



OPEN

## Identification, classification, and characterization of alpha and beta subunits of LVP1 protein from the venom gland of four Iranian scorpion species

Fatemeh Salabi<sup>1</sup>, Babak Vazirianzadeh<sup>2</sup> & Masoumeh Baradaran<sup>3</sup>✉

Scorpion venoms contain bioactive peptides and proteins. Some, can be used for pharmaceutical purposes. So, identification of venom proteins matters because, in addition to determining the function of the toxins can also be an excellent guide to developing new drugs. Here, we got transcriptome of venom glands from four Iranian scorpion species, including *Hemiscorpius lepturus*, *Mesobuthus eupeus*, *Androctonus crassicauda*, and *Hottentotta saulcyi* using cDNA library synthesis and high-throughput transcriptomic analysis of the venom glands. In a comparative way, we identified the cDNA encoding isoforms of subunits (alpha and beta) of BotLVP1/BmLVP1-like protein in the venom gland of three species except for *H. lepturus*. Characterization and structure determination of the LVP1-like proteins combined with gene map analysis provided evidence of the existence of some isoforms of LVP1-like proteins, encoded by genes with two exons and one intron, which can be classified in CSαβ superfamily in the venom gland of three Iranian scorpion species. According to the high similarity with BotLVP1 and BmLVP1, these proteins could also be potent to mediate cholesterol homeostasis. However, further research is needed to prove it, and this study just may lay the foundation lead to light up this way.

Scorpions and scorpion stings have taken an essential part in public health all over the world, including in Iran. The frequency of scorpion stings is very significant in Iran, especially in the southern and southwestern regions of the country<sup>1</sup>. Most cases of scorpion stings have been reported from six species of scorpions, including *Mesobuthus eupeus*, *Hemiscorpius lepturus*, *Hottentotta saulcyi*, *Odonthobuthus doriae*, *Androctonus crassicauda* and *Hottentotta schach*<sup>2</sup>. Therefore, the antivenom produced by the Razi Institute of Vaccine and Serum Production in Karaj, Iran, for the treating scorpion stings is prepared using a mixture of venom extracted from these six medically important species<sup>3</sup>.

Epithelial cells of the scorpion's venom gland synthesize components of the venom and release them into the lumen of the gland where stores them<sup>4</sup>. These compounds comprise peptides, proteins, enzymes, amines, and nucleotides<sup>5</sup>. Scorpions use these compounds for defense, feeding, reproduction, and overall survival<sup>6</sup>. Nevertheless, recent research on determining the function and potential of these compounds has shown that although some neurotoxins and cytotoxins in the venom gland of scorpions threaten human life, many venom peptides and proteins revealed a broad-spectrum bioactivity<sup>7,8</sup> in the treatment of many diseases including, cancers<sup>9,10</sup>, infections<sup>11</sup>, and cardiovascular disorders<sup>12</sup>, epilepsy<sup>13</sup>, autoimmune and inflammatory disease<sup>14</sup>.

There is also some evidence about the presence of the compounds interfering with lipid metabolism in the venom gland of scorpions. An increase in serum free fatty acids and phospholipids; and a reduction in total cholesterol in canine injected by scorpion venom<sup>15</sup>, inhibiting the activity of HMG-CoA reductase by a protein, Lipolysis-activating peptide 1-beta (BmLVP1\_beta, or bumarsin), isolated from *Mesobuthus martensii* venom gland<sup>16</sup>, and the presence of several Lipolysis activating peptides (LVP1s) in venom glands of some scorpion including Chinese *Buthus martensii*<sup>17</sup>, *Lychas mucronatus*<sup>18</sup>, and *Buthus occitanus tunetanus*<sup>19</sup> have been proved.

<sup>1</sup>Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Ahvaz, Iran. <sup>2</sup>Social Determinant of Health Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran. <sup>3</sup>Toxicology Research Center, Medical Basic Sciences Research Institute, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran. ✉email: mb.baradaran@gmail.com

LVP1 is a lipolytic peptide that was first isolated from the venom of *Buthus occitanus* and named BotLVP1. Some experiments on BotLVP1, revealed that it is a heterodimeric protein with alpha and beta subunits structurally related to Na<sup>+</sup> channel toxins but is not toxic and can stimulate lipolysis in rat adipocytes<sup>20,21</sup>.

Further research identified two cDNA encoding subunits  $\alpha$  and  $\beta$  of a BotLVP1-like peptide (BmLVP1) from the venom gland of *M. martensii*. The heterodimer structure determined for BmLVP1 further supported the relationship between BmLVP1 subunits and BotLVP1 subunits<sup>22</sup>.

More research indicated that two LVP1-beta subunits originated from *M. martensii* form a homodimer named bumarcin, which significantly inhibited HMG-CoA reductase compared to simvastatin<sup>16</sup>.

All in all, LVP1 was considered a candidate for the production of an alternative drug for the treatment of hyperlipidemia<sup>16</sup>. According to the importance of this protein, in this study, the venom gland transcriptome of four medically important scorpion species, including *H. saulcyi*, *M. eupeus*, *A. crassicauda*, and *H. lepturus* was analyzed to find possible isoforms of two subunits (alpha and beta) of LVP1. In addition, the characteristics of LVP1 subunits isoforms and their three-dimensional structures, were determined in different scorpion species. Furthermore, the exon–intron map of the gene encoding this protein was also determined.

## Materials and methods

### Sample preparation

Whatever will be explained here about *H. saulcyi* has previously been done for *H. lepturus*, *A. crassicauda*<sup>23</sup>, and *M. eupeus*<sup>24</sup> using the same methodology.

Scorpions of species *H. saulcyi* were collected from the deserts around Khuzestan province, southwest of Iran, and were transferred to the laboratory of the Toxicology research center located in Ahvaz Jundishapur University of Medical Sciences for identification and confirmation of the species according to the morphological properties<sup>25</sup>. Confirmed scorpion samples were milked by electrical shock, 72 h later, they were transferred to the molecular laboratory for RNA extraction.

### RNA extraction and cDNA library synthesis

A total of twenty specimens of *H. saulcyi* were captured in summer 2022. Considering that based on the general guidelines for RNA-Seq, at least three biological replicates are required to generate a reliable RNA-Seq analysis<sup>26</sup>, we selected six mature scorpions out of all captured specimens to construct cDNA libraries. The terminal segment (telson) of selected scorpions was separated and powdered with the help of the liquid nitrogen. Total RNA was extracted under sterile conditions using RNeasy Mini Kit (Cat. 74104) according to the manufacturer's instructions. The quality of extracted RNAs were evaluated using Nanodrop (co. Thermo, USA). Each of the three RNA samples were pooled in equal concentrations to form two samples prior to sequencing<sup>27</sup>. The RNA Integrity Number (RIN) of two pooled RNA samples was determined by Macrogen, Inc. (Seoul, Korea) using the Agilent 2100 Bioanalyzer System (Agilent Technologies, USA) according to the manufacturer's instructions. Two pooled RNA samples with appropriate RIN (A RIN value higher than 7) were used to prepare rRNA-depleted libraries using the TruSeq<sup>®</sup> Stranded mRNA Library Prep (cat. number 20020594) according to the manufacturer's instructions. Finally, the Illumina HiSeq 2000 platform was used to sequence cDNA libraries with 150 bp paired-end reads.

### Transcriptome assembly and annotation

First, the quality of the raw RNA-Seq reads was assessed with FastQC v0.11.5 (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>) using default parameters. Subsequently, reads were subjected to TRIMMOMATIC v0.36 to remove low-quality reads and any Illumina adapter remnants, with the following settings: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:25 TRAILING:25 MAXINFO:90:0.90 AVGQUAL:20 MINLEN:50. Finally, the FastQC was performed to verify the final quality of the dataset.

