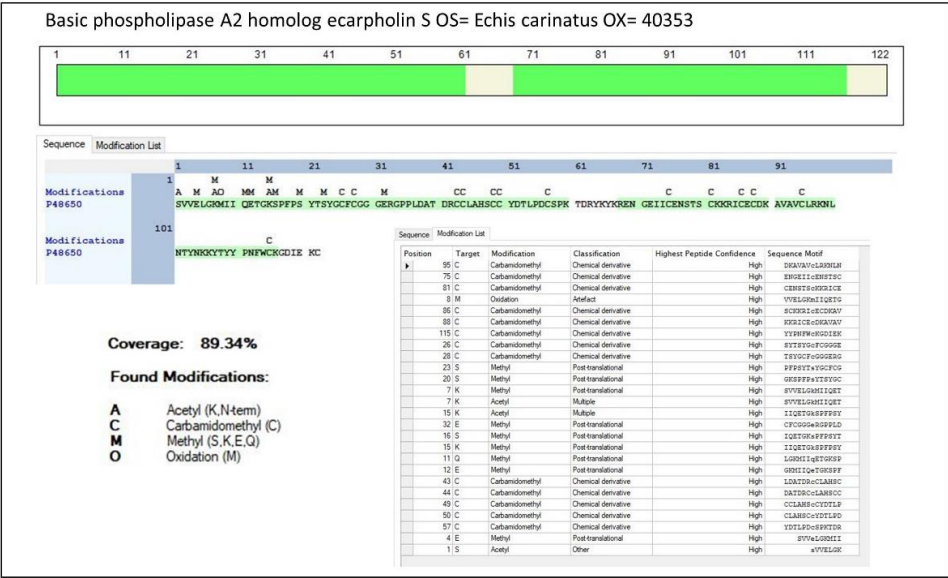
Figure S-17. PTMs identified by bottom-up proteomic approaches in PLA2 toxin family.

CHAPTER III

Venom Gland Mass Spectrometry Imaging of Saw-scaled Viper, *Echis carinatus sochureki*, **at High Lateral Resolution**

Parviz Ghezellou, Sven Heiles, Patrik Kadesch, Alireza Ghassempour, and Bernhard Spengler

Research Article

Venom Gland Mass Spectrometry Imaging of Saw-Scaled Viper, *Echis carinatus sochureki*, at High Lateral Resolution

Parviz Ghezellou, Sven Heiles, Patrik Kadesch, Alireza Ghassempour, and Bernhard Spengler*

Cite This: *J. Am. Soc. Mass Spectrom.* 2021, 32, 1105−1115

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information
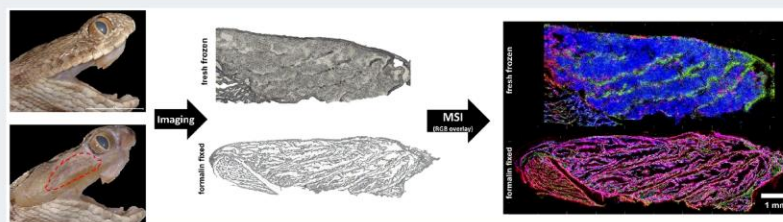
ABSTRACT: The snake venom gland is the place for the synthesis, storage, and secretion of a complex mixture of proteins and peptides, i.e., the venom. The morphology of the gland has been revealed by classical histology and microscopic studies. However, knowledge about the gland's cellular secretory and functional processes is still incomplete and has so far been neglected by the omics disciplines. We used autofocusing atmospheric-pressure matrix-assisted laser desorption/ionization (AP-SMALDI) mass spectrometry imaging (MSI) to investigate endogenous biomolecular distributions in the venom glands of the saw-scaled viper, *Echis carinatus sochureki*, employing different sample preparation methods. Fresh-freezing and formalin-fixation were tested for the gland to obtain intact tissue sections. Subsequently, MSI was conducted with 12 μm pixel resolution for both types of preparations, and the lateral distributions of the metabolites were identified. Experiments revealed that lipids belonging to the classes of PC, SM, PE, PS, PA, and TG are present in the venom gland. PC (32:0) and SM (36:1) were found to be specifically located in the areas where cells are present. The snake venom metalloprotease inhibitor pEKW (m/z 444.2233) was identified in the venom by top-down LC−MS/MS and localized by MALDI-MSI in the gland across secretory epithelial cells. The peptide can inhibit the venom's enzymatic activity during long-term storage within the venom gland. With a high degree of spectral similarities, we concluded that formalin-fixed tissue, in addition to its high ability to preserve tissue morphology, can be considered as an alternative method to fresh-frozen tissue in the case of lipid and peptide MS imaging in venom gland tissues.

■ INTRODUCTION

Since the introduction of matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) in 1994,[1] the method has further developed to become a powerful technique in bioanalytical sciences. Modern instruments combine the capabilities of high mass resolution, mass accuracy, and lateral resolution.[2,3] MALDI MSI has received considerable attention for the ionization and detection of molecules with a limit of detection in the attomole range, enabling the identification and localization of biological compounds such as metabolites, lipids, peptides, and proteins directly from tissue surfaces.[4] Despite the remarkable instrumental advances in MSI, obtaining high spatial resolution is the limiting factor to visualize the subcellular distribution of biomolecules.[4,5] However, some studies have improved the spatial resolution of MALDI MSI to 5−10 μm, enabling single-cell MSI.[6,7]

Most recently, our group developed an atmospheric-pressure scanning microprobe MALDI (AP-SMALDI)[8] MSI setup with a lateral resolution down to 1.4 μm at a mass resolution of more than 100 000[9] and autofocusing operation mode, which enables the chemical investigation of rough and three-dimensional surfaces.[10] The autofocusing mode allows for high lateral resolution MALDI experiments of nonplanar objects by eradicating topography-related signal artifacts. These technical advances opened the field to toxinology and biology by enabling the visualization of endogenous biomolecular compounds from the venom gland tissues of animals. Given that the venom gland is the place for the synthesis, storage, and secretion of venomous substances,[11] comprehensive molecular distribution topography

1105

could help explain the mechanisms behind tissue activation and toxin production on the cellular level. Trying to meet the goals have recently led to using MSI methods as an ideal imaging technique to interrogate the spatial distribution of venom components inside a few venomous animals' glands, e.g., centipedes (*Thereuopoda longicornis, Scolopendra morsitans,* and *Ethmostigmus rubripes*),[12] honeybee (*Apis mellifera*),[13] sea anemones (*Oulactis muscosa* and *Actinia tenebrosa*),[14,15] the brown forest cobra (*Naja subfulva*),[16] and the fire ant (*Solenopsis invicta*).[17]

One of the most challenging tasks for applying MSI to biological samples is maintaining the samples' integrity concerning morphology and molecular species distribution.[19] Sample preparation directly affects the identity, distribution, and abundance to be probed in MSI studies. In general, snap-freezing and chemical-fixation methods are used for sample preparation. The fixation of tissue samples by formalin is widely employed, particularly in clinical specimens, to prevent tissue destruction and degradation.[18] The preservation process leads to cross-links formation of proteins and amine-containing metabolites mainly by a formaldehyde reaction, known as methylene bridges.[20] Recent MSI- and LC−MS^E-based metabolic studies have shown that the molecular profiles obtained from the analysis of formalin-fixed tissues are considerably different from those commonly produced from fresh-frozen ones.[21−25] Differences were especially found in terms of the depletion of peptides and amine-containing phospholipids, PS and PE. Nevertheless, most metabolite species were still conserved and detectable by MS after formalin fixation, making the method suitable for lipidomics research.[21,22,26−29] In general, there is little information about the chemically induced modifications of biomolecules preserved in formalin-treated tissues compared with untreated-frozen ones.

In this study, the venom gland tissues of saw-scaled or carpet vipers, *Echis carinatus sochureki*, were used as a model organism to develop a sample preparation strategy of venomous animals. Vipers' venom apparatus consists of the bilaterally paired venom glands, primary ducts, accessory glands, and secondary ducts, which deliver venom to the hollow fang for injection into the victims (Figure S1). The main venom glands are located behind the snake's eyes and are surrounded by connective tissue. The glands are made up of the tubular cisternae, coated with secretory epithelial cells. These cells play an important role in the synthesis of the secretory products and their release *via* secretory vesicular exocytosis for storage in the ductules and main gland lumina.[11] Because the venom gland is a soft and fragile organ containing potentially active enzymatic compounds as a venom, it is extremely challenging to transport and prepare tissue samples without degradation. Therefore, efficient sample preparation is critical before collecting valid data of the gland by MALDI MSI.

Here, we developed a workflow to investigate the spatial arrangement of metabolites within the venom gland tissue using AP-SMALDI MSI. For this purpose, two sample preparation strategies, formalin-fixation and fresh-freezing, were implemented for the venom glands of *E. carinatus*. Resulting tissue sections were imaged with lateral resolutions down to 12 μm, allowing for localizing and identifying a bioactive metalloprotease inhibitor peptide and small metabolites from venom gland tissue. Besides, by comparing the MSI data obtained from both types of sample preparations as well as the supporting data achieved by LC−MS/MS of the fixed tissue

lipidome, we found that the formalin-fixed method, with its high ability to preserve tissue morphology, provides data regarding lipids and peptides imaging comparable to those obtained from the fresh-frozen tissues of the venom gland.

■ MATERIALS AND METHODS

Ethical Statement. Snake specimens used in this study were maintained under standard conditions, and all experiments were performed according to the international guiding principles involving animals for scientific research[30] and under approvals granted by the Ethics Committee of Shahid Beheshti University.

Samples. Field sampling of the saw-scaled viper, *Echis carinatus sochureki*, specimens took place in Iran. Venom milking was 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. The samples were pooled, lyophilized, and stored at −80 °C for future research.

Two specimens were milked and then anesthetized (after 1 week) with inhaled halothane prior to euthanasia and subsequent dissection of the venom gland *via* the use of a stereomicroscope (Leica; Wetzlar, Germany). One of the dissected venom glands was frozen in liquid nitrogen immediately and stored at −80 °C until sectioning, and the other gland was fixed in neutral buffered formalin (10% V/V) for 4 h at 4 °C and then kept at −80 °C until further analysis.

Materials. All MS-grade solvents including acetonitrile (ACN), water (H_2O), methanol (MeOH), and isopropanol (IPA) as well as glass microscope slides (ground edged, frosted, 75 × 25 × 1 mm^3) were purchased from VWR chemicals (VWR international, Darmstadt, Germany). 2,5-Dihydroxybenzoic acid (DHB, 98% purity), trifluoroacetic acid (TFA; LC−MS grade), formic acid (FA; LC−MS grade), methyl-*tert*-butyl ether (MTBE), ammonium formate (AF), and formalin solution (neutral buffered, 10% V/V) were purchased from Fluka Sigma-Aldrich (Sternheim, Germany).

Mass Spectrometry Imaging (MSI). Venom glands were placed into a cryochamber at −21 °C for 15 min and cut into longitudinal sections of 20 μm thickness with a cryotome (HM525 cryostat, Thermo Scientific, Bremen, Germany). The sections were freeze-mounted on glass slides and dried in a desiccator for 30 min to avoid water precipitation. Digital light microscopic images of venom gland sections were recorded (VHX-5000; Keyence, Osaka, Japan) prior to matrix application. A high-resolution matrix preparation system (SMALDI Prep, TransMIT GmbH, Giessen, Germany) was used for the homogeneous deposition of 70 μL of freshly prepared DHB matrix solution (30 mg/mL in 50:50 acetone/H_2O, 0.1% FA). The matrix solution was sprayed with a flow rate of 10 μL/min, and the samples were rotated at 500 rpm.

All MSI experiments were performed using an autofocusing AP-SMALDI5 AF ion source (TransMIT GmbH, Giessen, Germany) coupled to a Q Exactive HF mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). The venom gland tissue sections were scanned with step sizes of 12 μm. The desorption and ionization of the sample components were performed by applying 50 UV-laser pulses per pixel at a repetition rate of 100 Hz. The mass spectrometer was operated in positive-ion mode, the mass range was set to m/z 300−1200, and the mass resolution setting was 240 000 at m/z 200. A fixed injection time of 500 ms was set, combined with using a lock mass of m/z 716.12451 ([5 DHB − 4 H_2O + NH_4]$^+$) to yield a mass accuracy of less than 2 ppm. The S-lens level was set to 100

arbitrary units, the acceleration voltage was 3.0 kV, and the capillary temperature was 250 °C. Additionally, on-tissue MS/MS imaging experiments were performed to fragment the precursor ion at m/z 444.22 ± 0.2 Da in positive-ion mode with a mass resolution of 240 000 at m/z 200 and higher-energy collisional dissociation (HCD) at normalized collision energy (NCE) set to 25 arbitrary units for peptide identification.

MSI Data Analysis. MSI raw files were processed with the in-house developed imaging software package MIRION (v3.2.64.29)[31] with parameters set for a histogram bin width of 0.004 u and an absolute mass variance of spectra of 0.005 u. Alternatively, raw files were converted to imzML file format using imzML-converter (v1.3).[32] The mass spectra recorded at 12 μm pixel size were obtained with a mass measurement accuracy of less than 2 ppm. The red−green−blue (RGB) overlays and mass accuracy plots (signal intensity vs m/z) were generated using MSiReader (v1.0).[33] Additionally, molecular annotations were performed using the METASPACE[34] platform against the public Human Metabolome Database (HMDB)[35] and LipidMaps[36] database. The MS/MS spectra were analyzed by an in-house computer program, Peptide Composer version 1.2,[37] and compared with LC−MS/MS data.

Liquid Chromatography−Mass Spectrometry (LC−MS/MS) of the Venom Proteome. In the first step, the whole crude venom of *E. carinatus* was analyzed by top-down proteomics according to the protocol described before.[38] Subsequently, the crude venom was fractionated employing 10 kDa molecular weight cutoff (MWCO) Slide-A-Lyzer dialysis cassettes (Thermo Fisher Scientific, Pierce Biotechnology, Rockford, IL) and further washed and concentrated using an Amicon 10 kDa ultracentrifuge filter (Merck Millipore) for 15 min at 14 000g (4 °C). The sample was separated using an UltiMate 3000 RSLC UHPLC system (Thermo Fisher Scientific, Bremen, Germany) equipped with a Jupiter C18 (4.6 mm × 250 mm, 3 μm 300 Å particle size) column (Phenomenex, Torrance, CA) coupled to a Q Exactive (Thermo Fisher Scientific, Bremen, Germany) mass spectrometer. Chromatographic analysis was performed at a flow rate of 500 μL/min, using water/0.1% FA (mobile phase A) and ACN/0.1% FA (mobile phase B). The gradient of 70 min was applied as follows: isocratically (2% B) for 3 min, 2−40% B over 67 min, and re-equilibration in 2% B. The mass spectrometer was operated in data-dependent acquisition (top-3 DDA) with the following parameters in full MS scans: mass range m/z 300−500, mass resolution of 70 000 at m/z 200, AGC target of 1 × 10^6, and IT of 120 ms. MS/MS scans were performed for the targeted peptide, tripeptide pyroglutamate-lysine-tryptophan, (pEKW, m/z 444.22) in the inclusion list with parameters set to a mass resolution of 30 000 at m/z 200, AGC target of 1 × 10^5, IT of 120 ms, isolation window m/z ± 1.3, and NCE of 35.

LC−MS/MS of the Venom Gland Lipidome. The lipidome of formalin-fixed venom gland tissue was extracted using the MTBE extraction procedure[39] with some modification. Briefly, 300 μL of cold MeOH was added to 70 mg of sample, and the mixture vortexed for 5 min and incubated on ice for 10 min. Subsequently, 1 mL of cold MTBE was added to the solution, vortexed, and sonicated for 5 min with ice. The solution was incubated for 1 h on a Thermomixer (Eppendorf Thermomixer C, Hamburg, Germany) at 2 °C with 950 rpm. For the next step, 250 μL of cold water was added, incubated for 5 min, and cold centrifugated (Beckman Coulter) for 10 min at 12 000 rpm. The upper organic layer was transferred to a new precooled microtube, and the whole lipid extraction process was repeated again for the rest. The organic layers were vaporized in ice using a nitrogen flow, and the dry lipid sample was kept at −80 °C until further use.

The extracted lipidome was separated on an analytical Kinetex C18 (2.1 mm × 100 mm, 2.6 μm 100 Å particle size) column (Phenomenex, Torrance, CA) connected to a Thermo Scientific Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific, Bremen, Germany). Mobile phase A was a mixture of ACN/H$_2$O (60:40), and mobile phase B was a mixture of IPA/ACN/H$_2$O (90:8:2). Both mobile phases contained 10 mM AF and 0.1% FA. The flow rate was set to 250 μL/min, and the gradient elution started at 20% mobile phase B, rising to 30% B over 4 min, 45% B over 2 min, 60% B over 4 min, 65% B over 4 min and held for 4 more minutes, and 90% B over 13 min, and the column was re-equilibrated with 20% B for 10 min before the next injection.