A de novo assembly of clean reads was then performed using Trinity software (v2.10.0)<sup>28,29</sup> with a minimum contig length of 200 bp. Since no complementary genome of the studied scorpion species and no reference genome from a closely related species is available, the de novo transcriptome assembly method was used to determine the transcripts. The reads were mapped to transcriptome assemblies using Bowtie2 v2.3.5.1, under default parameters. Finally, TransDecoder v5.5.0 (<https://github.com/TransDecoder/TransDecoder/releases>) was employed to predict candidate coding regions in transcript sequences, discarding possible non-coding RNA and DNA contamination with “-t” option for Transdecoder.LongOrfs and Transdecoder.Predict<sup>29</sup>.

### Characterization of LVP1 subunits in scorpion species

In order to identify the proteins and nucleotides with the greatest sequence similarity to protein LVP1, alpha and beta subunits originated from scorpions and closely related species were manually searched. With Blastx and Blastp and the E value threshold of  $1e^{-3}$ , the TransDecoder-predicted proteins and nucleotide sequences were searched against the local LVP1 alpha and beta database. Finally, the homologous sequences related to LVP1 alpha and beta were selected from mRNA candidates of all four scorpion species. These sequences were classified based on similarity to already known families/classes. For this purpose, newly extracted sequences were searched against NCBI and UniProt databases. A sequence alignment was done using the EBI online server (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Amino acid alignment of isoforms was done using BioEdit software (<https://www.bioedit.com/>).

### Characterization of the gene and protein of LVP1-alpha and LVP1-beta

Characterization of the gene and protein of the predicted LVP1-alpha and LVP1-beta proteins not only increases the confidence in the sequences but also provides more complete information about the obtained proteins.

An intron–exon map of LVP1-alpha and LVP1-beta were predicted by comparing whole genome shotgun sequences with mRNAs, as described previously<sup>30</sup>. The mRNA sequence of KTXLP2 (alternative name: LVP1-beta) (AF155368.1) with its genomic DNA from contig339532 of whole genome shotgun sequence (AYEL01080250.1) from *Mesobuthus martensii* along with similar sequence originated from *Hottentota trilineatus* (SRX815889) were used as templates to compare with the LVP1-alpha and LVP1-beta mRNA sequences found in the current study. A sequence alignment was created with MAFT (high-speed multiple sequence alignment program) online server in EBI (<https://mafft.cbrc.jp/alignment/server/index.html>). In the following, potent signal peptides of all predicted HMG-CoA reductase inhibitors were predicted by SignalP-6.0 available on server <https://services.healthtech.dtu.dk/service.php?SignalP>. Physicochemical properties of mature proteins of LVP1-alpha and LVP1-beta, including molecular weight, Iso-electric point, half-life, water solubility, and net charge, were determined with the ProtParam tool from ExPasy (<https://web.expasy.org/protparam/>) and proteomics tool from INNOVAGEN (<http://pepcalc.com>). Disulfide bridge patterns were determined using the DISULFIND server (<https://bio.tools/disulfind>).

### Determination the 3D-structure of LVP1-alpha and LVP1-beta

The 3D-structure of the newly identified LVP1-alpha and LVP1-beta isoforms was determined via Homology modeling. For each protein, four models were generated using four different online modeling servers as follows: I-TASSER by multiple threading approaches, PHYRE2 with Poing<sup>2</sup> method, Robetta using RoseTTAFold, and SWISS-MODEL. Quality assessment of obtained structures was employed using the overall quality factor on SAVESv6.0 (<https://saves.mbi.ucla.edu/>) as well as by Z-score through the ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) to select the best model for each protein. A more negative Z-scores represents a more validated structure<sup>31</sup>. The selected protein structures were energetically minimized using the YASARA Minimization Server (<http://www.yasara.org/minimizationserver.htm>). Finally, to determine whether the predicted protein structures have the permitted torsional angles, the Ramachandran plot was drawn using the PSVS server (<https://montelionelab.chem.rpi.edu/PSVS/PSVS/>).

### Phylogenetic analysis

By searching newly identified LVP1-alpha and LVP1-beta proteins against the NCBI database using BlastP, homologous peptides of these proteins were extracted. Subsequently, a multispecies sequence alignment of LVP1-alpha and LVP1-beta amino acid sequences was created using the MUSCLE tool at Mega11 software<sup>32</sup>. LVP1-alpha and LVP1-beta originated from different scorpion species were used in this alignment. Ultimately, this multiple sequence alignment was utilized to generate a phylogenetic tree by the neighbor-joining algorithm<sup>32</sup>. The neighbor-joining tree was constructed under the Jones–Taylor–Thornton (JTT) model with 1000 bootstrap replicates.

### Ethics approval and consent to participate

Permission was obtained from the Environmental Protection Agency of Iran to collect scorpions of *A. crassicauda*, *H. saulcyi*, and *H. lepturus*. No animals were euthanized as part of this study, and all sample collection methods and experimental procedures described herein were rigorously reviewed and approved by the Ahvaz Jundishapur University of Medical Sciences (Ethical code: IR.AJUMS.REC.1400.557 and IR.AJUMS.REC.1398.785), the Institutional Animal Care Committee of Razi Vaccine and Serum Research Institute (Permit number IR.RVSRI.REC.1401.017) and AREEO protocols, which comply with Iran guidelines for work with animals. This study also adheres to the ARRIVE Guidelines for reporting animal research.

## Results

### Assembly and annotation of raw sequences obtained from RNA sequencing

The transcriptome of *H. saulcyi* obtained from RNA sequencing in this study was analyzed along with the transcriptome of three additional scorpions, *H. lepturus*, *A. crassicauda*<sup>30</sup>, and *M. eupeus*<sup>24</sup>, obtained using precisely the same methodology in the previous studies.

Through Illumina paired-end sequencing of venom glands of *H. saulcyi*, 97 million paired 150 bp clean reads were obtained after adapter and low-quality reads trimming. Using Trinity de novo assembler 191,150 contigs greater than 200 bp in length were produced, including different isoforms per contig. The contigs were assembled into 110,126 unigenes. Finally, we identified 98,365 potential coding sequences in venom glands of *H. saulcyi* transcriptome assemblies. Most of these protein-coding sequences (13,723) were matched to non-redundant proteins (Nrs), Swissprot, and Pfam databases. This value was 11,415 and 11,869 transcripts for *A. crassicauda* and *H. lepturus*, respectively<sup>23</sup>.

### Identification of the LVP1-alpha and LVP1-beta proteins

A local database of LVP1-alpha and LVP1-beta proteins was generated by collecting the known sequences of LVP1-alpha and LVP1-beta from different scorpion species (Table 1).

Using our local LVP1 database, we conducted exhaustive BLAST searches of the *H. saulcyi*, *H. lepturus*, *M. eupeus*, and *A. crassicauda* venom glands transcriptome and found some transcripts with high sequence similarity to LVP1-alpha and LVP1-beta. In order to increase the confidence of sequence identification and to classify the sequences, the obtained sequences were directly searched against Uniprot and NCBI databases. All sequences with high sequence identity with any previously classified LVP1-alpha and LVP1-beta sequences belonging to any recognized species were considered members of this group. Accordingly, we found that scorpions encode some isoforms of alpha and beta subunits of LVP1. The cDNA and protein sequences of all isoforms of LVP1-alpha and LVP1-beta reported here have been deposited in the GeneBank database under specific names and Accession

Taxon	Protein name	GenBank ID
<i>Buthus occitanus tunetanus</i>	Lipolysis activating peptide 1-alpha chain (BotLVP1-alpha)	P84809
<i>Mesobuthus martensii</i>	Lipolysis activating peptide 1-alpha chain (BmLVP1-alpha)	Q6WJF5
<i>Buthus occitanus israelis</i>	Lipolysis activating peptide 1-alpha chain (BoiLVP1-alpha) or Putative excitatory toxin Tx135 (BoiTx135)	B8XH01
<i>Lychas mucronatus</i>	Lipolysis activating peptide 1-alpha chain (LVP1-alpha or Neurotoxin LmNaTx7)	D9U299
<i>Buthus occitanus tunetanus</i>	Lipolysis activating peptide 1-beta chain (BoiLVP1-beta)	P84810
<i>Mesobuthus martensii</i>	Lipolysis activating peptide 1-beta chain (HMG-CoA reductase inhibitor or bumarsin)	Q95P90
<i>Buthus occitanus israelis</i>	Lipolysis activating peptide 1-beta chain (BoiLVP1-beta or beta-like toxin Tx457)	B8XGZ8
<i>Lychas mucronatus</i>	Lipolysis activating peptide 1-beta chain (LVP1-beta or Neurotoxin or LmNaTx19)	D9U2A2

**Table 1.** Lipolysis activating peptide sequences used to generate the multiple sequence alignment for detection the probable similar peptides based on the conserved sequences and phylogenetic analysis.

numbers (Table 2). The homologous sequence of alpha and beta subunits of LVP1 protein was not found in the venom gland transcriptome of *H. lepturus*. In contrast, in transcriptomes of *H. saulcyi*, *M. eupeus*, and *A. crassicauda* we detected a total of three isoforms of LVP1-beta subunits (OP612332, OP612333, and OP612334), two isoforms of alpha and beta subunits of LVP1 (KU513844 and KU513847), and five isoforms of alpha and beta subunits (two alpha and three beta; OP609899, OP609900, OP609901, OP609902, OP609903), respectively.

All identified proteins showed sequence similarities to the known alpha or beta-type of lipolysis activating peptide (LVP1) from other scorpions. BmLVP1-alpha and BmLVP1-beta from scorpion *Mesobuthus martensii* and BotLVP1-alpha and BotLVP1-beta from *Buthus occitanus* are well-studied scorpion LVP1s. Accordingly, the similarity of the identified proteins with these proteins was used as the criterion and is reported in Table 3.

Amino acids alignment of identified LVP1-alphas and LVP1-betas are shown in Fig. 1. This alignment clarified that except for AcLVP1\_alpha1 and AcLVP1\_alpha2, which consist of eight cysteine residues, all other proteins contain six cysteine residues (Fig. 1).

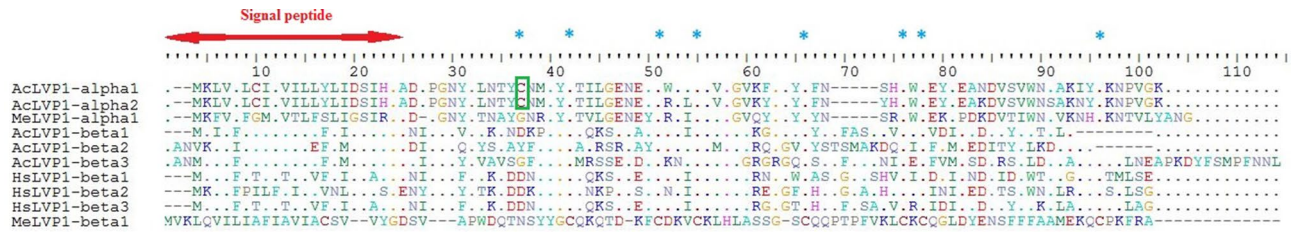
Scorpion species	Protein name	Accession number
<i>A. crassicauda</i>	LVP1-alpha_isoform1(AcLVP1-alpha_isoform1 or AcLVP1_alpha1)	OP609899
<i>A. crassicauda</i>	LVP1-alpha_isoform2(AcLVP1-alpha_isoform2 or AcLVP1_alpha2)	OP609900
<i>M. eupeus</i>	LVP1-alpha_isoform1(MeLVP1-alpha_isoform1 or MeLVP1_alpha1 or meuPep27)	KU513844
<i>A. crassicauda</i>	LVP1-beta_isoform1(AcLVP1-beta_isoform1 or AcLVP1_beta1)	OP609901
<i>A. crassicauda</i>	LVP1-beta_isoform2(AcLVP1-beta_isoform2 or AcLVP1_beta2)	OP609902
<i>A. crassicauda</i>	LVP1-beta_isoform3(AcLVP1-beta_isoform3 or AcLVP1_beta3)	OP609903
<i>M. eupeus</i>	LVP1-beta_isoform1(MeLVP1-beta_isoform1 or MeLVP1_beta1 or meuPep28)	KU513847
<i>H. saulcyi</i>	LVP1-beta_isoform1(HsLVP1-beta_isoform1 or HsLVP1_beta1)	OP612332
<i>H. saulcyi</i>	LVP1-beta_isoform2(HsLVP1-beta_isoform2 or HsLVP1_beta2)	OP612333
<i>H. saulcyi</i>	LVP1-beta_isoform3(HsLVP1-beta_isoform3 or HsLVP1_beta3)	OP612334
<i>H. lepturus</i>	-	-

**Table 2.** Identified and submitted isoforms of LVP1-alpha and LVP1-beta in GeneBank database.

Protein name	Identity with BmLVP1-alpha (%)	Identity with BotLVP1-alpha (%)	Identity with BmLVP1-beta (%)	Identity with BotLVP1-beta (%)
AcLVP1_alpha1	63.04	70.65	-	-
AcLVP1_alpha2	65.22	70.65	-	-
MeLVP1_alpha1	88.66	72.16	-	-
AcLVP1_beta1	-	-	67.47	59.52
AcLVP1_beta2	-	-	50	53.93
AcLVP1_beta3	-	-	58.51	58.06
HsLVP1_beta1	-	-	58.43	51.65
HsLVP1_beta2	-	-	55.43	56.52
HsLVP1_beta3	-	-	65.22	59.78
MeLVP1_beta1	-	-	75.53	61.05

**Table 3.** Identity of here identified proteins with well-known LVP1s subunit alpha and beta.





**Figure 1.** Amino acid alignment of identified LVP1-alphas and LVP1-betas from three scorpion species: *H. saulcyi*, *A. crassicauda*, and *M. eupeus*. Signal peptides are indicated with a red straight line leading to an arrow. Dots and dashes represent identical and deleted amino acids, respectively. Cysteine residues are indicated by blue asterisks. AcLVPI1\_alpha1 and AcLVPI1\_alpha2 consist of eight cysteine residues, while the other proteins consist of six cysteine residues. The green rectangle highlights two more cysteine residues in AcLVPI1\_alpha1 and AcLVPI1\_alpha2.

### Exon–intron map of scorpion LVP1s

In order to organization the exon–intron patterns of LVP1 subunits alpha and beta genes (Fig. 2), the mRNA sequences of LVP1 beta or alpha discovered in this study and previously identified similar sequences originated from *Hottentota trilineatus* were aligned with the genomic and mRNA sequences of LVP1 proteins from *M. martensii* using MAFT program. The total nucleotide sequence of both LVP1-alpha and LVP1-beta from START-codon (ATG) to STOP-codon (TAA) is around 665 bp (Fig. 2). The LVP1 gene was found to consist of 2 coding exons and one intron. Exon 1 runs from nucleotide 91 to 127 and consists of 36 nucleotides, while the more extended exon 2 spans from nucleotide 413 to 665 and comprises 253 nucleotides. The exons and intron of LVP1 genes follow the standard form of GT-AG.

### Tertiary structure and physicochemical properties of LVP1-alpha and LVP1-beta isoforms

By identifying and removing the signal peptide from the sequence of identified LVP1-alpha and LVP1-beta proteins, mature proteins were obtained. Then, the determination of physical and chemical characteristics, as well as the determination of the three-dimensional structure of mature proteins, were carried out.

Physicochemical properties, including molecular weight, theoretical pI, half-life in mammalian reticulocytes, water solubility, and instability index of all found proteins, are described in Table 4.

The tertiary structure of all identified LVP1-alpha and LVP1-beta proteins were predicted and illustrated in Fig. 3. For each protein, the structure with the best quality point and the most negative score among the structures made by four servers was selected as the structure of that protein. What Fig. 3 shows is the selected protein structures after energy minimization.