A heated electrospray ionization source (HESI II) was used for the ionization of the samples in positive-ion mode connected to a Q Exactive HF-X Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). The ion source was tuned as follows: spray voltage of 3.8 kV, source temperature of 325 °C, capillary temperature of 300 °C, sheath gas of 60, auxiliary gas of 20, and sweep gas of 0. Each sample was measured in the positive polarity using data-dependent acquisition (Top-10) with the following parameter values in full MS scan: mass range m/z 200−1200, mass resolution of 60 000 (@ m/z 200), AGC target of 5 × 10^6, IT of 75 ms, and MS/MS scans: mass resolution of 60 000 (@ m/z 200), AGC target of 5 × 10^6, IT of 175 ms, isolation window m/z ± 1, dynamic exclusion of 6 s, and stepped NCE levels of 20−25−30.

The raw files were converted to mzXML format using MSconvert,[40] and processing was performed by MZmine version 2.[41] Lipids were identified *via* LipidMatch R-based software.[42]

■ **RESULTS AND DISCUSSION**

The venom gland is soft and fragile, making the task of sample preparation very challenging for MSI and histological studies. Therefore, the qualified preparation steps are crucial to collecting valid data from the morphologically "original state" of gland sections. For this reason, we employed two sample preparation approaches to help minimize undesirable changes in the venom gland tissue. In a first approach, we applied snap-freezing to the venom gland, sinking it in liquid nitrogen immediately after removing the gland from the snake. The sample was cut into 20 μm thick longitudinal sections, and light microscopic images were recorded. Figure 1E (right) shows a fresh-frozen tissue section, with secretory epithelium (SE) and ductules (Du). Although in MSI studies, the snap-freezing of samples is preferred over chemical fixation due to concerns about the chemical process at tissues' molecular content;[26] the method causes a morphological distortion in most tissues.[18] Here, the rapid freezing disrupted the cell integrity of the gland, leading to poor morphological preservation by fragmenting the inner tissue during sectioning (for more details, see also Figure S2A). Therefore, another gland was fixed in neutral buffered formalin for a short time (4 h) quickly after dissecting the tissue to conserve the morphologic integrity of the substructure (SE and Du). Following 4 h of formalin fixation, the morphology of the venom gland was preserved better than seen on the snap-frozen sections, and the tissue integrity was maintained even after sectioning (Figure 1E, left, and Figure S2B). As can be seen from Figure 1E, implementing formalin treatment increased the
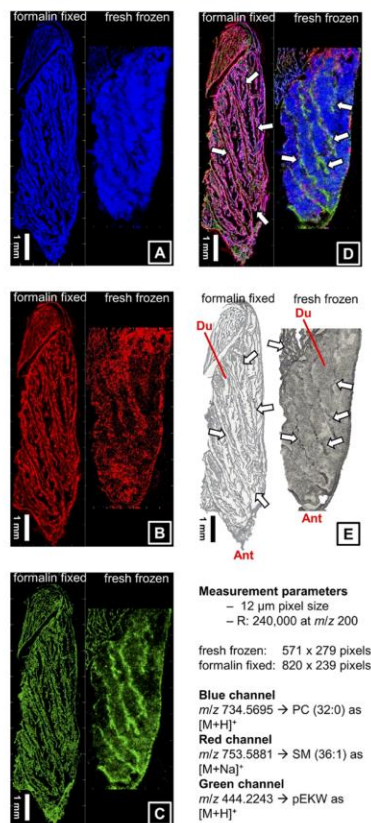
Figure 1. Snake venom gland mass spectrometry imaging (sVG-MSI). Positive-ion AP-SMALDI of glycerophospholipid, sphingolipid, and peptide from venom gland tissue sections of a saw-scaled viper (*Echis carinatus sochureki*) species. (A) AP-SMALDI image of the $[M + H]^+$ ion signal of phosphatidylcholine PC (32:0) with $\Delta m/m \leq 2$ ppm. (B) $[M + Na]^+$ signal of sphingomyelin SM (36:1) with $\Delta m/m \leq 2$ ppm (C) $[M + H]^+$ ion signal of the tripeptide pEKW with $\Delta m/m \leq 1$ ppm (fresh-frozen) and with $\Delta m/m \leq 5$ ppm (formalin fixation). (D) Image overlay of tripeptide (green), PC (blue), and SM (red). (E) Light microscopic images of fresh-frozen and formalin-fixed venom glands.

tissue section quality. For MSI analysis, the DHB-matrix was deposited onto the tissue by pneumatic spraying and, subsequently, the sample surfaces were scanned with a 12 μm step size in positive-ion mode. Because of the physical conditions of tissue sections, i.e., the large size (>1 cm length and >3 mm width) as well as the rough surface, we used pixel-by-pixel autofocusing to eliminate topography-related signal artifacts.

Fresh-frozen and formalin-fixed sections were analyzed with identical experimental parameters, which led to image sizes of 571×279 and 820×239 pixels, respectively (see Figure 1). To compare metabolite annotations between both procedures, we utilized the METASPACE online data repository for unsupervised signal annotation, using HMDB and LipidMaps databases (FDR = 10%). In total, 17/40 metabolites (HMDB/LipidMaps) were annotated for signals found in the fresh-frozen section and 58/100 compounds (HMDB/LipidMaps) were assigned after formalin-fixation. Mostly, lipids were detected belonging to the classes of phosphatidylcholine (PC), sphingomyelin (SM), phosphatidylethanolamine (PE), phosphatidic acid (PA), and triglyceride (TG). The complete annotated data sets from each venom gland sections are shown in the Tables S3−S6 and also can be accessed *via* the METASPACE online platform (https://metaspace2020.eu).[34] In addition, the Venn diagrams of annotated metabolites (Figure S3) display the overlap of assigned sum formulas between the two section types. Among them, five sum formulas were annotated in all four data sets and translated into possible molecules, including the lipid molecular ions at m/z 703.5748 ($[C_{39}H_{79}N_2O_6P + H]^+$ assigned to PE-Cer and SMs), m/z 725.5568 ($[C_{39}H_{79}N_2O_6P + Na]^+$ assigned to PE-Cer and SMs + Na$^+$), m/z 746.6058 ($[C_{42}H_{84}NO_7P + H]^+$ assigned to PEs and PCs), m/z 835.6663 ($[C_{47}H_{93}N_2O_6P + Na]^+$ assigned to SMs), and m/z 882.6194 ($[C_{47}H_{90}NO_{10}P + Na]^+$ assigned to PSs). The detailed information on them are presented in Table S1.

Besides assigning known metabolites automatically by METASPACE database searches, MS spectra were also interpreted to evaluate mass spectrometric images manually using the MIRION software package. Representative ion images are shown in Figure 1 for both fresh-frozen (right) and formalin-fixed (left) tissue sections. As illustrated in Figure 1A, the MS ion at m/z 734.5695 (blue) was assigned to the protonated molecular species of PC (32:0), a phosphatidylcholine with a total number of 32 carbon atoms in the fatty acyl chains and no double bond (for RMS plot, see Figures S4 and S5). PCs are common structural lipids of eukaryotic cell membranes.[43,44] The associated signal has a distinct distribution, visualizing all cell walls of secretory epithelium within the venom gland. Figure 1B shows the signal at m/z 753.5881 (red), assigned to a sphingomyelin lipid species SM (36:1) as sodiated molecular ion (for RMS plot, see Figures S4 and S5). Sphingomyelins are also known as a major constituent of the cellular membranes of animal tissues and play a critical role in signal transduction, cell differentiation, and metabolism.[44,45] Nonetheless, in-depth metabolomics investigation on snake venom glands can help to shed light onto the role of sphingomyelin in the gland.

As observed by optical microscopy, snap-freezing damaged the gland, leading to cell-wall tearing in the tissue section while cutting. Consequently, large amounts of cell debris were widely observed in most fresh-frozen tissue areas (Figure 1E). This is consistent with the MS images of structural lipids that appear to be located in ductules known to be empty cavities for venom storage (Figure 1, right). However, in the formalin-fixed tissue images, they were found only in the tubules' areas where the secretory epithelial cells are located (Figure 1, left). In an RGB overlay, the colocalization of the aforementioned ions becomes more apparent (see Figure 1D).

To evaluate the general quality of the mass spectra, the peak intensities and signal-to-noise ratios were compared between the recorded spectra of both tissue preparation methods, fresh-freezing vs formalin-fixation. As shown in Figure 2A,B, the common lipid species, of both type tissues, were selected for the assessment. We found that the intensities of lipid peaks were closely similar in the spectra independent of preparation types.
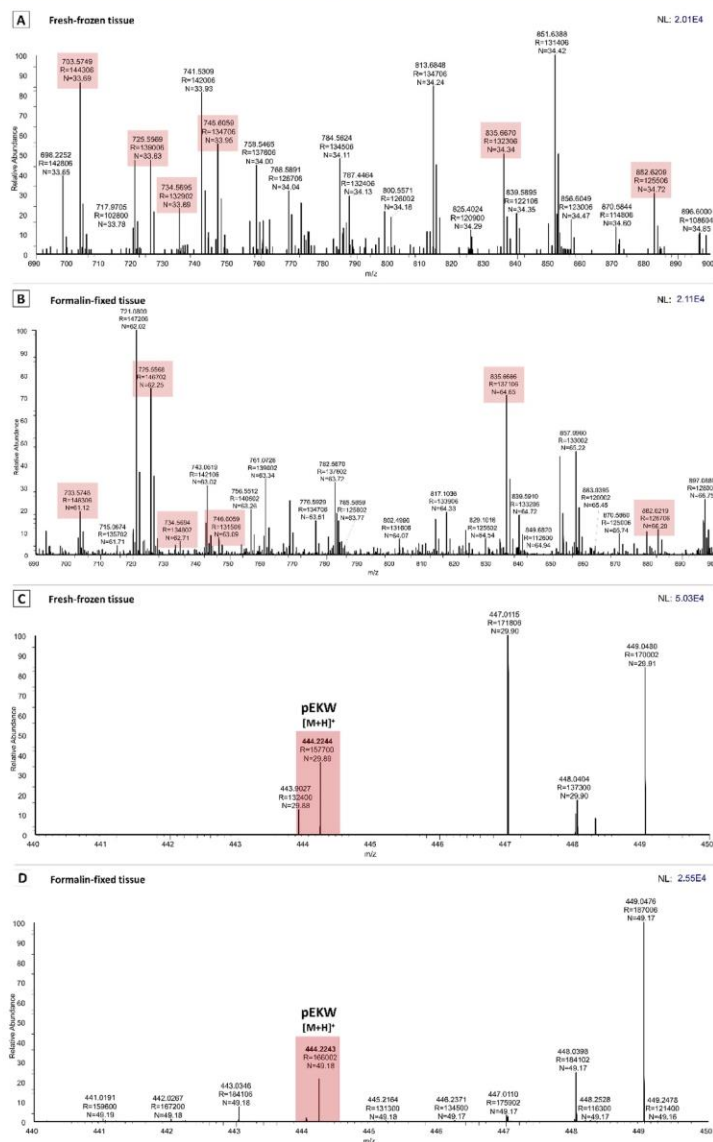
Figure 2. Positive-ion single-pixel (12 μm) mass spectra of saw-scaled viper, *Echis carinatus sochureki*, venom gland sections obtained by MALDI-MSI. Mass spectra in positive-ion mode at a mass range of m/z 690–900 are shown for (A) fresh-frozen and (B) formalin-fixed tissue sections (refer to text for lipid annotation of highlighted m/z peaks). Mass spectra in positive-ion mode of m/z 440–450 are illustrated for (C) fresh-frozen and (D) formalin-fixed tissue sections. R, resolution; N, signal-to-noise ratio; NL, base peak intensity (normalization level).

However, the corresponding signal-to-noise ratios of lipids in the formalin-fixed gland section were roughly 2-fold higher than in spectra of fresh-frozen tissue. This can be explained by ion suppression in freshly frozen tissue due to more numerous components[21] and dissolving and wash-out of some metabolites (mostly polar) in formalin solution, leading to the dilution of

71
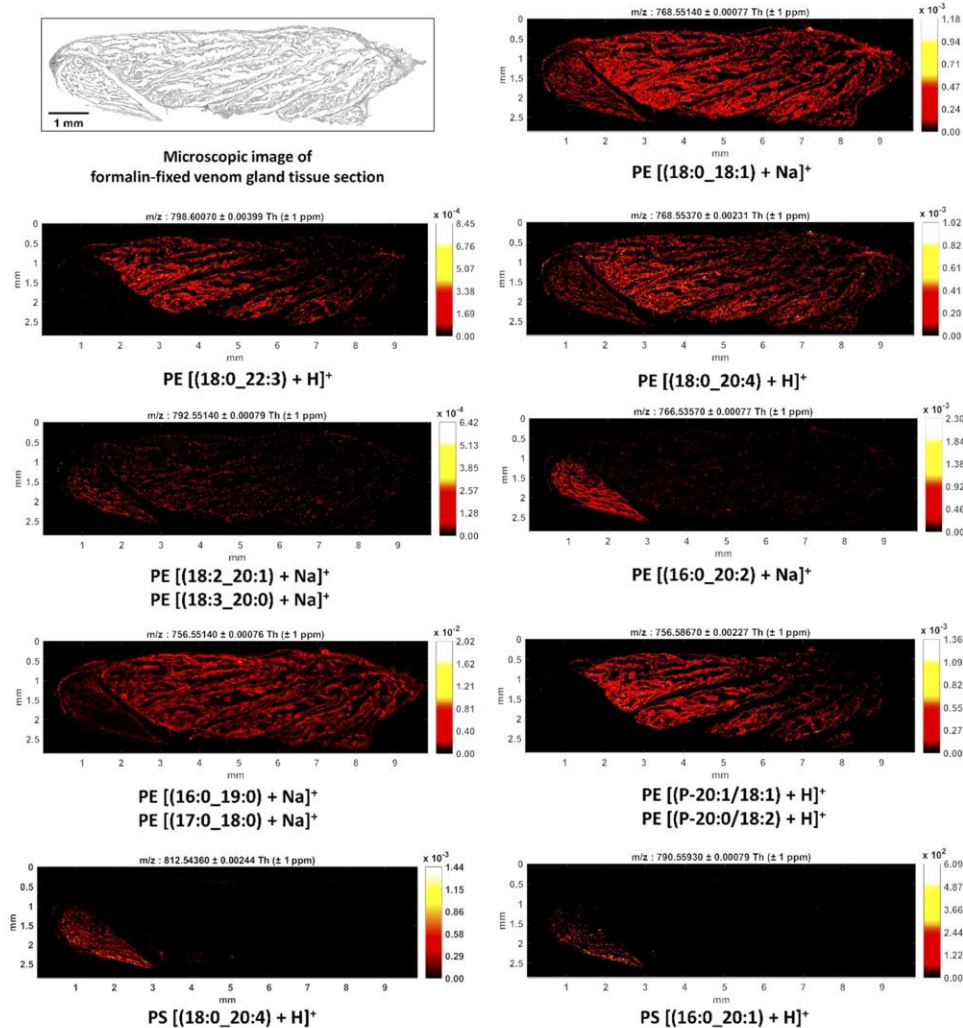
Figure 3. Positive-ion images of selected PE and PS species found *via* MALDI-MSI and confirmed by LC−MS/MS on formalin-fixed venom gland tissue section. Information on HCD fragmentation is present in Figures S9−S17.

compounds.[28] Additionally, since recent studies reported that no PE and PS detection occurred in formalin-fixed brain tissue[23] due to cross-linking with proteins and metabolites,[22] we expected a complete loss of MS detection of these species in the formalin-fixed venom gland. However, our results showed that it is possible to detect these lipids in the fixed tissue, although we experienced some reduction in their ion signals but not a complete loss, as one can see in the MSI images shown in Figure 3. To support the MSI annotations of PE and PS in the formalin-fixed tissue, we also applied untargeted high-resolution LC−MS/MS of the extracted lipidome from the fixed venom gland tissue. The positive-ion mode analysis revealed the presence of MSI-assigned lipids species using MS/MS rule-based annotation. The fragmentation data and lipid annotations are shown in Figures S9−S17. Data retrieved from the LC−MS/MS study, accompanied by data obtained from the MSI investigation, are displayed in Figure 3. It seems that fixing the snake venom gland in formalin with the herein reported workflow can be an option for profiling lipids in tissue, as alternative to fresh-freezing.