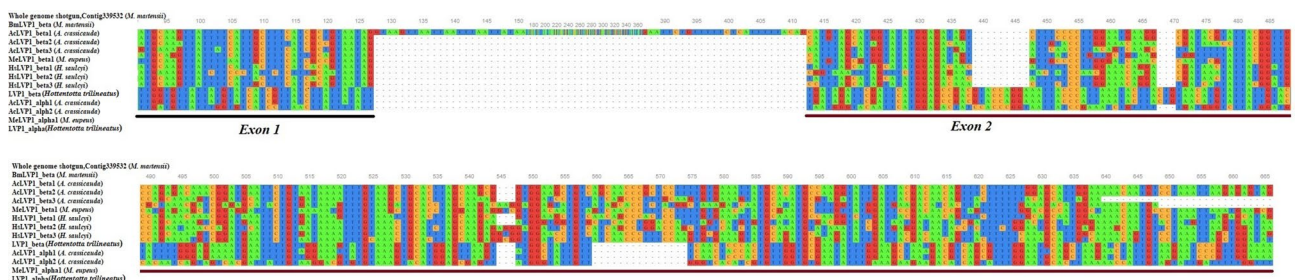
As seen in this figure, the LCN-type cysteine-stabilized alpha/beta (CS- $\alpha/\beta$ ) domain is a common domain seen in all LVP1-alpha and LVP1-beta proteins. This domain consists of one or two alpha helixes connected to a two or triple-stranded beta-sheet through three or four disulfide bonds<sup>33</sup>.

What is clear is that, except for MeLVPI1\_alpha1 and AcLVPI1\_alpha1 proteins, whose beta-sheet of CS- $\alpha/\beta$  domain is three-stranded, the rest of the identified proteins have a double-stranded beta-sheet.

As mentioned above, AcLVPI1\_alpha1 and AcLVPI1\_alpha2 consist of 8 cysteine residues, and the other LVP1-alpha and LVP1-beta isoforms contain six cysteine residues, which indicate the formation of 3 or 4 disulfide bonds in the tertiary structure of these proteins, respectively.

Ramachandran plot analysis (Fig. 4) revealed that more than 97% of all amino acids of predicted protein structures are located in the allowed regions, which verify the quality of the predicted models.

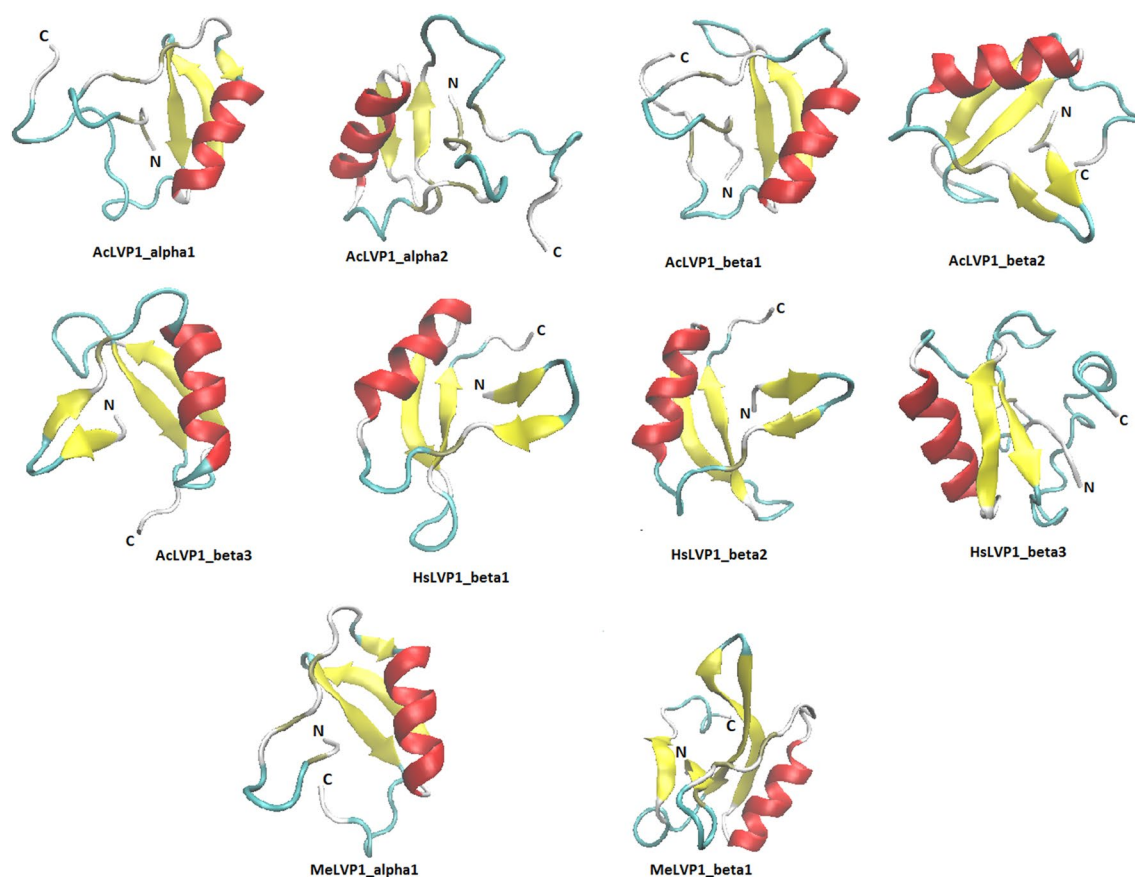
Members of the CS $\alpha\beta$  superfamily consist of six or eight cysteine residues. According to the number of disulfide bridges and the location of the fourth one, the CS $\alpha\beta$  superfamily is categorized into six groups (I–VI)



**Figure 2.** Exon–intron map of LVP1 alpha and beta gene. Accession numbers from top to bottom: AYEL01080250.1, AF155368.1, OP609901, OP609902, OP609903, KU513847, OP612332, OP612333, OP612334, SRX815889, OP609899, OP609900, KU513844, SRX815889. Coloured box parts are nucleotides of exons and the space between two exons is the only intron of these genes.

Scorpion species	Scorpion species	Mature peptide (aa)	Molecular weight (g/mol)	Theoretical pI	half-life in mammalian reticulocytes (h)	Water solubility	instability index
<i>A. crassicauda</i>	AcLVP1_alpha1	72 aa	8316.53	6.89	4.4	Poor	23.52 (stable)
	AcLVP1_alpha2	72aa	8423.65	8.16	4.4	Poor	21.95 (stable)
	AcLVP1_beta1	73aa	8511.55	8.20	1.1	Good	8.31 (stable)
	AcLVP1_beta2	64aa	7288.30	6.71	1.1	Good	28.78 (stable)
	AcLVP1_beta3	87aa	9894.32	6.72	1.1	Poor	72.26
<i>M. eupeus</i>	MeLVP1_alpha1	67aa	7819.94	6.72	1.1	Good	59.67 (unstable)
	MeLVP1_beta1	87aa	9894.37	6.72	1.1	Poor	72.26 (unstable)
<i>H. saulcyi</i>	HsLVP1_beta1	72aa	8198.33	8.22	1.1	Good	28.88 (stable)
	HsLVP1_beta2	72aa	8269.26	4.56	1.1	Good	48.35 (unstable)
	HsLVP1_beta3	73aa	8371.45	8.24	1	Good	46.50(unstable)