Figure 4. MS² spectra of the tripeptide pEKW with a precursor molecular weight of *m/z* 444.22. (A) LC−MS/MS fragment pattern of selected tripeptide precursor ion (*m/z* 444.22). (B) On-tissue (formalin-fixed venom gland) AP-SMALDI MS² spectra of selected *m/z* 444.22 ± 0.25 Da. (C) On-tissue (fresh-frozen venom gland) AP-SMALDI MS² spectra of selected *m/z* 444.22 ± 0.25 Da. The sequence-specific a-, b-, y-, and z- fragments were annotated manually, and the asterisk symbols (*) indicate immonium ions of respective amino acids. The chemical structure of the tripeptide pEKW was constructed using ChemDraw Professional software (version 17.1).

For the further characterization of the two different types of tissue preparations, we also searched for specific components.

For example, we found four ion signals across the fresh-frozen section, belonging to the class of lyso-PC (LPC). The ion images

73

of all LPC annotations are indicated in Table S2. LPC can be derived from the hydrolysis of phosphatidylcholines by the enzymatic action of phospholipase A2 (PLA2).[46] Given that PLA2 is one of the major toxin components of *E. carinatus sochureki* venom,[38,47] the presence of LPC in the frozen venom gland can be explained as an effect involving the hydrolysis of membrane phosphatidylcholine. These LPC signals did not show up in formalin-fixed tissue sections and, therefore, are assumed to be degradation-related artifacts resulting from insufficient sample preservation by snap-freezing alone. Nevertheless, the comprehensive lipidomic analyses of the two types of venom gland samples are required to validate this hypothesis.

It is interesting to note that MSI maped the spatial distribution of ions, which only appeared in the upper right region of the formalin-fixed venom gland section (Figure S6). We found that they belong to the group of Hexosylceramides (HexCer), including m/z 806.6480 [HexCer-NS(d18:1/22:0) + Na]$^+$, m/z 750.5854 [HexCer-NS[(d18:1/18:0) + Na]$^+$, m/z 834.6793 [HexCer-NS[(d18:1/24:0) + Na]$^+$, and m/z 832.6637 [HexCer-NS[(d18:1/24:1) + Na]$^+$. HexCers are amphiphathic molecules, a hydrophilic sugar part and a lipophilic ceramide anchor, present in all eukaryotic organisms and considered to be involved in many cellular activities, e.g., energy homeostasis, cellular interactions, and signaling.[48−50] They are also known as essential molecules for preventing epidermal water loss, protecting the skin from dryness.[57] However, their specific functions and mechanisms remain unclear.

Apart from lipid screening, we studied the snake venom proteome and the venom gland MSI for peptide profiling and the possible influence of the two tissue preparation methods on peptide imaging. In this respect, the whole and the fractionated crude venoms of *E. carinatus sochureki* were investigated by top-down LC−MS/MS, verifying the existence of an endogenous tripeptide pyroglutamate-lysine-tryptophan (pEKW, m/z 444.2233 [M + H]$^+$) (see Figures S7 and S8). The peak area of pEKW occupied roughly 6.68% of the total peak area, representing a significant concentration of the peptide inside the venom. This short peptide was also detected previously from closely related species, African saw-scaled viper (*Echis ocellatus*), venom and has been demonstrated to be a metalloprotease inhibitor.[51] Generally, venomous snakes are adapted to endure long periods without food, in which case their venom may even be stored in the gland for a long time without use.[58] Given that snake venom contains a variety of degradative enzymes such as phospholipase A2, metalloprotease, serine proteinase, and hyaluronidase, its enzymatic activities seem likely to be inhibited during storage in the venom gland, to prevent self-intoxication. For this reason, various physiological features like a high concentration of citrate, ionic strength, pH, and the production of endogenous peptides are thought to work together to inhibit venom enzymes and, therefore, protect against autotoxicity.[52] Recent studies have shown that the venom of vipers comprises significant amounts of short peptides known as pyroglutamic tripeptide inhibitors, pEKW (m/z 444.22), pEQW (m/z 444.18), pERW (m/z 472.22), and pENW (m/z 430.17).[53−56] These endogenous peptides serve as snake venom metalloprotease inhibitors (SVMPIs). Snake venom metalloproteinases (SVMPs) are among the most abundant Viperidae snake venom components, ergo severe hemorrhage and local tissue damage appear in the victims envenomated by vipers. We also found that SVMPs are the major toxin classes in the venom of

Iranian *E. carinatus sochureki*, jointly with the phospholipase A2 (PLA2) and snake venom serine protease (SVSP).[38]

In present study, positive-ion AP-SMALDI-MSI revealed the presence of pEWK, m/z 444.2243, in both types of venom gland sections (Figure 1C, in green; for RMS plot, see Figures S4 and S5). To prove the identity of the tripeptide directly from the venom gland, on-tissue AP-SMALDI MS^2 measurements were conducted, fragmenting the precursor ion (m/z 444.22 ± 0.25 Da) across both fresh-frozen and formalin-fixed tissues. The tandem mass spectra obtained by on-tissue MS/MS indicated product ions comparable to those observed by LC−MS/MS of the crude venom (Figure 4B,C) and, thus, provide confirmation of the tripeptide, pEKW, within the determined topography. Here, MALDI imaging demonstrates that the tripeptide, pEKW, is spatially distributed across the tissue in close proximity to secretory cells (Figure 1C,D), thus supporting the idea that the peptide protects secretory tissue from enzymatic proteolysis by SVMPs. This view can be expanded to give an idea of using the tripeptide protease inhibitor as a neutralizing agent to prevent SVMP-induced hemorrhage in viper bite victims.

By referring to Figure 2C,D, except for the signal-to-noise ratios, which are higher in the formalin-fixed gland section, the intensities of the peptide peaks were intimately comparable in the spectra of both types of preparation methods. As mentioned, this observation was also made for lipid species. Accordingly, with such a high degree of spectral similarities and its high ability to preserve tissue morphology, formalin-fixation can be proposed as an advantageous method for lipid and peptide imaging in venom gland tissues.

■ CONCLUSION

For the first time, we applied mass spectrometry imaging to snake venom gland tissues by implementing two different tissue preparation approaches, fresh-frozen and formalin-fixation. We propose to use formalin processing as a convenient conserving method for large and delicate snake venom gland tissues. Using formalin treatment enhanced the quality of venom gland tissue sectioning and provided compound distributions in the tissue without small molecule wash-out. The MSI study of saw-scaled viper, *Echis carinatus sochureki*, venom gland tissue revealed the localization of lipids as well as of tripeptide pEKW, a snake venom metalloprotease inhibitor, with high spatial and high mass resolution. Together, these results indicate that our autofocusing AP-SMALDI MSI strategy is well-suited for the distinctive mapping of lipids, metabolites, and peptides in venom gland tissues, shedding new light onto the biology of venom glands. Notably, AP-SMALDI as an innovative approach, in combination with LC−MS, provides a powerful platform for future venomics investigations with the potential to provide novel insights into the mechanistic biology of the production and secretion of toxins in the venom gland. In perspective, the technique offers great potential to map the spatial distribution of toxins and inhibitors, as well as previously unknown compounds, in the venom glands of other venomous animals such as scorpions and spiders.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.1c00042.

> Figures of venom apparatus of saw-scaled viper, *Echis carinatus sochureki*, representative microscopic images,

Venn diagrams comparing metabolite annotations, mass deviations for PC, SM, and pEKW signals, positive ion mode AP-SMALDI MSI images, total ion chromatograms, and HCD fragmentation patterns and tables of detailed information on common ion images and ion images annotated to lysophosphatidylcholine and META-SPACE annotation data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Bernhard Spengler − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University, 35392 Giessen, Germany;* ● orcid.org/0000-0003-0179-5653; Email: Bernhard.Spengler@anorg.chemie.uni-giessen.de

Authors

Parviz Ghezellou − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University, 35392 Giessen, Germany*

Sven Heiles − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University, 35392 Giessen, Germany;* ● orcid.org/0000-0003-3779-8071

Patrik Kadesch − *Institute of Inorganic and Analytical Chemistry, Justus Liebig University, 35392 Giessen, Germany*

Alireza Ghassempour − *Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, 1983969411 Tehran, Iran;* ● orcid.org/0000-0002-6435-9915

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.1c00042

Notes

The authors declare the following competing financial interest(s): B.S. is a consultant and P.G. is a part-time employee of TransMIT GmbH, Giessen, Germany.

■ REFERENCES

(1) Spengler, B.; Hubert, M.; Kaufmann, R. MALDI Ion Imaging and Biological Ion Imaging with a New Scanning UV-laser Microprobe. *Proceedings of the 42nd ASMS Conference on Mass Spectrometry and Allied Topics*, Chicago, IL, May 29–Jun 3, 1994; p 1041.

(2) Spengler, B. Mass Spectrometry Imaging of Biomolecular Information. *Anal. Chem.* **2015**, *87*, 64–82.

(3) Kadesch, P.; Quack, T.; Gerbig, S.; Grevelding, C. G.; Spengler, B. Lipid Topography in *Schistosoma mansoni* Cryosections, Revealed by Microembedding and High-Resolution Atmospheric-Pressure Matrix-Assisted Laser Desorption/Ionization (MALDI) Mass Spectrometry Imaging. *Anal. Chem.* **2019**, *91*, 4520–4528.

(4) Römpp, A.; Spengler, B. Mass Spectrometry Imaging with High Resolution in Mass and Space. *Histochem. Cell Biol.* **2013**, *139*, 759–783.

(5) Gilmore, I. S.; Heiles, S.; Pieterse, C. L. Metabolic Imaging at the Single-Cell Scale: Recent Advances in Mass Spectrometry Imaging. *Annu. Rev. Anal. Chem.* **2019**, *12*, 201–224.

(6) Korte, A. R.; Yandeau-Nelson, M. D.; Nikolau, B. J.; Lee, Y. Subcellular-level Resolution MALDI-MS Imaging of Maize Leaf Metabolites by MALDI-linear Ion Trap-Orbitrap Mass Spectrometer. *Anal. Bioanal. Chem.* **2015**, *407*, 2301–2309.

(7) Kettling, H.; Vens-Cappell, S.; Soltwisch, J.; Pirkl, A.; Haier, J.; Müthing, J.; Dreisewerd, K. MALDI Mass Spectrometry Imaging of Bioactive Lipids in Mouse Brain with a Synapt G2-S Mass Spectrometer Operated at Elevated Pressure: Improving the Analytical Sensitivity and the Lateral Resolution to Ten Micrometers. *Anal. Chem.* **2014**, *86*, 7798–7805.

(8) Koestler, M.; Kirsch, D.; Hester, A.; Leisner, A.; Guenther, S.; Spengler, B. a High-Resolution Scanning Microprobe Matrix-Assisted Laser Desorption/Ionization Ion Source for Imaging Analysis on an Ion Trap/Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 3275–3285.

(9) Kompauer, M.; Heiles, S.; Spengler, B. Atmospheric Pressure MALDI Mass Spectrometry Imaging of Tissues and Cells at 1.4-μm Lateral Resolution. *Nat. Methods* **2017**, *14*, 90–96.

(10) Kompauer, M.; Heiles, S.; Spengler, B. Autofocusing MALDI Mass Spectrometry Imaging of Tissue Sections and 3D Chemical Topography of Nonflat Surfaces. *Nat. Methods* **2017**, *14*, 1156–1158.

(11) Mackessy, S. P.; Baxter, L. M. Bioweapons Synthesis and Storage: The Venom Gland of Front-Fanged Snakes. *Zool. Anz.* **2006**, *245*, 147–159.

(12) Undheim, E. A. B.; Hamilton, B. R.; Kurniawan, N. D.; Bowlay, G.; Cribb, B. W.; Merritt, D. J.; Fry, B. G.; King, G. F.; Venter, D. J. Production and Packaging of a Biological Arsenal: Evolution of Centipede Venoms under Morphological Constraint. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4026–4031.

(13) Seppälä, U.; Francese, S.; Turillazzi, S.; Moneti, G.; Clench, M.; Barber, D. In Situ Imaging of Honeybee (*Apis Mellifera*) Venom Components from Aqueous and Aluminum Hydroxide–Adsorbed Venom Immunotherapy Preparations. *J. Allergy Clin. Immunol.* **2012**, *129*, 1314–1320.

(14) Mitchell, M. L.; Hamilton, B. R.; Madio, B.; Morales, R. A. V.; Tonkin-Hill, G. Q.; Papenfuss, A. T.; Purcell, A. W.; King, G. F.; Undheim, E. A. B.; Norton, R. S. The Use of Imaging Mass Spectrometry to Study Peptide Toxin Distribution in Australian Sea Anemones. *Aust. J. Chem.* **2017**, *70*, 1235–1237.

(15) Madio, B.; Peigneur, S.; Chin, Y. K. Y.; Hamilton, B. R.; Henriques, S. T.; Smith, J. J.; Cristofori-Armstrong, B.; Dekan, Z.; Boughton, B. A.; Alewood, P. F.; Tytgat, J.; King, G. F.; Undheim, E. A. B. PHAB Toxins: A Unique Family of Predatory Sea Anemone Toxins Evolving via Intra-Gene Concerted Evolution Defines a New Peptide Fold. *Cell. Mol. Life Sci.* **2018**, *75*, 4511–4524.

(16) Hamilton, B. R.; Marshall, D. L.; Casewell, N. R.; Harrison, R. A.; Blanksby, S. J.; Undheim, E. A. B. Mapping Enzyme Activity on Tissue by Functional Mass Spectrometry Imaging. *Angew. Chem., Int. Ed.* **2020**, *59*, 3855–3858.

(17) Das, T.; Alabi, I.; Colley, M.; Yan, F.; Griffith, W.; Bach, S.; Weintraub, S. T.; Renthal, R. Major Venom Proteins of the Fire Ant *Solenopsis Invicta*: Insights into Possible Pheromone-Binding Function from Mass Spectrometric Analysis. *Insect Mol. Biol.* **2018**, *27*, 505–511.

(18) Burry, R. W. *Immunocytochemistry*; Springer New York: New York, NY, 2010.

(19) Fletcher, J. S.; Rabbani, S.; Henderson, A.; Lockyer, N. P.; Vickerman, J. C. Three-Dimensional Mass Spectral Imaging of HeLa-M Cells - Sample Preparation, Data Interpretation and Visualisation. *Rapid Commun. Mass Spectrom.* **2011**, *25* (7), 925–932.

(20) Thavarajah, R.; Mudimbaimannar, V. K.; Elizabeth, J.; Rao, U. K.; Ranganathan, K. Chemical and physical basics of routine formaldehyde fixation. *J. Oral Maxillofac. Pathol.* **2012**, *16*, 400–405.

(21) Pietrowska, M.; Gawin, M.; Polańska, J.; Widłak, P. Tissue Fixed with Formalin and Processed without Paraffin Embedding Is Suitable for Imaging of Both Peptides and Lipids by MALDI-IMS. *Proteomics* **2016**, *16* (11–12), 1670–1677.

(22) Carter, C. L.; Jones, J. W.; Farese, A. M.; MacVittie, T. J.; Kane, M. A. Inflation-fixation method for lipidomic mapping of lung biopsies by matrix assisted laser desorption/ionization-mass spectrometry imaging. *Anal. Chem.* **2016**, *88*, 4788−4794.

(23) Gaudin, M.; Panchal, M.; Ayciriex, S.; Werner, E.; Brunelle, A.; Touboul, D.; Boursier-Neyret, C.; Auzeil, N.; Walther, B.; Duyckaerts, C.; Laprevote, O. Ultra-Performance Liquid Chromatography-Mass Spectrometry Studies of Formalin-induced Alterations of Human Brain Lipidome. *J. Mass Spectrom.* **2014**, *49*, 1035−1042.

(24) Vos, D. R. N.; Bowman, A. P.; Heeren, R. M. A.; Balluff, B.; Ellis, S. R. Class-Specific Depletion of Lipid Ion Signals in Tissues upon Formalin Fixation. *Int. J. Mass Spectrom.* **2019**, *446*, 116212.

(25) Addie, R. D.; Balluff, B.; Bovee, J. V.; Morreau, H.; McDonnell, L. A. Current state and future challenges of mass spectrometry imaging for clinical research. *Anal. Chem.* **2015**, *87*, 6426−6433.

(26) Buck, A.; Ly, A.; Balluff, B.; Sun, N.; Gorzolka, K.; Feuchtinger, A.; Janssen, K. P.; Kuppen, P. J. K.; Van De Velde, C. J. H.; Weirich, G.; Erlmeier, F.; Langer, R.; Aubele, M.; Zitzelsberger, H.; Aichler, M.; Walch, A. High-Resolution MALDI-FT-ICR MS Imaging for the Analysis of Metabolites from Formalin-Fixed, Paraffin-Embedded Clinical Tissue Samples. *J. Pathol.* **2015**, *237* (1), 123−132.