**Table 4.** Physicochemical properties of identified alpha and beta subunits of LVP1.

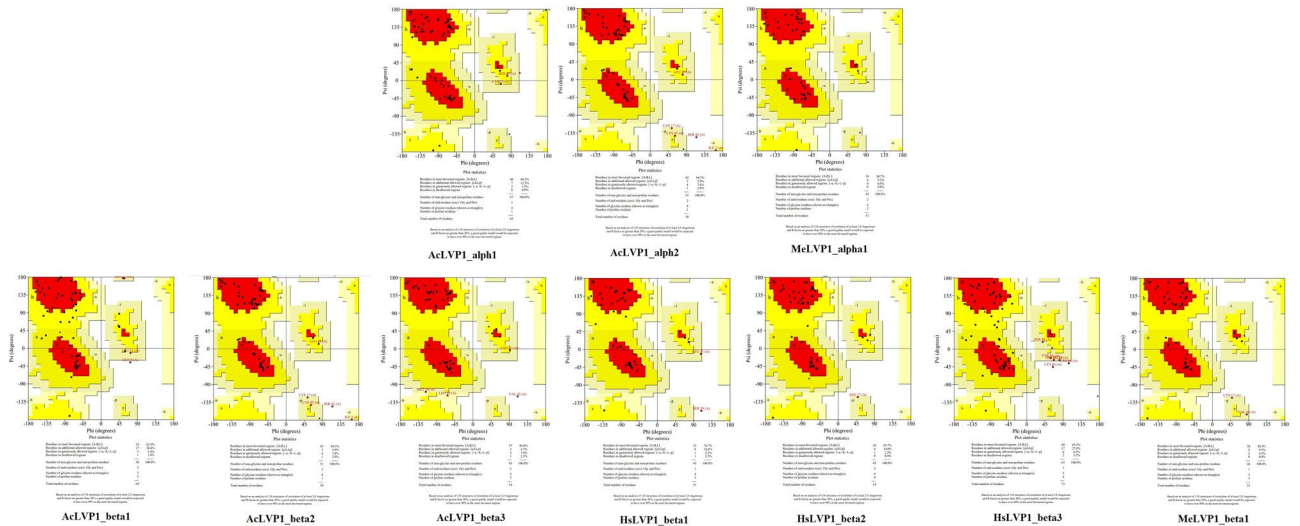


**Figure 3.** Three-dimensional structure of LVP1-alphas and LVP1-betas originated from *A. crassicauda*, *H. saulcyi*, and *M. eupeus* venom gland.

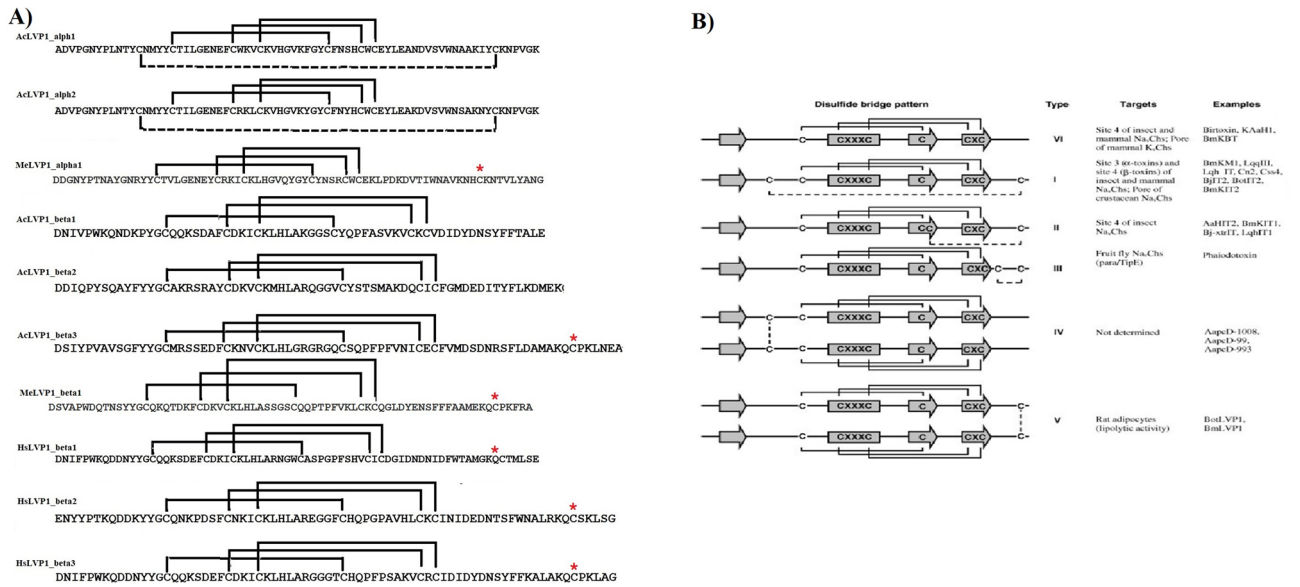
(Fig. 5)<sup>22</sup>. The disulfide bridge pattern of the identified peptides was shown in Fig. 5 to compare with the six identified groups of CSa $\beta$  superfamily proteins.

### Phylogenetic relationships based on LVP1

The phylogenetic relationships according to LVP1-alphas and betas are shown in Fig. 6. The phylogenetic tree is divided into two main monophyletic groups, A and B. The proteins in group A (light green) are all LVP1-alphas, and the members of group B (light brown) are all LVP1-betas. Group A is subdivided into two clades. Clade 1 is composed of LVP1-alphas related to species of *Centruroides*, *Leiurus*, and *Tityus* genus, and clade 2 encompass LVP1-alphas of *Lyctas mucronatus*, *A. crassicauda* (here identified proteins plus previously identified proteins), *Buthus occitanus*, *M. martensii* and here identified MeLVP1\_alpha1 from *M. eupeus*. None of the identified LVP1-alphas belong to clade 1. In clade 2, two isoforms of LVP1-alphas from *A. crassicauda* (AcLVP1\_alpha1 and AcLVP1\_alpha2) clustered as a sister group which were diverged from an ancestor from which toxin Acra



**Figure 4.** Ramachandran plot of predicted protein structures. LVP1-alfas are shown in above and LVP1-betas in bottom.

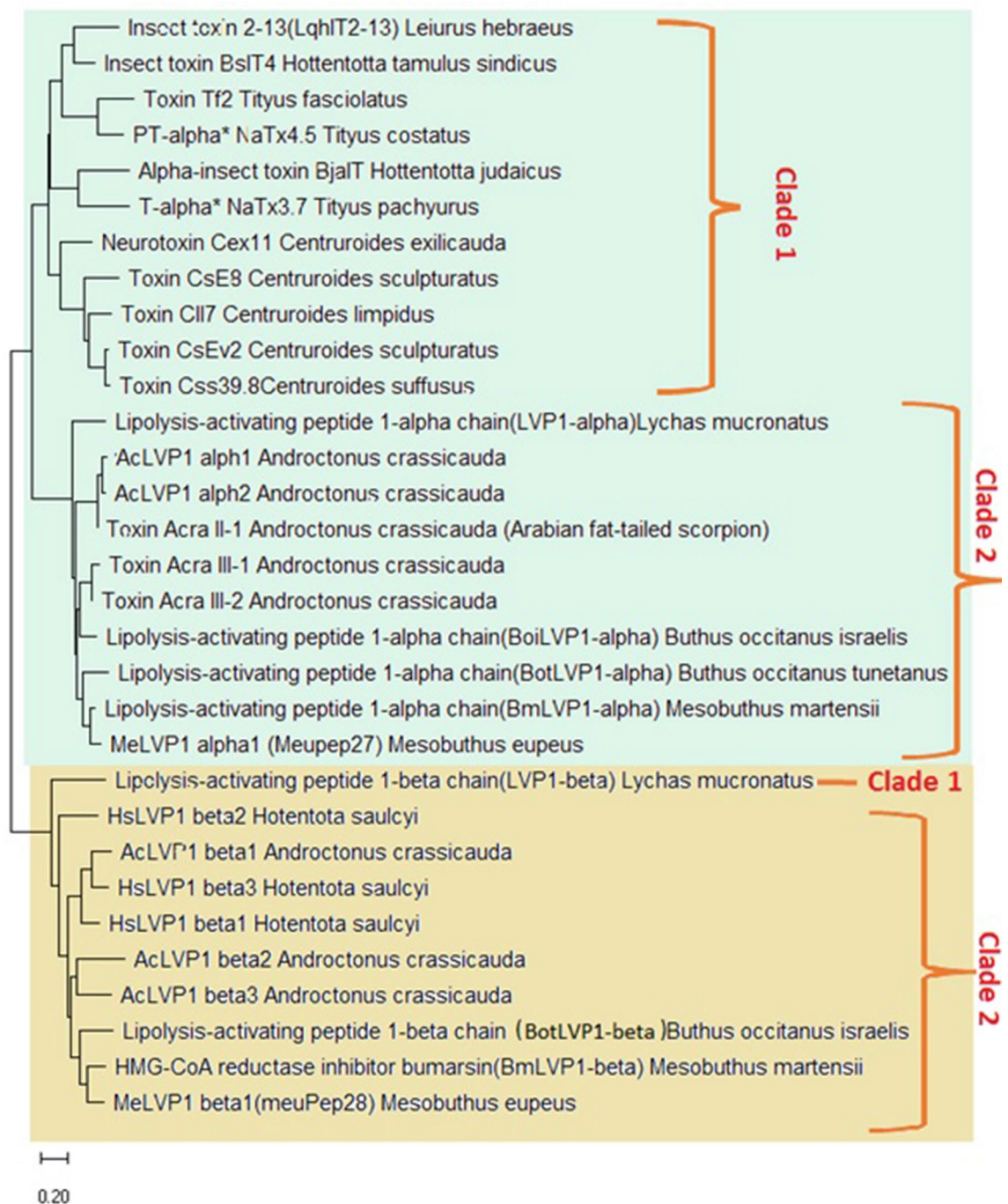


**Figure 5.** (A) Disulfide bridge pattern of identified peptides. The name of each peptide are deposited above the related sequence. The unbounded cysteine residues are shown with red stars. (B) Categorization of CSaβ superfamily according to the disulfide bridge pattern scorpion proteins taken from<sup>17</sup>.

II-1 (from Arabian *A. crassicauda*) was deviated. The other here identified LVP1-alpha, MeLVP1\_alpha1 (or meuPep27) from *M. eupeus* formed a sister group with BmLVP1-alpha from *M. martensii*. This could be because *M. eupeus* and *M. martensii* both belong to the same genus. This group has a common ancestor with BotLVP1-alpha from *B. occitanus*. Since *B. occitanus* is a species of the Buthidae family like *M. eupeus* and *M. martensii*, such a position in the phylogenetic tree is justified. Although *A. crassicauda* and *M. eupeus* are also in the same family, MeLVP1\_alpha1 and AclLVP1\_alpha1,2 proteins are more distant in clade 2. AclLVP1\_alpha1,2 contain eight cysteine residues and four disulfide bridges, while MeLVP1\_alpha1 contains seven cysteine residues. A possible scenario is that these proteins diverged from a common ancestor. However, during evolution, eight cysteine residues have changed to another amino acid, and a free cysteine has constructed to form a dimer protein.

Monophyletic group B is also subdivided into two main clades. Clade 1 is LVP1-beta from *L. mucronatus*, and clade two is subdivided further into two branches. One includes HsLVP1\_beta2, and the other includes the ancestor of LVP1-betas of *B. occitanus*, *M. martensii*, and all the other LVP1-betas identified here. It means HsLVP1\_beta2 diverged from the same ancestral protein that gave rise to BmLVP1-beta, BotLVP1-beta, AclLVP1\_beta1,2,3, HsLVP1\_beta1,2,3, and MeLVP1\_beta. It is noteworthy that HsLVP1\_beta3 and AclLVP1\_beta1, forming a sister group, are diverged from a common ancestor along with HsLVP1\_beta1. It means the HsLVP1\_beta3





**Figure 6.** Phylogenetic tree inferred from amino acid sequences of scorpion LVP1-alpha and LVP1-beta. Results from the neighbour-joining bootstrap analyses were mapped. Monophyletic groups A and B are shown in light green and light brown, respectively.

isoform deviated from a common ancestor with AcLVP1\_beta1 and is at a further evolutionary distance than the HsLVP1\_beta1 isoform.

Like the MeLVP1\_alpha1, MeLVP1\_beta1 (or meuPep28) clustered in a sister group with a protein from *M. martensii*, BmLVP1-beta. This group diverged from the same ancestor that gave rise to BotLVP1-beta from *B. occitanus*. AcLVP1\_beta2 and AcLVP1\_beta3 are also clustered in a sister group, whose ancestor together with the ancestor of MeLVP1\_beta1, BmLVP1-beta, and BotLVP1-beta have diverged from the same ancestor.

## Discussion

Scorpion venom is a rich source of biologically active peptides and proteins that some of which have pharmaceutical properties. So, the identification of venom molecules can be extremely beneficial. In the current study, we aimed to find different isoforms of subunits of a medicinal potent scorpion protein, LVP1, in four medically



important scorpion species, *H. saulcyi*, *M. eupeus*, *A. crassicauda*, and *H. lepturus*, using the RNA-sequencing approach. Analysis of the obtained transcriptome revealed that there are some isoforms of alpha and beta subunits of LVP1 in the venom gland of *A. crassicauda* and *M. eupeus*. In contrast, only the beta isoform is expressed in the venom gland of *H. saulcyi*, and none of the subunits of LVP1 were found in the venom gland of *H. lepturus*.

*H. saulcyi*, *M. eupeus*, and *A. crassicauda* species belong to the Buthidea family, while *H. lepturus* is a member of the Hemiscorpiidae family<sup>34</sup>. So the absence of LVP1 in the venom gland of scorpion *H. lepturus* may be attributed to interspecifically variation in gene expression and venom components, related to differences in evolutionary ancestry of *H. lepturus* compared to the other scorpion species examined in this study. However, it must be pointed out that such variation in the venomous animals also occur intraspecifically<sup>35</sup>. Some studies considered these differences to be due to natural selection for the optimization of venom to diet<sup>36,37</sup>. Even low-expression toxins with significant consequences have been found<sup>38</sup>. Some other studies have explored the effect of climate variation<sup>39</sup> and geographical differences in intraspecific variation in venom composition<sup>1</sup>.

All identified proteins here presented 50% or higher sequence identity with the referral proteins (BmLVP1-alpha, BotLVP1-alpha, BmLVP1-beta, BotLVP1-beta). It is believed that when two proteins share sequence identity, it indicates similar structure and function. Lower levels of sequence similarity between protein sequences may indicate some relationship between their structures and functions<sup>40</sup>. The threshold of sequence similarity sufficient for structural homology depends strongly on the length of the alignment. As a rule of thumb, for protein sequences longer than 80 amino acids, > 24.8% sequence identity shows similar structure and function<sup>41</sup>. However, some structure alignments have interestingly identified homologous protein pairs with less than 10% sequence identity<sup>42</sup>. The length of all identified proteins here is more than 82 amino acids. So, the calculated sequence identity (> 50%) can verify the similarity of the structure and function of referral proteins with the corresponding identified proteins.

Among the identified proteins, MeLVP1\_alpha1 and MeLVP1\_beta1 have the highest sequence similarity with BmLVP1 and BotLVP1, i.e., MeLVP1\_alpha 88.66% similar to BmLVP1-alpha and MeLVP1\_beta1 75.53% similar to BmLVP1\_beta. However, Intron–exon pattern analysis of LVP1 subunit genes led to the perdition of two exons and introns with equal size in all identified proteins.

In addition to the fact that the alignment of the proteins showed a significant similarity between BmLVP1-alpha, BotLVP1-alpha, AcLVP1\_alpha1, AcLVP1\_alpha2, and MeLVP1\_alpha1, the phylogenetic analysis has also determined a common ancestor for them. The same story happened for LVP1\_beta isoforms, i.e., BmLVP1-beta and BotLVP1-beta have significant amino acids identity and common ancestor with AcLVP1\_beta1, AcLVP1\_beta2, AcLVP1\_beta3, MeLVP1\_beta1, HsLVP1\_beta1, HsLVP1\_beta2, and HsLVP1\_beta3.