(27) Ly, A.; Buck, A.; Balluff, B.; Sun, N.; Gorzolka, K.; Feuchtinger, A.; Janssen, K. P.; Kuppen, P. J. K.; Van De Velde, C. J. H.; Weirich, G.; Erlmeier, F.; Langer, R.; Aubele, M.; Zitzelsberger, H.; McDonnell, L.; Aichler, M.; Walch, A. High-Mass-Resolution MALDI Mass Spectrometry Imaging of Metabolites from Formalin-Fixed Paraffin-Embedded Tissue. *Nat. Protoc.* **2016**, *11* (8), 1428−1443.

(28) Cacciatore, S.; Zadra, G.; Bango, C.; Penney, K. L.; Tyekucheva, S.; Yanes, O.; Loda, M. Metabolic profiling in formalin-fixed and paraffin-embedded prostate cancer tissues. *Mol. Cancer Res.* **2017**, *15*, 439−447.

(29) Carter, C. L.; McLeod, C. W.; Bunch, J. Imaging of phospholipids in formalin fixed rat brain sections by matrix assisted laser desorption/ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1991−1998.

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

(31) Paschke, C.; Leisner, A.; Hester, A.; Maass, K.; Guenther, S.; Bouschen, W.; Spengler, B. Mirion—A Software Package for Automatic Processing of Mass Spectrometric Images. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1296−1306.

(32) Race, A. M.; Styles, I. B.; Bunch, J. Inclusive Sharing of Mass Spectrometry Imaging Data Requires a Converter for All. *J. Proteomics* **2012**, *75*, 5111−5112.

(33) Bokhart, M. T.; Nazari, M.; Garrard, K. P.; Muddiman, D. C. MSiReader v1.0: Evolving Open-Source Mass Spectrometry Imaging Software for Targeted and Untargeted Analyses. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 8−16.

(34) Palmer, A.; Phapale, P.; Chernyavsky, I.; Lavigne, R.; Fay, D.; Tarasov, A.; Kovalev, V.; Fuchser, J.; Nikolenko, S.; Pineau, C.; Becker, M.; Alexandrov, T. FDR-Controlled Metabolite Annotation for High-Resolution Imaging Mass Spectrometry. *Nat. Methods* **2017**, *14*, 57−60.

(35) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vazquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: the Human Metabolome Database for 2018. *Nucleic Acids Res.* **2018**, *46*, 608−617.

(36) O'Donnell, V. B.; Dennis, E. A.; Wakelam, M. J. O.; Subramaniam, S. LIPID MAPS: Serving the Next Generation of Lipid Researchers with Tools, Resources, Data, and Training. *Sci. Signaling* **2019**, *12*, No. eaaw2964.

(37) Spengler, B. De Novo Sequencing, Peptide Composition Analysis, and Composition-Based Sequencing: A New Strategy Employing Accurate Mass Determination by Fourier Transform Ion Cyclotron Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 703−714.

(38) Ghezellou, P.; Albuquerque, W.; Garikapati, V.; Caswell, N. R.; Kazemi, S. M.; Ghassempour, A.; Spengler, B. Integrating Top-Down and Bottom-Up Mass Spectrometric Strategies for Proteomic Profiling of Iranian Saw-Scaled Viper, *Echis Carinatus Sochureki* Venom. *J. Proteome Res.* **2021**, *20*, 895−908.

(39) Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. Lipid Extraction by Methyl-Tert-Butyl Ether for High-Throughput Lipidomics. *J. Lipid Res.* **2008**, *49* (5), 1137−1146.

(40) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30* (10), 918−920.

(41) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinf.* **2010**, *11* (1), 395.

(42) Koelmel, J. P.; Kroeger, N. M.; Ulmer, C. Z.; Bowden, J. A.; Patterson, R. E.; Cochran, J. A.; Beecher, C. W. W.; Garrett, T. J.; Yost, R. A. LipidMatch: An Automated Workflow for Rule-Based Lipid Identification Using Untargeted High-Resolution Tandem Mass Spectrometry Data. *BMC Bioinformatics* **2017**, *18* (1), 331.

(43) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane Lipids: Where They are and How They Behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112−124.

(44) Yeagle, P. L. *The Structure of Biological Membranes*, 3rd ed.; CRC Press: New York, 2011; pp 27−56.

(45) Breslow, D. K.; Weissman, J. S. Membranes in Balance: Mechanisms of Sphingolipid Homeostasis. *Mol. Cell* **2010**, *40*, 267−279.

(46) Grainger, D. W.; Reichert, H.; Ringsdorf, H.; Salesse, C. An Enzyme Caught in Action: Direct Imaging of Hydrolytic Function and Domain Formation of Phospholipase A2 in Phosphatidylcholine Monolayers. *FEBS Lett.* **1989**, *252*, 73−82.

(47) 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.* **2019**, *33*, 20−27.

(48) Ishibashi, Y.; Kohyama-Koganeya, A.; Hirabayashi, Y. New Insights on Glucosylated Lipids: Metabolism and Functions. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2013**, *1831* (9), 1475−1485.

(49) Jennemann, R.; Gröne, H.-J. Cell-Specific in Vivo Functions of Glycosphingolipids: Lessons from Genetic Deletions of Enzymes Involved in Glycosphingolipid Synthesis. *Prog. Lipid Res.* **2013**, *52* (2), 231−248.

(50) Hirabayashi, Y. A World of Sphingolipids and Glycolipids in the Brain Novel Functions of Simple Lipids Modified with Glucose. *Proc. Jpn. Acad., Ser. B* **2012**, *88* (4), 129−143.

(51) Wagstaff, R.; Favreau, P.; Cheneval, O.; Laing, G. D.; Wilkinson, M. C.; Miller, R. L.; Stocklin, R.; Harrison, R. A. Molecular Characterizatioin of Endogenous Snake Venom Metalloproteinase Inhibitors. *Biochem. Biophys. Res. Commun.* **2008**, *365*, 650−656.

(52) Francis, B.; Kaiser, I. I. Inhibition of Metalloproteinases in *Bothrops asper* Venom by Endogenous Peptides. *Toxicon* **1993**, *31*, 889−899.

(53) Huang, K. F.; Hung, C. C.; Wu, S. H.; Chiou, S. H. Characterization of Three Endogenous Peptide Inhibitors for Multiple Metalloproteinases with Fibrinogenolytic Activity from the Venom of Taiwan Habu (*Trimeresurus mucrosquamatus*). *Biochem. Biophys. Res. Commun.* **1998**, *248*, 562−568.

(54) Yee, K.; Pitts, M.; Tongyoo, P.; Rojnuckarin, P.; Wilkinson, M. Snake Venom Metalloproteinases and Their Peptide Inhibitors from Myanmar Russell's Viper Venom. *Toxins* **2017**, *9* (1), 15.

(55) Hempel, B. F.; Damm, M.; Gocmen, B.; Karis, M.; Oguz, M. A.; Nalbantsoy, A.; Süssmuth, R. D. Comparative Venomics of the *Vipera ammodytes transcaucasiana* and *Vipera ammodytes montandoni* from Turkey Provides Insights into Kinship. *Toxins* **2018**, *10*, 23.

(56) Aoki-Shioi, N.; Koh, C. Y.; Kini, R. M. Natural Inhibitors of Snake Venom Metalloproteinases. *Aust. J. Chem.* **2020**, *73* (4), 277−286.

(57) Hamanaka, S.; Suzuki, A.; Hara, M.; Nishio, H.; Otsuka, F.; Uchida, Y. Human Epidermal Glucosylceramides Are Major Precursors of Stratum Corneum Ceramides. *J. Invest. Dermatol.* **2002**, *119* (2), 416−423.

(58) Munekiyo, S. M.; Mackessy, S. P. Presence of Peptide Inhibitors in Rattlesnake Venoms and Their Effects on Endogenous Metalloproteases. *Toxicon* **2005**, *45*, 255−263.

Supplementary Information

Venom Gland Mass Spectrometry Imaging of Saw-scaled Viper, *Echis carinatus sochureki*, at High Lateral Resolution

Parviz Ghezellou, Sven Heiles, Patrik Kadesch, Alireza Ghassempour, and Bernhard Spengler

Table of contents

Figure S1. The venom apparatus of *Echis carinatus sochureki*. A) The head of saw-scaled viper. The bilaterally paired venom glands rest just under the skin and wrapped in a connective tissue capsule. B) The skin has been removed from the head. The venom apparatus of *E. carinatus* consists of the main venom glands (VG), primary duct (PD), an accessory gland (AG) and a secondary duct (SD) which connects the glands to the base of a hollow specialized teeth named fang (FG). In the image, the fang is placed backwards along the roof of the mouth inside a sheath.
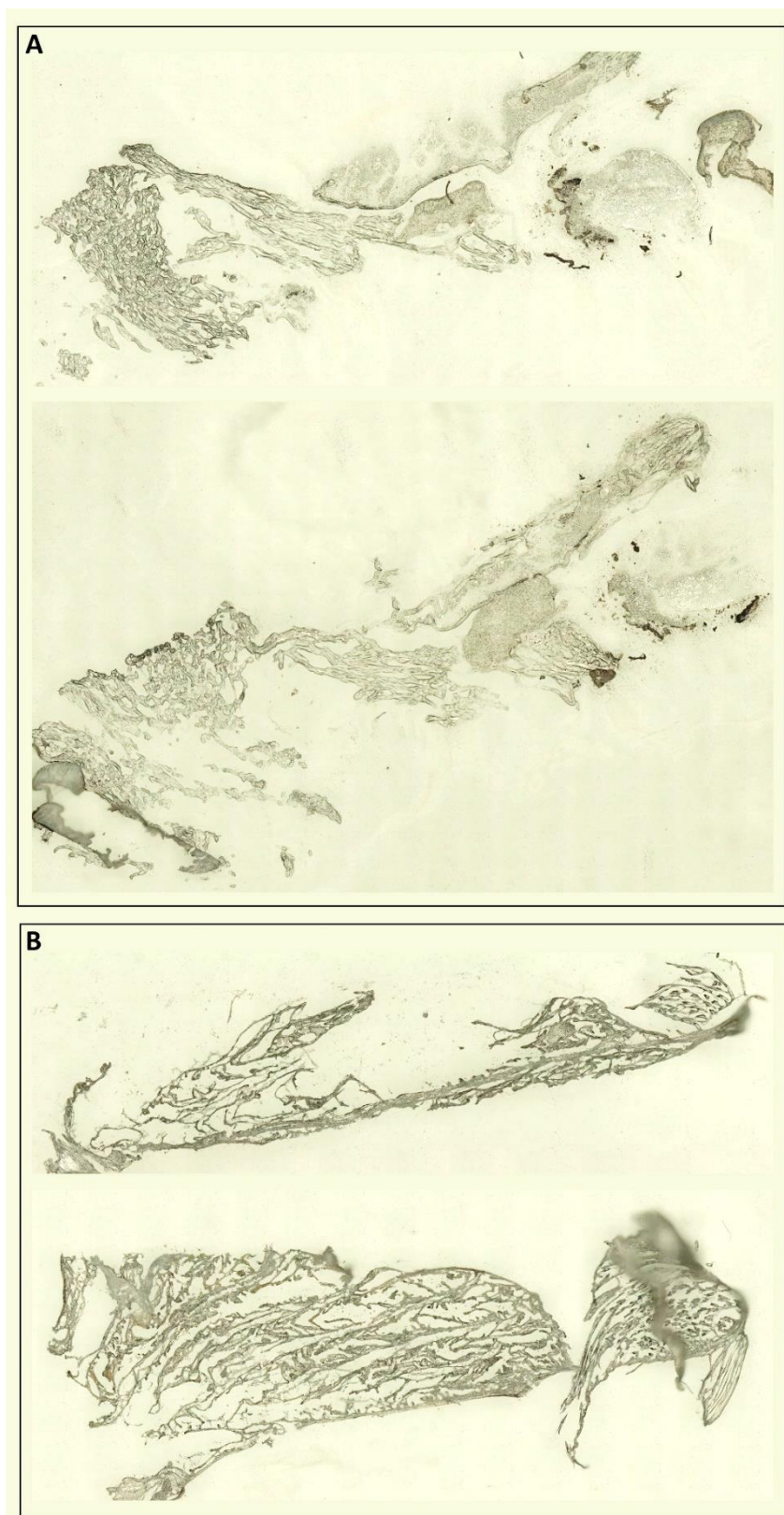
Figure S2. Representative microscopic images obtained from sections of *Echis carinatus sochureki* venom glands. A) Fresh-frozen tissue sections, B) formalin-fixed tissue sections.
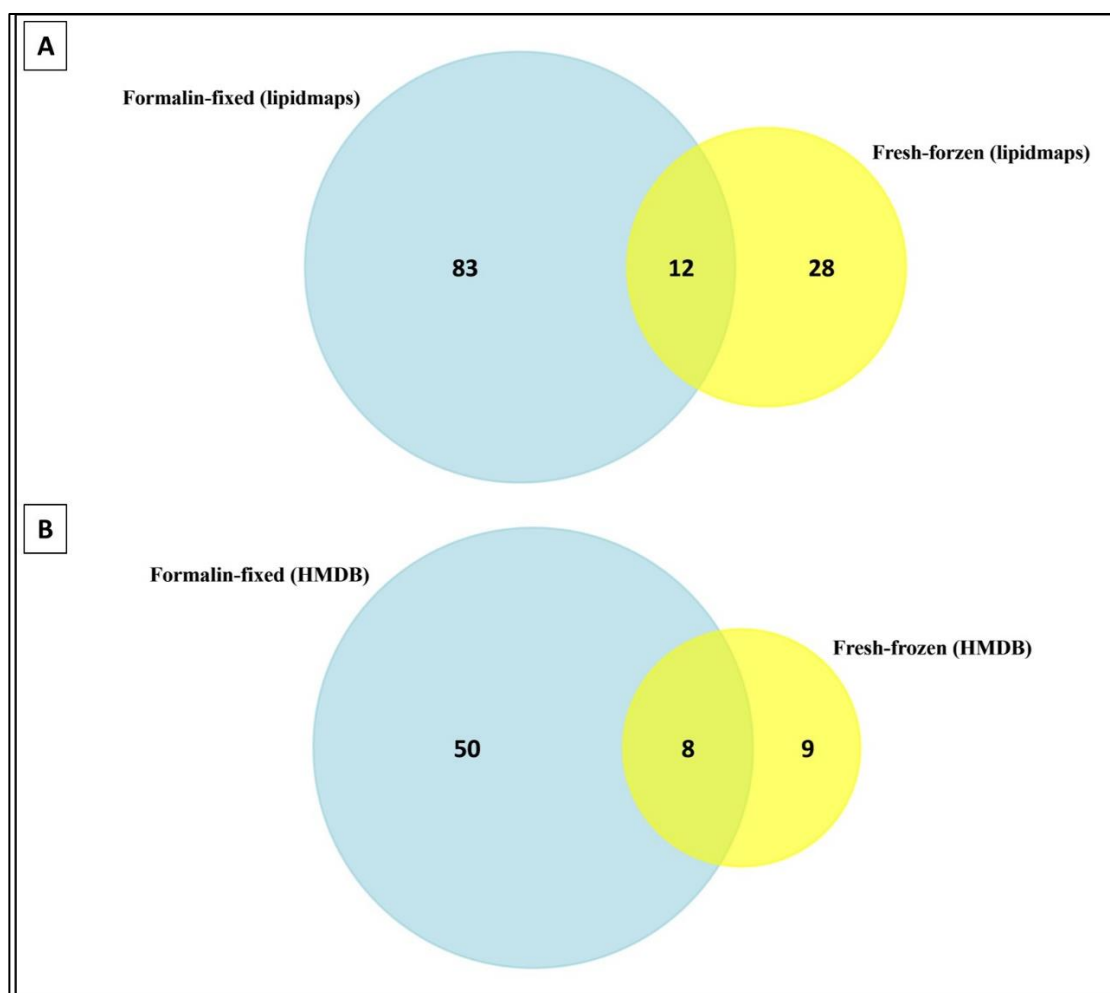
Figure S3. Comparing signal annotations obtained by METASPACE platform of both formalin-fixed and fresh-frozen tissue sections against A) HMDB and B) LipidMaps databases with the desired FDR of 0.1.
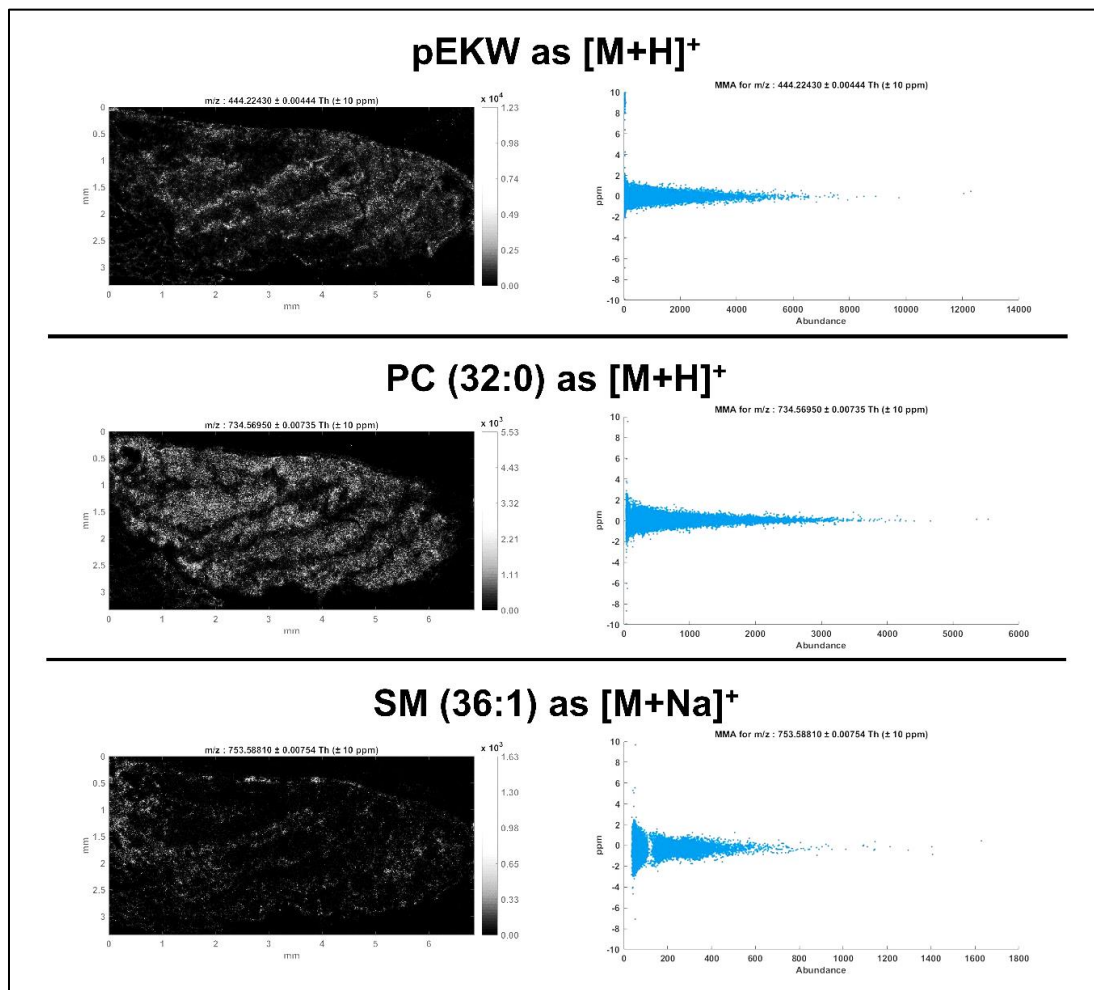
Figure S4. Mass deviations for PC, SM and pEKW signals across fresh-frozen venom gland tissue as a function of signal intensities and resulting root mean square (RMS) values in parts per million (ppm).
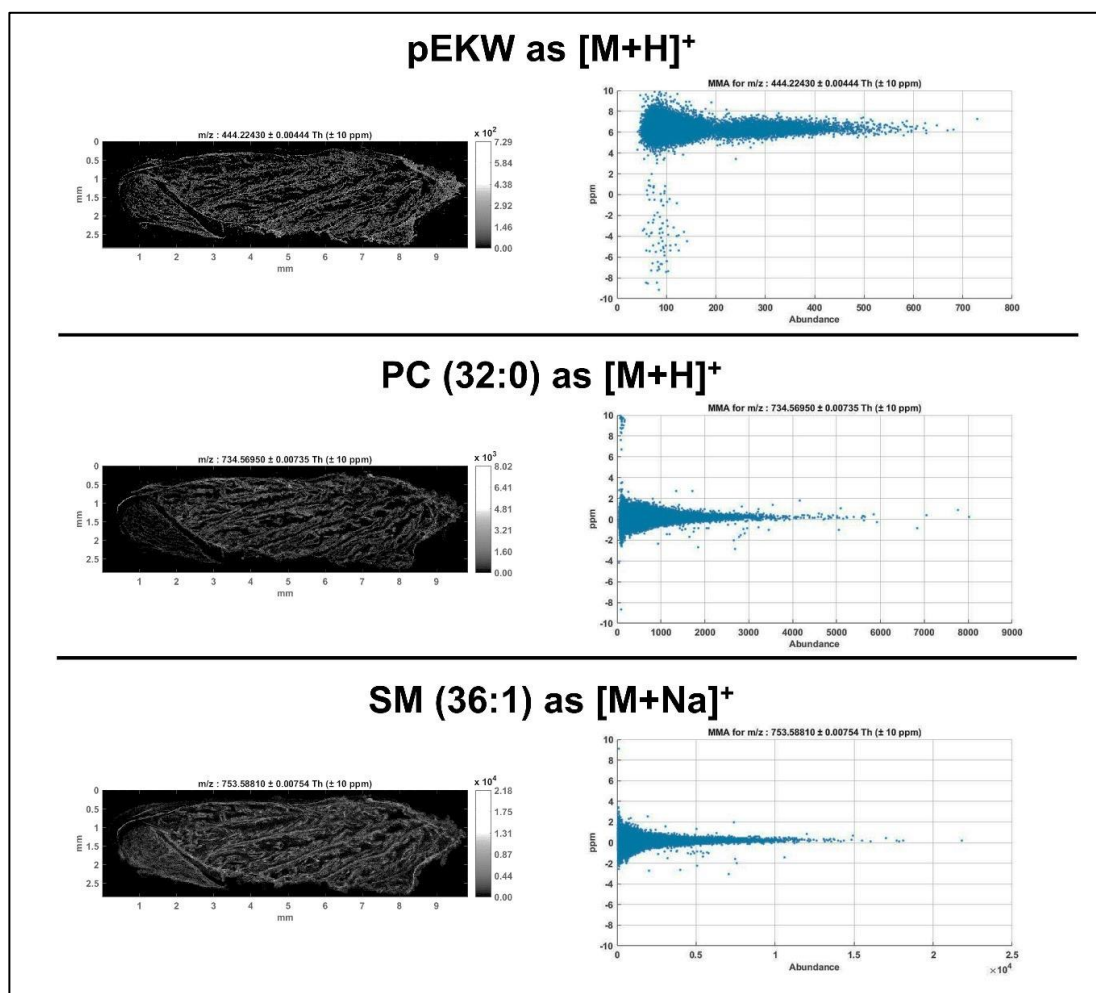
Figure S5. Mass deviations for PC, SM and pEKW signals across formalin-fixed venom gland tissue as a function of signal intensities and resulting root mean square (RMS) values in parts per million (ppm).
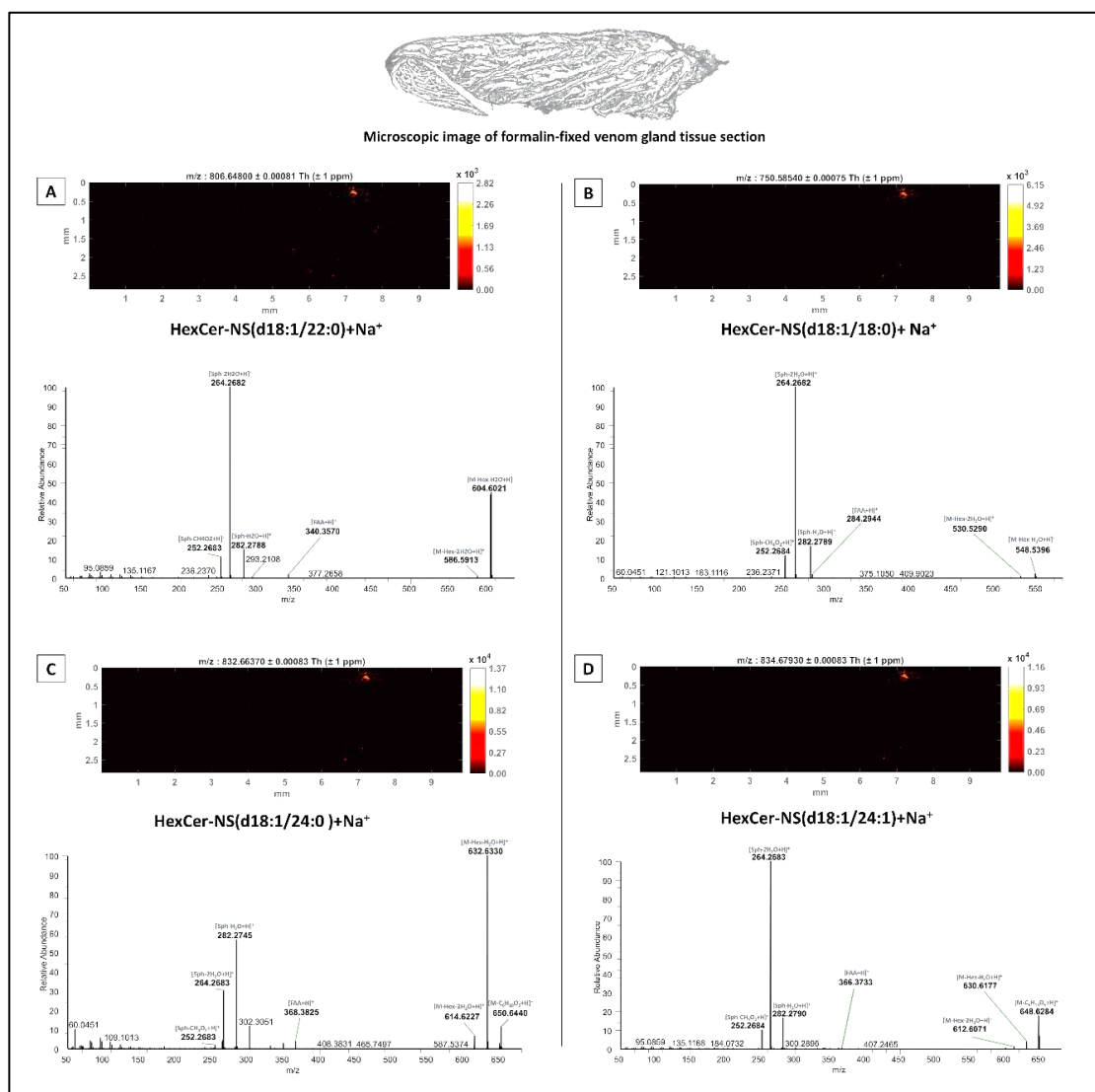
Figure S6. Positive ion mode MALDI-MSI images of Hexosylceramides (HexCer) species ions at A) *m/z* 806.6480, HexCer-NS[(d18:1/22:0)+Na]$^+$; B) *m/z* 750.5854, HexCer-NS[(d18:1/18:0)+ Na]$^+$; C) *m/z* 834.6793, HexCer-NS[(d18:1/24:0)+Na]$^+$; D) *m/z* 832.6637, HexCer-NS[(d18:1/24:1)+Na]$^+$. HCD fragmentation patterns of each species are obtained from LC-MS/MS of the venom gland lipidome and presented below the MSI images.
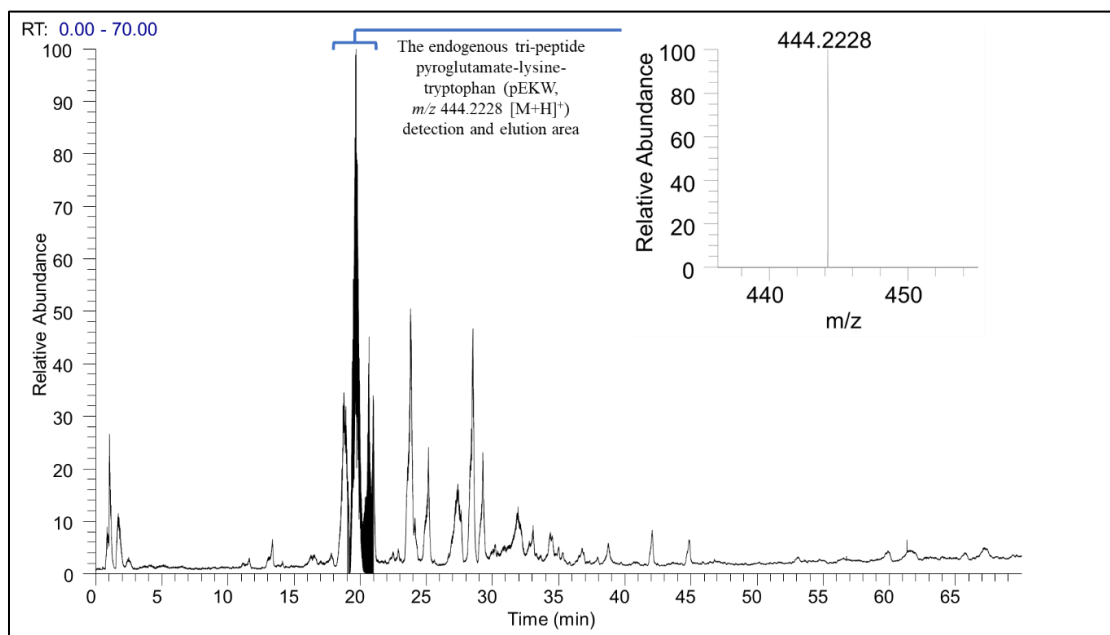
Figure S7. Total ion chromatogram (TIC) of the saw scaled viper, *Echis carinatus sochureki*, venom. The crude venom applied to LC-HR-MS/MS analysis after initial filtering using 10 kDa molecular-weight cut-off (MWCO). The targeted peptide (pEKW) was detected in the retention time between 19 to 21 min by setting the inclusion list and narrow mass rang in data-dependent MS^2 acquisition mode.
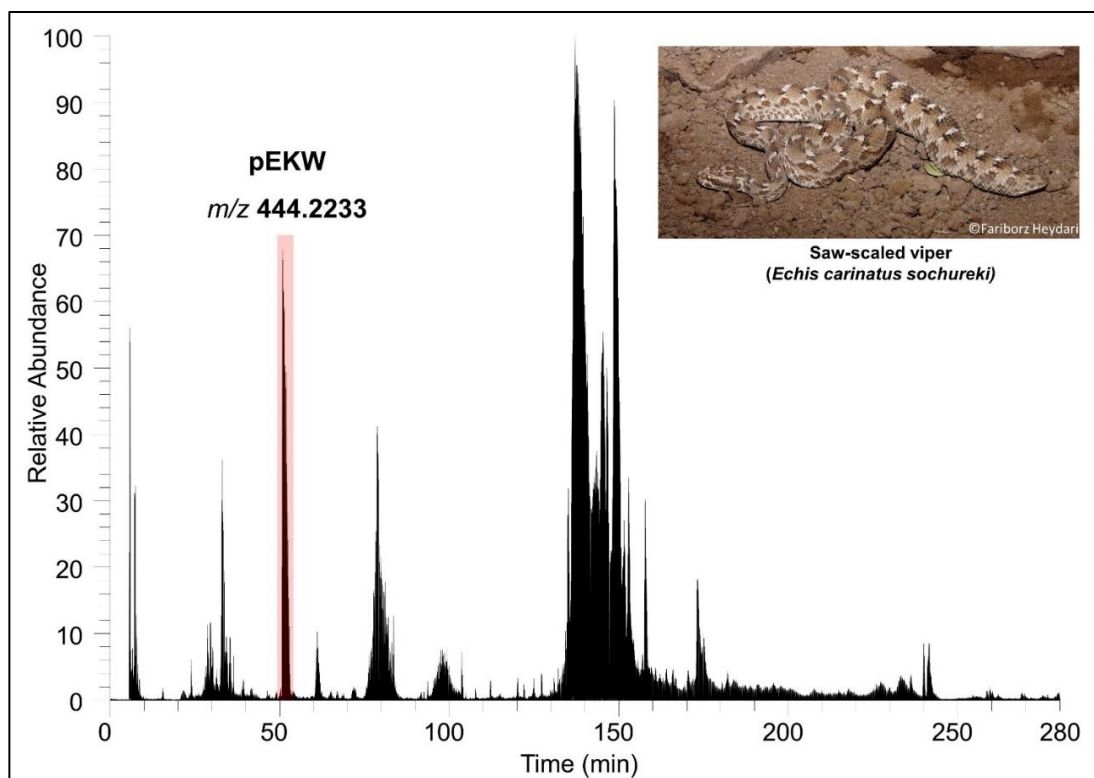
Figure S8. TIC chromatogram of Saw-scaled viper, *Echis carinatus sochureki*, venom from top-down proteomics analysis. The peak highlighted in red (at 50.5 min to 53.2 min) represent the elution profile of the tri-peptide pyroglutamate-lysine-tryptophan (pEKW, m/z 444.2233 [M+H]+). The peak area of tri-peptide (1.07×10^{12}) is about 6.68% of the total peak area (1.62×10^{13}).
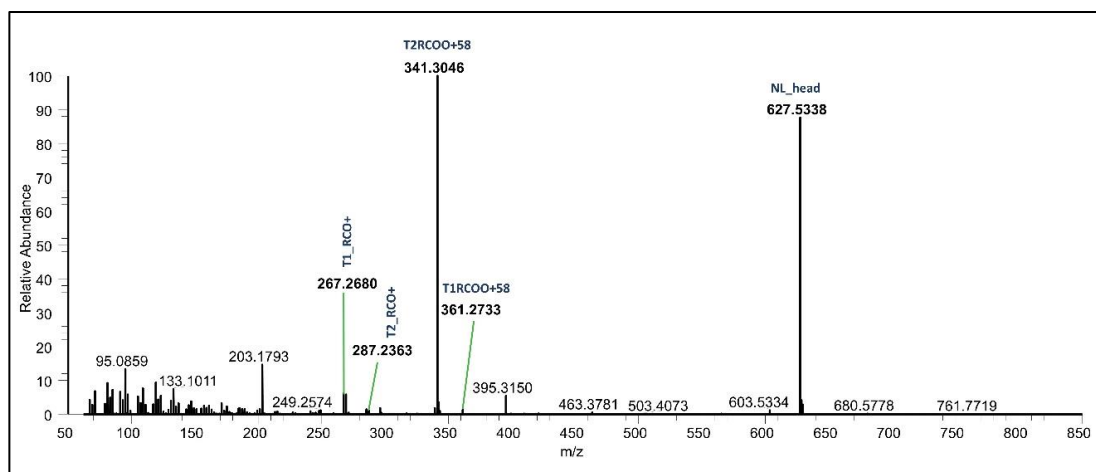
Figure S9. HCD fragmentation pattern of PS [(18:0_20:4) + H]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.
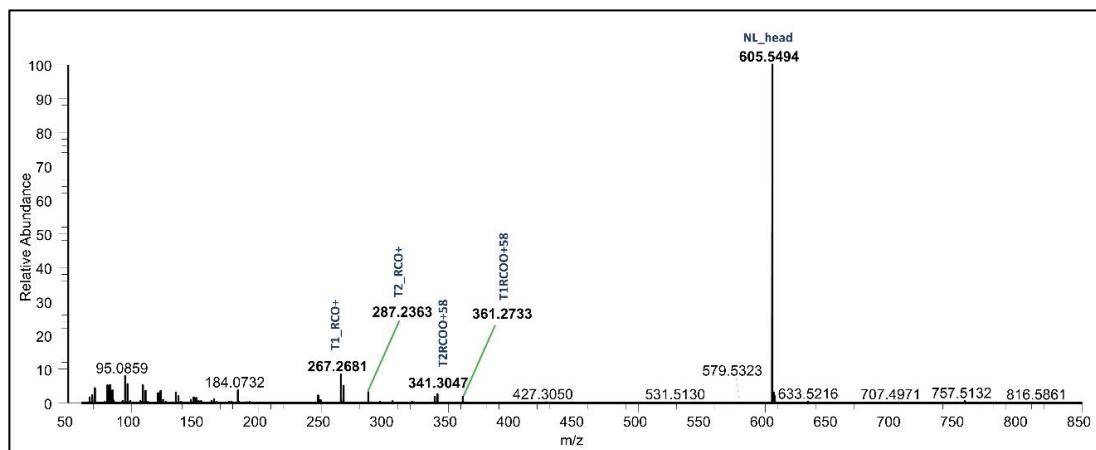
Figure S10. HCD fragmentation pattern of PS [(16:0_20:1) + H]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.
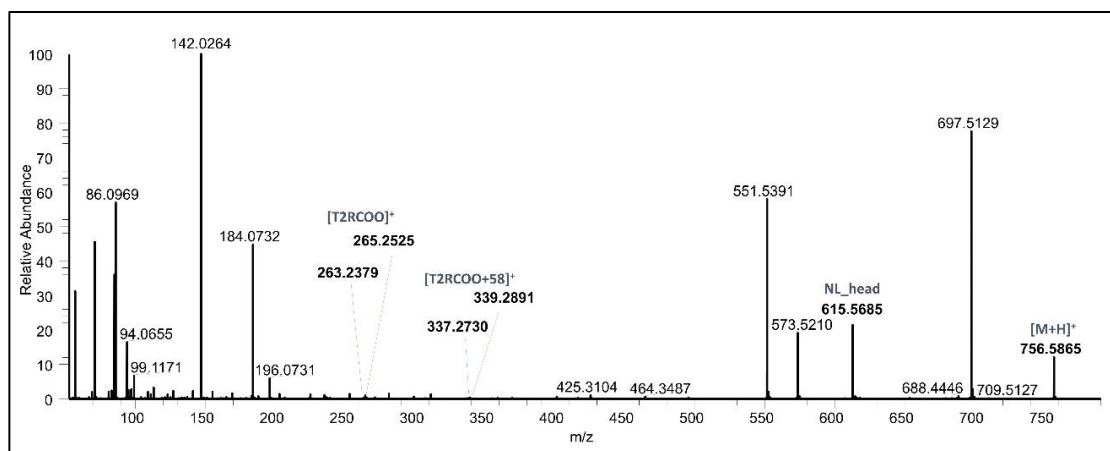
Figure S11. HCD fragmentation pattern of PE [(P-20:1/18:1) + H]⁺ and PE [(P-20:0/18:2) + H]⁺ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.