The molecular weight, ranging from 7288 g/mol to 9894 g/mol, was calculated for the alpha and beta subunits of the identified LVP1s. In the previous studies, the molecular weight of 8877 g/mol was obtained for the alpha subunit of BotLVP1, and 8807 g/mol and 8132 g/mol were determined for the beta subunits of BotLVP1 and BmLVP1, respectively<sup>16,19</sup>.

Scorpion venom peptides have revealed different functions, but a common, typical folding, CSaβ, has been found in the structure of some of them. These venom peptides were structurally categorized in a group as CSaβ superfamily<sup>43,44</sup>. All here identified isoforms of LVP1-alpha and LVP1-beta contain also determinants of CSaβ folding. However, they have differences in the numbers of the strands of the beta-sheet. AcLVP1\_alpha1 and MeLVP1\_alpha1 consist of an antiparallel double-stranded beta-sheet, while the others contain a triple-stranded beta-sheet.

Members of the CSaβ superfamily contain six or eight cysteine residues, which, by forming disulfide bridges, stabilize the corresponding protein structure. Three disulfide bridges are a conserved feature in this family, but the position of the fourth disulfide bridge varies among the different members of this family. Considering the number of disulfide bridges and the location of the fourth one, CSaβ proteins are categorized into six groups (Fig. 5)<sup>22</sup>. AcLVP1\_alpha1, 2 with eight cysteines are categorized in group I, while AcLVP1\_beta1, 2 with six cysteine residues are classified in group VI. MeLVP1\_alpha1, MeLVP1\_beta1, AcLVP1\_beta3, and HsLVP1\_beta1, 2, 3 all encompassed seven cysteine residues and are assorted in group V. From these three groups, just group V has the potential of forming the dimer because of having an extra cysteine, which is not in bond with any other cysteine. So, MeLVP1\_alpha1, MeLVP1\_beta1, HsLVP1\_beta1, 2, 3 can potentially be a dimer.

As described above, LVP1 exists in two forms: heterodimer (containing alpha and beta subunits) and homodimer (containing two beta subunits). LVP1 in the heterodimer form was a simulator of lipolysis in mouse adipocytes<sup>20,21</sup>, and the homodimer form of LVP1-beta inhibited the activity of HMG-CoA reductase<sup>16</sup>. Both BotLVP1 and BmLVP1, along with MeLVP1\_alpha1, MeLVP1\_beta1, AcLVP1\_beta3, and HsLVP1\_beta1, 2, 3 took in the V category, in which six of cysteine residues involved in the formation of three intermolecular disulfide bridges and the extra cysteine take into an intramolecular disulfide bridge to form a dimer molecule, either a homodimer or a heterodimer<sup>16,19,22</sup>. So, MeLVP1\_alpha1, MeLVP1\_beta1, AcLVP1\_beta3, and HsLVP1\_beta1, 2, 3 can potentially be a dimer.

There is some evidence that RNA editing is responsible for altering the fourth cysteine residue. RNA editing involves modifying transcripts or inserting and deleting specific sequences to produce alternative protein products with different functions<sup>45</sup>. Similarly, Zhu and Gao interpreted the functional switch of BmLVP1 from adipocyte lipolysis to neurotoxicity by altering the disulfide bridge pattern of the peptides<sup>17</sup>. It seems such a process is responsible for the different patterns seen in the structure of the identified peptides.

## Conclusion

Some studies have identified LVP1s from the scorpion venom as a lipolysis inducer and suggested them be used as new candidates for hyperlipidemia treatment. Due to the pharmacological potential of LVP1, this study provided relevant findings to some new isoforms related to subunits of this protein from three scorpion species.

Here, we reported the information related to the transcriptome of *H. saulcyi*, *A. crassicauda*, and *M. eupeus*, with a focus on the identification and characterization of gene and protein of LVP1, which can provide insight into the identification of the similar proteins in the insects. According to similarity, genes encoding isoforms of LVP1s in scorpions are classified into two groups: alpha and beta. We found both subunits in the examined scorpion species, except for *H. lepturus*. The structure of all identified proteins was determined. Understanding the structures of proteins could give us hints about the family of a protein and how a protein works, which can allow us to create hypotheses about how to use them in pharmacological fields or modify them to make more beneficial proteins. The difference in the number of cysteine residues in the identified proteins has caused these proteins to be placed in different groups with different tonalities in dimer formation. AcLVP1\_alpha1, AcLVP1\_alpha2, AcLVP1\_beta1, and AcLVP1\_beta2 cannot form a dimer, while MeLVP1\_alpha1, AcLVP1\_beta3, HsLVP1\_beta1, HsLVP1\_beta2, HsLVP1\_beta3, and MeLVP1\_beta1 can form a dimer through a free cysteine residue. In addition, variability in amino acid composition in the identified proteins and similar proteins in other species, can also lead to changes in the structure and eventually changes in the function (ability and adaptability) of the proteins. Taken together, physico-chemical characterization and protein structure determination of LVP1s, along with phylogenetic analysis, may lay the foundation and shed light on assessing the potential of the scorpion-derived LVP1s for developing novel functional medicine against hyperlipidemia.

## Data availability

The mRNA and protein sequences reported in this paper are available in the NCBI under the accession numbers OP609899, OP609900, OP609901, OP609902, OP609903, OP612332, OP612333, OP612334, KU513844, and KU513847.