Figure S12. HCD fragmentation pattern of PE [(18:2_20:1) + Na]⁺ and PE [(18:3_20:0) + Na]⁺ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.

Figure S13. HCD fragmentation pattern of PE [(18:0_22:3) + H]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.
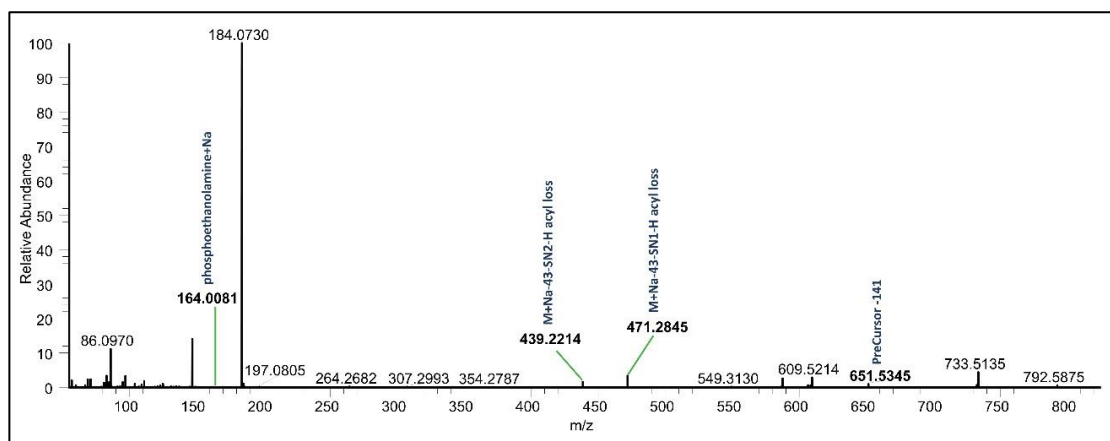

Figure S14. HCD fragmentation pattern of PE [(18:0_20:4) + H]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.
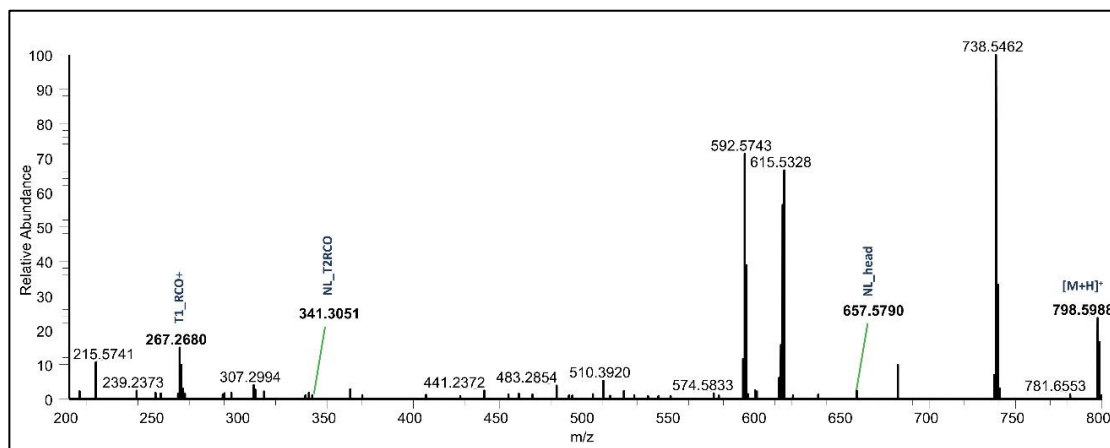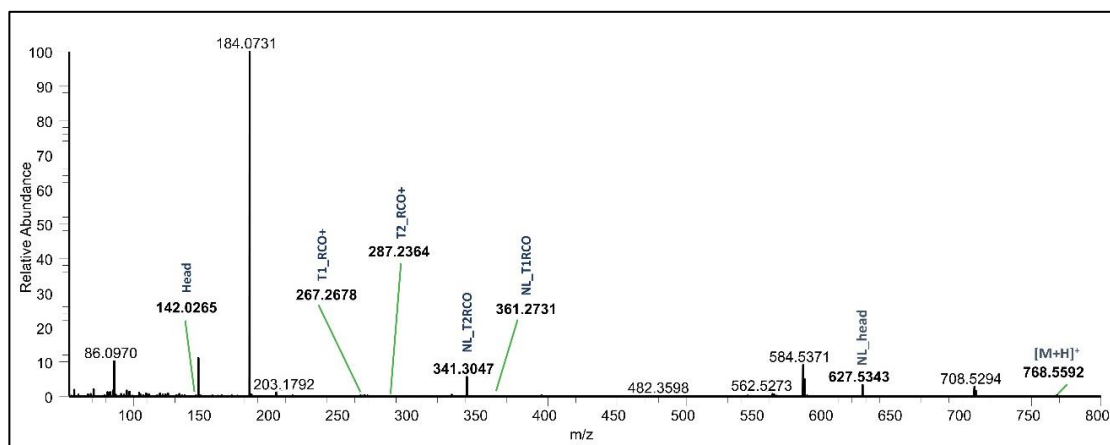

Figure S15. HCD fragmentation pattern of PE [(18:0_18:1) + Na]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.
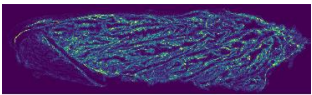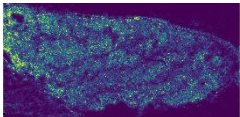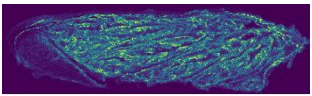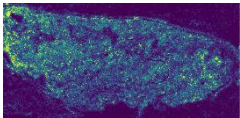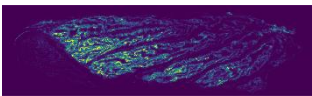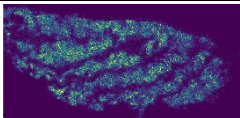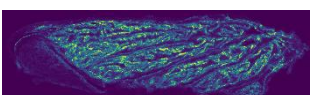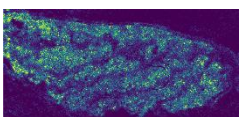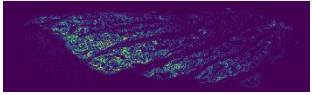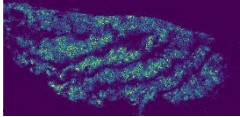
Figure S16. HCD fragmentation pattern of PE [(16:0_20:2) + Na]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.
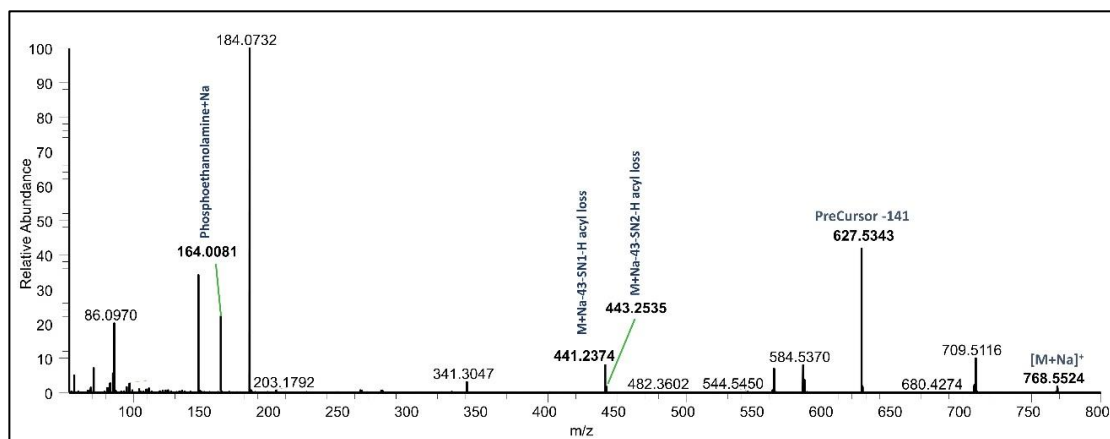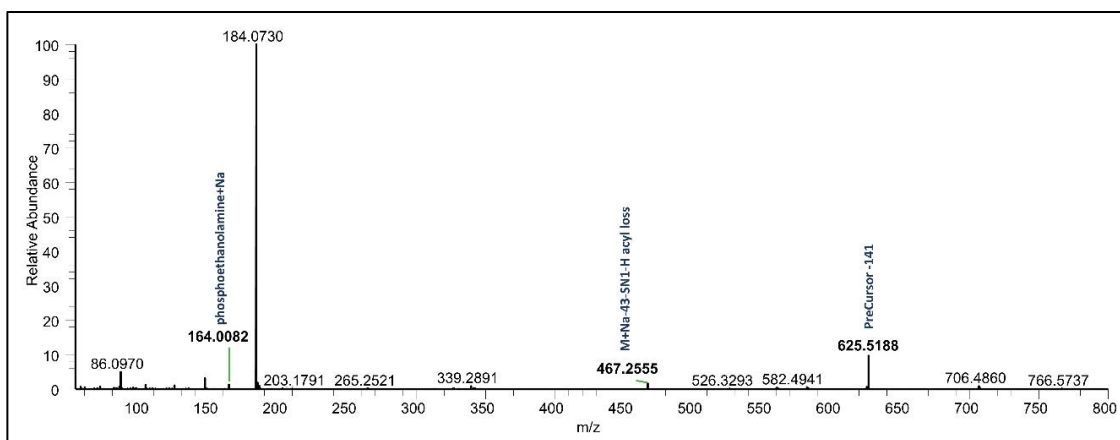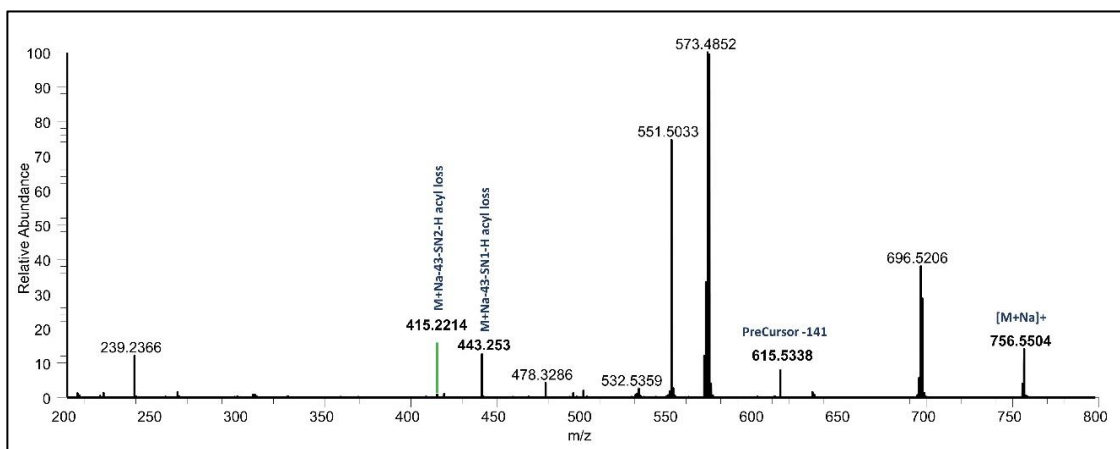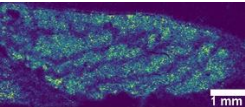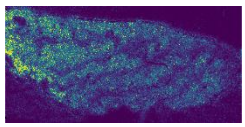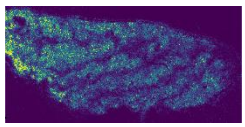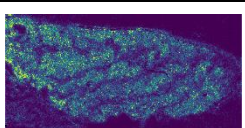

Figure S17. HCD fragmentation pattern of PE [(16:0_19:0) + Na]$^+$ and PE [(17:0_18:0) + Na]$^+$ obtained by LC-MS/MS of the formalin-fixed tissue lipidome in positive-ion mode.

Table S2. Detailed information on the ion images annotated in both types of tissue sections (formalin-fixed and fresh-frozen) using METASPACE search platform against HMDB and LipidMaps database. Blue color is minimum intensity and yellow is maximum intensity.

m/z	Sum formulas + ion adduct	Annotated metabolites	Formalin-fixed tissue	Fresh-frozen tissue
703.5748	$[C_{39}H_{79}N_2O_6P + H]^+$	PE-Cer(d37:1) SM(d34:1)		
725.5568	$[C_{39}H_{79}N_2O_6P + Na]^+$	PE-Cer(d37:1) SM(d34:1)		
746.6058	$[C_{42}H_{84}NO_7P + H]^+$	PE(P-37:0) PE(O-37:1) PC(O-34:1) PC(P-34:0)		
835.6663	$[C_{47}H_{93}N_2O_6P + Na]^+$	SM(d42:2)		
882.6194	$[C_{47}H_{90}NO_{10}P + Na]^+$	PS(41:1)		

Table S3. Detailed information on the ion images annotated to lysophosphatidylcholine (LPC) from fresh-frozen tissue section.

m/z	Sum formulas + ion adduct	Annotated metabolites	Fresh-frozen tissue
496.3397	$[C_{24}H_{50}NO_7P + H]^+$	LPC(16:0)	
508.3761	$[C_{26}H_{54}NO_6P + H]^+$	LPC(P-18:0)	
510.3917	$[C_{26}H_{56}NO_6P + H]^+$	LPC(0-18:0)	
524.3710	$[C_{26}H_{54}NO_7P + H]^+$	LPC(18:0)	

Table S4. METASPACE annotation of fresh-frozen venom gland tissue section against HMDB database (version 2.5) at the desired FDR level of 0.1 and MSM-score value.

formula	adduct	m/z	msm[1]	fdr[2]	molecule names
C47H93N2O6P	M+H	813.68438	0.963872	0.05	SM(d42:2)
C39H79N2O6P	M+H	703.574829	0.951515	0.05	SM(d34:1)
C42H84NO7P	M+H	746.605795	0.930229	0.05	PC(P-34:0), PC(o-34:1)
C39H79N2O6P	M+K	741.530711	0.907533	0.05	SM(d34:1)
C26H54NO6P	M+H	508.376129	0.901582	0.05	LPC(P-18:0)
C24H50NO7P	M+H	496.339744	0.890348	0.05	LPC(16:0)
C39H79N2O6P	M+Na	725.556774	0.889093	0.05	SM(d34:1)
C29H47NO4	M+Na	496.339708	0.88767	0.05	Docosa-4,7,10,13,16-pentaenoyl carnitine, Clupanodonyl carnitine, 23-Acetoxysoladulcidine
C47H93N2O6P	M+K	851.640262	0.879352	0.05	SM(d42:2)
C47H93N2O6P	M+Na	835.666324	0.876111	0.05	SM(d42:2)
C26H54NO7P	M+H	524.371044	0.867737	0.05	LPC(18:0)
C40H80NO8P	M+H	734.56941	0.851538	0.1	PC(32:0), PE(35:0), PE-NMe(34:0), PE-NMe2(33:0)
C26H56NO6P	M+H	510.391779	0.831463	0.1	LPC(O-18:0)
C42H82NO8P	M+H	760.58506	0.79077	0.1	PC(34:1), PE(37:1), PE-NMe(36:1), PE-NMe2(35:1)
C42H86NO7P	M+H	748.621445	0.771179	0.1	PC(O-34:0)

[1] Metabolite-Signal Match (MSM) score; [2] False Discovery Rate (FDR)

Table S5. METASPACE annotation of formalin-fixed venom gland tissue section against HMDB database (version 2.5) at the desired FDR level of 0.1 and MSM-score value.

formula	adduct	m/z	msm[1]	fdr[2]	molecule names
C42H84NO7P	M+Na	768.58774	0.979674	0.05	PC(P-34:0), PC(o-34:1)
C39H79N2O6P	M+Na	725.556774	0.974808	0.05	SM(d34:1)
C47H93N2O6P	M+Na	835.666324	0.969944	0.05	SM(d42:2)
C42H82NO8P	M+Na	782.567004	0.966001	0.05	PC(34:1), PE(37:1), PE-NMe(36:1), PE-NMe2(35:1)
C39H79N2O6P	M+H	703.574829	0.953887	0.05	SM(d34:1), Palmitoyl sphingomyelin
C42H80NO8P	M+Na	780.551354	0.948187	0.05	PC(34:2), PE(37:2), PE-NMe(36:2), PE-NMe2(35:2)
C42H84NO7P	M+H	746.605795	0.947698	0.05	PC(P-34:0), PC(o-34:1)
C44H84NO6P	M+Na	776.592825	0.939133	0.05	PC(P-36:1)
C47H93N2O6P	M+H	813.68438	0.929958	0.05	SM(d42:1)
C42H80NO8P	M+H	758.56941	0.929014	0.05	PC(34:2), PE(37:2), PE-NMe(36:2), PE-NMe2(35:2)
C45H91N2O6P	M+Na	809.650674	0.926924	0.05	SM(d40:1)
C44H84NO8P	M+Na	808.582654	0.92398	0.05	PC(36:2), 1,2-dioleoyl-sn-glycero-3-phosphocholine, PE-NMe(38:2), PE-NMe2(37:2)
C39H71O8P	M+Na	721.477855	0.922759	0.05	PA(36:3)
C42H82NO8P	M+H	760.58506	0.919222	0.05	PC(34:1), PE(37:1), PE-NMe(36:1), PE-NMe2(35:1)
C44H86NO7P	M+Na	794.60339	0.91342	0.05	PC(P-36:1), PC(o-36:2)
C39H73O8P	M+Na	723.493505	0.913023	0.05	PA(36:2)
C40H80NO8P	M+Na	756.551354	0.909536	0.05	PC(32:0), PE(35:0), PE-NMe(34:0), PE-NMe2(33:0)
C47H95N2O6P	M+Na	837.681974	0.907259	0.05	SM(d42:1)
C23H36N2O11	M+Na	539.221109	0.899669	0.05	Perindoprilat glucuronide
C41H83N2O6P	M+Na	753.588074	0.898766	0.05	SM(d36:1), stearoyl sphingomyelin
C22H42O4	M+Na	393.297509	0.895549	0.05	Dioctyl hexanedioate, Diethylhexyl adipate
C44H82NO7P	M+Na	790.57209	0.894894	0.05	PC(P-36:3)
C29H47NO4	M+Na	496.339708	0.891097	0.05	Docosa-4,7,10,13,16-pentaenoyl carnitine, Clupanodonyl carnitine, 23-Acetoxysoladulcidine
C53H98O6	M+Na	853.72554	0.881524	0.1	TG(50:2)
C27H46O4	M+Na	457.328809	0.880309	0.1	3a,7a,12a-Trihydroxy-5b-cholestan-26-al, 3alpha,7alpha,24(S)-trihydroxy-5beta-cholestan-27-al, 3a,7a-Dihydroxycoprostanic acid
C55H102O6	M+Na	881.75684	0.866331	0.1	TG(52:2)
C53H100O6	M+Na	855.74119	0.865801	0.1	TG(50:1)
C26H54NO7P	M+Na	546.352989	0.856679	0.1	2-acetyl-1-alkyl-sn-glycero-3-phosphocholine
C44H86NO8P	M+Na	810.598304	0.856649	0.1	PC(36:1), PE(39:1), PE-NMe(38:1), PE-NMe2(37:1)
C37H66O8	M+Na	661.464968	0.85525	0.1	Annoglaucin, Muricatin C, Rollidecin A, Bullatetrocin, Rollitacin, Mucocin, Glabracin A, 12,15-cis-Squamostatin A, Rollimusin, 20,23-cis-Bullatalicinone, Bullatanocin, Annonin XIV, 27-Hydroxybullatacin, 9-Hydroxyasimicinone, Purpureacin 2, Purpurenin
C39H81N2O6P	M+Na	727.572424	0.853388	0.1	N-hexadecanoylsphinganine-1-phosphocholine
C42H81NO8	M+Na	750.585418	0.851993	0.1	Glucosylceramide (d36:1), Galactosylceramide (d36:1)
C48H93NO8	M+Na	834.679318	0.846072	0.1	Glucosylceramide (d42:1)
C44H80NO8P	M+Na	804.551354	0.844162	0.1	PC(36:4), PE-NMe(38:4), PE-NMe2(37:4)
C37H75N2O6P	M+Na	697.525474	0.840476	0.1	SM(d32:1)

C44H88NO7P	M+Na	796.61904	0.8331	0.1	PC(P-36:0)
C48H91NO8	M+Na	832.663668	0.827591	0.1	Glucosylceramide (d42:2), Galactosylceramide (d42:2)
C46H93N2O6P	M+Na	823.666324	0.823621	0.1	SM(d41:1)
C44H82NO8P	M+Na	806.567004	0.820203	0.1	PC(36:3), PE-NMe(38:3)
C57H104O6	M+Na	907.77249	0.815608	0.1	TG(54:3)
C51H98O6	M+Na	829.72554	0.801083	0.1	TG(48:0)
C46H89NO8	M+Na	806.648018	0.798915	0.1	Glucosylceramide (d18:1/22:0), Galactosylceramide (d18:1/22:0)
C46H80NO8P	M+Na	828.551354	0.788569	0.1	PC(38:6), PE-NMe(40:6)
C41H75O8P	M+Na	749.509155	0.787066	0.1	PA(38:3)
C42H84NO8P	M+Na	784.582654	0.785837	0.1	PC(34:0), PE(37:0), PE-NMe(36:0), PE-NMe2(35:0)
C55H100O6	M+Na	879.74119	0.784019	0.1	TG(52:3)
C42H82NO7P	M+Na	766.57209	0.783294	0.1	PC(P-34:1)
C44H84NO7P	M+Na	792.58774	0.779868	0.1	PC(P-36:2)

[1] Metabolite-Signal Match (MSM) score; [2] False Discovery Rate (FDR)

Table S6. METASPACE annotation of fresh-frozen venom gland tissue section against LipidMaps database at the desired FDR level of 0.1 and MSM-score value.

formula	adduct	m/z	msm[1]	fdr[2]	molecule names
C47H93N2O6P	M+H	813.68438	0.963872	0.05	SM(d42:1)
C39H79N2O6P	M+H	703.574829	0.951515	0.05	PE-Cer(d37:1), SM(d34:1)
C42H84NO7P	M+H	746.605795	0.930229	0.05	PE(P-37:0), PE(O-37:1), PC(O-34:1), PC(P-34:0)
C24H52NO6P	M+H	482.360479	0.930003	0.05	PC(O-16:0)
C40H82NO7P	M+H	720.590145	0.910592	0.05	PC(O-32:0), PE(O-35:0)
C24H52NO6P	M+K	520.316361	0.908893	0.05	PC(O-16:0)
C39H79N2O6P	M+K	741.530711	0.907533	0.05	PE-Cer(d37:1), SM(d34:1)
C26H54NO6P	M+H	508.376129	0.901582	0.05	PC(O-18:1), PC(P-18:0)
C24H50NO7P	M+H	496.339744	0.890348	0.05	LPC(16:0)
C39H79N2O6P	M+Na	725.556774	0.889093	0.05	PE-Cer(d37:1), SM(d34:1)
C29H47NO4	M+Na	496.339708	0.88767	0.05	Clupanodonyl carnitine, Docosa-4,7,10,13,16-pentaenoyl carnitine, (7Z,10Z,13Z,16Z,19Z)-docosapentaenoylcarnitine
C47H93N2O6P	M+K	851.640262	0.879352	0.05	SM(d42:2)
C47H93N2O6P	M+Na	835.666324	0.876111	0.05	SM(d42:2)
C26H54NO7P	M+H	524.371044	0.867737	0.05	PC(O-18:0), PC(18:0), PE(21:0)
C47H90NO10P	M+Na	882.619434	0.851776	0.05	PS(41:1)
C40H80NO8P	M+H	734.56941	0.851538	0.05	PE(35:0), PC(32:0)
C45H88NO10P	M+Na	856.603784	0.836782	0.1	PS(39:0)
C24H50NO7P	M+K	534.295626	0.836594	0.05	LPC(16:0)
C26H56NO6P	M+H	510.391779	0.831463	0.1	PC(O-18:0)
C24H52NO6P	M+Na	504.342424	0.827575	0.1	PC(O-16:0)
C47H86NO10P	M+H	856.606189	0.825054	0.1	PS(41:3)
C42H84NO7P	M+K	784.561677	0.807992	0.05	PE(P-37:0), PE(O-37:1), PC(O-34:1), PC(P-34:0)
C40H82NO7P	M+K	758.546027	0.80659	0.05	PE(O-35:0), PC(O-32:0)
C26H54NO6P	M+K	546.332011	0.794447	0.05	PC(O-18:1), PC(P-18:0)
C49H88NO10P	M+H	882.621839	0.793157	0.1	PS(22:4(7Z,10Z,13Z,16Z)/21:0), PS(21:0/22:4(7Z,10Z,13Z,16Z))
C42H84NO7P	M+Na	768.58774	0.790919	0.1	PE(P-37:0), PE(O-37:1), PC(O-34:1), PC(P-34:0)
C42H82NO8P	M+H	760.58506	0.79077	0.1	PC(34:1), PE(37:1)
C38H68O15	M+Na	787.445021	0.783402	0.1	13-sophorosyloxydocosanoate 6',6''-diacetate
C29H56NO10P	M+Na	632.353383	0.78049	0.1	PC(21:0(COOH))

C42H86NO7P	M+H	748.621445	0.771179	0.1	PC(O-34:0), PE(O-37:0)
C44H88NO7P	M+H	774.637095	0.765303	0.1	PC(P-36:0), PC(O-36:1), PE(P-39:0)
C42H84NO8P	M+H	762.60071	0.752239	0.1	PC(34:0), PE(37:0), PE-NMe(36:0), PE(37:0)
C47H95N2O6P	M+H	815.70003	0.748999	0.1	SM(d32:1)
C31H60NO10P	M+Na	660.384683	0.746407	0.1	PS(25:0)
C24H50NO7P	M+Na	518.321689	0.745287	0.1	LPC(16:0)
C44H82NO7P	M+H	768.590145	0.738053	0.1	PC(O-36:4), PC(P-36:3)
C43H76NO7P	M+Na	772.525139	0.732245	0.1	PE(O-38:6), PE(P-38:5)
C40H82NO7P	M+Na	742.57209	0.730442	0.1	PE(O-35:0), PC(O-32:0), PE(O-35:0)
C45H91N2O6P	M+H	787.66873	0.724454	0.1	SM(d40:1)
C41H83N2O6P	M+H	731.606129	0.702646	0.1	PE-Cer(d39:1), SM(d36:1)

[1] Metabolite-Signal Match (MSM) score; [2] False Discovery Rate (FDR)

Table S7. METASPACE annotation of formalin-fixed venom gland tissue section against LipidMaps database at the desired FDR level of 0.1 and MSM-score value.