Received: 11 June 2023; Accepted: 9 December 2023

Published online: 14 December 2023

## References

- Safdarian, M., Vazirianzadeh, B., Ghorbani, A., Pashmforoosh, N. & Baradaran, M. Intraspecific differences in *Androctonus crassicauda* venom and envenomation symptoms. *EXCLI J.* **21**, 1222–1230 (2022).
- Saghafipour, A., Jesri, N., Noroozi, M. & Mostafavi, R. Epidemiologic status of scorpion sting in Qom, Iran 2004–2013. *Arch. Hyg. Sci.* **6**(2), 145–151 (2017).
- Dehghani, R., Kamiabi, F. & Mohammadi, M. Scorpionism by *Hemiscorpius* spp. in Iran: A review. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **24**, 8 (2018).
- Mackessy, S. P. Venom production and secretion in reptiles. *J. Exp. Biol.* **225**(7), 66 (2022).
- Jalali, A. *et al.* Molecular characterization of a cDNA encoding of an anionic cysteine-free antimicrobial peptide from the Iranian Scorpion *Odontobuthus Doriae* Venom Glands. *Pharm. Biomed. Res.* **8**(3), 199–204 (2022).
- Simone, Y. & van der Meijden, A. Armed stem to stinger: A review of the ecological roles of scorpion weapons. *J. Venom. Anim. Toxins Include. Trop. Dis.* **27**, e20210002 (2021).
- Suhas, R. Structure, function and mechanistic aspects of scorpion venom peptides—A boon for the development of novel therapeutics. *Eur. J. Med. Chem. Rep.* **6**, 100068 (2022).
- Khodayar, M. J., Mahdavinia, M., Baradaran, M. & Jalali, A. Explanation of structure and function of kv1.3 potent blocker from *Mesobuthus eupeus* venom gland: A new promise in drug development. *Jundishapur J. Nat. Pharm. Prod.* **17**(3), e120271 (2022).
- Ding, J., Chua, P. J., Bay, B. H. & Gopalakrishnakone, P. Scorpion venoms as a potential source of novel cancer therapeutic compounds. *Exp. Biol. Med.* **239**(4), 387–393 (2014).
- Kampo, S. *et al.* Scorpion venom analgesic peptide, BmK AGAP inhibits stemness, and epithelial-mesenchymal transition by down-regulating PTX3 in breast cancer. *Front. Oncol.* **9**, 21 (2019).
- Harrison, P. L., Abdel-Rahman, M. A., Miller, K. & Strong, P. N. Antimicrobial peptides from scorpion venoms. *Toxicon* **88**, 115–137 (2014).
- Uzair, B. *et al.* Scorpion venom peptides as a potential source for human drug candidates. *Protein Pept. Lett.* **25**(7), 702–708 (2018).
- Hu, Y. *et al.* Scorpion venom peptide HsTx2 suppressed PTZ-induced seizures in mice via the circ\_0001293/miR-8114/TGF- $\beta$ 2 axis. *J. Neuroinflamm.* **19**(1), 284 (2022).
- Wu, X. F. *et al.* Scorpion venom heat-resistant peptide attenuates microglia activation and neuroinflammation. *Front. Pharmacol.* **12**, 704715 (2021).
- Murthy, R. K. & Medh, J. D. Increase in serum free fatty acids, phospholipids and reduction in total cholesterol in acute myocarditis produced by scorpion (*Buthus tamulus*) venom injection in dogs. *Indian Heart J.* **38**(5), 369–372 (1986).
- Chai, S. C., Armugam, A., Strong, P. N. & Jeyaseelan, K. Characterization of bumarsin, a 3-hydroxy-3-methylglutaryl-coenzyme reductase inhibitor from *Mesobuthus martensii* Karsch venom. *Toxicon* **60**(3), 272–279 (2012).
- Zhu, S. & Gao, B. Molecular characterization of a new scorpion venom lipolysis activating peptide: Evidence for disulfide bridge-mediated functional switch of peptides. *FEBS Lett.* **580**(30), 6825–6836 (2006).
- Ruiming, Z. *et al.* Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. *BMC Genom.* **11**, 452 (2010).
- Soudani, N. *et al.* Isolation and molecular characterization of LVP1 lipolysis activating peptide from scorpion *Buthus occitanus tunetanus*. *Biochimica et Biophysica Acta* **1747**(1), 47–56 (2005).
- Soudani, N. *et al.* Isolation and molecular characterization of LVP1 lipolysis activating peptide from scorpion *Buthus occitanus tunetanus*. *Biochimica et Biophysica Acta BBA Proteins Proteom.* **1747**(1), 47–56 (2005).
- Drira-Chaabane, S., el Ayeb, M., Torresani, J. & Gharbi-Chihi, J. Lipolytic action of *Buthus occitanus tunetanus* venom: Involvement of the beta adrenergic pathway. *Biochem. Biophys. Res. Commun.* **226**(1), 280–286 (1996).
- Zhu, S. & Gao, B. Molecular characterization of a new scorpion venom lipolysis activating peptide: Evidence for disulfide bridge-mediated functional switch of peptides. *FEBS Lett.* **580**(30), 6825–6836 (2006).
- Salabi, F. & Jafari, H. New insights about scorpion venom hyaluronidase; isoforms, expression and phylogeny. *Toxin Rev.* **66**, 1–16 (2022).
- Baradaran, M., Jalali, A., Naderi-Soorki, M., Jokar, M. & Galehdari, H. First transcriptome analysis of Iranian scorpion, *Mesobuthus eupeus* venom gland. *Iran. J. Pharm. Res.* **17**(4), 1488–1502 (2018).
- Barahoei, H., Navidpour, S., Aliabadian, M., Siahsharvie, R. & Mirshamsi, O. Scorpions of Iran (Arachnida: Scorpiones): Annotated checklist, DELTA database and identification key. *J. Insect Biodivers. Syst.* **6**(4), 375–474 (2020).
- Soni, P. *et al.* Transcriptome analysis identified coordinated control of key pathways regulating cellular physiology and metabolism upon *Aspergillus flavus* infection resulting in reduced aflatoxin production in groundnut. *J. Fungi* **6**(4), 66 (2020).

27. Santibáñez-López, C. E., Cid-Urbe, J. I., Batista, C. V., Ortiz, E. & Possani, L. D. Venom gland transcriptomic and proteomic analyses of the enigmatic scorpion *Superstitionia donensis* (Scorpiones: Superstitioniidae), with Insights on the evolution of its venom components. *Toxins* **8**(12), 66 (2016).
28. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**(7), 644–652 (2011).
29. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protocols* **8**(8), 1494–1512 (2013).
30. Salabi, F. & Jafari, H. Differential venom gland gene expression analysis of juvenile and adult scorpions *Androctonus crassicauda*. *BMC Genomics* **23**(1), 636 (2022).
31. Gupta, C. L. *et al.* Computational modeling and validation studies of 3-D structure of neuraminidase protein of H1N1 influenza A virus and subsequent in silico elucidation of piceid analogues as its potent inhibitors. *EXCLI J.* **12**, 215–225 (2013).
32. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis Version 11. *Mol. Biol. Evol.* **38**(7), 3022–3027 (2021).
33. Zhu, S. *et al.* Experimental conversion of a defensin into a neurotoxin: Implications for origin of toxic function. *Mol. Biol. Evol.* **31**(3), 546–559 (2014).
34. Dehghani, R. & Fathi, B. Scorpion sting in Iran: A review. *Toxicon* **60**(5), 919–933 (2012).
35. Casewell, N. R., Jackson, T. N. W., Laustsen, A. H. & Sunagar, K. Causes and consequences of snake venom variation. *Trends Pharm Sci.* **41**(8), 570–581 (2020).
36. Casewell, N. R., Wüster, W., Vonk, F. J., Harrison, R. A. & Fry, B. G. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol. Evol.* **28**(4), 219–229 (2013).
37. Smiley-Walters, S. A., Farrell, T. M. & Gibbs, H. L. Evaluating local adaptation of a complex phenotype: Reciprocal tests of pigmy rattlesnake venoms on treefrog prey. *Oecologia* **184**(4), 739–748 (2017).
38. Margres, M. J. *et al.* Quantity, not quality: Rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. *Mol. Biol. Evol.* **34**(12), 3099–3110 (2017).
39. Ghorbani, A., Mansouri, B. & Baradaran, M. Effects of climate variables on the incidence of scorpion stings in Iran for five years. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **27**, e20200110 (2021).
40. Koehl, P. & Levitt, M. Sequence variations within protein families are linearly related to structural variations. *J. Mol. Biol.* **323**(3), 551–562 (2002).
41. Sander, C. & Schneider, R. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**(1), 56–68 (1991).
42. Rost, B. Protein structures sustain evolutionary drift. *Fold. Des.* **2**(3), S19–24 (1997).
43. Zhu, S., Gao, B. & Tytgat, J. Phylogenetic distribution, functional epitopes and evolution of the CSalphabeta superfamily. *Cell. Mol. Life Sci.* **62**(19–20), 2257–2269 (2005).
44. Park, B. G. *et al.* Solution structure and functional analysis of HelaTx1: The first toxin member of the  $\kappa$ -KTx5 subfamily. *BMB Rep.* **53**(5), 260–265 (2020).
45. Hoopengardner, B., Bhalla, T., Staber, C. & Reenan, R. Nervous system targets of RNA editing identified by comparative genomics. *Science* **301**(5634), 832–836 (2003).

## Acknowledgements

We wish to thank Ahvaz Jundishapur University of Medical Sciences (Grant Number: TRC-0052 and TRC-9819) and Razi Vaccine and Serum Research Institute of Iran for their financial support and our colleagues for many stimulating discussions.

## Author contributions

Conceived and designed the experiments: M.B., F.S. and B.V. collect and identify the scorpion specimens; M.B. and F.S. Analyzed the data; M.B. wrote the manuscript; All authors have read and approved the manuscript.

## Funding

Financial support for the all authors was provided by Ahvaz Jundishapur University of Medical Sciences (Grant Number: TRC-0052 and TRC-9819) and Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran (960354-020-18-83-2 and 011062-062-1851-83-3).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023