formula	adduct	m/z	msm[1]	fdr[2]	molecule names
C42H84NO7P	M+Na	768.58774	0.979674	0.05	PE(P-37:0), PE(O-37:1), PC(O-34:1), PC(P-34:0)
C39H79N2O6P	M+Na	725.556774	0.974808	0.05	PE-Cer(d37:1), SM(d34:1)
C47H93N2O6P	M+Na	835.666324	0.969944	0.05	SM(d42:2), SM(d42:1)
C42H82NO8P	M+Na	782.567004	0.966001	0.05	PC(34:1)
C49H88NO10P	M+H	882.621839	0.961266	0.05	PS(43:4)
C39H79N2O6P	M+H	703.574829	0.953887	0.05	PE-Cer(d37:1), SM(d34:1)
C42H80NO8P	M+Na	780.551354	0.948187	0.05	PE(37:2), PC(34:2), PE-NMe(36:2)
C42H84NO7P	M+H	746.605795	0.947698	0.05	PE(P-37:0), PE(O-37:1), PC(O-34:1), PC(P-34:0)
C44H84NO6P	M+Na	776.592825	0.939133	0.05	
C39H75O7P	M+Na	709.51424	0.930898	0.05	PA(O-36:2), PA(P-36:1), PA(O-36:2)
C45H91N2O6P	M+Na	809.650674	0.926924	0.05	SM(d40:1)
C44H84NO8P	M+Na	808.582654	0.92398	0.05	PC(36:2), PE(39:2)
C39H71O8P	M+Na	721.477855	0.922759	0.05	PA(36:3)
C26H50O4	M+Na	449.360109	0.920057	0.05	Hexacosanedioic acid
C24H50NO7P	M+Na	518.321689	0.915299	0.05	LPC (16:0)
C44H86NO7P	M+Na	794.60339	0.91342	0.05	PC(P-36:1), PC(O-36:2)
C39H73O8P	M+Na	723.493505	0.913023	0.05	PA(36:2)
C40H80NO8P	M+Na	756.551354	0.909536	0.05	PE(35:0), PC(32:0)
C47H95N2O6P	M+Na	837.681974	0.907259	0.05	SM(d42:1)
C43H74NO9P	M+Na	802.499319	0.902932	0.05	PE(38:6)
C41H83N2O6P	M+Na	753.588074	0.898766	0.05	PE-Cer(d39:1), SM(d36:1)
C33H64NO9P	M+Na	672.421068	0.897499	0.05	PC(25:0)
C22H42O4	M+Na	393.297509	0.895549	0.05	Phellogenic acid, Dioctyl hexanedioate, 3-Acetoxy-eicosanoic acid
C44H82NO7P	M+Na	790.57209	0.894894	0.05	PC(O-36:4), PC(P-36:3)
C29H47NO4	M+Na	496.339708	0.891097	0.05	Clupanodonyl carnitine, Docosa-4,7,10,13,16-pentaenoyl carnitine, (7Z,10Z,13Z,16Z,19Z)-docosapentaenoylcarnitine
C53H98O6	M+Na	853.72554	0.881524	0.05	TG(50:2)
C27H46O4	M+Na	457.328809	0.880309	0.1	1alpha,25-dihydroxy-2beta-hydroxymethyl-19-norcholecalciferol, 1alpha,25-dihydroxy-2alpha-hydroxymethyl-19-nor-20-epivitamin D3
C55H102O6	M+Na	881.75684	0.866331	0.1	TG(52:2)
C53H100O6	M+Na	855.74119	0.865801	0.1	TG(50:1)
C26H54NO7P	M+Na	546.352989	0.856679	0.1	PC(O-18:0), PC(18:0), PE(21:0)

C44H86NO8P	M+Na	810.598304	0.856649	0.1	PC(36:1), PE(39:1)
C39H81N2O6P	M+Na	727.572424	0.853388	0.1	SM(d34:0)
C42H81NO8	M+Na	750.585418	0.851993	0.1	GlcCer(d36:1), DGCC(32:0)
C48H93NO8	M+Na	834.679318	0.846072	0.1	GlcCer(d42:1), GalCer(d42:1)
C44H80NO8P	M+Na	804.551354	0.844162	0.1	PC(36:4), PE(39:4)
C37H75N2O6P	M+Na	697.525474	0.840476	0.1	PE-Cer(d35:1), SM(d32:1)
C44H88NO7P	M+Na	796.61904	0.8331	0.1	PC(P-36:0), PC(O-36:1), PE(O-39:1), PE(P-39:0)
C48H91NO8	M+Na	832.663668	0.827591	0.1	GalCer(d42:2), GlcCer(d42:2)
C46H93N2O6P	M+Na	823.666324	0.823621	0.1	SM(d41:1)
C40H82NO7P	M+Na	742.57209	0.821305	0.1	PE(O-35:0), PC(O-32:0)
C44H82NO8P	M+Na	806.567004	0.820203	0.1	PE(39:3), PC(36:3)
C57H104O6	M+Na	907.77249	0.815608	0.1	TG(54:3)
C46H91NO8	M+Na	808.663668	0.815074	0.1	GalCer(d40:0), GlcCer(d40:0)
C39H73O7P	M+Na	707.49859	0.807237	0.1	PA(O-36:3), PA(P-36:2)
C51H98O6	M+Na	829.72554	0.801083	0.1	TG(48:0)
C46H89NO8	M+Na	806.648018	0.798915	0.1	GlcCer(d40:1), GalCer(d40:1)
C47H90NO10P	M+Na	882.619434	0.788957	0.1	PS(41:1)
C46H80NO8P	M+Na	828.551354	0.788569	0.1	PC(38:6)
C41H75O8P	M+Na	749.509155	0.787066	0.1	PA(38:3)
C42H84NO8P	M+Na	784.582654	0.785837	0.1	PC(34:0), PE(37:0), PE-NMe(36:0)
C55H100O6	M+Na	879.74119	0.784019	0.1	TG(52:3)
C42H82NO7P	M+Na	766.57209	0.783294	0.1	PC(P-34:1), PC(O-34:2), PE(O-37:2), PE(P-37:1)
C41H77O7P	M+Na	735.52989	0.780568	0.1	PA(O-38:3), PA(P-38:2)
C44H84NO7P	M+Na	792.58774	0.779868	0.1	PC(P-36:2), PC(O-36:3)
C41H73O7P	M+Na	731.49859	0.773718	0.1	PA(O-38:5), PA(P-38:4)
C43H87N2O6P	M+Na	781.619374	0.761498	0.1	SM(d38:1)
C53H96O6	M+Na	851.70989	0.743591	0.1	TG(50:3)
C29H45NO3	M+Na	478.329143	0.742581	0.1	Ecalcidene
C46H84NO7P	M+Na	816.58774	0.733849	0.1	PC(P-38:4), PC(O-38:5)
C42H83NO8	M+Na	752.601068	0.731581	0.1	GlcCer(d36:0)
C41H78NO9P	M+Na	782.530619	0.730948	0.1	PS(P-35:1), PS(O-35:2)
C19H40O3	M+Na	339.286944	0.728464	0.1	1,2,4-Nonadecanetriol, 1-Hexadecylglycerol
C36H68NO9P	M+Na	712.452368	0.727593	0.1	PS(P-30:1)
C24H52NO6P	M+Na	504.342424	0.727309	0.1	PC(O-16:0)
C26H52NO7P	M+Na	544.337339	0.722469	0.1	PC(18:1), PC(P-18:0), LPC(18:1)
C49H88NO10P	M+Na	904.603784	0.721871	0.1	PS(43:4)
C41H85N2O6P	M+Na	755.603724	0.717293	0.1	SM(d36:0)
C33H66NO8P	M+Na	658.441804	0.715527	0.1	PC(25:0), PE(28:0)
C51H96O6	M+Na	827.70989	0.711571	0.1	TG(48:1)
C42H86NO7P	M+Na	770.60339	0.711498	0.1	PE(O-37:0), PC(O-34:0)
C57H102O6	M+Na	905.75684	0.705417	0.1	TG(54:4)
C35H68NO9P	M+Na	700.452368	0.704487	0.1	PS(P-29:0)
C55H98O6	M+Na	877.72554	0.702278	0.1	TG(52:4)
C21H41O7P	M+Na	459.248189	0.700384	0.1	PA(18:1)
C37H71O8P	M+Na	697.477855	0.696264	0.1	PA(34:1)
C46H91N2O6P	M+Na	821.650674	0.69337	0.1	SM(d41:2)
C41H73O8P	M+Na	747.493505	0.693113	0.1	PA(38:4)
C33H64NO10P	M+Na	688.415983	0.692319	0.1	PC(25:0(COOH)), PS(27:0)
C45H89N2O6P	M+Na	807.635024	0.691701	0.1	SM(d40:2)
C47H91N2O6P	M+Na	833.650674	0.686062	0.1	SM(d42:3)
C42H78NO8P	M+Na	778.535704	0.681561	0.1	PE(37:3), PC(34:3)
C57H100O6	M+Na	903.74119	0.679773	0.1	TG(54:5)
C29H56NO9P	M+Na	616.358468	0.679204	0.1	PC(21:0(CHO))
C51H94O6	M+Na	825.69424	0.675573	0.1	TG(48:2)

[1] Metabolite-Signal Match (MSM) score; [2] False Discovery Rate (FDR)

ACKNOWLEDGEMENT

I would like to thank the following people, without whom I would not have been able to complete this research, and without whom I would not have made it through my PhD degree.

My deepest appreciation to my supervisor, Prof. Dr. Bernhard Spengler, who with continued support and knowledge, steered me through this research.

I owe a deep sense of gratitude to Prof. Alireza Ghassempour for his keen interest on me at every stage of my research.

And special thanks to AG Spengler's group members; Dr. Sven Heiles, Dr. Stefanie Gerbig, Dr. Karl-Christian Schäfer, Dr. Bernd Commerscheidt, Ulrike Lenz, Vannuruswamy Garikapati, Dr. Dhaka Ram Bhandari, Lilli Walz, Patrik Kadesch, David Lüke, Simon Becher, Fabian Wäldchen and Max Müller whose support allowed my studies to go the extra mile.

I am extremely grateful to my colleague and friend Wendell Albuquerque, who has supported me and had to put up with my stresses and moans for the past four years of study! My sincere thanks also goes to Prof. Dr. Holger Zorn for offering me the opportunities to work in his group on Food Chemistry and Food Biotechnology, leading me working on diverse exciting projects.

And finally my biggest thanks to my family for all the support you have shown me through this research. And for my best friends, Jahangir, Ghazal, and Ali thanks for all your support.
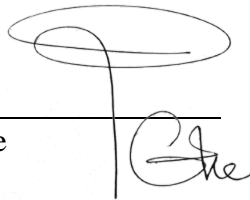
DECLARATION

I declare that I have completed this dissertation without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and referenced all text passages that are derived literally from or are based on the content of published or unpublished work of others, and all information that relates to verbal communications. I have abided the principles of good scientific conduct laid down in the charter of the Justus Liebig University of Giessen in carrying out the investigations described in the dissertation.

Gießen, 12.10.2021

Place, date and signature

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten.

Gießen, 12.10.2021

Ort, Datum und Unterschrift

CURRICULUM VITAE

Der Lebenslauf wurde aus der elektronischen Version der Arbeit entfernt.

The curriculum vitae was removed from the electronic version of the paper.

CURRICULUM VITAE

Publications

1. (2021) **Ghezellou, P.**, Heiles, S., Kadesch, P., Ghassempour, A., Spengler, B. Venom Gland Mass Spectrometry Imaging of Saw-Scaled Viper, *Echis carinatus sochureki*, at High Lateral Resolution. *J. Am. Soc. Mass Spectrom.* 32, 1105-1115.

2. (2021) **Ghezellou, P.**, Albuquerque, W., Garikapati, V., Casewell, N. R., Kazemi, S. M., Ghassempour, A., Spengler, B. Integrating Top-Down and Bottom-Up Mass Spectrometric Strategies for Proteomic Profiling of Iranian Saw-Scaled Viper, *Echis carinatus sochureki*, Venom. *J. Proteome Res.* 20 (1), 895-908.

3. (2021) Albuquerque, W., **Ghezellou, P.**, Li, B., Spengler, B., Will, F., Zorn, H., Gand, M. Identification of Intact Peptides by Top-Down Peptidomics Reveals Cleavage Spots in Thermolabile Wine Proteins. *Food Chem.* 363, 130437.

4. (2021) op den Brouw, B., Coimbra, F. C. P., Bourke, L. A., Huynh, T. M., Vlecken, D. H. W., **Ghezellou, P.**, Visser, J. C., Dobson, J. S., Fernandez-Rojo, M. A., Ikonomopoulou, M. P., Casewell, N. R., Ali, S., Fathinia, B., Hodgson, W. C., Fry, B. G. Extensive Variation in the Activities of Pseudocerastes and Eristicophis Viper Venoms Suggests Divergent Envenoming Strategies Are Used for Prey Capture. *Toxins*. 13 (2), 112.

5. (2021) Derbali, W., Manaa, A., Spengler, B., Goussi, R., Abideen, Z., **Ghezellou, P.** Abdelly, C., Forreiter, C., Koyro, H. Comparative Proteomics Approach to Study the Salinity Effect on the Growth of Two Contrasting Quinoa Genotypes. *Plant Physiol. Biochem.* 163, 215-229.

6. (2021) Rezaie-Atagholipour, M., Imani, F., **Ghezellou, P.**, Seminoff, J. A. Feeding Ecology of Juvenile Green Turtules in Food-Poor Habitats of the Persian Gulf. *Mar Biol.* 168, 4.

7. (2019) **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.* 33 (S1), 20-27.

8. (2018) Nekouie, M., **Ghezellou, P.**, Ghassempour, A., Kiaei, M. Changes in biophysical characteristics of PFN1 due to mutation causing amyotrophic lateral sclerosis. *Metab Brain Dis.* 33 (6), 1975-1984.

9. (2017) Hosseinzadeh, M.S., **Ghezellou, P.**, Kazemi, S.M. Predicating the potential distribution of the endemic Zebra Snake, *Spalerosophis microlepis*, Jan 1865 (Ophidia: Colubridae) in the Zagros Mountains, western Iran. *Salamandra* 53(2), 294-298

10. (2016) Rezaie-Atagholipour, **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* 622, 129-164

11. (2016) Ukuwela, K.D.B., Lee, M.S.Y., Rasmussen, A.R., de Silva, A., Mumpuni, Fry, B.G., **Ghezellou, P.**, Rezaie-Atagholipour, M., Sanders, K.L. Evaluating the drivers of Indo-Pacific biodiversity: speciation and dispersal of sea snakes (Elapidae: Hydrophiinae). *J. Biogeogr.* 43, 243-255.

12. (2015) Khorjestan, S.M., Abtahi, B., Ranei Siadat, S.O., Motevalli, M., Rezadoost, H., **Ghezellou, P.**, Ghassempour, A. Analysis of Annulated Sea Snake Venom , *Hydrophis Cyanocinctus* , Using Liquid Chromatography and MALDI-TOF / TOF. *Curr. Proteomics* 12, 45-55.

13. (2014) Abtahi, B., Mosafer Khorjestan, S., **Ghezellou, P.**, Aliahmadi, A., Ranaei Siadat, S.O., Kazemi, S.M., Ghassempour, A., Fathinia, B. Effects of Iranian Snakes Venom ; True Sea and Terrestrial Snakes on Some Bacterial Cultures. *J. Persian Gulf* 5.

14. (2012) Calvete, J.J., **Ghezellou, P.**, Paiva, O., Matainaho, T., Ghassempour, A., Goudarzi, H., Kraus, F., Sanz, L., Williams, D.J. Snake Venomics of Two Poorly known Hydrophiinae: Comparative Proteomics of the Venoms of Terrestrial *Toxicocalamus longissimus* and Marine *Hydrophis cyanocinctus*. *J Proteomics* 75, 4091–4101.

15. (2013) Samiei, J.V., Dab, K., **Ghezellou, P.**, Shirvani, A. Some Scleractinian Corals (Scleractinia: Anthozoa) of Larak Island, Persian Gulf. *Zootaxa* 3636, 101–143.

16. (2012) Rezaie-Atagholipour, M., Riyahi-Bakhtiari, A., Rajabizadeh, M., **Ghezellou, P.** Status of the Annulated Sea Snake, *Hydrophis cyanocinctus* , in the Hara Protected Area of the Persian Gulf. *Zool. Middle East* 57, 53–60.

17. (2012) Rezaie-Atagholipour, M., Riyahi-Bakhtiari, A., Sajjadi, M., Yap, C.K., Ghaffari, S., Ebrahimi-Sirizi, Z., **Ghezellou, P.** Metal concentrations in selected tissues and main prey species of the annulated sea snake (*Hydrophis cyanocinctus*) in the Hara Protected Area, northeastern coast of the Persian Gulf, Iran. *Mar Pollut Bull* 64, 416-421.

Conferences

1. (2020) 53rd Annual DGMS Conference,
   01-04 March 2020, Münster, Germany. *Poster presentation*

2. (2019) 12th Annual GGL Conference,
   04-05 September 2019, Giessen, Germany. *Oral presentation*

3. (2018) 52nd Annual DGMS Conference,
   10-13 March 2019, Rostock, Germany. *Oral presentation*

4. (2018) 22nd International Mass Spectrometry Conference,
   26-31 August 2018, Florence, Italy. *Poster presentation*

5. (2018) Snakebite Conference – From Science to Society,
   21-22 June 2018, Leiden, Netherlands. *Oral presentation*

6. (2018) 51st Annual DGMS Conference,
   11-15 March 2018, Saarbrüchen, Germany. *Oral presentation*

7. (2014) NASIC Workshop on Modern Spectroscopic Techniques and Their Applications in Structure Determination, 1-3 December 2014, University of Karachi, Pakistan.

8. (2014) 1th Iranian Peptide Symposium and Workshop on Peptide Purification, 27 November 2014, Shahid Beheshti University, Tehran, Iran.

9. (2011) 2nd International Congress on Analytical Proteomics, 17-20 July 2011, Ourense, Spain. *Poster presentation*

10. (2011) The First Regional Congress on Venomous Animals and Toxin, 22-24 November 2011, Karaj, Iran. *Oral presentation*